ROUTLEDGE

Public Policy Analysis

EDITION NO. 05

WILLIAM N. DUNN

Public Policy Analysis

Fifth Edition

WILLIAM N. DUNN

Graduate School of Public and International Affairs
University of Pittsburgh

Routledge
Taylor & Francis Group
LONDON AND NEW YORK

First published 2012, 2008, 2004, 1994 by Pearson Education, Inc.

Published 2016 by Routledge
2 Park Square, Milton Park, Abingdon, Oxon OX14 4RN
711 Third Avenue, New York, NY, 10017, USA

Routledge is an imprint of the Taylor & Francis Group, an informa business

ISBN: 9780205252572 (pbk)

Cover Designer: Suzanne Duda

Library of Congress Cataloging-in-Publication Data
Dunn, William N.
Public policy analysis : an introduction / William N. Dunn.—5th ed.
p. cm.
ISBN-13: 978-0-205-25257-2
ISBN-10: 0-205-25257-5
1. Policy sciences. 2. Political planning—Evaluation. I. Title.
H61.D882 2011
320.6—dc23

2011028590

In memory of Donald T. Campbell

This page intentionally left blank

BRIEF CONTENTS

BOX CONTENTS

USEFUL CONCEPTS, GUIDES, AND CHECKLISTS

CASE STUDY CONTENTS

LESSONS FROM PRACTICE

VISUAL DISPLAY CONTENTS

FIGURES FOR ORGANIZING THINKING

TABLES FOR ORGANIZING THINKING

CONTENTS

PREFACE

My aim in writing the several editions of this book has been to produce a critical synthesis of the field, while at the same time offering students, instructors, and practitioners a body of knowledge and skills that is applicable to real-world problems. This is not a text in the narrow and conventional sense of the term.

Ever since the publication of the first edition of *Public Policy Analysis* more than thirty years ago, I have become more and more convinced that the methodology of policy analysis rests or should rest on epistemological foundations that differ from those of the disciplines of which policy analysis is composed. For this reason, I continue to define policy analysis as an applied social science discipline that employs multiple methods of inquiry to solve practical problems.

What this means is that the methodology of policy analysis cannot be reduced to the theories and analytical routines of microeconomics, because solutions for practical problems demand much more than the analysis of rational choice, expected utility, and opportunity costs. By the same token, the methodology of policy analysis cannot be reduced to the study of politics, because solutions for practical problems require more than the analysis of power, rule, and authority or who gets what, when, and how. Much more is involved. Finally, because a principal aim of policy analysis is to improve policies, the methodology of policy analysis cannot be reduced to an academic spectator sport in which knowledge is prized for its own sake.

NEW TO THIS EDITION

This broadly accessible 5th edition employs a simplified style of writing, cases based on real-world analytical practices, and visual displays that make complex ideas understandable. Examples and case materials now include issues in foreign policy and international security as well as domestic issues including environmental justice, urban economics, transportation, and public safety. The inclusion of advanced graphics software for mapping and evaluating policy arguments helps cultivate critical thinking skills in areas ranging from qualitative forecasting and statistical analysis to theories of justice. The book also provides students with practical and marketable skills in communicating policy analysis by writing policy memos, position papers, and other forms of structured analytical writing. Finally, new study suggestions and demonstration exercises emphasize active rather than passive learning.

Special instructional devices and learning strategies are again employed throughout the book:

- *Advance organizers.* The book uses advance organizers, especially visual displays, to introduce students to the logic and organization of methods. The advance organizer for the book as a whole is the information-processing model of policy analysis presented in the first chapter.

- *Learning objectives.* At the beginning of each chapter is an explicit statement of the learning objectives that students should be able to attain by reading the chapter and completing its study questions and demonstration exercises. I have tried to state these objectives in terms of the acquisition of knowledge and skills to do or perform something—that is, behaviorally defined learning objectives that involve recognizing, defining, understanding, explaining, predicting, evaluating, and applying. By stating objectives in this way, the emphasis is on active rather than passive learning, on application rather than regurgitation.
- *Review questions.* Knowledge and skills must be reinforced. For this reason, review questions are provided at the end of each chapter. The review questions address *higher-order* knowledge and skills (e.g., explaining or applying) as well as *lower-order* knowledge and skills (e.g., calculating or estimating). Review questions may be used by students for self-study and by instructors who are developing written assignments, examinations, and tests.
- *Demonstration exercises.* Knowledge and skills are not acquired or retained without frequent opportunities for application to real-world problems. For this reason, each chapter contains opportunities to demonstrate the application of knowledge and skills to significant practical problems. The attempt is to draw students away from "blackboard policy analysis" into the real world of messy problems and provisional solutions.
- *Cases.* Cases in policy analysis are the focus of the demonstration exercises. Cases span a number of issue areas including foreign policy and security, transportation policy, occupational health and safety, and urban policy. Some cases are primarily conceptual; most are methodological in nature.
- *Bibliographies.* In addition to literature cited in footnotes, each chapter is accompanied by a list of readings that are keyed to the subject matter of that chapter. I have attempted to include literature that is representative of many of the most important developments in public policy analysis.
- *Guidelines for written and oral communication.* Students who master methods almost always face difficulties when they must communicate the results of analysis through policy memoranda, position papers, issue papers, and oral briefings. To help overcome these difficulties, appendices present step-by-step guidelines and checklists.
- *Argument maps.* Analysis is about making and understanding policy arguments. This edition uses numerous argument maps created with *Rationale,* an exceptionally useful and innovative computer program available at www.austhink.com. Elsewhere I have expressed my gratitude to Professor Tim van Gelder, one of the originators of the program, for his help.

SUPPLEMENTS

WEBSITE A special website (www.policyonline.org) supports users of this book by providing slides keyed to each chapter and data sets related to cases covered in the chapters. Many of the slides can serve as teaching notes.

ARGUMENT MAPPING SOFTWARE Users of this book may purchase *Rationale* at a student and instructor discount by visiting www.austhinkconsulting.com/dunn.

ACKNOWLEDGEMENTS

I would not have begun this ambitious multidisciplinary project without the encouragement of the late Paul F. Lazarsfeld, who challenged me to investigate and write about what he called the "policy sciences movement." Lazarsfeld, one of a handful of premier applied social scientists of the twentieth century, was skeptical of the enterprise, as it had been sketched by Lasswell, Kaplan, Lerner, and others. Its aims seemed to him unwisely all-encompassing and grand, an assessment that holds some truth today. Lazarsfeld did not self-identify as a sociologist, but as an applied social scientist and, for this reason, he was University Professor of Social Science at the University of Pittsburgh.

Some ten years later, we made an ultimately unsuccessful attempt to fill Lazarsfeld's vacant chair with another premier applied social scientist, Donald T. Campbell, who virtually revolutionized the methodology of the applied social sciences in the twentieth century. Campbell, like Lazarsfeld, did not self-identify with his original discipline, which was social psychology. Instead, he viewed himself as a multidisciplinary applied social scientist specializing in program and policy evaluation and the philosophy and sociology of science. Campbell's mentorship has had a profound effect on the way I think about the strengths and limitations of policy analysis and program evaluation. His imprint can be seen throughout this book.

At about the same time, I was asked to join a team of faculty who were developing curricular materials on policy analysis for practitioners in local governments. The group had wisely contracted specialists in learning theory and curriculum development, including Doris Gow and Jyotsna Vasudev. I learned from them the important pedagogical lesson that abstract subjects such as policy analysis can be more effectively taught by focusing on behaviorally defined learning objectives. I am grateful to them for making me see that much of the literature we assign in courses is not easily or successfully tied to learning outcomes, which means that we often cannot say why we want students to read the materials we assign. This was a revelation.

I have been fortunate to meet and work with colleagues who changed my mind about many things, including the important role that the philosophy and sociology of science play in the applied social sciences. These colleagues include Ian I. Mitroff, Burkart Holzner, Gerald Zaltman, and my former student and now professor, Bahman Fozouni, who introduced me to yet other colleagues in the Center for History and Philosophy of Science at the University of Pittsburgh. I learned much about the pragmatic importance of policy analysis and the applied social sciences from Robert F. Rich, Thomas D. Cook, and Carol H. Weiss, all of whom were and are committed to investigating and deliberately changing the conditions under which the social sciences may be used to solve practical problems.

Faculty and students in the Graduate School of Public and International Affairs (GSPIA), University of Pittsburgh, have challenged my thinking, writing, and teaching. They include Alex Weilenman, Louise Comfort, Tom Pavlak, John Mendeloff, Hector Correa, Michael Sabath, Soumana Sako, Sam Overman, Tony Cahill, Kevin Kearns, Dave Miller, Mary Jo Dukes, Ralph Bangs, Jan Jernigan, and Andrea Hegedus, all of whom were affiliated with the doctoral program in Public Policy Research and Analysis at GSPIA. Most recently, two new colleagues have

been important to me in writing this edition. The first is my new junior colleague, Ilia Murtazashvili, whose critiques of many of the ideas in this book have helped me question and improve my thinking and writing. I am also indebted to Tim van Gelder of the University of Melbourne and Austhink Consulting for introducing me to *Rationale*, a brilliant new computer program that enables users to map policy arguments, probe the assumptions underlying policy claims, and practice critical thinking. Argument maps created with *Rationale* appear throughout this edition of *Public Policy Analysis*.

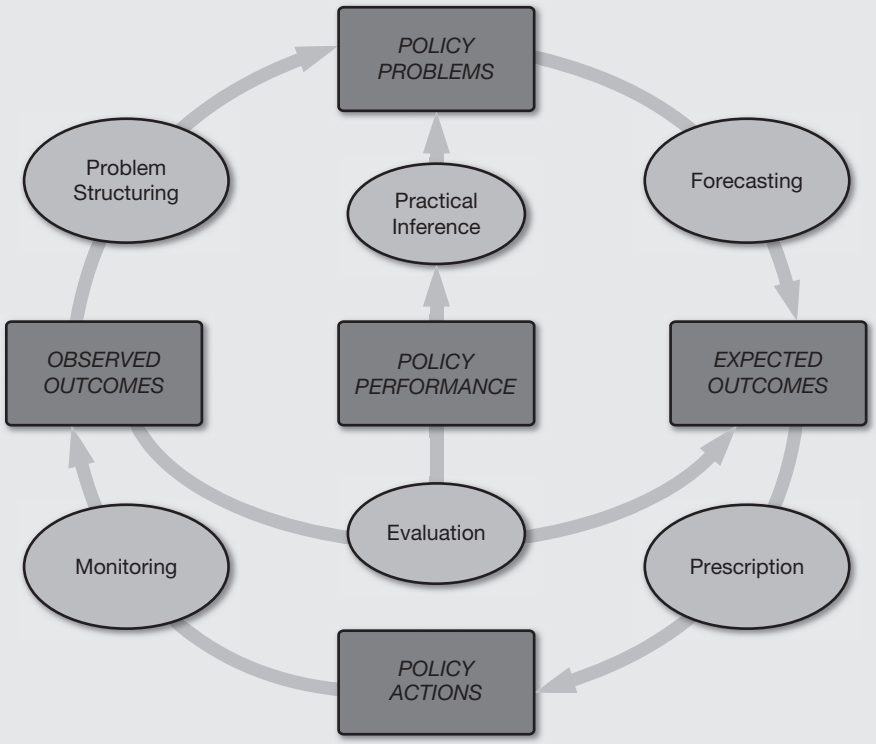
In the past thirty years and more, this book has been used in degree and certificate programs in universities, think tanks, and governments in this country and abroad. Translations into Arabic, Chinese, Indonesian, Korean, Macedonian, Romanian, and Spanish are completed or under way. The book has been used in training programs and projects in countries of the European Union, Southeastern Europe, the Middle East, North Africa, and Latin America. The revisions incorporated in this edition reflect much of what I have learned from the participants and organizers of these programs.

I am also deeply grateful to students and faculty at the Graduate Center for Public Policy and Management in Skopje, Macedonia, a great but now inactive institution with some of the best students I have taught. In addition, my friends and colleagues at the Institute for Environmental Studies at the Free University of Amsterdam, particularly Matthijs Hisschemoeller and Eefje Cuppen, have been a constant source of ideas about participative methods of policy analysis. I also want to thank reviewers of this and previous editions. In addition to anonymous reviewers, they include David Nice of Washington State University, David Houston of the University of Tennessee at Knoxville, and Louise Comfort of the University of Pittsburgh. Sheila Kelly, Lien Rung-Kao, Sujatha Raman, Ranjan Chaudhury, Eric Sevigny, Kate Freed, Bojana Aceva-Andonova, Erin McGrath, Alla Golovina Khadka, and Jessica Reyes assisted in preparing materials for this and previous editions. Finally, I am grateful for the splendid editorial assistance of Integra Software Services Pvt Ltd, in Pondicherry, India, and to Stephanie Chaisson and other editors at Pearson. Special mention goes to Toni Magyar, political science editor at Pearson, who provided masterful editorial support and exceptionally creative guidance in helping me with this fifth edition.

William N. Dunn
Graduate School of Public and International Affairs
University of Pittsburgh

PART I

Methodology of Policy Analysis



CHAPTER 1

The Process of Policy Analysis

OBJECTIVES

By studying this chapter, you should be able to

- Define and illustrate phases of policy analysis.
- Describe elements of integrated policy analysis.
- Distinguish four strategies of policy analysis.
- Contrast reconstructed logic and logic-in-use.
- Distinguish prospective and retrospective policy analysis.
- Describe the structure of a policy argument and its elements.
- Understand the role of argument mapping in critical thinking.
- Interpret scorecards, spreadsheets, influence diagrams, decision trees, and argument maps.

Policy analysis is a process of multidisciplinary inquiry aiming at the creation, critical assessment, and communication of policy-relevant information. As a problem-solving discipline, it draws on social science methods, theories, and substantive findings to solve practical problems.[1]

[1]For a sample of alternative definitions see Harold D. Lasswell, *A Pre-view of Policy Sciences* (New York: American Elsevier Publishing, 1971); Yehezkel Dror, *Ventures in Policy Sciences: Concepts and Applications* (New York: American Elsevier Publishing, 1971); Edward S. Quade, *Analysis for Public Decisions,* 3d rev. ed., ed. Grace M. Carter (New York: North Holland Publishing, 1989); David L. Weimer and Aidan R. Vining, *Policy Analysis: Concepts and Practice,* 2d ed. (Englewood Cliffs, NJ: Prentice Hall, Inc., 1992); Duncan Mac Rae Jr., *The Social Function of Social Science* (New Haven, CT: Yale University Press, 1976).

METHODOLOGY OF POLICY ANALYSIS

As used here, the word *methodology* refers to a process of reasoned inquiry aimed at finding solutions to practical problems. The aim of methodology is to help us understand not only the products of policy inquiry but also the processes employed to create these products.[2] The methodology of policy analysis is not confined to the analytical routines of specialized social science fields—for example, benefit-cost analysis in economics or implementation analysis in political science—because none of these holds a privileged place in policy inquiry. Nor is the methodology of policy analysis constrained by the doctrines and principles of obsolescent philosophies of science such as logical positivism, which mistakenly claimed that scientific knowledge, properly understood, is objective, value free, and quantitative.[3] On the contrary, policy analysis is methodologically eclectic; its practitioners are free to choose among a wide range of scientific methods, qualitative as well as quantitative, as long as these yield reliable knowledge. In this context, policy analysis includes art, craft, and reasoned persuasion, all of which are scientific to the extent that they succeed in producing reliable knowledge.[4] Ordinary commonsense knowing and well-winnowed practical wisdom—both products of evolutionary learning across generations of

---

[2]Abraham Kaplan, *The Conduct of Inquiry: Methodology for Behavioral Science* (San Francisco, CA: Chandler Publishing Company, 1964), pp. 23–24.

[3]Logical positivism (or logical empiricism) was abandoned by most philosophers of science more than 50 years ago, although its epistemological pillars—the correspondence theory of truth, the empirical criterion of meaning, and quantificationism—are still venerated by many social scientists. For alternatives to logical positivism in economics and political science see Daniel Bromley, *Sufficient Reason: Volitional Pragmatism and the Meaning of Economic Institutions* (Princeton, NJ: Princeton University Press, 2006); Henry E. Brady and David Collier, eds. *Rethinking Social Inquiry: Diverse Tools, Shared Standards* (Lanham, MD: Rowman Littlefield, 2004); Deirdre N. McCloskey, *The Rhetoric of Economics*, 2nd ed., Madison: University of Wisconsin Press, 1998; Stephen Thomas Ziliak and Deirdre N. McCloskey, *The Cult of Statistical Significance: How the Standard Error Costs Us Jobs, Justice, and Lives*. Ann Arbor: University of Michigan Press, 2008. Paul Diesing, *How Does Social Science Work? Reflections on Practice* (Pittsburgh, PA: University of Pittsburgh Press, 1991); and Mary Hawkesworth, *Theoretical Issues in Policy Analysis* (Albany: State University of New York Press, 1988).

[4]Larry Laudan has argued that the demarcation between science and non-science, including art and craft, is a pseudo-problem that should be replaced by focusing on the distinction between reliable and unreliable knowledge. It is not necessary to ask whether knowledge is "scientific," only whether it is reliable. "The Demise of the Demarcation Problem," in R.S. Cohen and L. Laudan, *Physics, Philosophy and Psychoanalysis: Essays in Honor of Adolf Grünbaum*. Boston Studies in the Philosophy of Science, Vol.76 (Dordrecht: D. Reidel, 1983), pp. 111–127. Aaron Wildavsky and others have used the terms *art* and *craft* to characterize policy analysis. See Aaron Wildavsky, *Speaking Truth to Power: The Art and Craft of Policy Analysis* (Boston, MA: Little Brown, 1979); and Iris Geva-May and Aaron Wildavsky, *An Operational Approach to Policy Analysis: The Craft, Prescriptions for Better Analysis* (Boston, MA: Kluwer, 1997). The term *policy science(s)* is Harold Lasswell's. See the short methodological history of the policy sciences in Ronald Brunner, "The Policy Movement as a Policy Problem," in *Advances in Policy Studies since 1950,* vol. 10, *Policy Studies Review Annual,* ed. W. N. Dunn and R. M. Kelly (New Brunswick, NJ: Transaction Books, 1992), pp. 155–97 and contributions to Michael Moran, Martin Rein, and Robert E. Goodin, eds. *The Oxford Handbook of Public Policy* (Oxford: Oxford University Press, 2006).

problem solvers—often permit conclusions that are more trustworthy and reliable than those produced by means of policy analysis and other specialized forms of professional and scientific inquiry.[5]

The rationale for policy analysis is pragmatic. For this reason, it is unmistakably different from social science disciplines that prize knowledge for its own sake. The policy-relevance of these disciplines depends not on their status as sciences but on the extent to which they are successful in illuminating and alleviating practical problems, problems that come in complex bundles that are at once economic, political, cultural, ethical, and more. Practical problems do not arrive in separate disciplinary packages addressed to departments of economics and political science—to name two of the most important policy disciplines. In today's world, multidisciplinary policy analysis seems to provide the best fit with the manifold complexity of public policy making.

POLICY ANALYSIS—A MULTIDISCIPLINARY FRAMEWORK

Policy analysis is partly *descriptive*. It relies on traditional social science disciplines to describe and explain the causes and consequences of policies. But it is also *normative*, a term that refers to value judgments about what *ought* to be, in contrast to descriptive statements about what *is*.[6] To investigate problems of efficiency and fairness, policy analysis draws on normative economics and decision analysis as well as ethics and other branches of social and political philosophy—all of which are about what *ought* to be. This normative orientation stems from the fact that analyzing policies demands that we choose among desired consequences (ends) and preferred courses of action (means). The choice of ends and means requires continuing trade-offs among competing values of efficiency, equity, security, liberty, and democracy.[7] The importance of normative reasoning in policy analysis was well stated by a former undersecretary in the Department of Housing and Urban Development: "Our problem is not to do what is right. Our problem is to know what is right."[8]

---

[5]On the contrasts between scientific and professional knowledge on one hand, and ordinary commonsense knowing on the other, see Charles E. Lindblom and David K. Cohen, *Usable Knowledge: Social Science and Social Problem Solving* (New Haven, CT: Yale University Press, 1979). On the frequent soundness of evolved practical knowledge—but the periodic need for supplemental scientific testing—see Donald T. Campbell, "Evolutionary Epistemology," in *Methodology and Epistemology for Social Science: Selected Papers*, ed. E. S. Overman (Chicago: University of Chicago Press, 1989).

[6]One classic statement of the difference between positive and normative knowledge in economics is Milton Friedman, *Essays in Positive Economics* (Chicago, IL: University of Chicago Press, 1953). This same positive-normative distinction is present throughout the social sciences.

[7]Deborah Stone, *Policy Paradox: The Art of Political Decision Making,* rev ed. (New York: W. W. Norton, 2001).

[8]Robert C. Wood, "Foreword" to *The Study of Policy Formation,* ed. Raymond A. Bauer and Kenneth J. Gergen (New York: Free Press, 1968), p. v. Wood is quoting President Lyndon Johnson.

Policy-Relevant Information

Policy analysis is designed to provide policy-relevant information about five types of questions:

- *Policy problems.* What is the problem for which a potential solution is sought? Is global warming a human-made consequence of aircraft and motor vehicle emissions? Or is global warming a consequence of periodic fluctuations in the temperature of the atmosphere? What alternatives are available to mitigate global warming? What are the potential outcomes of these alternatives and what is their value or utility?
- *Expected policy outcomes.* What are the expected outcomes of policies designed to reduce harmful emissions? Because periodic natural fluctuations are difficult or impossible to control, what is the likelihood that emissions can be reduced by raising the price of gasoline and diesel fuel, compared with requiring that aircraft and motor vehicles use biofuels?
- *Preferred policies.* Which policies should be chosen, considering not only their expected outcomes in reducing harmful emissions, but also the value of reduced emissions in terms of economic costs and benefits? Should distributional criteria involving environmental justice be used along with criteria of economic efficiency?
- *Observed policy outcomes.* What policy outcomes are observed, as distinguished from the outcomes expected before a preferred policy is implemented? Did the preferred policy actually result in reduced emissions? Were other factors such as political opposition to governmental regulation responsible for the limited achievement of emissions targets?
- *Policy performance.* To what extent do observed policy outcomes contribute to the reduction of global warming through emissions controls? What are the benefits and costs of government regulation to present and future generations?

Answers to these questions yield five types of information, which are *policy-informational components*. These components are shown as rectangles in Figure 1.1.[9]

A *policy problem* is an unrealized need, value, or opportunity for improvement attainable through public action.[10] Knowledge of what problem to solve requires information about a problem's antecedent conditions (e.g., school dropouts as an antecedent condition of unemployment), as well as information about values (e.g., safe schools or a living wage) whose achievement may lead to the problem's solution. Information about policy problems plays a critical role in policy analysis,

[9]The framework was originally suggested by Walter Wallace, *The Logic of Science in Sociology* (Chicago: Aldine Books, 1971). Wallace's framework addresses research methodology in sociology, whereas Figure 1.1 addresses the methodology of policy analysis.

[10]Compare James A. Anderson, *Public Policymaking: An Introduction*, 7th ed. (Boston, MA: Wadsworth, 2011); Charles O. Jones, *An Introduction to the Study of Public Policy,* 2d ed. (North Scituate, MA: Duxbury Press, 1977), p. 15; and David Dery, *Problem Definition in Policy Analysis* (Lawrence: University of Kansas Press, 1984).

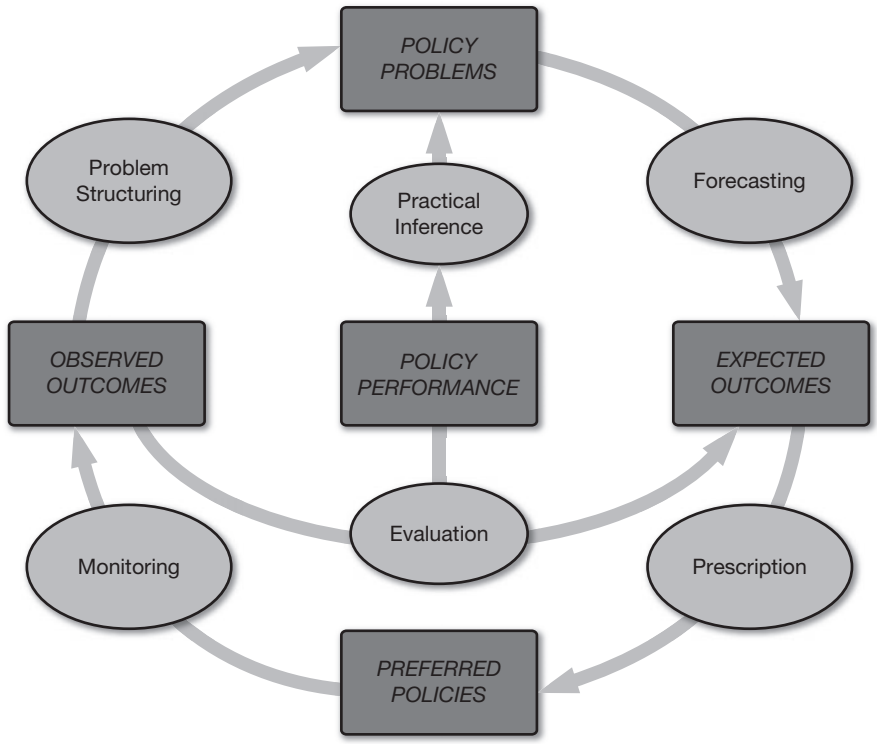


FIGURE 1.1
The process of integrated analysis

because the way a problem is defined shapes the search for available solutions. Inadequate or faulty information may result in a fatal error: defining the wrong problem.[11]

Expected policy outcomes are likely consequences of one or more policy alternatives designed to solve a problem. Information about the circumstances that gave rise to a problem is essential for producing information about expected policy outcomes. Such information is often insufficient, however, because the past does not repeat itself completely, and the values that shape behavior may change in the future. For this reason, information about expected policy outcomes is not "given" by the existing situation. To produce such information may require creativity, insight, and the use of tacit knowledge.[12]

[11]Defining the wrong problem is a type III error, as contrasted with type I and type II errors committed when the level of statistical significance (*alpha*) is set too high or too low in testing the null hypothesis. An early statement of this contrast is Ian I. Mitroff and Thomas R. Featheringham, "On Systematic Problem Solving and the Error of the Third Kind," *Behavioral Sciences* 19, no. 6 (1974): 383–93.
[12]Dror, *Ventures in Policy Sciences*; Sir Geoffrey Vickers, *The Art of Judgment: A Study of Policy Making* (New York: Basic Books, 1965); and C. West Churchman, *The Design of Inquiring Systems; Basic Concepts of Systems and Organization* (New York: Basic Books, 1971).

A *preferred policy* is a potential solution to a problem. To select a preferred policy, it is necessary to have information about expected policy outcomes as well as information about the value or utility of these expected outcomes. Another way to say this is that factual as well as value premises are required for policy prescriptions. Fact alone—for example, the fact that one policy produces more of some quantity than another—do not justify the choice of a preferred policy. Factual premises must be joined with value premises involving efficiency, equality, security, democracy, or some other value.

An *observed policy outcome* is a present or past consequence of implementing a preferred policy. It is sometimes unclear whether an outcome is actually an effect of a policy, because some effects are not *policy* outcomes; many outcomes are the result of other, extra-policy factors. It is important to recognize that the consequences of action cannot be fully stated or known in advance, which means that many consequences are neither anticipated nor intended. Fortunately, information about such consequences can be produced *ex post* (after policies have been implemented), not only *ex ante* (before policies are implemented).

Policy performance is the degree to which an observed policy outcome contributes to the solution of a problem. In practice, policy performance is never perfect. Problems are rarely "solved"; most often, problems are resolved, reformulated, and even "unsolved."[13] To know whether a problem has been solved, resolved, reformulated, or unsolved requires information about observed policy outcomes, as well as information about the extent to which these outcomes contribute to the opportunities for improvement that gave rise to a problem.

Policy-Informational Transformations

The five types of policy-relevant information are interdependent. The arrows connecting each pair of components represent *policy-informational transformations*, whereby one type of information is changed into another, so that the creation of information at any point depends on information produced in an adjacent phase. Information about policy performance, for example, depends on the transformation of prior information about observed policy outcomes. The reason for this dependence is that any assessment of how well a policy achieves its objectives assumes that we already have reliable information about the outcomes of that policy. The other types of policy-relevant information are dependent in the same way.

Information about policy problems is a special case. Information about policy problems usually includes some problem elements—for example, potential solutions or expected outcomes—and excludes others. What is included or excluded affects which policies are eventually prescribed, which values are appropriate as criteria of policy performance, and which potentially predictable outcomes warrant or do not warrant attention. At the risk of being overly repetitious, it is worth stressing again that a fatal error of policy analysis is a type III error—defining the wrong problem.[14]

---

[13]Russell L. Ackoff, "Beyond Problem Solving," *General Systems* 19 (1974): 237–39.

[14]Type I and type II errors are also known as false positives and false negatives. Other sources on type III errors include A. W. Kimball, "Errors of the Third Kind in Statistical Consulting," *Journal of the American Statistical Association* 52 (1957): 133–42; Howard Raiffa, *Decision Analysis* (Reading, MA: Addison-Wesley, 1968), p. 264; and Ian I. Mitroff, *The Subjective Side of Science* (New York: Elsevier, 1974).

Policy-Analytic Methods

The five types of policy-relevant information are produced and transformed by using policy-analytic methods. All methods involve judgments of different kinds:[15] judgments to accept or reject an explanation, to affirm or dispute the rightness of an action, to prescribe or not prescribe a policy, to accept or reject a prediction, and to formulate a problem in one way rather than another.

In policy analysis, these procedures have special names:

- *Problem structuring.* Problem-structuring methods are employed to produce information about which problem to solve. One example of problem-structuring methods is the influence diagram and decision tree presented in Case 1.3 of this chapter (The Influence Diagram and Decision Tree—Structuring Problems of Energy Policy and International Security). Other examples of problem-structuring methods include critical thinking tools such as argument mapping (Case 1.4: The Argument Map—Problem Structuring in National Defense and Energy Policy). Chapter 3 of this book covers problem-structuring methods and their application.
- *Forecasting.* Forecasting methods are used to produce information about expected policy outcomes. Although many kinds of forecasting methods are covered in Chapter 4, an example of a simple forecasting tool is the scorecard described in Case 1.1 (The Goeller Scorecard—Monitoring and Forecasting Technological Impacts). Scorecards, which are based on the judgments of experts, are particularly useful in identifying expected outcomes of science and technology policies.
- *Prescription.* Methods of prescription are employed to create information about preferred policies. An example of a prescriptive method is the spreadsheet (Case 1.2: The Spreadsheet—Evaluating the Benefits and Costs of Energy Policies). The spreadsheet goes beyond the identification of expected policy outcomes by expressing consequences in terms of monetary benefits and costs. Benefit-cost analysis and other methods of prescription are presented in Chapter 5.
- *Monitoring.* Methods of monitoring are employed to produce information about observed policy outcomes. The scorecard (Case 1.1) is a simple method for monitoring observed policy outcomes as well as for forecasting expected policy outcomes. Chapter 6 covers methods of monitoring in detail.
- *Evaluation.* Evaluation methods are used to produce information about the value or utility of observed policy outcomes and their contributions to policy performance. Although evaluation methods are covered more fully in Chapter 7, the spreadsheet (Case 1.2) may be used for evaluation as well as prescription.

The first method, problem structuring, is about the other methods. For this reason, it is a *metamethod* (method of methods). In the course of structuring a problem, analysts typically experience a "troubled, perplexed, trying situation, where the difficulty is, as

[15] John O'Shaughnessy, *Inquiry and Decision* (London: George Allen & Unwin, 1972).

it were, spread throughout the entire situation, infecting it as a whole."[16] *Problem situations* are not problems; problems are representations of problem situations. Hence, problems are not "out there" in the world, but they stem from the interaction of thought and external environments. Imagine a graph showing the growth of defense expenditures as a percentage of gross domestic product. The graph represents a problem situation, not a problem, because one analyst will see the graph as evidence of increasing national security (more of the budget is allocated to defense), while another interprets the graph as an indication of a declining budget for social welfare (less of the budget can be allocated to social services). Problem structuring, a procedure for testing different representations of a problem situation, is the central guidance system of policy analysis.

Policy-analytic methods are interdependent. It is not possible to use one method without first having used others. Thus, although it is possible to monitor past policies without forecasting their future consequences, it is usually not possible to forecast policies without first monitoring them.[17] Similarly, analysts can monitor policy outcomes without evaluating them, but it is not possible to evaluate an outcome without first establishing that it is an outcome in the first place. Finally, to select a preferred policy requires that analysts have already monitored, evaluated, and forecasted outcomes.[18] This is yet one more way of saying that policy prescription is based on factual as well as value premises.

Figure 1.1 supplied a framework for integrating methods from different policy-relevant disciplines. Some methods are used solely or primarily in some disciplines, and not others. Program evaluation, for example, employs monitoring to investigate whether a policy is causally relevant to an observed policy outcome. Although program evaluation has made extensive use of interrupted time-series analysis, regression discontinuity analysis, causal modeling, and other techniques associated with the design and analysis of field experiments,[19] implementation research within political science has not. Instead, implementation researchers have relied mainly on techniques of case study analysis.[20] Another example comes from forecasting. Although forecasting is central to both economics and systems analysis, economics has drawn almost exclusively on econometric techniques. Systems analysis has made greater use of qualitative forecasting techniques for synthesizing expert judgment, for example, the Delphi technique.[21]

---

[16]John Dewey, *How We Think* (Boston, MA: D.C. Heath and Company, 1933), p. 108. The original statement of the difference between a problem and a problem situation is attributable to philosophical pragmatists including Charles Sanders Peirce.

[17]An exception is predictions made on the basis of expert judgment. The explanation of a policy is not necessary for predicting its future consequences. Strictly speaking, a prediction is a causal inference, whereas a projection, extrapolation, or "rational forecast" is not.

[18]Causation may be assumed but not understood. Recipes claim only that a desired result is a consequence of action. Joseph L. Bower, "Descriptive Decision Theory from the 'Administrative' Viewpoint," in *The Study of Policy Formation,* ed. Bauer and Gergen, p. 10.

[19]See, for example, William R. Shadish, Thomas D. Cook, and Donald T. Campbell, *Experimental and Quasi-Experimental Designs for Generalized Causal Inference* (Boston, MA: Houghton Mifflin, 2002).

[20]Paul A. Sabatier and Hank C. Jenkins-Smith, "The Advocacy Coalition Framework: An Assessment," in *Theories of the Policy Process,* ed. P. A. Sabatier (Boulder, CO: Westview Press, 1999), pp. 117–66.

[21]See Chapter 5.

FOUR STRATEGIES OF ANALYSIS

Relationships among policy-informational components, policy-analytic methods, and policy-informational transformations provide a basis for contrasting four strategies of policy analysis (Figure 1.2).

Prospective and Retrospective Analysis

Prospective policy analysis involves the production and transformation of information *before* policy actions are taken. This strategy of *ex ante* analysis, shown as the right half of Figure 1.2, typifies the operating styles of economists, systems analysts, operations researchers, and decision analysts.

The prospective strategy is what Williams means by policy *analysis*.[22] Policy analysis is "a means of synthesizing information to draw from it policy alternatives and preferences stated in comparable, predicted quantitative and qualitative terms as a basis or guide for policy decisions; conceptually, it does not include the *gathering* of information [emphasis in original]." Policy *research*, by contrast,

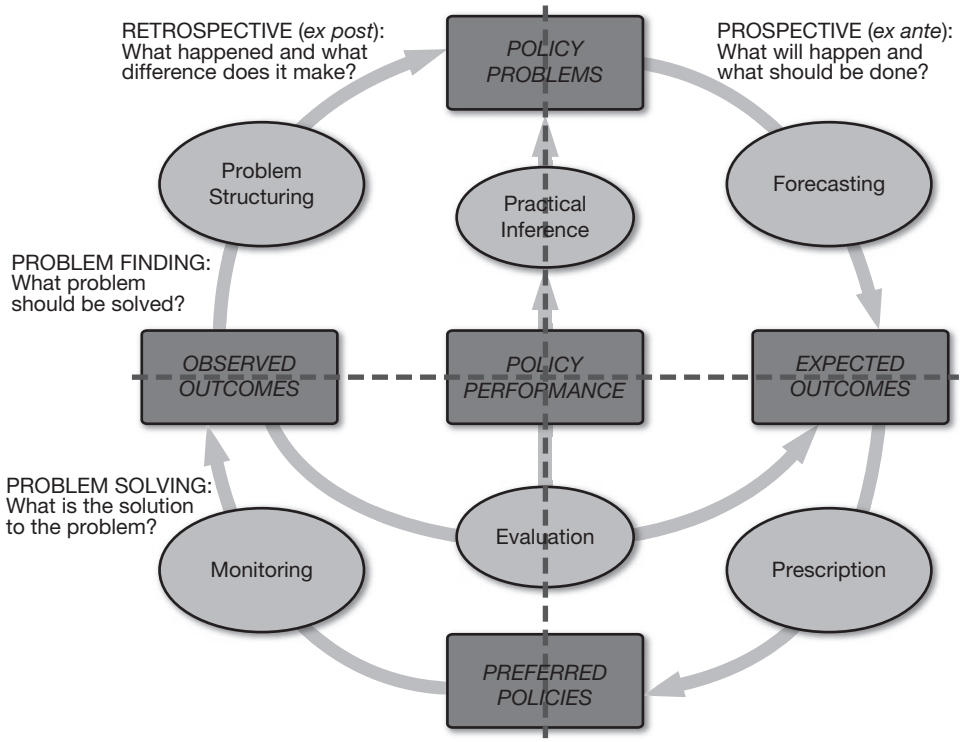


FIGURE 1.2
Forms strategies of policy analysis

[22]Walter Williams, *Social Policy Research and Analysis: The Experience in the Federal Social Agencies* (New York: American Elsevier, 1971), p. 8.

refers to "all studies using scientific methodologies to describe phenomena and/or determine relationships among them." Prospective analysis often creates wide gaps between preferred solutions and actual efforts to implement them. Perhaps no more than 10 percent of the work actually required to achieve a desired set of policy outcomes is carried out *before* policies are implemented: "It is not that we have too many good analytic solutions to problems. It is, rather, that we have more good solutions than we have appropriate actions."[23]

Retrospective policy analysis is displayed as the left half of Figure 1.2. This strategy of *ex post* analysis involves the production and transformation of information *after* policies have been implemented. Retrospective analysis characterizes the operating styles of three groups of analysts:

- *Discipline-oriented analysts.* This group, composed mainly of political scientists, economists, and sociologists, seeks to develop and test discipline-based theories that describe the causes and consequences of policies. This group is not concerned with the identification of specific policy goals or with distinctions between "policy" variables that are subject to policy manipulation and those that are not.[24] For example, the analysis of the effects of party competition on government expenditures provides no information about specific policy goals; nor is party competition a variable that policy makers can manipulate to change public expenditures.
- *Problem-oriented analysts.* This group, again composed mainly of political scientists, economists, and sociologists, seeks to describe the causes and consequences of policies. Problem-oriented analysts, however, are less concerned with the development and testing of theories believed to be important in social science disciplines than with identifying variables that may explain a problem. Problem-oriented analysts are not overly concerned with specific goals and objectives, primarily because the practical problems they analyze are usually general in nature. For example, the analysis of aggregate data on the effects of gender, ethnicity, and social inequality on national achievement test scores provides information that helps explain a problem (e.g., inadequate test performance) but does not provide information about policy variables that can be manipulated.
- *Applications-oriented analysts.* A third group includes applied economists, applied sociologists, applied psychologists, and applied anthropologists, as well as analysts from professions such as public administration, social work, and evaluation research. This group also seeks to describe the causes and consequences of public policies and programs and is not concerned with the development and testing of discipline-based theories. This group is concerned not only with manipulable policy variables but also with the identification of specific policy goals and objectives. Information about specific goals and objectives provides a basis for monitoring and evaluating outcomes and

[23]Graham T. Allison, *Essence of Decision: Explaining the Cuban Missile Crisis* (Boston, MA: Little, Brown, 1971), pp. 267–68.

[24]James S. Coleman, "Problems of Conceptualization and Measurement in Studying Policy Impacts," in *Public Policy Evaluation,* ed. Kenneth M. Dolbeare (Beverly Hills and London: Sage Publications, 1975), p. 25.

impacts of policies. For example, applications-oriented analysts may address early childhood reading readiness programs that can be manipulated in order to achieve higher scores on reading tests.

The operating styles of the three groups reflect their characteristic strengths and limitations. Discipline-oriented as well as problem-oriented analysts seldom produce information that is directly useful to policy makers. Even when problem-oriented analysts investigate important problems such as educational opportunity, energy conservation, crime control, or national security, the resultant information is often *macronegative*. Macronegative information describes the basic (or "root") causes and consequences of policies, usually by employing aggregate data to show why policies do *not* work. By contrast, *micropositive* information shows what policies and programs *do* work under specified conditions.[25] It is of little practical value to policy makers to know that the crime rate is higher in urban than rural areas, but it is practically important to know that a specific form of gun control reduces the commission of serious crimes or that intensive police patrolling is a deterrent.

Even when applications-oriented analysts provide micropositive information, they may find it difficult to communicate with practitioners of *ex ante* policy analysis, who in most cases are professional economists. In agency settings, *ex ante* analysts, whose job it is to find optimally efficient solutions, often have limited access to information about policy outcomes produced through retrospective analysis. For their part, practitioners of *ex ante* analysis often fail to specify in sufficient detail the kinds of policy-relevant information that will be most useful for monitoring, evaluating, and implementing their recommendations. Often, the intended outcomes of a policy are so vague that "almost any evaluation of it may be regarded as irrelevant because it missed the 'problem' toward which the policy was directed."[26] Legislators, for example, usually formulate problems in general terms in order to gain acceptance, forestall opposition, or maintain neutrality.

Contrasts among the operating styles of policy analysts suggest that discipline-oriented and problem-oriented analysis are inherently less useful than applications-oriented analysis—that retrospective (*ex post*) analysis as a whole is perhaps less effective in solving problems than prospective (*ex ante*) analysis. Although this conclusion may have merit from the point of view of policy makers who want advice on what actions to take, it overlooks several important benefits of retrospective analysis. Retrospective analysis, whatever its shortcomings, places primary emphasis on the results of action and is not content with information about expected policy outcomes, as is the case with prospective analysis. Discipline-oriented and problem-oriented analysis may offer new frameworks for understanding policy-making processes, challenging conventional formulations of problems, questioning social and economic myths, and shaping the climate of opinion in a community or society. Retrospective analysis,

[25]Williams, *Social Policy Research and Analysis,* p. 8.
[26]Ibid. p. 13; and Alice Rivlin, *Systematic Thinking for Social Action* (Washington, DC: Brookings, 1971).

however, "has been most important in its impact on intellectual priorities and understandings, and not nearly so effective in offering solutions for specific political problems."[27]

Descriptive and Normative Analysis

Figure 1.2 also captures another important contrast, the distinction between descriptive and normative strategies of policy analysis. Descriptive policy analysis parallels *descriptive decision theory*, which refers to a set of logically consistent propositions that describe or explain action.[28] Descriptive decision theories may be tested against observations obtained through monitoring and forecasting. Descriptive theories, models, and conceptual frameworks originate for the most part in political science, sociology, and economics. The main function of these theories, models, and frameworks is to explain, understand, and predict policies by identifying patterns of causality. The principal function of approaches to monitoring such as field experimentation is to establish the approximate validity of causal inferences relating policies to their presumed outcomes.[29] In Figure 1.2, the descriptive form of policy analysis can be visualized as an axis moving from the lower left (monitoring) to the upper right (forecasting).

Normative policy analysis parallels *normative decision theory*, which refers to a set of logically consistent propositions that evaluate or prescribe action.[30] In Figure 1.2, the normative strategy of policy analysis can be visualized as an axis running from the lower right (prescription) to upper left (evaluation). Different kinds of information are required to test normative and descriptive decision theories. Methods of evaluation and prescription provide information about policy performance and preferred policies, for example, policies that have been or will be optimally efficient because benefits outweigh costs or optimally equitable because those most in need are made better off. One of the most important features of normative policy analysis is that its propositions rest on disagreements about values such as efficiency, equity, responsiveness, liberty, and security.

Problem Finding and Problem Solving

The upper and lower halves of Figure 1.2 provide another important distinction. The upper half points to methods that are designed for *problem finding*, whereas the lower designates methods for *problem solving*. The problem-finding strategy has to do with the discovery of elements that go into the definition of problems, and not to their solution. How well do we understand the problem? Who are the most important *stakeholders* who affect and are affected by the problem? Have the appropriate objectives been identified? Which alternatives are available to achieve objectives?

---

[27]Janet A. Weiss, "Using Social Science for Social Policy," *Policy Studies Journal* 4, (Spring 1976): 237.
[28]Bower, "Descriptive Decision Theory," p. 104.
[29]See Thomas D. Cook and Donald T. Campbell, *Quasi-Experimentation: Design and Analysis Issues for Field Settings* (Boston, MA: Houghton Mifflin, 1979); Shadish, Cook, and Campbell, *Experimental and Quasi-Experimental Designs for Generalized Causal Inference.*
[30]Bower, "Descriptive Decision Theory," pp. 104–05.

Which uncertain events should be taken into account? Are we solving the "right" problem rather than the "wrong" one?

Problem-solving methods, located in the lower half of Figure 1.2, are designed to solve rather than find problems. The problem-solving strategy is primarily technical in nature, in contrast to problem finding, which is more conceptual. Problem-solving methods such as econometrics are useful in answering questions about policy causation, statistical estimation, and optimization. How much of the variance in a policy outcome is explained by one or more independent variables? What is the probability of obtaining a coefficient as large as that obtained? Another problem-solving method is benefit-cost analysis. What are the net benefits of different policies? What is their expected utility or payoff?

Segmented and Integrated Analysis

Integrated policy analysis links the four strategies of analysis displayed in Figure 1.2. Retrospective and prospective strategies are joined in one continuous process. Descriptive and normative strategies are also linked, as are methods designed to find as well as solve problems. Practically speaking, this means that policy analysts bridge the several main pillars of multidisciplinary policy analysis, especially economics and political science. Today, this need is not being properly met by specialized social science disciplines, which tend to practice *segmented policy analysis*. The job of bridging segmented disciplines—to convert intellectual knowledge into practical knowledge—is carried out by multidisciplinary professions including public administration, planning, management, and policy analysis. The American Society for Public Administration (ASPA), the National Association of Schools of Public Affairs and Administration (NASPAA), the American Planning Association (APA), the International Association of Schools and Institutes of Administration (IASIA), the Academy of Management (AM), the Operations Research Society of America (ORSA), and the Association for Public Policy and Management (APPAM) are organizations that represent these professions. So far, these professions have been more open to the disciplines of economics and political science than those disciplines have been open to them, notwithstanding a consensus among policy scholars and practitioners that the substance and methods of these and other disciplines are essential for producing policy-relevant information.

In summary, the framework for integrated policy analysis (Figure 1.1) helps examine the assumptions, strengths, and limitations of methods employed in disciplines that tend to be overly segmented and excessively specialized to be useful in practical problem solving. The framework identifies and relates major elements of policy analysis—policy-informational components, policy-analytic methods, and policy-informational transformations—enabling us to see the particular roles performed by methods of problem structuring, monitoring, evaluation, forecasting, and prescription. The framework (Figure 1.2) identifies different strategies of policy analysis: prospective (*ex ante*) and retrospective (*ex post*), descriptive and normative, and problem finding and problem solving. The framework integrates these strategies of analysis and explains why we have defined policy analysis as a problem-solving discipline that links social science theories, methods, and substantive findings to solve practical problems.

THE PRACTICE OF POLICY ANALYSIS

Reconstructed Logic versus Logic-in-Use

The process of integrated policy analysis is a *logical reconstruction (reconstructed logic)*. The process of actually doing policy analysis never completely conforms to this reconstruction, because all logical reconstructions are abstract representations of stylized practices endorsed by the scientific community.[31] By contrast, the *logic-in-use* of practicing analysts, as distinguished from the logical reconstruction of their use of reason and evidence to solve practical problems, always varies from methodological "best practices" due to personal characteristics of analysts, their professional socialization, and the institutional settings in which they work.

- *Cognitive styles.* The personal cognitive styles of analysts predispose them toward different modes of acquiring, interpreting, and using information.[32] Corporations, nonprofit organizations, and public agencies such as the U.S. Department of Corrections and the National Science Foundation use the Myers-Briggs test as a training and personnel selection diagnostic.
- *Analytic roles.* In agency settings, most analysts are largely insulated from politics. As such, they are primarily "technicians." Others perform roles that, in addition to technical content, are political. These "politicians" are actively committed to advancing the interests of political leaders or officials to whom they report. Other activist analysts are "entrepreneurs" who seek greater influence in policy making.[33]
- *Institutional incentive systems.* Policy "think tanks" encourage different orientations toward analysis, including the "humanistic-value-critical" and the "scientific."[34] Institutional rewards and punishments affect the validity of conclusions and recommendations.[35]
- *Institutional time constraints.* Analysts working in governmental settings are often subject to tight institutional time constraints (three to seven days is typical). They work with much greater speed, and perhaps greater efficiency, than analysts in academic settings or think tanks. Understandably, government analysts rarely collect original data; nor do they employ complex and time-consuming techniques.[36]

---

[31]On reconstructed logic and logic-in-use, see Kaplan, *Conduct of Inquiry,* pp. 3–11.

[32]Studies using the Myers-Briggs type indicator (Jungian personality types) suggest different cognitive styles among scientists, managers, and analysts. References provided by the Myers and Briggs Foundation at www.myersbriggs.org. See also Ian I. Mitroff and Ralph H. Kilmann, *Methodological Approaches to Social Science* (San Francisco: Jossey-Bass, 1978).

[33]Arnold Meltsner, Policy *Analysts in the Bureaucracy* (Berkeley: University of California Press, 1976); Robert A. Heineman, William T. Bluhm, Steven A. Peterson, and Edward N. Kearney, *The World of the Policy Analyst.* Chatham, NJ: Chatham House, 1990.

[34]Pamela Doty, "Values in Policy Research," in *Values, Ethics, and the Practice of Policy Analysis,* ed. William N. Dunn (Lexington, MA: D.C. Heath, 1983).

[35]Donald T. Campbell, "Guidelines for Monitoring the Scientific Competence of Preventive Intervention Research Centers: An Exercise in the Sociology of Scientific Validity," *Knowledge: Creation, Diffusion, Utilization* 8, no. 3 (1987): 389–430.

[36]See P. J. Cook and J. W. Vaupel, "What Policy Analysts Do: Three Research Styles," *Journal of Policy Analysis and Management* 4, no. 3 (1985): 427–28.

- *Professional socialization.* The different disciplines and professions that make up policy analysis socialize their members into different norms and values. Analyses of published papers suggest that analysts employ formal-quantitative as well as informal-narrative approaches, although sound policy recommendations sometimes require formal-quantitative procedures.[37]
- *Multidisciplinary teamwork.* Much of the analysis conducted in public agencies is carried out by multidisciplinary teams. Some members have primary responsibility for the particular types of analysis displayed in Figure 1.2. Team members trained in economics and decision analysis are typically more qualified to perform prospective (*ex ante*) analysis, whereas team members trained in applied sociology, applied political science, and program evaluation are usually better at retrospective (*ex post*) analysis. The effectiveness of teams depends on everyone acquiring an operational understanding of analytic methods employed throughout the process of integrated policy analysis.

Methodological Opportunity Costs

Integrated analysis has opportunity costs. Given limited time and resources, it is difficult to conduct systematic economic, political, and organizational analyses simultaneously. *Multiple triangulation*,[38] or what Cook calls *critical multiplism*,[39] responds to some of the inadequacies of logical positivism.[40] Positivism now appears as a one-sided methodology and epistemology claiming that true statements about the world must be logically and empirically verifiable, expressed in a formal (ideal) language such as mathematical statistics, and confirmed by means of statements that correspond to objective reality. Objective reality, rather than a reality constituted by subjective meaningful actions and institutions, is the foundation of true statements. Logical positivism, as Cook argues, was the dominant methodology of policy analysis and program evaluation during the era of President Lyndon Johnson's War on Poverty. The advantage of critical multiplism over logical positivism is that multiplism provides a better approximation of what is true by employing procedures that triangulate from a variety of perspectives on what is worth knowing and what is known about policies.[41]

---

[37]An early but representative overview of approaches is Janet A. Schneider, Nancy J. Stevens, and Louis G. Tornatzky, "Policy Research and Analysis: An Empirical Profile, 1975–1980," *Policy Sciences* 15 (1982): 99–114.

[38]The methodology of triangulation is analogous to practices employed in geodesic surveys; cartography; navigation; and, more recently, satellite tracking. The position or location of an object is found by means of bearings from two or more fixed points or electronic signals a known distance apart.

[39]Cook advanced critical multiplism as an alternative to logical positivism. See Thomas D. Cook, "Postpositivist Critical Multiplism," in *Social Science and Social Policy,* ed. R. Lane Shotland and Melvin M. Mark (Beverly Hills, CA: Sage Publications, 1985), pp. 21–62.

[40]A critical assessment of logical positivism is Mary E. Hawkesworth, "Epistemology and Policy Analysis," in *Advances in Policy Studies since 1950,* ed. Dunn and Kelly, pp. 293–328; and Hawkesworth, *Theoretical Issues in Policy Analysis.*

[41]Cook, "Postpositivist Critical Multiplism," p. 57.

A disadvantage of multiplism lies in its costs. Triangulation among multiple disciplinary perspectives, along with the use of multiple methods, measures, and data sources, involves trade-offs and opportunity costs.[42] When single methods such as econometric modeling are employed to achieve measurement precision and statistical generalizability, analysts forgo opportunities to acquire a deeper understanding of policies that is possible through ethnographic interviews, case studies, and other qualitative methods. A leading econometrician, noting that economists are unique among social scientists because they are trained only to analyze data, not to collect it, observes that "empirical work can be greatly enhanced by being sensitive to the context of the problem (the data-generating process) and knowing a lot about one's data."[43] Similar trade-offs apply to methods of research synthesis, or meta-analysis, which purchase measurement precision and generalized policy causation at the expense of a deeper understanding of contexts of policy-making.[44]

Ethnographic interviews, by contrast, involve high information costs because they require the collection of substantial primary data through interviews. However, they also lack precision and seldom permit the generalization of policy causation to other settings. Although greater precision and generalizability can be obtained by means of field studies and field experiments, these are expensive, especially when they are employed in conjunction with mixed (quantitative and qualitative) methods. To be sure, triangulation among convergent (and divergent) perspectives, methods, and measures may enhance the validity of policy analysis and other applied social sciences.[45] But the time and financial constraints make trade-offs inevitable.

CRITICAL THINKING AND PUBLIC POLICY

The world of the policy analyst is complex. Analysts must sift through and evaluate a large volume of available quantitative and qualitative data, make difficult choices among sources of information, select appropriate methods and techniques, and employ effective strategies for communicating the results of analysis through oral briefings and documents (see Chapter 9). These practical challenges place a premium on *critical thinking*—that is, the capacity to organize, synthesize, and evaluate diverse sources of reasoning and evidence. One method available for this

[42]See David Brinberg and Joseph E. McGrath, *Validity and the Research Process* (Beverly Hills, CA: Sage Publications, 1985). For Brinberg and McGrath and other methodological pragmatists, the choice of methods is similar to an optimization problem in decision analysis. See C. West Churchman, *Prediction and Optimal Decision: Philosophical Issues of a Science of Values* (Englewood Cliffs, NJ: Prentice Hall, 1961); and Russell Ackoff, *Scientific Method: Optimizing Applied Research Decisions* (New York: John Wiley, 1962).

[43]Peter Kennedy, *A Guide to Econometrics*, 4th ed. (Cambridge, MA: MIT Press, 1998), pp. 83–84.

[44]See Lawrence Rudner, Gene V. Glass, David L. Evartt, and Patrick J. Emery, *A User's Guide to the Meta-Analysis of Research Studies*. ERIC Clearinghouse on Assessment and Evaluation, University of Maryland, College Park, 2002. http://echo.edres.org

[45]The case for triangulation in its many forms is found in Campbell, *Methodology and Epistemology for Social Science*, ed. Overman.

purpose is the analysis of *policy arguments*. By analyzing policy arguments, we are able to identify and probe the assumptions underlying competing policy claims, recognize and evaluate objections to these claims, and synthesize policy-relevant information from different sources.

The Structure of Policy Arguments

Policy arguments are the main vehicle carrying debates about public policies.[46] Although social scientists may rightly pride themselves on methodological specialization, they too often forget that "public policy is made of language. Whether in written or oral form, argument is central to all stages of the policy process."[47]

The structure of a policy argument can be represented as a set of seven elements (Figure 1.3):[48]

Policy claim (C). A policy claim is the conclusion of a policy argument. Arguments also include other elements, including policy-relevant information (I), warrants (W), backings (B), qualifiers (Q), objections (O), and rebuttals (R). The movement from policy-relevant information to claim implies *therefore, thus*, or *so*. Policy claims are of different types. Some are normative: "Congress should pass the amendments to the Fair Employment Practices Act." Some are descriptive: "The use of the Internet will double in the next ten years."

Policy-relevant information (I). Policy-relevant information provides the grounds for a policy claim. These grounds may be statistical data, experimental findings, expert testimony, common sense, or political judgments. Policy-relevant information is a response to the question: *What information is relevant to the claim?* Information is the starting point of a new argument and the end of a previous one. Policy arguments may lead to complex argument chains, trees, or cycles.

Warrant (W). The warrant is a reason to support a claim. Warrants may be economic theories, ethical principles, political ideas, professional authority, and so forth.[49] A warrant answers the question: *Why does this reason support the claim?* Different types of warrants are related to arguments

---

[46]See Frank Fischer and John Forester, ed., *The Argumentative Turn in Policy Analysis and Planning* (Durham, NC: Duke University Press, 1993). Earlier works on policy argumentation are Ian I. Mitroff and Richard O. Mason, *Creating a Dialectical Social Science* (Boston: D. Reidel, 1981); William N. Dunn, "Reforms as Arguments," *Knowledge: Creation, Diffusion, Utilization* 3 (1982): 293–326; Donald T. Campbell, "Experiments as Arguments," *Knowledge: Creation, Diffusion, Utilization* 3 (1982): 327–47; Giandomenico Majone, *Evidence, Argument, and Persuasion in the Policy Process* (New Haven, CT: Yale University Press, 1989); and Stone, *Policy Paradox and Political Reason.*

[47]Majone, *Evidence, Argument, and Persuasion,* p. 1.

[48]This structural model of argument is part of the computer software called *Rationale 2*, which was developed by Tim van Gelder and his colleagues in Australia. URL: www.austhink.com. The classic structural model is presented in Stephen Toulmin, *The Uses of Argument* (Cambridge: Cambridge University Press, 1958); and Stephen Toulmin, A. Rieke, and A. Janik, *An Introduction to Reasoning* (New York: Macmillan, 1984).

[49]Different kinds of warrants yield the "modes" of policy argument presented in Chapter 8.

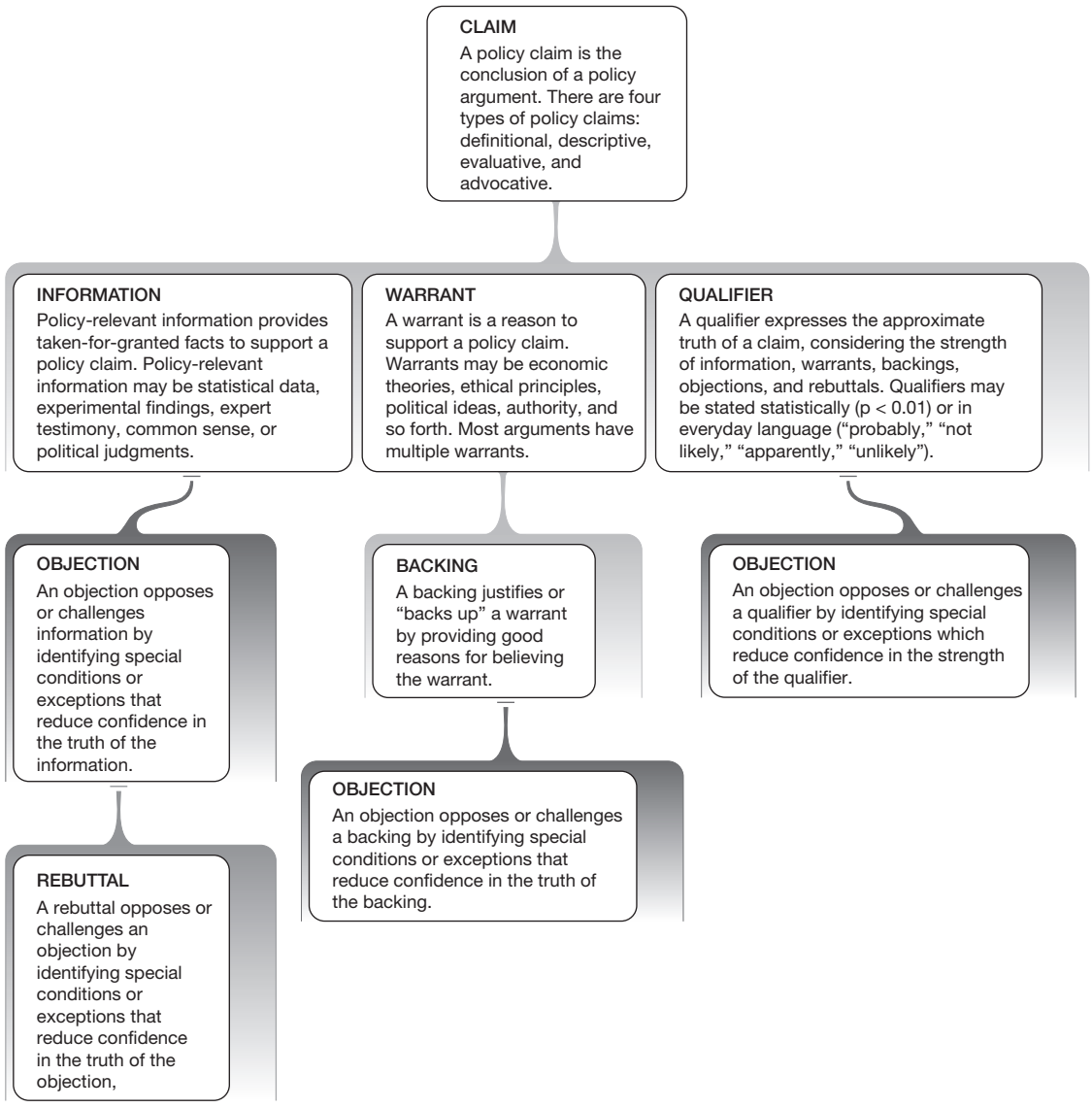


FIGURE 1.3
Elements of a policy argument

Source: Created with *Rationale 2.* Melbourne: Austhink Consulting, 2010. www.austhink.com

made in different disciplines and professions. For example, law uses case comparisons and rules of evidence, whereas economics uses theories and component laws such as the law of diminishing utility of money. Policy makers as well as social scientists employ causal warrants such as "Ethnic cleansing will be deterred by air strikes that establish NATO's credibility in

the region." The warrant, which provides a justification for accepting a claim, answers the question: *Considering the information, what reasons make the claim true?*

Qualifier(Q). The qualifier expresses the degree to which a claim is approximately true, given the strength of the information, warrants, and backings, as well as objections and rebuttals. Although social scientists may state qualifiers in the language of formal probability ($p = 0.01$ or $t = 2.24$), ordinary language is the normal mode of qualifying claims with such terms as *certainly, absolutely, necessarily, probably, in all likelihood, presumably, apparently,* and *barring unforeseen circumstances*. The qualifier answers the question: *How strong or credible is the claim?* It is primarily through processes of argumentation and debate that policy makers, policy analysts, and other policy stakeholders adjust or even abandon arguments. Such changes, when they occur, are motivated by the strength of objections and rebuttals offered by those who have a stake in policies.

Backing (B). The backing is an additional reason to support or "back up" the warrant. The backing answers the question: *Why does the warrant support the claim?* with a more general reason, assumption, or argument that begins with *because*. Different kinds of backings are characteristically employed by members of different disciplines and professions. Backings may be scientific laws, appeals to the authority of experts, or ethical and moral principles. For example, consider the warrant presented earlier: "Ethnic cleansing will be deterred by air strikes that establish NATO's credibility in the region." The backing for warrants advocating the use of coercive force is frequently an informal statement of the law of diminishing utility: "The greater the cost of an alternative, the less likely it will be pursued."

Objection (O). An objection opposes or challenges the information, warrant, backing, or qualifier by identifying special conditions or exceptions that reduce confidence in the truth of the information, warrant, backing, or qualifier. An objection answers the question: *Are there special circumstances or exceptions that threaten the credibility of the warrant?* Analysts who pay attention to objections are more likely to take a critical perspective toward a policy argument, identifying weak or hidden assumptions, anticipating unintended consequences, or questioning possible rebuttals to objections. Thereby, analysts can be self-critical, challenging their own assumptions and arguments.

Rebuttal(R). A rebuttal is an objection to an objection. Rebuttals oppose or challenge objections by identifying special conditions or exceptions that reduce confidence in the truth of the objection. Rebuttals answer the question: *Are there special circumstances or exceptions that threaten the credibility of the objection?* Most policy arguments have objections and rebuttals, because policy making involves bargaining, negotiation, competition, and compromise among opponents and proponents of policies.

The frames of reference, perspectives, and reasons of policy makers and analysts are found in their underlying warrants, backings, objections, and rebuttals. Therefore, identical policy-relevant information is interpreted in distinctly different

ways. A decrease in crime rates in urban areas may be welcomed by the urban poor, viewed with skepticism by owners of central city businesses, rejected by criminologists who attribute urban crime rates to changes in unemployment and homelessness, and hailed as an achievement by elected officials. By examining contending arguments and their underlying assumptions, analysts can uncover and critically assess reasoning and evidence that otherwise goes unnoticed. Equally important is what it brings to analysts themselves—they can probe their own assumptions by examining the objections, qualifications, and exceptions to their own conclusions.

CHAPTER SUMMARY

This chapter has provided a framework for policy analysis that identifies the role of policy-analytic methods in creating and transforming policy-relevant information. The four regions of this framework call attention to similarities and differences among methods of policy analysis and point to the origins of these methods in different social science disciplines and professions, thus clarifying the meaning of multidisciplinary inquiry. No one methodology is appropriate for all or most problems. Given the need to choose among methods, methodological choices can be viewed as an optimization problem involving trade-offs and opportunity costs. The actual work of practicing analysts demands critical thinking. The analysis of policy arguments is well suited for this purpose.

REVIEW QUESTIONS

1. What does it mean to define policy analysis as a process of inquiry as distinguished from a set of methods?
2. Describe the dynamics of policy-informational components, policy-analytic methods, and policy-informational transformations.
3. Contrast segmented and integrated policy analysis. Give examples.
4. How does normative decision theory differ from descriptive decision theory?
5. List some of the key differences between problem solving and problem finding.
6. Contrast retrospective and prospective analysis. Which social science disciplines tend to specialize in prospective analysis? Retrospective analysis?
7. Discuss the strengths and limitations of critical multiplism.
8. Contrast the "logic-in-use" and the "reconstructed logic" of policy analysis. Provide examples.
9. How can argumentation mapping assist analysts to become critical thinkers?

DEMONSTRATION EXERCISES

1. Scorecards provide a useful overview of the observed and expected outcomes of different policies. When using the scorecard for monitoring and forecasting the outcomes of the two speed limits (Case 1.2), the 55 mph speed limit seems preferable to the 65 mph speed limit. Compare the scorecard (Figure C1.2) with the spreadsheet (Figure C1.3). Does the comparison change your conclusions about the performance of the 55 mph speed limit? What does this tell us about the differences between monitoring and evaluation in policy analysis? What are the implications for the distinctions among different types of policy analysis?
2. Influence diagrams and decision trees are useful methods for structuring policy problems. The diagram and tree displayed in Figure 1.2 help identify policy stakeholders, policy alternatives, uncertain outcomes and events, probabilities of these outcomes and events, and valued

outcomes (valued outcomes are objectives). Consider the influence diagram that represents the problem of energy supply in 1973 and 1974, when the OPEC oil embargo posed significant challenges both to U.S. energy supply and national security. How does Figure C1.3 help us formulate the problem? What do the arrows suggest about the causes of the energy shortage as well as the causes of the decline in traffic fatalities? What does the influence diagram suggest about the conditions that gave rise to the 55 mph speed limit and its effectiveness? After comparing the influence diagram with the decision tree (which is based on the influence diagram), describe why these two problem representations are good examples of descriptive and normative decision theory.

3. Create an argument map based on the influence diagram presented in Case 1.3. Begin with the following claim: "The United States should return to the 55 mph speed limit in order to conserve fuel and save lives." Include in your map as many warrants, backings, objections, and rebuttals as you can. Assuming that the original qualifier was *certainly*, indicate whether the qualifier changes as we move from a simple, static, uncontested argument to a complex, dynamic, and contested argument.

BIBLIOGRAPHY

Campbell, Donald T. *Methodology and Epistemology for Social Science: Selected Papers*. Edited by E. Samuel Overman. Chicago: University of Chicago Press, 1988.

Diesing, Paul. *How Social Science Works: Reflections on Practice*. Pittsburgh, PA: Pittsburgh University Press, 1991.

Dunn, William N., and Rita Mae Kelly. *Advances in Policy Studies since 1950*. New Brunswick, NJ: Transactions Books, 1992.

Fischer, Frank, and John Forester. *The Argumentative Turn in Policy Analysis and Planning*. Durham, NC: Duke University Press, 1993.

Hawkesworth, Mary E. *Theoretical Issues in Policy Analysis*. Albany: State University of New York Press, 1988.

Kaplan, Abraham. *The Conduct of Inquiry: Methodology for Behavioral Science*. San Francisco, CA: Chandler, 1964.

Mac Rae, Duncan Jr. *The Social Function of Social Science*. New Haven, CT: Yale University Press, 1976.

Stone, Deborah. *Policy Paradox: The Art of Political Decision Making*. Rev Ed. New York: W. W. Norton, 2001.

Toulmin, Stephen R. *Return to Reason*. Cambridge, MA: Harvard University Press, 2001.

Van Gelder, Tim. "The Rationale for Rationale." Law, Probability, and Risk 6 (2007): 23–42.

CASE 1.1 THE GOELLER SCORECARD—MONITORING AND FORECASTING TECHNOLOGICAL IMPACTS

When advanced technologies are used to achieve policy goals, sociotechnical systems of considerable complexity is created. Although it is analytically tempting to prepare a comprehensive economic analysis of the costs and benefits of such policies, most practicing analysts do not have the time or the resources to do so. Given the time constraints of policy making, many analyses are completed in a period of several days to a month, and in most cases policy analyses do not involve the collection and analysis of new data. Early on in a project, policy makers and their staffs typically want an

overview of the problem situation and the potential impacts of alternative policies. Under these circumstances, the scorecard is appropriate.

The Goeller *scorecard,* named after Bruce Goeller of the RAND Corporation, is appropriate for this purpose. Table C1.1 shows the impacts of alternative transportation systems. Some of the impacts involve transportation services used by members of the community, whereas others involve impacts on low-income groups. In this case, as Quade observes, the large number of diverse impacts are difficult to value in dollar terms, making a benefit-cost analysis impractical and even impossible.[50] Other impacts involve financial and economic questions such as investments, jobs created, sales, and tax revenues. Other impacts are distributional because they involve the differential effects of transportation. ■

TABLE C1.1

Scorecard

Social Impacts	CTOL	VTOL	TACV
TRANSPORTATION			
Passengers (million miles)	7	4	9
Per trip time (hours)	2	1.5	2.5
Per trip cost ($)	$17	$28	$20
Reduced congestion (%)	0%	5%	10%
FINANCIAL			
Investment ($ millions)	$150	$200	$200
Annual subsidy ($ millions)	0	0	90
ECONOMIC			
Added jobs (thousands)	20	25	100
Added sales ($millions)	50	88	500
COMMUNITY			
Noise (households)	10	1	20
Added air pollution (%)	3%	9%	1%
Petroleum savings (%)	0%	–20%	30%
Displaced households	0	20	500
Taxes lost ($millions)	0	0.2	2
Landmarks destroyed	None	None	Fort X
DISTRIBUTIONAL			
Low-income trips (%)	7%	1%	20%
Low-income household Noise annoyance (%)	2%	16%	40%

Source: Goeller (1974); Quade, *Analysis for Public Decisions* (1975), p. 60.

Note: Conventional takeoff and landing aircraft (CTOL); vertical takeoff and landing aircraft (VTOL); tracked air-cushion vehicle (TACV).

[50]E.S. Quade, *Analysis for Public Decisions* (New York: American Elsevier, 1975), p. 65.

CASE 1.2 THE SPREADSHEET—EVALUATING THE BENEFITS AND COSTS OF ENERGY POLICIES

In 1972 and 1973, the United States and other petroleum-dependent countries experienced the first of several oil crises precipitated by a dramatic increase in the price of crude oil by the Organization of Petroleum Exporting Countries (OPEC). The response of American and European leaders was to adopt maximum speed limits of 55 mph and 90 kph, respectively. In the United States, the National Maximum Speed Limit (NMSL) was designed to reduce the consumption of gasoline by requiring that all vehicles on interstate highways travel at a maximum of 55 mph, a speed that would maximize fuel efficiency for most vehicles.

Soon after the implementation of the 55 mph speed limit, it was discovered that the new policy not only reduced fuel consumption, but apparently caused a dramatic decline in traffic fatalities and injuries as well. Therefore, long after the OPEC oil crisis was over, the speed limit was retained, although it was no longer needed to respond to the energy crisis that prompted its passage in 1973. Indeed, the 55 mph speed limit was retained for more than 20 years until it was officially repealed in November 1995.[51]

Heated debates preceded the repeal. Senator John C. Danforth of Missouri, an influential advocate of the policy, argued that the repeal would save one minute per day per driver but result in an additional 600 to 1,000 deaths. The *Washington Post* and the *New York Times* joined the opposition, reporting that, although fatalities would surely rise, the savings in time was trivial. Later, Secretary of Transportation Pena announced that the Clinton administration was firmly opposed to abandoning the speed limit.

This was the right moment for an evaluation of the benefits and costs of the NMSL. A spreadsheet is a simple but powerful tool for doing so. The scorecard, as we saw in Case 1.1, is a useful tool for monitoring and forecasting impacts when benefit-cost analysis is not feasible or desirable. On the scorecard, policy alternatives are arrayed in columns along the top of the matrix and policy impacts are listed in each row. Spreadsheets, by contrast, are appropriate and useful for prescribing preferred policies and evaluating their outcomes. Spreadsheets display the benefits and costs of observed or expected policy outcomes, creating information about policy performance as well as preferred policies (see Figure 1.1).

Table C1.2 displays a spreadsheet used to evaluate the effects of the 55 mph speed limit at the end of 1974, one year after the policy was implemented. To show the differences between the spreadsheet and the scorecard, Table C1.2 also displays the same information as a scorecard. ■

[51]On April 2, 1987, Congress enacted the Surface Transportation and Uniform Relocation Assistance Act, permitting 40 states to experiment with speed limits up to 65 mph.

TABLE C1.2

Scorecard and Spreadsheet

(a) Scorecard

OUTCOMES	65 MPH (Base Case)*	55 MPH
Fatalities	54,052	45,196
Miles traveled (billions)	1,313	1,281
Hours driving (billions)	20.2	21.9
Gallons fuel consumed (billions)	46.8	43.3
Fuel efficiency (mpg)	14.9	16.1
Traffic citations (millions)	5,711	7,425
Property damage (million cases)	25.8	23.1

*The base case is the policy against which the new policy is compared.

(b) Spreadsheet

OBJECTIVES	65MPH	55MPH	Difference	Value	$ Billions
I Fatalities (000s)	54.1	45.2	8.856	$240,000.00	$ 2.13
II Hours driving (billions)	20.2	21.9	−1.7	5.05	−8.59
III Gallons fuel consumed (billions)	46.8	43.3	3.5	0.53	1.86
IV Traffic citations (000s)	5,711	7,425	−1,714	3.94	−0.0068
V Property damage cases (000s)	25,800	23,100	2,700	363.00	0.98
Benefits (I + III + V)					4.97
Costs (II + IV)					−8.60
Net Benefits (B–C)					$ −3.63

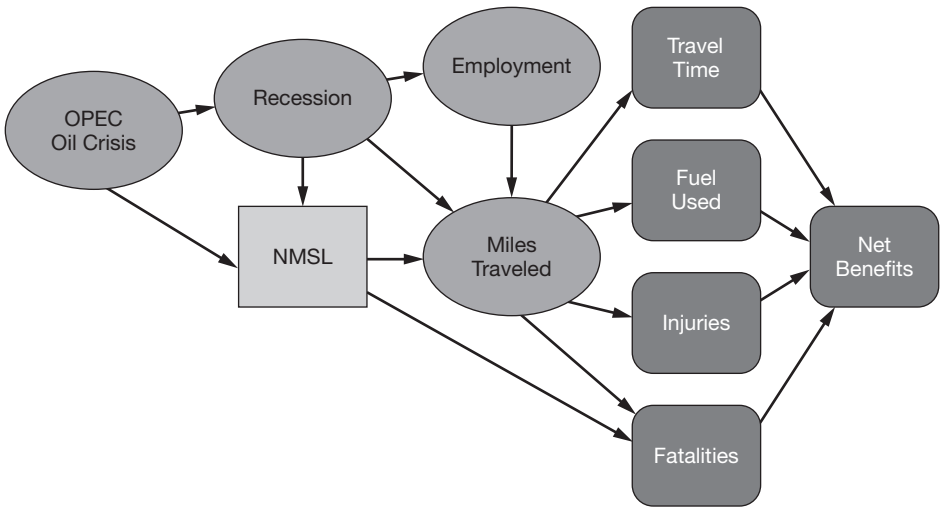
CASE 1.3 THE INFLUENCE DIAGRAM AND DECISION TREE—STRUCTURING PROBLEMS OF ENERGY POLICY AND INTERNATIONAL SECURITY

Along with other policy-analytic methods discussed earlier in this chapter (Figure 1.1), the *influence diagram* and *decision tree* are useful tools for structuring policy problems.[52] The influence diagram (Figure C1.3) displays the policy, the National Maximum Speed Limit, as a rectangle. A rectangle always refers to a policy choice or decision node, which in this case is the choice

(*continued*)

[52]The diagram and tree were created with the Decision Programming Language (DPL), which is available from Syncopation Software at http://www.syncopation.com. Educational, professional, and commercial versions of DPL 7.0 are available.

(a) Influence Diagram



(b) Decision Tree

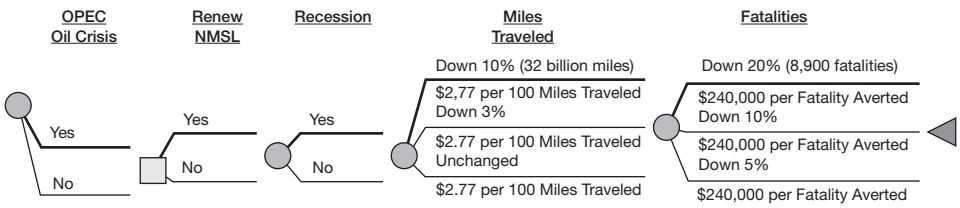


FIGURE C1.3
Influence diagram and decision tree

between adopting and not adopting the national maximum speed limit of 55 mph. To the right and above the decision node are uncertain events, represented as ovals, which are connected to the decision node with arrows showing how the speed limit affects or is affected by them. The rectangles with shaved corners represent valued policy outcomes or objectives. The objectives are to lower fuel consumption, reduce travel time, reduce injuries, and avert traffic fatalities. To the right of the objectives is another shaved rectangle, which designates the net benefits (benefits less costs) of the four objectives. The surprising result of using the influence diagram for problem structuring is the discovery of causally relevant economic events, such as the recession and unemployment, which affect miles driven, which in turn affect all four objectives. The "root cause" appears to be the OPEC oil embargo.

The decision tree is another representation of the influence diagram. Whereas the influence diagram shows how policy choices and uncertain events affect the achievement of objectives, the decision tree displays the monetary value of these objectives. In this abridged and simplified decision tree, there are two branches that represent the alternatives but also the OPEC oil embargo, the recession, the costs of miles traveled, and the dollar benefits of reducing fatalities. The bolded branches show the events with the greatest likelihood of occurring or that already have occurred. ■

CASE 1.4 THE ARGUMENT MAP—PROBLEM STRUCTURING IN NATIONAL DEFENSE AND TRANSPORTATION POLICY

The role of causal arguments in transforming policy-relevant information into policy claims may be illustrated by Allison's well-known study of foreign policy decision making during the Cuban missile crisis of October 1962.[53] Showing how different explanatory models yield different conclusions, Allison argues that government policy analysts think about problems of foreign policy in terms of implicit conceptual models that shape their thought; most analysts explain the behavior of governments in terms of a model that assumes the rationality of political choices (*rational actor model*); alternative models, including those that emphasize organizational processes (*organizational process model*) and bureaucratic politics (*bureaucratic politics model*), provide a basis for improved explanations.

In 1962, the policy alternatives open to the United States ranged from no action and diplomatic pressure to secret negotiations, invasion, surgical air strikes, and blockade. Among the several claims made at the time of the Cuban missile crisis, let us consider the policy actually adopted by the United States: "The United States should blockade Cuba." In this case, the policy-relevant information (I) is "The Soviet Union is placing offensive missiles in Cuba." The warrant states that "the blockade will force the withdrawal of missiles by showing the Russians that the United States is determined to use force." In providing reasons to accept the warrant, the backing (B) supports the warrant by stating that "an increase in the cost of an alternative reduces the likelihood of that alternative being chosen."[54] The backing (B) represents a general theoretical proposition, or law, within the rational policy model. After the objection (O) has successfully challenged the warrant, the qualifier (Q) changes from *absolutely* to *doubtful.*

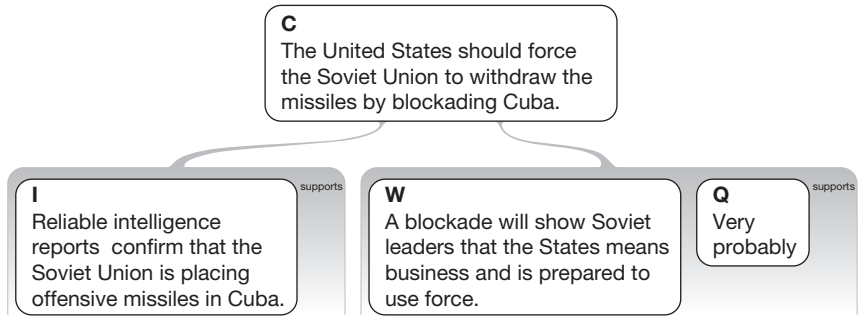
Allison's account shows how the use of multiple competing explanations can facilitate critical thinking. The use of multiple competing models moves the analysis from a simple uncontested argument (Figure C1.4.1) to a new argument that is complex, contested, and dynamic (Figure C1.4.2). This change occurs because a serious objection has been raised about the warrant and the backing of the claim. The objection states: "But Soviet leaders may fail to convince their naval units to depart from established organizational routines." The warrant for this objection is: "The bulk of research on organizations shows that major lines of organizational behavior tend to be straight. Behavior at time *t+1* differs little from behavior at time *t.*[55] The blockade will not work." The warrant for the objection is again a general proposition or law within the organizational process model, otherwise known as the disjointed incremental theory of policy change.

Simple uncontested maps of arguments about the 55 mph speed limit are presented alongside of the arguments about the Cuban missile crisis (Figures C1.4.1a and C1.4.1b). The comparisons show that the simple argument maps represent *uncritical* thinking. By contrast, the complex, dynamic, and contested maps of the same crises (Figures C1.4.2a and C1.4.2b) illustrate what is meant by *critical thinking.* ■

(*continued*)

[53]Graham T. Allison, "Conceptual Models and the Cuban Missile Crisis," *American Political Science Review* 3002, no. 3 (1969): 689–718.
[54]Allison, "Conceptual Models," p. 694.
[55]Ibid., p. 702.

(a) The Cuban Missile Crisis—Simple Uncontested Argument



(b) The 55 mph Speed Limit—Simple Uncontested Argument

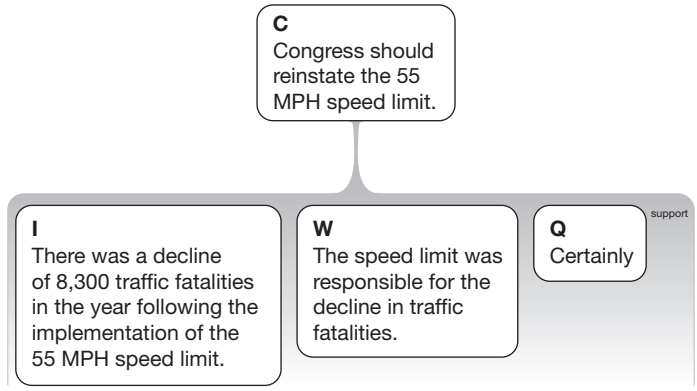
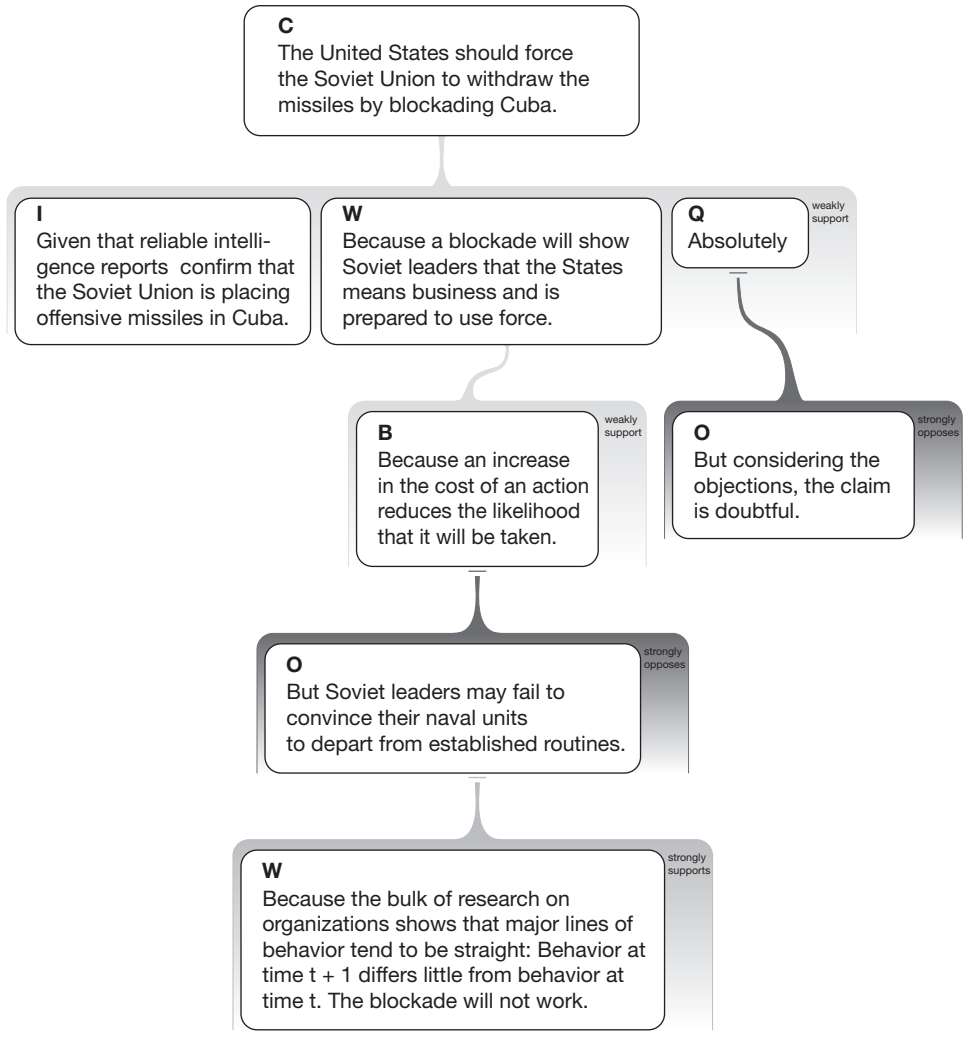


FIGURE C1.4.1
Simple argument maps are static and uncontested

(a) The Cuban Missile Crisis—Dynamic Contested Argument



(*continued*)

(b) The 55 mph Speed Limit—Simple Uncontested Argument

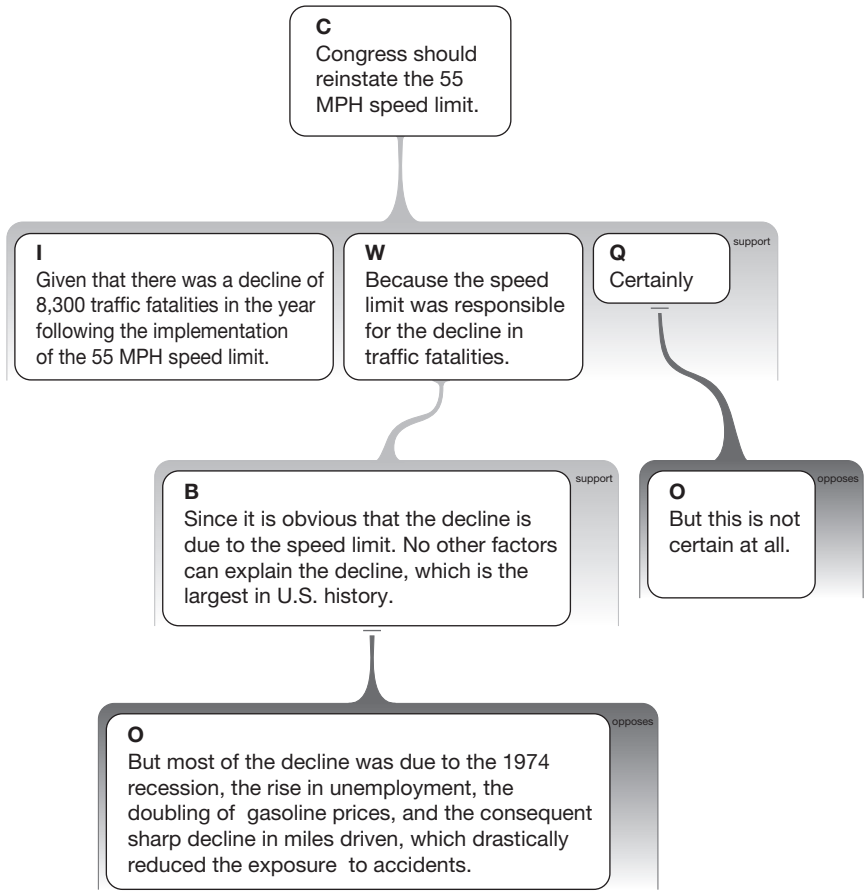


FIGURE C1.4.2
Complex argument maps are dynamic and contested

References

1 Chapter 1 The Process of Policy Analysis

Campbell, Donald T. Methodology and Epistemology for Social Science: Selected Papers. Edited by E. Samuel Overman. Chicago: University of Chicago Press, 1988.

Diesing, Paul. How Social Science Works: Reflections on Practice. Pittsburgh, PA: Pittsburgh University Press, 1991.

Dunn, William N., and Rita Mae Kelly. Advances in Policy Studies since 1950. New Brunswick, NJ: Transactions Books, 1992.

Fischer, Frank, and John Forester. The Argumentative Turn in Policy Analysis and Planning. Durham, NC: Duke University Press, 1993.

Hawkesworth, Mary E. Theoretical Issues in Policy Analysis. Albany: State University of New York Press, 1988. Kaplan, Abraham. The Conduct of Inquiry: Methodology for Behavioral Science. San Francisco, CA: Chandler, 1964. Mac Rae, Duncan Jr. The Social Function of Social Science. New Haven, CT: Yale University Press, 1976. Stone, Deborah. Policy Paradox: The Art of Political Decision Making. Rev Ed. New York: W. W. Norton, 2001. Toulmin, Stephen R. Return to Reason. Cambridge, MA: Harvard University Press, 2001. Van Gelder, Tim. “The Rationale for Rationale.” Law, Probability, and Risk 6 (2007): 23–42.

When advanced technologies are used to achieve

policy goals, sociotechnical systems of considerable

complexity is created. Although it is analytically

tempting to prepare a comprehensive economic

analysis of the costs and benefits of such policies,

most practicing analysts do not have the time or the resources to do so. Given the time constraints of policy making, many analyses are completed in a period of several days to a month, and in most cases policy analyses do not involve the collection and analysis of new data. Early on in a project, policy makers and their staffs typically want an

CASE 1.1 THE GOELLER SCORECARD—

MONITORING AND FORECASTING

TECHNOLOGICAL IMPACTS 50 E.S. Quade, Analysis for Public Decisions (New York: American Elsevier, 1975), p. 65. TABLE C1.1 Scorecard Social Impacts CTOL VTOL TACV TRANSPORTATION Passengers (million miles) 7 4 9 Per trip time (hours) 2 1.5 2.5 Per trip cost ($) $17 $28 $20 Reduced congestion (%) 0% 5% 10% FINANCIAL Investment ($ millions) $150 $200 $200 Annual subsidy ($ millions) 0 0 90 ECONOMIC Added jobs (thousands) 20 25 100 Added sales ($millions) 50 88 500 COMMUNITY Noise (households) 10 1 20 Added air pollution (%) 3% 9% 1% Petroleum savings (%) 0% −20% 30% Displaced households 0 20 500 Taxes lost ($millions) 0 0.2 2 Landmarks destroyed None None Fort X DISTRIBUTIONAL Low-income trips (%) 7% 1% 20% Low-income household Noise annoyance (%) 2% 16% 40% SOurce: Goeller (1974); Quade, Analysis for Public Decisions (1975), p. 60. NOte: Conventional takeoff and landing aircraft (CTOL); vertical takeoff and landing aircraft (VTOL); tracked air-cushion vehicle (TACV). overview of the problem situation and the potential impacts of alternative policies. Under these circumstances, the scorecard is appropriate. The Goeller scorecard, named after Bruce Goeller of the RAND Corporation, is appropriate for this purpose.Table C1.1 shows the impacts of alternative transportation systems. Some of the impacts involve transportation services used by members of the community, whereas others involve impacts on low-income groups. In this case, as Quade observes, the large number of diverse impacts are difficult to value in dollar terms, making a benefit-cost analysis impractical and even impossible. 50 Other impacts involve financial and economic questions such as investments, jobs created, sales, and tax revenues. Other impacts are distributional because they involve the differential effects of transportation.

In 1972 and 1973, the United States and other

petroleum-dependent countries experienced the

first of several oil crises precipitated by a

dramatic increase in the price of crude oil by the

Organization of Petroleum Exporting Countries

(OPEC).The response of American and European leaders was to adopt maximum speed limits of 55 mph and 90 kph, respectively. In the United States, the National Maximum Speed Limit (NMSL) was designed to reduce the consumption of gasoline by requiring that all vehicles on interstate highways travel at a maximum of 55 mph, a speed that would maximize fuel efficiency for most vehicles. Soon after the implementation of the 55 mph speed limit, it was discovered that the new policy not only reduced fuel consumption, but apparently caused a dramatic decline in traffic fatalities and injuries as well. Therefore, long after the OPEC oil crisis was over, the speed limit was retained, although it was no longer needed to respond to the energy crisis that prompted its passage in 1973. Indeed, the 55 mph speed limit was retained for more than 20 years until it was officially repealed in November 1995. 51 Heated debates preceded the repeal. Senator John C. Danforth of Missouri, an influential advocate of the policy, argued that the repeal would save one minute per day per driver but result in an additional 600 to 1,000 deaths. The Washington Post and the New York Times joined the opposition, reporting that, although fatalities would surely rise, the savings in time was trivial. Later, Secretary of Transportation Pena announced that the Clinton administration was firmly

opposed to abandoning the speed limit. This was the right moment for an evaluation of the benefits and costs of the NMSL. A spreadsheet is a simple but powerful tool for doing so. The scorecard, as we saw in Case 1.1, is a useful tool for monitoring and forecasting impacts when benefitcost analysis is not feasible or desirable. On the scorecard, policy alternatives are arrayed in columns along the top of the matrix and policy impacts are listed in each row. Spreadsheets, by contrast, are appropriate and useful for prescribing preferred policies and evaluating their outcomes. Spreadsheets display the benefits and costs of observed or expected policy outcomes, creating information about policy performance as well as preferred policies (see Figure 1.1). Table C1.2 displays a spreadsheet used to evaluate the effects of the 55 mph speed limit at the end of 1974, one year after the policy was implemented. To show the differences between the spreadsheet and the scorecard, Table C1.2 also displays the same information as a scorecard.

51 On April 2, 1987, Congress enacted the Surface Transportation and Uniform Relocation Assistance

Act, permitting 40 states to experiment with speed limits up to 65 mph.

CASE 1.2 THE SPREADSHEET—EVALUATING THE

BENEFITS AND COSTS OF ENERGY POLICIES (continued) TABLE C1.2 Scorecard and Spreadsheet (a) Scorecard 65 MPH OUTCOMES (Base Case)* 55 MPH Fatalities 54,052 45,196 Miles traveled (billions) 1,313 1,281 Hours driving (billions) 20.2 21.9 Gallons fuel consumed (billions) 46.8 43.3 Fuel efficiency (mpg) 14.9 16.1 Traffic citations (millions) 5,711 7,425 Property damage (million cases) 25.8 23.1 *The base case is the policy against which the new policy is compared. (b) Spreadsheet OBJECTIVES 65MPH 55MPH Difference Value $ Billions I Fatalities (000s) 54.1 45.2 8.856 $240,000.00 $ 2.13 II Hours driving (billions) 20.2 21.9 1.7 5.05 ▯8.59 III Gallons fuel consumed (billions) 46.8 43.3 3.5 0.53 1.86 IV Traffic citations (000s) 5,711 7,425 1,714 3.94 ▯0.0068 V Property damage cases (000s) 25,800 23,100 2,700 363.00 0.98 Benefits (I ▯ III ▯ V) 4.97 Costs (II ▯ IV) ▯8.60 Net Benefits (B-C) $ ▯3.63 Along with other policy-analytic methods discussed earlier in this chapter (Figure 1.1), the influence diagram and decision tree are useful tools for structuring policy problems. 52 The influence diagram (Figure C1.3) displays the policy, the National Maximum Speed Limit, as a rectangle. A rectangle always refers to a policy choice or decision node, which in

this case is the choice 52 The diagram and tree were created with the Decision Programming Language (DPL), which is available from Syncopation Software at http://www.syncopation.com. Educational, professional, and commercial versions of DPL 7.0 are available. CASE 1.3 THE INFLUENCE DIAGRAM AND DECISION TREE—STRUCTURING PROBLEMS OF ENERGY POLICY AND INTERNATIONAL SECURITY Travel Time Fuel Used Net Benefits Injuries Fatalities Miles Traveled Employment Recession OPEC Oil Crisis NMSL FIGURE C1.3 Influence diagram and decision tree

between adopting and not adopting the national

maximum speed limit of 55 mph. To the right and

above the decision node are uncertain events,

represented as ovals, which are connected to the

decision node with arrows showing how the speed

limit affects or is affected by them. The rectangles

with shaved corners represent valued policy

outcomes or objectives. The objectives are to lower

fuel consumption, reduce travel time, reduce

injuries, and avert traffic fatalities. To the right of

the objectives is another shaved rectangle, which

designates the net benefits (benefits less costs) of

the four objectives.The surprising result of using

the influence diagram for problem structuring is

the discovery of causally relevant economic events, such as the recession and unemployment, which affect miles driven, which in turn affect all four objectives. The “root cause” appears to be the OPEC oil embargo. The decision tree is another representation of the influence diagram. Whereas the influence diagram shows how policy choices and uncertain events affect the achievement of objectives, the decision tree displays the monetary value of these objectives. In this abridged and simplified decision tree, there are two branches that represent the alternatives but also the OPEC oil embargo, the recession, the costs of

miles traveled, and the dollar benefits of reducing fatalities. The bolded branches show the events with the greatest likelihood of occurring or that already have occurred. Yes Yes OPEC Oil Crisis Renew NMSL Recession FatalitiesMiles Traveled Yes Down 10% (32 billion miles) $2.77 per 100 Miles Traveled $2,77 per 100 Miles Traveled Down 3% $2.77 per 100 Miles Traveled Unchanged Down 20% (8,900 fatalities) $240,000 per Fatality Averted $240,000 per Fatality Averted Down 5% $240,000 per Fatality Averted Down 10% No No No (b) Decision Tree (a) Influence Diagram CASE 1.4 THE ARGUMENT MAP—PROBLEM STRUCTURING IN NATIONAL DEFENSE AND TRANSPORTATION POLICY The role of causal arguments in transforming policyrelevant information into policy claims may be illustrated by Allison's well-known study of foreign policy decision making during the Cuban missile crisis of October 1962. 53 Showing how different explanatory models yield different conclusions, Allison argues that government policy analysts think about problems of foreign policy in terms of implicit conceptual models that shape their thought; most analysts explain the behavior of governments in terms of a model that assumes the rationality of political choices (rational actor model); alternative models, including those that emphasize organizational processes (organizational process model) and bureaucratic politics (bureaucratic politics model), provide a basis for improved explanations. In 1962, the policy alternatives open to the United States ranged from no action and diplomatic pressure to secret negotiations, invasion, surgical air strikes, and blockade. Among the several claims made at the time of the Cuban missile crisis, let us consider the policy actually adopted by the United States: "The United States should blockade Cuba." In this case, the policy-relevant information (I) is "The Soviet Union is placing offensive missiles in Cuba." The warrant states that "the blockade will force the withdrawal of missiles by showing the Russians that the United States is determined to use force." In providing reasons to accept the warrant, the backing (B) supports the warrant by stating that "an increase in the cost of an alternative reduces the likelihood of that alternative being chosen." 54 The backing (B) represents a general theoretical proposition, or law, within the rational policy model. After the objection (O) has successfully challenged the warrant, the qualifier (Q) changes from absolutely to doubtful. Allison's account shows how the use of multiple competing explanations can facilitate critical thinking. The use of multiple competing models moves the analysis from a simple uncontested argument (Figure C1.4.1) to a new argument that is complex, contested, and dynamic (Figure C1.4.2). This change occurs because a serious objection has

been raised about the warrant and the backing of the claim. The objection states: “But Soviet leaders may fail to convince their naval units to depart from established organizational routines.” The warrant for this objection is: “The bulk of research on organizations shows that major lines of organizational behavior tend to be straight. Behavior at time t+1 differs little from behavior at time t. 55 The blockade will not work.” The warrant for the objection is again a general proposition or law within the organizational process model, otherwise known as the disjointed incremental theory of policy change. Simple uncontested maps of arguments about the 55 mph speed limit are presented alongside of the arguments about the Cuban missile crisis (Figures C1.4.1a and C1.4.1b).The comparisons show that the simple argument maps represent uncritical thinking. By contrast, the complex, dynamic, and contested maps of the same crises (Figures C1.4.2a and C1.4.2b) illustrate what is meant by critical thinking. 53 Graham T. Allison, “Conceptual Models and the Cuban Missile Crisis,” American Political Science Review 3002, no. 3 (1969): 689–718. 54 Allison, “Conceptual Models,” p. 694. 55 Ibid., p. 702. (continued) 27 FIGURE C1.4.1 Simple argument maps are static and uncontested Reliable intelligence reports confirm that the Soviet Union is placing offensive missiles in Cuba. I A blockade will show Soviet leaders that the States means business and is prepared to use force. W Very probably Q supports supports The United States should force the Soviet Union to withdraw the missiles by blockading Cuba. C (a) The Cuban Missile Crisis—Simple Uncontested Argument There was a decline of 8,300 traffic fatalities in the year following the implementation of the 55 MPH speed limit. I The speed limit was responsible for the decline in traffic fatalities. W Certainly Q support Congress should reinstate the 55 MPH speed limit. C (b) The 55 mph Speed Limit—Simple Uncontested Argument The United States should force the Soviet Union to withdraw the missiles by blockading Cuba. C Because an increase in the cost of an action reduces the likelihood that it will be taken. B But considering the objections, the claim is doubtful. O Given that reliable intelligence reports confirm that the Soviet Union is placing offensive missiles in Cuba. I Because a blockade will show Soviet leaders that the States means business and is prepared to use force. W Absolutely Q weakly support weakly support strongly opposes But Soviet leaders may fail to convince their naval units to depart from established routines. O strongly opposes Because the bulk of research on organizations shows that major lines of behavior tend to be straight: Behavior at time t + 1 differs little from behavior at time t. The blockade will not work. W strongly

supports (a) The Cuban Missile Crisis—Dynamic Contested Argument (continued) FIGURE C1.4.2 Complex argument maps are dynamic and contested Congress should reinstate the 55 MPH speed limit. C Since it is obvious that the decline is due to the speed limit. No other factors can explain the decline, which is the largest in U.S. history. B But this is not certain at all. O Given that there was a decline of 8,300 traffic fatalities in the year following the implementation of the 55 MPH speed limit. I Because the speed limit was responsible for the decline in traffic fatalities. W Certainly Q support support opposes But most of the decline was due to the 1974 recession, the rise in unemployment, the doubling of gasoline prices, and the consequent sharp decline in miles driven, which drastically reduced the exposure to accidents. O opposes (b) The 55 mph Speed Limit—Simple Uncontested Argument

76 Rowland Egger, “The Period of Crises: 1933 to 1945,” in American Public Administration: Past,

Present, and Future, ed. Frederick C. Mosher (Tuscaloosa: University of Alabama Press, 1975),

pp. 49–96.

77 Fritz Machlup, The Production and Distribution of Knowledge in the United States (Princeton, NJ:

Princeton University Press, 1962).

78 Daniel Bell, The Coming of Post-Industrial Society: A Venture in Social Forecasting (New York: Basic

Books, 1976).

79 Jurgen Schmandt and James E. Katz, “The Scientific State: A Theory with Hypotheses,” Science,

Technology, and Human Values 11 (1986): 40–50.

80 Bell, Coming of Post-Industrial Society, pp. xvi–xviii. industrial society, is increasingly dominated by an educated professional-technical class. Many of the characteristics of a postindustrial society are important to consider when we ponder the historical evolution and significance of policy analysis: 80 Centrality of theoretical knowledge. Although the earliest civilizations relied on specialized information, only in the late twentieth and twenty-first centuries are innovations in “soft” (social) and “hard” (physical) technologies directly dependent on the codification of theoretical knowledge provided by the social, physical, and biological sciences. Creation of new intellectual technologies. The improvement in mathematical and statistical techniques has made it possible to use advanced modeling, simulation, and various forms of systems analysis to find more efficient and “rational” solutions to public problems. Spread of a knowledge class. The most rapidly expanding occupational group in the United States is composed of technicians and professionals. This group, including professional managers, represented 28 percent of employed persons in 1975, 28.3 percent in 1995, and 30.3 percent in 2000. Persons in

nonfarm occupations grew from 62.5 to 98.4 percent between 1900 and 1990. Shift from goods to services. In 1975, more than 65 percent of the active labor force was engaged in the production and delivery of services, a figure that exceeded 80 percent by 2000. Services are

In the first half of the twentieth century, the

involvement of social and natural scientists in policy

making was largely a response to social, economic,

and politico-military crises. 76 Members of the

sciences and professions were ad hoc and irregular

participants in government in the period of

depression, dislocation, and war that stretched from

the early 1930s to the late 1940s. In succeeding

years, however, the social sciences and social

professions—political science, sociology, economics,

public administration, business management, social

work, planning—became one of the main sources of

full-time governmental staff, advisers, and

consultants. For the first time, public agencies

recruited on a regular basis specialists who had

received their training from and had been certified

by their respective professional organizations. For many observers, this new era signified the

advent of a new form of social and political

organization in which public policy making and

societal guidance were critically dependent on the

specialized knowledge and technologies of the

sciences and professions. Emphasizing its social

aspect, some referred to this new form of

organization as the “knowledge society” 77 or the

“postindustrial society.” 78 Others, focusing on its

political aspect, spoke of the transition from the

“administrative state” to the “scientific state.” 79

Postindustrial Society and the Scientific State

Postindustrial society, an extension of patterns of

policy making and socioeconomic organization of

CASE 2.1 ARE POLICY ANALYSTS

TECHNOCRATS? 81 See Jeffrey D. Straussman, “Technocratic Counsel and Societal Guidance,” in Politics and the Future of Industrial Society, ed. Leon N. Lindberg (New York: David McKay, 1976), pp. 126–66. 82 For example, Amitai Etzioni, The Active Society (New York: Free Press, 1968); and Donald T. Campbell, “The Experimenting Society,” in The Experimenting Society: Essays in Honor of Donald T. Campbell, ed. William N. Dunn (New Brunswick, NJ: Transaction Books, 1998), pp. 35–68. (continued) primarily human (e.g., health, education, social welfare) and professional-technical (e.g., information services, computer programming, telecommunications).
Instrumentalization of sciences. Although the natural and social sciences have long been used as instruments for controlling the human and material environment, they have been largely open to internal criticism and protected by norms of scientific inquiry. In the present period, science and technology have become increasingly bureaucratized, subordinated to government goals, and assessed in terms of their payoffs to society. The vast bulk of funded research is applied research on practical problems of business, commerce, and government. Scarcity of scientific and technical information. Information has increasingly become one of society’s most scarce resources. Information is a collective, not a private, good, and cooperative rather than competitive strategies for its production and use are required if optimal results are to be achieved. The rapid exponential growth of the Internet and World Wide Web are counteracting the scarcity of information demanded by governments, businesses, and nonprofit organizations. The development of policy analysis and other applied sciences

seems to be part of an emerging postindustrial society and scientific state. Yet, what does this development signify for issues of democratic governance, civil liberties, and privacy? How much influence do producers of policy-relevant information have? Although theoretical knowledge clearly seems to be more central than it has been in other historical periods, what does this mean for the distribution of political power? Does the proliferation of new intellectual technologies mean also that structures and processes of policy making have changed? Is the spread of a "knowledge class" commensurate with its power and influence? Considering that the aim of policy analysis and other applied sciences is to produce information for practical purposes, whose purposes are being served? In short, how are we to interpret the historical transformation of policy-oriented inquiry from earlier times to the present-day "knowledge society" and "scientific state"? Technocratic Guidance versus Technocratic Counsel There are several contending perspectives of these questions, two of which we consider here. 81 One perspective holds that the professionalization of policy analysis and other applied social sciences is—or should be—shifting power and responsibility from policy makers to policy analysts. 82 This perspective, technocratic guidance, is associated with the "analycentric turn" described earlier in this chapter. The basic premise of the analycentric turn is that "the surest way to improve the quality of public choice is to have more analysts producing more analyses." 83 By contrast, a contending perspective, technocratic counsel, holds that the professionalization of policy analysis and other applied social sciences signifies new and more effective ways to enhance the influence of policy makers and other dominant groups whose positions continue to rest on power, wealth, and privilege. 84 Strictly speaking, neither of these perspectives is wholly accurate in its interpretation of events surrounding the movement toward postindustrial society; each contains a one-sided emphasis on particular characteristics of contemporary society to the exclusion of others. Yet just as the development of nineteenth-century policy analysis was a practical response to problems of the day as viewed by 83 Schick, "Beyond Analysis, p. 259. 84 See, for example, Guy Benveniste, The Politics of Expertise (Berkeley, CA: Glendessary Press, 1972); and Frank Fischer, Technocracy and the Politics of Expertise (Newbury Park, CA: Sage Publications, 1990).

dominant groups, so too is contemporary policy

analysis a consequence of changes in the structure

and role of government as it has attempted to grapple with new problems.“The contemporary boom in policy analysis, ” writes Schick,“has its primary source in the huge growth of American governments, not in the intellectual development of the social sciences.” 85 Throughout the twentieth century, the expansion of government was followed by demands for more policy-relevant information. As government has grown, so too has the market for policy analysis. 86 The development of policy analysis in the past century followed fits and starts in the growth of federal executive agencies. 87 In 1900, there were some ninety federal executive agencies. By 1940, this figure had grown to 196, and, by 1973, the total number of federal executive agencies had jumped to 394.The greatest percentage increase, 35 percent, occurred after 1960.The rapid growth of the federal government was mainly a response to new problems in areas of national defense, transportation, housing, health, education, welfare, crime, energy, and the environment. There were many new “challenges” to policy making: 88 participation overload, mobilization of service sector workers, transformation of basic social values, realignment of interest groups, growing

cleavages between urban and suburban citizens, and recurrent fiscal crises. Some observers, notably policy scientist Yehezkel Dror, went so far as to formulate the relationship between such problems and policy analysis in the form of a general law: "While the difficulties and dangers of problems tend to increase at a geometric rate, the number of persons qualified to deal with these problems tends to increase at an arithmetic rate." 89 Thus, the growth of policy analysis seems to be a consequence, not a cause, of changes in the structure of government and the nature and scope of social problems. The technocratic guidance perspective asserts that policy-relevant knowledge is an increasingly scarce resource whose possession enhances the power and influence of policy analysts. The technocratic guidance perspective may be summarized in terms of five major propositions: 90 The growing interdependencies, complexity, and pace of change of contemporary society make existing knowledge obsolete, thus increasing the demand for new forms of policy-relevant knowledge. Problems of contemporary society may be resolved with specialized knowledge produced by policy analysts. The technical complexity of policy choices prompts higher levels of direct involvement of policy analysts and other applied scientists. Higher levels of direct involvement enhance the power of professional analysts to influence policy making, making politicians increasingly dependent on them. The growing dependence of politicians on professional policy analysts erodes their political power. A rival perspective, that of technocratic counsel, begins from the assumption that professional policy analysts operate in settings in which policy makers, as the consumers of specialized knowledge, determine to a large extent the activities of producers. In this context, the primary role of analysts is to legitimize—that is, justify in scientific and technical terms—policy decisions made by the real holders of power. The technocratic counsel

85 Schick, "Beyond Analysis," p. 258.

86 Schmandt and Katz, "Scientific State."

87 Herbert Kaufman, Are Government Organizations Immortal?

(Washington, DC: Brookings

Institution, 1976), pp. 34–63.

88 Samuel P. Huntington, “Postindustrial Politics; How Benign Will It Be?” Comparative Politics 6

(January 1974): 163–92; Ronald Inglehart, “The Silent Revolution in Europe: Intergenerational Change

in Postindustrial Societies,” American Political Science Review, 65 (December 1971): 991–1017; and

James O’Connor, The Fiscal Crisis of the State (New York: St. Martin’s Press, 1973).

89 Dror, Ventures in Policy Sciences, p. 2.

90 Straussman, “Technocratic Counsel,” pp. 150–51. perspective may also be summarized in terms of several key propositions: 91 Major policy alternatives reflect conflicting values held by different segments of the community. Value conflicts are associated with disparities of political power. The choice of a given policy alternative symbolizes the victory of one segment of the community over another. Policy makers use scientific and technical justifications produced by analysts to suppress conflicts and legitimize choices that are actually made on political grounds. The credibility of scientific and technical justifications requires that policy analysis and other applied sciences be presented as value neutral, impartial, and apolitical. Professional analysts, although they are a source of scientific and technical justifications, are expendable. They also serve as convenient scapegoats for policies that fail. If we pick our cases carefully, we can show that analysts exercised a large measure of “technocratic guidance.” For example, the New York City-RAND Corporation’s analysis of factors influencing the response time of city firefighters to reported fires has been credited with unusual success, even though other RAND efforts have been less effective in shaping 91 Ibid., pp. 151–52. 92 On this and other cases, see Greenberger, Crenson, and Crissey, Models in the Policy Process, pp. 231–318. 93 Nathan Caplan, “Factors Associated with Knowledge Use among Federal Executives,” Policy Studies Journal 4, no. 3 (1976): 233. See also Robert F. Rich, Social Science Information and Public Policy Making (San Francisco, CA: Jossey-Bass, 1981); and David J. Webber, “The Production and Use of Knowledge in the Policy Process,” in Advances in Policy Studies since 1950, ed.

Dunn and Kelly, pp. 415–41. 94 Nathan Caplan et al., The Use of Social Science Knowledge in Policy Decisions at the National Level (Ann Arbor, MI: Institute for Social Research, Center for Research on the Utilization of Scientific Knowledge, 1975). decisions. 92 On the other hand, the use of specialized knowledge to make policy choices appears to be highly uneven. For example, in a study of 204 federal executives conducted in 1973–74, social science research was relatively unimportant as a basis for choosing among policy alternatives Moreover, the extent to which research was used depended on essentially nontechnical factors:“the level of knowledge utilization is not so much the result of a slow flow of relevant and valid information from knowledge producers to policy makers, but is due more to factors involving values, ideology, and decisionmaking styles.” 93 , 94 Similar findings are reported in recent issues of journals such as Science Communication, Implementation Science, and the British Medical Journal. The finding that policy analysis is used for political purposes adds a measure of credibility to the technocratic counsel perspective. So does the apparent conservative character of many analyses. In this context, policy analysis has been characterized as a conservative and superficial kind of social science that fails to pose radical questions about basic social values and institutions, and neglects policy alternatives that depart significantly from existing practices. 95 Moreover, some observers see in policy analysis an ideology in disguise that suppresses ethical and value questions in the name of science. 96 Under such conditions, it is understandable that analysts might be used as instruments of everyday politics. 95 Charles E. Lindblom, “Integration of Economics and the Other Social Sciences through Policy Analysis,” in Integration of the Social Sciences through Policy Analysis, ed. James C. Charlesworth (Philadelphia: American Academy of Political and Social Sciences, 1972), p. 1. For an elaboration of this critique, see Charles E. Lindblom, Inquiry and Change: The Troubled Attempt to Understand and Change Society (New Haven, CT: Yale University Press, 1990). 96 Laurence H. Tribe, “Policy Science: Analysis or Ideology?” Philosophy and Public Affairs 2, no. 1 (1972): 66–110. j. supports your own position on the issue (CUE) k. provides evidence to back up recommendations (TQ) l. contains politically feasible implications (AO) m. agrees with previous knowledge about the issue (TQ) n. can be implemented under the present circumstances (AO) o. challenges existing assumptions and institutional arrangements (CSQ) p. raises new issues or perspectives (CSQ) q. is inexpensive to implement (AO) r. was available at the time a decision had to be made (AO) s. contains

novel or unexpected findings (negative coding) t. provides quantitative data (TQ) u. can be generalized to similar settings or persons (TQ) v. addresses in a thorough manner different factors that could explain outcomes (TQ) w. uses statistically sophisticated analyses (TQ) x. demonstrates high technical competence (TQ) y. presents internally consistent and unambiguous findings (TQ) z. takes an objective and unbiased approach (TQ) Using the following categories, sum up the scaled responses (scale numbers 1 through 5), and then average them (divide by the number of questions answered).The capital letters after each question (e.g.,TQ) refer to the categories. Technical Quality (TQ): Conformity to User Expectations (CUE): Action Orientation (AO): Challenge Status Quo (CSQ):

The following profile, drawn from the work of Carol

H. Weiss at Harvard Univer sity, is helpful in learning

about the uses of research in policymaking. 97 Please

use the following scale for responses to questions about

a policy paper or report published by a government

agency or a think tank.You may obtain the paper or

report on the Internet from a website such as

www.publicagenda.org, www.rand.org, or www.aei.org. 1 very much 2 quite a bit 3 to some extent 4 not much 5 not at all

1. To what extent does the analysis contain information that is relevant to your work?

2. To what extent do you think the analysis is reliable (could be repeated with similar conclusions) and valid (plausibly explains what it set out to explain)?

3. To what extent do the conclusions agree with your own views about the issue?

4. Please indicate the extent to which the analysis a. deals with a high priority issue (AO) b. adds to theoretical knowledge in the field (TQ) c. adds to practical knowledge about policies or programs (AO) d. agrees with your ideas or values (CUE) e. identifies outcomes that policy makers can do something about (AO) f. suggests potential courses of action (AO) g. implies the

need for major policy changes (CSQ) h. focuses on specific policy outcomes (AO) i. contains explicit recommendations for action (AO)

CASE 2.2 UNDERSTANDING THE USE OF POLICY

ANALYSIS

97 These questions are drawn from Carol H. Weiss and Michael J. Bucuvalas, “The Challenge of Social Research to

Decision Making,” in Using Social Research in Public Policy Making, ed. Weiss, Appendix 15A and 15B, pp. 231–33. I am

grateful to Carol Weiss for suggesting modifications in questions.

3 Chapter 3 Structuring Policy Problems

Ackoff, Russell L. The Art of Problem Solving:

Accompanied by Ackoff's Fables. New York:

John Wiley, 1978.

_____. Redesigning the Future.

New York: Wiley, 1974.

Adams, James L. Conceptual Blockbusting.

Stanford, CA: Stanford Alumni Association,

1974.

Adelman, L., T. R. Stewart, and K. R. Hammond.

"A Case History of the Application of Social

Judgment Theory to Policy Formulation."

Policy Sciences 6 (1975): 137-59.

Bobrow, Davis B., and John S. Dryzek. Policy

Analysis by Design. Pittsburgh, PA: University

of Pittsburgh Press, 1986.

Churchman, C. West. The Design of Inquiring

Systems. New York: Basic Books, 1971.

Dery, David. Problem Definition in Policy

Analysis. Lawrence: University Press of Kansas,

1984.

Dror, Yehezkel. Public Policy Making Reexamined.

Rev Ed. New Brunswick. NJ: Transaction

Books, 1983.

Dryzek, John S. "Policy Analysis as a

Hermeneutic Activity.” Policy Sciences 14

(1982): 309–29.

Dunn, William N. “Methods of the Second Type:

Coping with the Wilderness of Conventional

Policy Analysis.” Policy Studies Review 7, no. 4

(1988): 720–37.

Dunn, William N., and Ari Ginsberg. “A

Sociocognitive Approach to Organizational

Analysis.” Human Relations 39, no. 11

(1986): 955–75.

Fischhoff, Baruch. “Clinical Policy Analysis.” In

Policy Analysis: Perspectives, Concepts, and

Methods. Edited by William N. Dunn.

Greenwich, CT: JAI Press, 1986.

———. “Cost-Benefit Analysis and the Art of

Motorcycle Maintenance.” Policy Sciences 8

(1977): 177–202. George, Alexander. “Criteria for Evaluation of Foreign Policy Decision Making.” Global Perspectives 2 (1984): 58–69. Hammond, Kenneth R. “Introduction to Brunswickian Theory and Methods.” New Directions for Methodology of Social and Behavioral Science 3 (1980): 1–12. Hofstadter, Richard. Godel, Escher, Bach. New York: Random House, 1979. Hogwood, Brian W., and B. Guy Peters. The Pathology of Public Policy. Oxford: Clarendon Press, 1985. Linder, Stephen H., and B. Guy Peters. “A Metatheoretic Analysis of Policy Design.” In Advances in Policy Studies since 1950. Vol. 10 of Policy Studies Review Annual. Edited by William N. Dunn and Rita Mae Kelly. New Brunswick, NJ: Transaction Books, 1992. Linstone, Harold A. Multiple Perspectives for Decision Making. New York: North-Holland, 1984. Mason, Richard O., and Ian I. Mitroff. Challenging Strategic Planning Assumptions. New York: Wiley, 1981. Meehan, Eugene J. The Thinking Game. Chatham, NJ: Chatham House, 1988. Mitroff, Ian I., Richard O. Mason,

and Vincent P. Barabba. The 1980 Census: Policymaking amid Turbulence. Lexington, MA: D. C. Heath, 1983. Saaty, Thomas L. The Analytic Hierarchy Process. New Y ork: McGraw-Hill, 1980. Schon, Donald A. The Reflective Practitioner. New York: Basic Books, 1983. Sieber, Sam. Fatal Remedies. New York: Plenum Press, 1981. Warfield, John N. Societal Systems: Planning, Policy, and Complexity. New York: Wiley, 1976. Watzlawick, Paul, John Weakland, and Richard Fisch. Change: Principles of Problem Formation and Problem Resolution. New York: W. W. Norton, 1974. BOX 3.0—CONDUCTING A STAKEHOLDER ANALYSIS Definition A stakeholder is a person who speaks for or represents a group that is affected by or affects a policy. Stakeholders include the president of a legislative assembly or parliament, a chairperson of a legislative committee, or an executive director or members of an organized interest or advocacy group such as the National Rifle Association, the Sierra Club, or Human Rights Watch. Policy analysts and their employers are stakeholders, as are clients who commission a policy analysis. Persons or groups who do not have a stake in a policy (e.g., an uninvolved college professor) are not stakeholders. Assumptions • Stakeholders are best identified by policy issue area. A policy issue area is a domain in which stakeholders disagree or quarrel about policies. Housing, welfare, education, and international security are policy issue areas. • Stakeholders have specific names and titles— for example, State Senator Xanadi; Mr.Young, chairperson of the House Finance Committee; or Ms. Ziegler, a spokesperson for the National Organization of Women (NOW). • A sociometric or “snowball” sample such as that described next is an effective way to estimate the “population” of stakeholders. STEP 1: Using Google or a reference book such as The Encyclopedia of Associations, identify and list about ten stakeholders who have taken a public position on a policy. Make the initial list as heterogeneous as possible by sampling opponents as well as supporters. STEP 2: For each stakeholder, obtain a policy document (e.g., a report, news article, e-mail, or telephone interview) that describes the position of each stakeholder. STEP 3: Beginning with the first statement of the first stakeholder, list other stakeholders mentioned as opponents or proponents of the policy. STEP 4: For each remaining statement, list the new stakeholders mentioned. Do not repeat. STEP 5: Draw a graph that displays statements 1, 2, . . . n on the horizontal axis. On the vertical axis, display the cumulative frequency of new stakeholders mentioned in the statements. The graph will gradually flatten out, with no new stakeholders mentioned. If this does not occur before reaching the last stakeholder on the initial list, repeat

steps 2 to 4. Add to the graph the new statements and the new stakeholders. STEP 6: Add to the estimate stakeholders who should be included because of their formal positions (organization charts show such positions) or because they are involved in one or more policy-making activities: agenda setting, policy formulation, policy adoption, policy implementation, policy evaluation, and policy adaptation, succession or termination. Retain the full list for further analysis. You now have an estimate of the "population" of key stakeholders who are affected by and affect the policy, along with a description of their positions on an issue. This is a good basis for structuring the problem.

79 The architecture analogy is from Herbert Simon.

CASE 3.1 STRUCTURING PROBLEMS OF RISK IN

MINING AND TRANSPORTATION

Complex problems must be structured before they

can be solved. The process of structuring a policy

problem is the search for and specification of

problem elements and how they are the elements are

Policy stakeholders. Which stakeholders affect or are affected by a problem?

Policy alternatives. What alternative courses of action may be taken to solve the problem?

Policy actions. Which of these alternatives should be acted on to solve the problem?

Policy outcomes. What are the probable outcomes of action and are they part of the solution to the problem?

Policy values (utilities). Are some outcomes more valuable than others in solving the problem?

Most policy problems are messy or ill-structured.

For this reason, one or more problem elements can

be incorrectly omitted from the definition of a

problem. Even when problem elements are correctly

specified, relations among the elements may be unknown or obscure. This makes it difficult or impossible to determine the strength and significance, practical as well as statistical, of causal relations. For example, many causal processes that are believed to govern relations among atmospheric pollution, global warming, and climate change are obscure. The obscurity of these processes stems not only from the complexity of "nature" but also from the conflicting beliefs of stakeholders who disagree, often intensely, about the definition of problems and their potential solutions. For this reason, the possible combinations and permutations of problem elements—that is, stakeholders, alternatives, actions, outcomes, values—appear to be unmanageably huge.

Under these conditions, standard methods of decision theory (e.g., risk-benefit analysis), applied economics (e.g., benefit-cost analysis), and political science (e.g., policy implementation analysis) are of limited value until the problem has been satisfactorily defined. This is so because an adequate definition of the problem must be constructed before the problem can be solved with these and other standard methods. Standard methods are useful in solving relatively well-structured (deterministic) problems involving certainty, for example, problems represented as fixed quantities in a spreadsheet. Standard methods are also useful in solving moderately structured (probabilistic) problems involving uncertainty, for example, problems represented as policy outcomes with different probabilities. However, ill-structured problems are of a different order. Estimates of uncertainty, or risk, cannot be made because we do not even know the

outcomes to which we might attach probabilities. Here, the analyst is much like an architect who has been commissioned to design a custom building for which there is no standard plan. 79 The adoption of a standard plan, if such existed, would almost certainly result in a type III error: solving the wrong problem. Public policies are deliberate attempts to change complex systems. The process of making and implementing policies occurs in social systems in which many contingencies lie beyond the control of policy makers. It is these unmanageable contingencies that are usually responsible for the success and failure of policies in achieving their objectives. The contingencies are rival hypotheses that can challenge claims that a policy (the presumed cause) produced one or more policy outcomes (the presumed effects). In such cases, it is usually desirable to test, and when possible eliminate, these rival hypotheses through a process of eliminative induction. Eliminative induction takes this general form:“Repeated observations of policy x and outcome y confirm that x is causally relevant to the occurrence of y. However, additional observations of x, z, and y confirm that unmanageable contingency z and not policy x is responsible for the

112 occurrence of y.” By contrast, enumerative induction takes this general form:“Repeated observations confirm that the policy x is causally relevant to the occurrence of policy outcome y.” Eliminative induction permits a critical examination of contingencies that are beyond the control of policy makers. Because the number of these contingencies is potentially unlimited, the process of identifying and testing rival explanations is never complete. Yet, precisely for this reason, it seems impossible to identify and test an unmanageably huge number of potential rival hypotheses. How is this to be done? One answer is creativity and imagination. But creativity and imagination are impossible to teach, because there are no rules governing the replication of creative or imaginative solutions. Another answer is an appeal to well-established theories. However, the bulk of theories in the social sciences are disputed and controversial. “Well-established” theories are typically “well-defended” theories, and rival hypotheses are rarely considered seriously, let alone tested. A more appropriate alternative is the use of boundary analysis and estimation to structure problems involving a large number of rival hypotheses. Boundary analysis and estimation look for rival hypotheses in the naturally occurring policy quarrels that take place among stakeholders. In addition to the policy analyst, these stakeholders include scientists, policy makers, and

organized citizen groups. The aim of boundary estimation is to obtain a relatively complete set of rival hypotheses in a given policy context. Although boundary estimation strives to be comprehensive, it does not attempt the hopeless task of identifying and testing all plausible rival hypotheses. Although the range of rival hypotheses is never complete, it is possible to estimate the probable limit of this range. Assessing the Impact of National Maximum Speed Limit A boundary analysis was conducted with documents prepared by thirty-eight state officials responsible for reporting on the effects of the original 55 mph and later 65 mph speed limits in their states. As expected, there were sharp disagreements among many of the thirty-eight stakeholders. For example, some states were tenaciously committed to the hypothesis that speed limits are causally related to fatalities (e.g., Pennsylvania and New Jersey). Others were just as firmly opposed (e.g., Illinois, Washington, Idaho). Of direct importance to boundary estimation is that 718 plausible rival hypotheses were used by thirty-eight stakeholders to affirm or dispute the effectiveness of the 55 mph speed limit in saving lives. Of this total, 109 hypotheses were unique, in that they did not duplicate hypotheses advanced by any other stakeholder. Here, it is important to note that from the standpoint of communications theory and language, the information content of a hypothesis is inversely related to its relative frequency or probability of occurrence. Hypotheses that are mentioned more frequently—those on which there is greater consensus—have less probative value than rarely mentioned hypotheses, because highly probable or predictable hypotheses do not challenge accepted knowledge claims. The rival hypotheses were analyzed according to the cumulative frequency of unique (nonduplicate) causal hypotheses. As Figure C3.1 shows, the cumulative frequency curve of unique rival hypotheses flattens out after the twenty-second stakeholder. Although the total number of rival hypotheses continues to increase without apparent limit, the boundary of unique rival hypotheses is reached within a small and affordable number of observations. This indicates that a satisfactory definition of the problem has probably been achieved. Indeed, of the 109 unique rival hypotheses, several variables related to the state of the economy—unemployment, the international price of oil, industrial production— explain the rise and fall of traffic fatalities better than average highway speeds and the 55 mph speed limit. (continued)

Evaluating Research on Risk in Mine Safety

and Health

In 1997, a branch of the U.S. Office of Mine Safety and Health Research began a process of strategic planning. The aim of the process was to reach consensus, if possible, on the prioritization of research projects that address different aspects of risk associated with the safety and health of miners. Priority setting in this and other research organizations typically seeks to build consensus under conditions in which researchers, research managers, and external stakeholders use conflicting criteria to evaluate the relative merits of their own and other research projects. Even when reliable and valid data are available—for example, quantitative data from large sample studies of the probabilities of different kinds of mine injuries and deaths—it is often unclear whether "objective" measures of risk, by themselves, provide a sufficient basis for prioritizing research. The difficulty is that extra-scientific as well as scientific factors affect judgments about the relative merits of research on risk.

For example, the probabilities of the occurrence of "black lung" disease and other high-risk conditions have been intensively investigated. Despite the importance of the black lung problem, additional research is not a priority. Accordingly, data on high expected severity (probability ¥ severity) do not alone provide a sufficient basis for prioritizing research problems. Because judgments about research priorities are based on multiple, hidden, and frequently conflicting criteria, it is important to uncover

and evaluate these criteria as part of the process of problem structuring. This is a classic ill-structured problem for which the fatal error is defining the wrong problem. The problem-structuring process had three major objectives. The first was to uncover hidden sources of agreement and disagreement, recognizing that disagreement is an opportunity for identifying alternative approaches to priority setting. Second, the process was designed to generate from stakeholders the criteria they use to evaluate research on risks affecting mine safety and health. Third, the process employed graphs, matrices, and other visual displays in order to externalize the criteria underlying individual 120 100 80 60 R i v a l C a u s e s 40 20 0 1 4 7 10 13 Stakeholders 16 2322 0% 25% 50% 75% P e r c e n t 100% Boundary Estimate FIGURE C3.1 Pareto chart—cumulative frequency of rival causes of traffic fatalities judgments. The priority-setting process enabled each team member to understand and debate the varied reasons underlying the priorities used to evaluate research on risk. Stakeholders were prompted to state the criteria they use to distinguish among twenty-five accident and health problems. Each stakeholder was presented with a form listing pre-randomized sets of three problems. One stakeholder was chosen at random to present criteria to the group. The first presented 14 criteria for evaluating research on risk, which included two criteria—severity and catastrophic potential of accidents and diseases— which were based on large-sample relative frequency (probability) data. The second randomly chosen team member presented 17 additional (new) criteria. The third team member added 11, the fourth 12, the fifth 6, and so on, until no additional criteria could be offered that were not mentioned before. In all, 84 criteria were generated, none after the eleventh presenter. The approximate boundary of the problem was displayed with a Pareto chart to show the cumulative frequency of constructs (Figure C3.2). The process of interactive group priority setting not only addressed the “objective” side of risk research but also captured “subjective” dimensions that go by such labels as “perceived risk,” “acceptable risk,” “researchable risk,” and “actionable risk.” The problem was successfully bounded by means of an open problem-structuring process of generating and discussing criteria that were otherwise tacit, concealed, or implicit. The method of boundary analysis and estimation provides a “stop rule” that avoids the infinite search for problems. 100 80 60 R i s k C r i t e r i a 40 20 0 1 2 3 4 5 6 7 8 9 10 11 12 13 0% 20% 40% 60% P e r c e n t 80% 100% 120% Boundary Estimate Stakeholder FIGURE C3.2 Pareto chart—cumulative frequency of criteria for evaluating research on risk

80 Russell Ackoff, The Art of Problem Solving (New York: Wiley-Interscience, 1978), “Fable 3.2 An Ups-and-Downs Story,” pp. 53–4.

81 Another example of a (probable) type III error is the aphorism popularized by Aaron Wildavsky: “No solution, no problem.” See Wildavsky’s preface to Speaking Truth to Power: The Art and Craft of Policy Analysis (Boston, MA: Little Brown, 1979).

CASE 3.2 BRAINSTORMING ACROSS DISCIPLINES:THE ELEVATOR PROBLEM

This case is based on a story by Russell Ackoff, a master at solving ill-structured problems, or what he called “messes.” 80 The story, a classic of unknown origin that has been in circulation for many years, shows why methods such as benefit-cost analysis, although appropriate for well-structured and moderately structured problems, are not appropriate for ill-structured problems.

The manager of a large office building had been receiving a large and increasing number of complaints about the elevator service, particularly during rush hours. When tenants threatened to move out unless the service was improved, the manager decided that a solution for the problem must be found.

The manager first called on a group of consulting

engineers who specialized in the design and construction of elevators. After looking into the situation, the engineers identified three options: (1) add more elevators, (2) replace the existing elevators with faster ones, or (3) add a computerized control system so that elevator service could be optimally routed for faster service.

In collaboration with an economist, the engineers conducted benefit-cost analyses of the three options. They found that investing in additional elevators, replacement elevators, or a computer control system would yield significant improvements of service. However, they found that the cost of any of the three alternatives was not justified by the monetary and intangible benefits of renting the building. Because none of the alternatives was acceptable, the manager faced a dilemma. He was desperate. He called a meeting of his staff and presented the problem to them as part of a brainstorming session similar to the one described earlier in this chapter. Many suggestions were made and then discarded. The discussion slowed down and stalled. During a break, a young new assistant in the human resources department, who had been silent to this point, quietly made a suggestion that was immediately embraced by everyone present. A few weeks later, the manager made a modest investment and the problem disappeared. Full-length mirrors had been installed on all the walls of the elevator lobbies on each floor. The young psychologist in the human resources department observed that the waiting times were actually quite short. He reasoned that the complaints came from the boredom of waiting for the elevators. The time only seemed long. He gave people something to do: Look at themselves and others (particularly those of the opposite sex) without appearing to do so. This kept them occupied. The consulting engineers and the economist had committed a type III error:They defined the wrong problem. Consequently, they falsely concluded that there was no

solution. 81

4 Chapter 4 Forecasting Expected Policy Outcomes

Dror, Yehezkel. Policymaking under Adversity. New Brunswick, NJ: Transaction Books, 1986.

Finsterbusch, Kurt, and C. P. Wolf, eds. Methodology of Social Impact Assessment. Stroudsburg, PA: Dowden, Hutchinson & Ross, 1977.

Gass, Saul I., and Roger L. Sisson, eds. A Guide to Models in Governmental Planning and Operations. Washington, DC: U.S. Environmental Protection Agency, 1974.

Guess, George M., and Paul G. Farnham. Cases in Public Policy Analysis. New York: Longman, 1989, ch. 3, “Forecasting Policy Options,” pp. 49–67.

Harrison, Daniel P. Social Forecasting Methodology. New York: Russell Sage Foundation, 1976.

Liner, Charles D. “Projecting Local Government Revenue.” In Budget Management: A Reader in Local Government Financial Management. Edited by W. Bartley Hildreth and Gerald J. Miller. Athens: University of Georgia Press, 1983, pp. 83–92.

Linstone, Harold A., and Murray Turoff, eds. The Delphi Method: Techniques and Applications. Reading, MA: Addison-Wesley, 1975.

Marien, Michael. Future Survey Annual: A Guide

to the Recent Literature of Trends, Forecasts,

and Policy Proposals. Bethesda, MD: World

Future Society, published annually.

—— “The Scope of Policy Studies: Reclaiming

Lasswell’s Lost Vision.” In Advances in

Policy Studies since 1950, Vol. 10 of Policy Studies Review Annual. Edited by William N. Dunn and Rita Mae Kelly. New Brunswick, NJ: Transaction Books, 1992, pp. 445–88. McNown, Robert. “On the Use of Econometric Models: A Guide for Policy Makers.” Policy Sciences 19 (1986): 360–80. O’Leary, Michael K., and William D. Coplin. Everyman’s “Prince.” North Scituate, MA: Duxbury Press, 1976. Schroeder, Larry D., and Roy Bahl. “The Role of Multi-year Forecasting in the Annual Budgeting Process for Local Governments.” Public Budgeting and Finance 4, no. 1 (1984): 3–14. Thomopoulos, Nick T. Applied Forecasting Methods. Englewood Cliffs, NJ: Prentice Hall, 1980. Toulmin, Llewellyn M., and Glendal E. Wright, “Expenditure Forecasting.” In Handbook on Public Budgeting and Financial Management. Edited by Jack Rabin and Thomas D. Lynch. New York: Marcel Dekker, 1983, pp. 209–87. Tufte, Edward R. The Visual Display of Quantitative Information. Cheshire, CN: The Graphics Press, 1983. U.S. General Accounting Office. Prospective Evaluation Methods: The Prospective Evaluation Synthesis. Washington, DC: U.S. General Accounting Office, Program Evaluation and Methodology Division, July 1989. The four scatterplots in Figure C4.1 show different patterns among the observations on the graph. Each set of observations of X and Y yields identical correlation coefficients (r), coefficients of determination (r 2), and p-values. CASE 4.1 CONCEPTUAL ERRORS IN REGRESSION ANALYSIS:THE ANSCOMBE QUARTET (a) X 1614121086420 Y 10 9 8 7 6 5 4 3 2 1 0 (b) X 1614121086420 Y 14 12 10 8 6 4 2 0 r = .82 r 2 = .67 p = .001 r = .82 r 2 = .67 p = .001 r = .82 r 2 = .67 p = .001 r = .82 r 2 = .67 p = .001 (c) X 14121086420 Y 12 10 8 6 4 2 0 (d) X 20181614121086420 Y 14 12 10 8 6 4 2 0 FIGURE C4.1 The Anscombe Quartet 183

184

In the case that follows, policy makers responsible

for the performance of the Metropolitan Atlanta

Regional Planning Authority (MARTA) hire

consultants who use forecasting methods to project

future revenues generated by sales and use taxes.

Even with the reasonably accurate estimates of

future revenues from sales tax receipts and user

fees (a reasonably accurate forecast will lie within

;10 percent of the observed value), MARTA policy makers could not avoid making politically unpopular and controversial policy decisions that involved raising fares and cutting back services. In the MARTA case, the accuracy of forecasts is directly related to issues of environmental justice raised primarily by relatively poor and disadvantaged minorities in the Atlanta region. Box 4.1 provides an overview of many of these issues.

CASE 4.2 REVENUE FORECASTING AND

ENVIRONMENTAL JUSTICE BOX 4.1 TRANSIT EQUITY: A LOOK AT MARTA*

Why Is Transportation

Important? Other than housing, Americans spend more on transportation than on any other household expense. The average American household spends one-fifth of its income on transportation.

What Counties Support MARTA? The Metropolitan Atlanta Rapid Transit Authority (MARTA) serves just two counties, Fulton and DeKalb, in the ten-county Atlanta region. In the 1960s, MARTA was hailed as the solution to the region's growing traffic and pollution problems. The first referendum to create a five-county rapid rail system failed in 1968. However, in 1971, the city of Atlanta, Fulton County, and DeKalb County approved a referendum for a 1 percent sales tax to support a rapid rail and feeder bus system. Cobb County and Gwinnett County voters rejected the MARTA system.

Who Pays for MARTA? MARTA's operating budget comes from sales tax (46%), fares (34%), and the Federal Transit Administration and other sources (20%). Only Fulton and DeKalb County residents pay for the upkeep and expansion of the system with a one-cent MARTA sales tax. Revenues from

bus fares generated $5 million more revenue than taken in by rail in 1997. In 1999, the regular one-way fare on MARTA was $1.50, up from $1 in 1992. Who Rides MARTA? A recent rider survey revealed that 78 percent of MARTA's rail and bus riders are African American and other people of color.Whites make up 22 percent of MARTA riders. Where Do MARTA Riders Live? More than 45 percent of MARTA riders live in the city of Atlanta, 14 percent live in the remainder of Fulton County, 25 percent live in DeKalb County, and 16 percent of MARTA riders live outside MARTA's service area. Where Are Weekday MARTA Riders Headed? The majority (58%) of MARTA's weekday riders are on their way to work. The second highest use of MARTA was for getting to medical centers and other services (21%). Other MARTA riders use the system for attending special events (8%), shopping (7%), and school. How Much Is MARTA's Proposed Fare Increase? MARTA proposed raising one-way fares from $1.50 to $1.75, a 17 percent increase. The increase is proposed to offset a $10 million shortfall associated with the openings of the Sandy Springs and North Springs stations. The proposal also calls for increasing the weekly transit pass from $12 to $13 and the monthly pass from $45 to $52.50. Who Would Be Most Impacted by the Proposed MARTA Fare Increase? Although the increase of $7.50 a month may not seem like a lot at first glance, it could do irreparable harm to a $5.25 per hour minimum-wage transit user. These fare increases would fall heaviest on the transitdependent, low-income households and people of color who make up the greatest percentage of MARTA users. How Can the Public Comment on the Proposed MARTA Fare Increase? Because MARTA receives federal transportation dollars, it is required to hold public hearings before any fare increase takes effect. How Has MARTA Grown? MARTA has grown from thirteen rail stations in 1979 to thirty-six rail stations in 2000. Two additional stations (Sandy Springs and North Springs) along the North Line were under construction. These two new northern stations opened in early 2001.With its $270.4 million annual budget, MARTA operates 700 buses and 240 rail cars. The system handles more than 534,000 passengers on an average weekday. MARTA operates 154 bus routes that cover 1,531 miles and carry 275,000 passengers on an average weekday. MARTA's rail lines cover 46 miles with rail cars carrying 259,000 passengers on an average weekday. Who Uses MARTA's Parking Spaces? MARTA provides nearly 21,000 parking spaces at twenty-three of its thirty-six transit stations. Parking at MARTA lots is free except for the overnight lots that cost $3 per day. MARTA provides 1,342 spaces in four overnight lots. All of the overnight lots are on MARTA's North Line. It is becoming increasingly difficult to find a parking

space in some MARTA lots. A recent license tag survey, “Who Parks-and-Rides,” covering the period 1988–97, revealed that 44 percent of the cars parked at MARTA lots were from outside the MARTA Fulton/DeKalb County service area. TABLE 4.14 MARTA Receipts, 1973–2010 Year Receipts 1973 $43,820 1974 $50,501 1975 $50,946 1976 $52,819 1977 $57,933 1978 $66,120 1979 $75,472 1980 $88,342 1981 $99,836 1982 $104,685 1983 $112,008 1984 $123,407 1985 $134,902 1986 $147,149 1987 $148,582 1988 $158,549 1989 $162,543 1990 $165,722 1991 $168,085

Following are statistics on MARTA receipts from sales tax and user fees from 1973 to 2010.

What Are the Similarities

Between Atlanta and Los

Angeles? A similar transit proposal in Los Angles sparked a grassroots movement. In 1996, the Labor Community Strategy Center and the Bus Riders Union (a grassroots group of transit users) sued the Los Angeles MTA over its plan to raise bus fares and build an expensive rail system at the expense of bus riders, who made up 95 percent of transit users. The MTA bus system, used largely by low-income persons and people of color, received only 30 percent of the MTA’s transit dollars. Grassroots organizing and the Bus Riders Union’s legal victory resulted in $1.5 billion for new clean-fuel buses, service improvements, lower fares, a landmark Civil Rights Consent Decree, and a vibrant multiracial grassroots organization of more than 2,000 dues-paying members.

Where Can I Get More

Information on Transportation

Equity? Contact the Environmental Justice Resource Center at Clark Atlanta University, 223 James P. Brawley Drive, Atlanta, GA 30314 (404) 880-6911 (ph), (404) 880-6909 (fx), E-mail: ejrc@cau.edu. Website: http://www.ejrc.cau.edu.

* Prepared by the Environmental Justice Resource Center at Clark Atlanta University under its Atlanta Transportation

Equity Project (ATEP).The ATEP is made possible by grants from the Turner Foundation and the Ford Foundation.

See http://www.ejrc.cau.edu. Year Receipts 1992 $167,016 1993 $181,345 1994 $198,490 1995 $222,475 1996 $229,510

1997 $236,545 1998 $243,579 1999 $277,309 2000 $299,103 2001 $303,562 2002 $284,427 2003 $271,006 2004 $283,381 2005 $307,227 2006 $334,486 2007 $350,526 2008 $349,668 2009 $312,704 2010 $307,525 Source: Guess and Farnham (2000), p. 196; and 2010 Comprehensive Annual Financial Report, Metropolitan Atlanta Regional Planning Authority, 2010. Receipts for 1996 and 1997 are linear interpolations.

5 Chapter 5 Prescribing Preferred Policies

Boardman, Anthony E., David H. Greenberg, Aidan

R. Vining, and David L. Weimer. Cost-Benefit

Analysis: Concepts and Practice. 4th ed. Upper

Saddle River, NJ: Pearson Prentice Hall, 2011.

Fischhoff, Baruch. “Cost-Benefit Analysis and the

Art of Motorcycle Maintenance,” Policy

Sciences 8, (1977): 177–202.

Gramlich, Edward M. Benefit-Cost Analysis of

Public Programs. 2d ed. Englewood Cliffs, NJ:

Prentice Hall, 1990.

Haveman, Robert H., and Julius Margolis, eds.

Public Expenditure and Policy Analysis. 3d ed.

Boston: Houghton Mifflin, 1988. Miller, Roger Le Roy, Daniel K. Benjamin, and Douglass C. North. The Economics of Public Issues. 16th ed. New York: Addison-Wesley, 2010. Mishan, Edward J. Cost-Benefit Analysis. New York: Praeger, 1976. Moore, Mark A., Anthony E. Boardman, Aidan R. Vining, David L. Weimer, and David H. Greenberg “ ‘Just Give Me a Number!’ Practical Values for the Social Discount Rate.” Journal of Policy Analysis and Management, 23, 4 (2004): 789–812.

Conducting a benefit-cost analysis is not only

a technical matter of economic analysis. It is

also a matter of identifying, and if necessary

challenging, the assumptions on which benefit-cost

analysis is based. This can be seen if we examine

the case of the National Maximum Speed Limit

of 1974. Table 5.11 describes steps in conducting a benefit-cost analysis and a critique of the assumptions

underlying the analysis. The case shows, among other things, that all steps in conducting a benefit-cost are sensitive to these assumptions.

CASE 5.1 OPPORTUNITY COSTS OF SAVING

LIVES—THE 55 MPH SPEED LIMIT 41 TABLE 5.11 Measuring the Costs and Benefits of the 55 mph Speed Limit:A Critical Appraisal Steps Critique Costs 1.The major cost of the National Maximum Speed Law (NMSL) was the additional time spent driving as a result of slower speeds. To calculate the number of hours spent driving in 1973, divide the total number of vehicle miles traveled on interstate highways by the average highway speed (65 mph) and then multiply by the average occupancy rate per vehicle, which is approximately 1.77 persons. Why use 1973 mileage without any adjustment? The average growth rate in travel before 1973 was 4 percent. Therefore, the formula should be Next, find the number of hours spent driving in 1974 by dividing total vehicle miles traveled on interstate highways by the average highway speed in 1974 (58 mph).The NMSL caused some people to cancel trips and others to find alternative modes of transportation; as a result, time calculations based on 1974 mileage would be an underestimate. Therefore, we should use the 1973 mileage of 525 million miles. H = a VM 1973 S 1974 VM 1973 S 1973 b ∗ R Using the following formula, where VM is vehicle miles, S is average speed, R is average occupancy rate, and H is the number of hours lost, H = a 1.04VM 1973 S 1974 VM 1973 S 1973 b ∗ R The number of hours lost driving in 1974, based on this equation, is estimated to be 1.72 billion. Using the above formula, the estimated number of hours lost should be 1.95 billion—not 1.72 billion. 2.To estimate the value of this time, begin with the average wage rate for all members of the labor force in 1974—$5.05.The value of one hour’s travel is not $5.05 per hour because very few persons would pay this sum to avoid an hour of travel. We estimate that the people will pay up to 33 percent of their average hourly wage rate to avoid an hour of commuting. The value of time spent traveling is therefore about $1.68 per hour. Why take a percentage of the $5.05 figure based on what commuters would pay to avoid an hour of travel? We should avoid reducing the value of people’s time for two reasons. First, the value of time in cost to society is equal to what society will pay for productive use of that time. Time’s value is not what a commuter will pay to avoid commuting because commuting has other benefits, such as solitude for thinking or the advantages of suburban living. Second, the value of time spent driving for a trucker is many times the industrial wage rate. Discounting would

greatly underestimate the value of commercial drivers. 3. Application of the cost figure ($1.68) to the time lost figure (1.72 billion hours) results in an estimated travel cost of $2.89 billion. Applying the value of one hour's time to the hours lost as calculated above (1.95 billion) results in an estimated travel cost of $9.85 billion. (continued)

4.The NMSL also has some enforcement costs. Total enforcement costs for signs, advertising, and patrolling are about $810,000. Total enforcement cost should be about $12 million—not $810,000. a. New signs were posted. Cost estimates from 25 states for modification of speed limit signs totaled $707,000; for 50 states, this results in an estimated $1.23 million. Spread out over the three-year life of traffic signs, we get an estimate of $410,000. OK. b. The federal government engaged in an advertising campaign encouraging compliance. The Federal Highway Administration's advertising budget for 1974 was $2 million. About 10 percent of this, or $200,000, was spent to encourage compliance with the NMSL. Assume that an additional amount of public service advertising time was donated, for a total of $400,000. Not OK. The Federal Highway Administration does other advertising. Public service advertising estimate also seems low. c. Compliance costs are difficult to estimate. The cost of highway patrols cannot be used because these persons were patrolling highways before the NMSL. Assume that states did not hire additional personnel solely for enforcement of the NMSL. Therefore, we assume that enforcement of the NMSL will not entail any additional costs above enforcement of previous speed limits. Compliance costs pose some problems, but they can be estimated. In 1973, some 5,711,617 traffic citations jumped by 1,713,636 to over 7.4 million. Each additional traffic citation includes an opportunity cost to society. If a law enforcement officer were not issuing traffic tickets, he or she could be solving other crimes. Assuming that it requires 15 minutes for a law enforcement officer to issue a speeding ticket, the total cost of law enforcement is $2.9 million. This figure is based on the average cost of placing a law enforcement officer on the streets at $6.75 per hour. This figure is clearly an underestimate because it does not count time lost waiting to catch speeders. Approximately 10 percent of all speeders will demand a court hearing. Estimating an average of 30 minutes for each hearing and an hourly court cost of $45 results in an additional cost to society of $3.8 million for 171,000 cases. Given the overloaded court dockets, this opportunity cost may be even higher.

Benefits

1.The most apparent benefit of the NMSL is the amount of gasoline saved. The average gasoline economy improves from 14.9 miles per gallon at 65 miles per hour to 16.1 at 58 miles per hour. Use this information to estimate the number of gallons of gasoline saved by traveling at lower speeds. Gallons saved will be calculated by the following formula, where VMT is vehicle miles traveled on interstate highways (not all highways) and MPG is miles per gallon. Why estimate gasoline saved by comparing 1973 and 1974 miles-per-gallon figures in relation to vehicle miles traveled? The federal figures for average miles per hour are estimates based on several assumptions. Given the conflict between industry estimates, Environmental Protection Agency estimates, and Energy Department estimates, any milesper-hour estimate must be considered unreliable. The number of vehicle miles traveled is also based on gallons of fuel sold multiplied by average miles per hour. Hence, this figure is also subject to error. = 3.487 billion gallons G = VM 1973 MPG 1973 VM 1973 MPG 1974 Studies of the efficiency of gasoline engines show that the effect of reducing the average speed of free-flow interstate highways would be to save 2.57 percent of the normal gas used. In 1979, American motorists consumed 106.3 billion gallons of gasoline. Saving 2.57 percent would total 2.73 billion gallons. In 1974, the average price of gasoline was 52.8 cents per gallon. This market price, however, does not reflect the social cost of gasoline, due to government price controls on domestic oil. The marginal (or replacement) cost of crude oil is the price of foreign oil. Therefore, the price of gasoline must reflect the higher cost of foreign oil. Use the market price of gasoline in the absence of price controls, which is about 71.8 cents per gallon. This figure yields an estimate of $2.50 billion in benefits through gasoline saved. Why use the market price? There is no way to determine whether a marginal gallon of gasoline will be imported or come from domestic reserves. In addition, the costs and benefits of the NMSL should not be distorted simply because the U.S. government does not have a market-oriented energy policy. In 1974, gasoline cost 52.8 cents per gallon, and therefore, a gallon of gasoline saved was worth 52.8 cents. 2. A major second-order benefit of the 55 mph limit was a large drop in traffic fatalities, from 55,087 in 1973 to 46,049 in 1974. Part of the gain must be attributable to reduction in traffic speeds. Studies by the National Safety Council estimate that up to 59 percent of the decline in fatalities was the result of the speed limit. Applying this proportion to the decline in fatalities provides an

estimated 5,332 lives saved. The consensus of several studies is that a traffic fatality costs $240,000 in 1974 dollars. Using this figure, the value of lives saved in 1974 is estimated at $1,279.7 million. OK. 3.The NMSL also resulted in a reduction of nonfatal injuries. Use the 59 percent figure found in the fatality studies. Between 1973 and 1974, nonfatal traffic injuries declined by 182,626. Applying the estimated percentages results in 107,749 injuries avoided. Generally, three levels of injuries are indentified: (1) permanent total disability, (2) permanent partial disability and permanent disfigurement, and (3) nonpermanent injury. In 1971, the proportion of traffic injuries that accounted for injuries in each category was 0.2 percent, 6.5 percent, and 93.3 percent, respectively. The National Highway Traffic Safety Administration estimated that in OK. (continued)

Benefits 1971, the average cost of each type of injury was $260,300, $67,100, and $2,465, respectively. The average injury, therefore, cost $8,745 in 1974 dollars. Applying this figure to our injury estimate results in $942.3 million as the social benefit of injury reduction.

4.The final benefit of the reduction in property damage fell from 25.8 million to 23.1 million. About 50 percent of this reduction was the result of lower speeds. The NMSL saved 1.3 million cases of property damage at an average cost of $363.Therefore, the total benefit from property damage prevented is $472 million. OK.

Conclusion

The first estimate of the costs and benefits of the

NMSL results in the following figures (in millions): Using different assumptions, the second estimate of the costs and benefits of the NMSL is as follows (in millions):

Costs Costs Time spent traveling $2,890.0 Time spent traveling $9,848.0 Enforcement .8 Enforcement 12.0 $2,890.8 $9,860.0

Benefits Benefits Gasoline saved $2,500.0 Gasoline save $1,442.0 Lives saved 1.297.7 Lives saved 998.0 Injuries prevented 942.3 Injuries prevented 722.0 Property damage averted 472.0 Property damage adverted 236.0 $5,212.0 $3,398.0 Net benefits: $2,321.2 million Net benefits: -$6,462 million Benefits to costs ratio: 1.8 Benefits to costs ratio: .345

Source: The steps and data were suggested by Charles T. Clotfelter and John C. Hahn, “Assessing the National 55 m.p.h. Speed

Limit,” Policy Sciences 9 (1978): 281–94.The critical comments are based on Charles A. Lave, “The Costs of Going 55,” Car and

Driver. May 1978, p. 12.

6 Chapter 6 Monitoring Observed Policy Outcomes

306

306 CHAPTER 2 Sociology as a Passion to Understand CASE 6.1 UNDERSTANDING POLICY OUTCOMES:

THE POLITICAL ECONOMY OF TRAFFIC

FATALITIES IN EUROPE AND THE UNITED STATES YEAR FAT BIMI UEMR EMP POP IIP 1966 50894 926 3.2 72895000 195578000 32.50 1967 50724 964 3.1 74372000 197457000 34.30 1968 52725 1016 2.9 75920000 199399000 35.40 1969 53543 1062 2.8 77902000 201385000 37.30 1970 52627 1110 4.4 78627000 203984000 36.90 1971 52542 1179 5.3 79120000 206827000 36.50 1972 54589 1260 5.0 81702000 209284000 39.00 1973 54052 1313 4.2 84409000 211357000 43.00

The 55 mph speed limit was adopted in 1974 as a

means to reduce gasoline consumption during the

1973–74 OPEC oil embargo. Unexpectedly, the

policy was followed by a sharp decline in traffic

fatalities. Between January 1, 1974, and December

31, 1975, there was a decline of 9,100 fatalities, a

16.8 percent drop. Despite continuing opposition

from rural Western states and the trucking industry,

there was broad political support for the policy. And,

clearly, the problem is not unimportant or trivial.

Traffic fatalities represent the leading cause of death

among persons age 35 and younger. Average annual

highway deaths since 1974 are equivalent to a fully

loaded 767 aircraft crashing with no survivors every

third day of the week.

On April 2, 1987, Congress enacted the Surface

Transportation and Uniform Relocation Assistance Act of 1987, overriding President Reagan's veto. Provisions of this bill permitted individual states to experiment with speed limits up to 65 mph on rural interstate highways. By July 1988, forty states had raised the speed limit to 60 or 65 mph on 89 percent of rural interstate roads. Senator John C. Danforth, an influential supporter of the 55 mph speed limit, argued against the new policy. The 65 mph speed limit, said Danforth, would save an average of one minute per day per driver but result in an annual increase of 600 to 1,000 deaths. The Washington Post joined the opponents: "The equation is in minutes versus lives. It's not even close. . . . A hundred miles at 55 mph take about 17 minutes longer than at 65. That's the price of those lives. It ought to be the easiest vote the House takes this year." Eight years later, in a November 1995 press release, U.S. Secretary of Transportation Federico Pena reaffirmed the Clinton administration's opposition to the higher speed limits (see Chapter 1, Case 1).The National Maximum Speed Limit was officially abandoned the same year. Many analysts who have evaluated the effects of the intervention, along with elected officials from the ten northeastern states that retained the 55 mph speed limit until its repeal, affirm that the 1974 law was responsible for the decline in traffic fatalities. However, the evidence suggests that they may have failed to consider rival hypotheses—which had they been identified and tested would have resulted in a different explanation of traffic deaths and, consequently, a different policy recommendation. Here are data on U.S. traffic fatalities and other variables for 1966 to 2010 (Exhibit 6.3). In addition, fatality rates in the United States and European countries are provided in a separate listing (Exhibit 6.4). Use the data for the demonstration exercise. YEAR FAT BIMI UEMR EMP POP IIP 1974 45196 1281 4.9 85936000 213342000 44.40 1975 44525 1328 7.9 84783000 215465000 40.60 1976 45523 1402 7.1 87485000 217561000 42.20 1977 47878 1467 6.3 90546000 219758000 44.70 1978 50331 1545 5.3 94373000 222093000 47.00 1979 51093 1529 5.1

96945000 224569000 50.50 1980 51091 1527 6.9 97722000
227255000 51.10 1981 49301 1553 7.4 98878000 229637000
50.50 1982 43945 1595 9.9 98047000 231996000 48.70 1983
42589 1653 9.9 99347000 234284000 47.70 1984 44257 1720 7.4
103413000 236477000 53.00 1985 43825 1774 7.0 105613000
237924000 54.40 1986 46087 1835 6.9 107860000 240133000
55.70 1987 46390 1921 6.2 110634000 242289000 55.90 1988
47087 2026 5.5 113153000 244499000 60.10 1989 45582 2096
5.2 115453000 246819000 61.70 1990 44599 2144 5.6 116660000
249403000 61.20 1991 41508 2172 7.0 115657000 252138000
60.70 1992 39250 2247 7.8 116395000 255039000 60.90 1993
40150 2297 7.2 118149000 257800000 63.90 1994 40716 2360
6.1 121999000 260350000 66.10 1995 41817 2423 5.6 124766000
262755000 70.40 1996 41907 2469 5.4 125311000 265284000
71.60 1997 42013 2560 4.9 129558000 267744000 76.50 1998
41471 2619 4.5 131463000 270299000 82.90 1999 41611 2691
4.2 133488000 272691000 86.10 2000 41945 2750 4.0 135208000
282172000 90.20 2001 42196 2797 4.7 136933000 285082000
90.70 2002 43005 2856 5.8 136485000 287804000 86.80 2003
42884 2890 6.0 137736000 290326000 89.50 2004 42836 2965
5.5 139252000 293046000 90.40 2005 43510 2989 5.1 141730000
295753000 94.10 2006 42708 3014 4.6 144427000 298593000
95.50 2007 41059 3030 4.6 146047000 301580000 97.50 2008
37423 2926 5.8 145362000 304375000 99.80 2009 33808 2991
9.3 139877000 307007000 87.40 2010 31000 2980 9.6 132000000
308400408 87.60 Note: FAT = fatalities; BIMI = billions of
miles traveled; UEMR = unemployment rate; EMP = employment;
POP = population; IIP = Index of Industrial Production.
YEAR US EUREXP EURC UK FIN FR DEN 1970 25.8 21.9 24.3 14.3
24.8 23.5 25 1971 25.4 22.8 24.7 14.8 25.9 25.4 25 1972
26.1 22.8 25.0 14.6 25.2 27.4 24 1973 25.6 22.1 23.5 14.7
23.8 25.6 24 1974 21.2 17.4 21.3 13.0 18.0 22.5 16 1975
20.1 17.6 20.2 11.9 19.0 24.0 17 1976 20.9 18.3 18.8 12.0
20.0 26.0 17 Note: EUREXP = European “experimental” states
that adopted a 90 kph (54 mph) speed limit in 1972; EURCON
European “control” states. Other names refer to countries.
EXHIBIT 6.4

7 Chapter 7 Evaluating Policy Performance

Bok, Sissela. Lying: Moral Choice in Public and Private Life. New York: Pantheon Books, 1978.

Dasgupta, Partha. “Valuation and Evaluation: Measuring the Quality of Life and Evaluating Policy.” Background paper prepared for the World Bank’s annual World Development Report (2000).

Mac Rae, Duncan, Jr. The Social Function of Social Science. New Haven, CT: Yale University Press, 1976. Nozick, Robert. Anarchy, State, and Utopia. New York: Basic Books, 1974. Rawls, John. A Theory of Justice. Cambridge, MA: Harvard University Press, 1971. Sen, Amartya. Inequality Reexamined. Oxford: Clarendon Press, 1992. Stone, Deborah. Policy Paradox: The Art of Political Decision Making. Rev ed. New York: W. W. Norton, 2002. CASE 7.1 THE ECONOMICS OF MORAL JUDGEMENT: EVALUATING LIVING WAGE POLICIES The living wage movement is considered by many to be the most important grassroots political mobilization since the U.S. civil rights movement. Issues surrounding the adoption of a living wage are economic as well as moral and ethical. Economists are divided on the effects of the living wage on poverty, inequality, and employment; the majority of religious and community leaders have supported the living wage for more than 100 years. The following calculations have been extracted from a website (www.livingwage.geog.psu.edu) developed by Dr. Amy K. Glasmeier of the Massachusetts Institute of Technology. The living wage shown is the hourly rate in 2011 that a full-time employed person (2,080 hours per year) in New York City (Queens County) must earn to support families of different sizes. In New York City, the minimum wage of $7.25 per hour is the same for all persons, regardless of the number of dependents. Although the poverty rate is typically quoted as gross annual income, for purposes of comparison it has been converted into an hourly wage. Wages that are less than the living wage are italicized. Municipalities such as San Francisco, Detroit, and Chicago set the minimum wage higher than the federally mandated 2011 minimum wage of $7.25. The minimum living wage for an adult with one child and larger families also differs. In Pittsburgh, the living wage for an adult with one child is $34,244 per year, whereas in Los Angeles it is $45,235. Hourly Wages One Adult One Adult, One Child Two Adults Two Adults, One Child Two Adults, Two Children Living Wage $11.86 $19.66 $16.29 $24.10 $30.30 Poverty Wage $5.04 $6.68 $6.49 $7.81 $9.83 Minimum Wage $7.25 $7.25 $7.25 $7.25 $7.25 Typical Expenses These figures were used to calculate

the 2011 living wage in New York City. Numbers vary by family size. For typical expenses, living wages, and salaries for different occupations across the United States, see www.livingwage.geog.psu.ed. 335 Monthly Expenses One Adult One Adult, One Child Two Adults Two Adults, One Child Two Adults, Two Children Food $232 $378 $448 $594 $740 Child Care $0 $572 $0 $572 $1,012 Medical $76 $151 $152 $227 $302 Housing $1,185 $1,318 $1,185 $1,318 $1,318 Transportation $232 $397 $464 $629 $794 Other $188 $369 $376 $557 $738 Monthly AfterTax Income That's Required $1,913 $3,185 $2,625 $3,897 $4,904 Annual AfterTax Income That's Required $22,956 $38,220 $31,500 $46,764 $58,853 Annual Taxes $1,706 $2,678 $2,387 $3,373 $4,161 Annual Before Tax Income That's Required $24,662 $40,898 $33,887 $50,137 $63,014 Source: © 2011 Dr. Amy K. Glasmeier and The Pennsylvania State University Site created by West Arete Computing

Roe, E. Narrative Policy Analysis. Durham, NC:

Duke University Press, 1994.

Scriven, M. Reasoning. New York: McGraw-Hill, 1976.

Stone, D. Policy Paradox: The Art of Political

Decision Making. Rev Ed. New York: W. W.

Norton, 2002. Toulmin, S., R. Rieke, and A. Janik. An Introduction to Reasoning. 2d ed. New York: Macmillan, 1984. Van Gelder, Tim. “The Rationale for Rationale.” Law, Probability & Risk 6, nos. 1-4 (2007): 23-42.

CASE 8.1 PROS AND CONS OF BALKAN

INTERVENTION 58

“Must the agony of Bosnia-Herzegovina be

regarded, with whatever regrets, as somebody else’s

trouble? We don’t think so, but the arguments on

behalf of that view deserve an answer. Among them

are the following:

The Balkan conflict is a civil war and unlikely to

spread beyond the borders of the former

Yugoslavia.Wrong. Belgrade has missiles trained

on Vienna. Tito’s Yugoslavia claimed, by way of

Macedonia, that northern Greece as far south as

Thessaloniki belonged under its sovereignty.

Those claims may return. “Civil” war pitting

non-Slavic Albanians against Serbs could spread

to Albania,Turkey, Bulgaria, and Greece.

The United States has no strategic interest in the

Balkans. Wrong. No peace, no peace dividend.

Unless the West can impose the view that ethnic

purity can no longer be the basis for national

sovereignty, then endless national wars will

replace the Cold War. This threat has appeared

in genocidal form in Bosnia. If it cannot be

contained here, it will erupt elsewhere, and the

Clinton administration's domestic agenda will be

an early casualty.

If the West intervenes on behalf of the Bosnians,

the Russians will do so on behalf of the Serbs, and the Cold War will be reborn. Wrong. The Russians have more to fear from "ethnic cleansing" than any people on Earth. Nothing would reassure them better than a new, post-Cold War Western policy of massive, early response against the persecution of national minorities, including the Russian minorities found in every post-Soviet republic. The Russian right may favor the Serbs, but Russian self-interest lies elsewhere. The Serbs also have their grievances. Wrong. They do, but their way of responding to these grievances, according to the State Department's annual human rights report, issued this past week,"dwarfs anything seen in Europe since Nazi times." Via the Genocide Convention, armed intervention is legal as well as justified. The UN peace plan is the only alternative. Wrong. Incredibly, the plan proposes the reorganization of Bosnia-Herzegovina followed by a cease-fire. A better first step would be a UN declaration that any nation or ethnic group proceeding to statehood on the principle of ethnic purity is an outlaw state and will be treated as such. As now drafted, the UN peace plan, with a map of provinces that not one party to the conflict accepts, is really a plan for continued 'ethnic cleansing.'" 58 From the Los Angeles Times, January 24, 1993, p. A7. CASE 8.2 IMAGES,ARGUMENTS,AND THE SECOND PERSIAN GULF CRISIS, 1990–91 The analysis of policy arguments can be employed to investigate the ways that policy makers represent or structure problems (Chapter 3).We can thereby identify the images, or problem representations, that shape processes of making and

justifying decisions. For example, during times of crisis, the images that U.S. policy makers have of another country affect deliberations about the use of peacekeeping and negotiation, the imposition of economic sanctions, or the use of deadly force. This case looks at the deliberations surrounding the U.S. decision to use military force to produce an Iraqi withdrawal from Kuwait during the Second Persian Gulf Crisis of 1990–91. It is important to recognize that there have been three Persian Gulf crises since 1980.The First Persian Gulf Crisis, which involved eight years of war between Iraq and Iran, lasted from September 1980 to August1988. An estimated half-million civilians and military combatants died in the conflict. The Third Gulf Crisis began in 2003 and appeared to be coming to a conclusion in 2011, as occupying U.S. and coalition troops continued a steady withdrawal from Iraq. Between 2003 and 2010, it is estimated that between 700,000 and 1.5 million Iraqi civilians died as a direct and indirect result of the war, and some 11,800 members of the Iraqi security forces and police were killed. 59 There were nearly 4,500 U.S. military deaths and another 220 among coalition forces, about one-half of whom were British. The Second Persian Gulf Crisis may be conveniently dated to August 2, 1990, when Iraq invaded Kuwait. Seven months later, on March 3, 1991, a cease-fire was signed. 60 When Iraq invaded Kuwait, the United States began sending troops and Patriot missiles to defend Saudi Arabia and Israel. The purpose was to defend against Iraqi SCUD missile attacks on Riyadh, Tel Aviv, and Haifa. An economic embargo was also put in place along with a U.S.-UN Security Council demand for an immediate withdrawal from Kuwait. Over the next five months, Saddam Hussein failed to comply. At that point, a Security Council resolution was passed giving Iraq until January 15, 1991, to withdraw. The resolution authorized intervention by a U.S.-led coalition after that date. Prior to the deadline for the withdrawal from Kuwait, the U.S. Senate debated the merits of U.S. intervention. Two policy alternatives were presented. One of these, the Dole-Warner resolution, supported the Bush administration and the United Nations by authorizing the use of military force to compel a withdrawal. The other resolution, the Mitchell-Nunn resolution, called for the continued use of sanctions. The vote in favor of the Dole-Warner resolution was 52–47. Political scientist Richard Cottam and cognitive political psychologist James Voss show how the images 59 Hannah Fischer, “Iraq Casualties: U.S. Military Forces and Iraqi Civilians, Police, and Security Forces.” U.S. Congressional Research Service Report (June 11, 2010). 60 Earlier, on February 15, the Shiite cleric Muqtada al-Sadr led a revolt against Saddam Hussein’s regime, partly in

response to President George H. W. Bush's earlier call on February 2 for Iraqis to overthrow Hussein's government. In the same month, the Kurds in Northern Iraq also rose up in an attempt to overthrow Saddam's regime. 377

of another state can be investigated by focusing on indicators of the perceived motivation, capability, culture, and decision processes of that state: 61

Ally Image. The image of the ally is known by finding in foreign policy discourse indicators referring to the mutually beneficial goals of the ally, to the ally's adequate but often less than possible military and economic capabilities, the ally's comparably civilized culture, and the ally's well-managed and popularly supported process of governmental decision making.

Enemy Image. The image of the enemy state is known by finding in foreign policy discourse indicators referring to the aggressive, evil, and expansionistic motivations of the enemy; to the enemy's comparable but penetrable military and economic capabilities; to the enemy's comparably civilized culture; and to the enemy's monolithic and undemocratic process of governmental decision making.

Radical Image. The image of the radical state is known by finding in foreign policy discourse indicators referring to fanatic or

extremist motivations, to the radical state's use of terror to compensate for its inferior military and economic capabilities, to the radical state's less than civilized culture, and to a decision process that is well organized and clever.

Degenerate Image. The image of the degenerate state is known by finding in foreign policy discourse indicators referring to leaders motivated by the accumulation and preservation of personal power, to the degenerate state's inferior and declining military and economic capabilities, to the degenerate state's less than civilized culture, and to a confused and disorganized decision process. Imperial Image. The image of the imperial state is known by finding in foreign policy discourse indicators referring to motivations based on nationalism and modernization in the interests of the people, to the imperial state's need for external help because of inferior military and economic capabilities, to the imperial state's less than civilized culture, and to poorly managed decision processes that require military and economic assistance. In the Second Persian Gulf Crisis, Senate debates were held on January 10, 11, and 12 of 1991.The arguments put forth in the debates contain indicators of the images just described. 62 For example, one supporter of the Dole-Warner resolution, whose position was nearly identical to that of the White House, was Senator Bryan (D-Nebraska). He argued in terms indicating a traditional enemy image of Iraq: "Hussein now has under arms more men than Hitler when the German Army marched into the Rhineland . . . more tanks than when the Panzer Divisions crushed France . . . and most chilling of all, much closer to having a nuclear weapon than Adolf Hitler ever was." This is not the image of a radical or terrorist state, but that of a traditional aggressive and expansionistic state. Senator Roth (D-Delaware) reinforced

this image: “Hussein has demonstrated that with the Cold War fading, the real threat to freedom-loving nations is the proliferation of arms in the hands of despotic dictators. Intercontinental missiles, chemical, biological, and nuclear arms can turn unstable Third World nations into first-rate military powers.” By contrast, Senator Cohen (R-Maine), another supporter of Dole-Warner, used an argument suggesting the image of a radical or terroristic state: “Not one of us . . . is safe from the violence currently being inflicted in Kuwait . . . Our 61 The images are based on Richard W. Cottam, Foreign Policy Motivation: A General Theory and a Case Study (Pittsburgh, PA: Pittsburgh University Press, 1977). The descriptions of images and use of quotations are based on Voss et al., “Representations of the Gulf Crisis as Derived from the U.S. Senate Debate,” pp. 279–302 in Donald A. Sylvan and James F. Voss, eds. Problem Representation in Foreign Policy Decision Making (Cambridge: Cambridge University Press, 1998). 62 All statements were extracted from The Congressional Record (January 9, 10,11) by Voss et al., “Representations of the Gulf Crisis as Derived from the U.S. Senate Debate.” The Senate should authorize the use of force under the DoleWarner resolution. CLAIM (52) The Dole-Warner resolution supports U.N. Security Council Resolution 678 of November 29, 1990 authorizing U.S. military intervention to force an Iraqi withdrawal from Kuwait. INFORMATION (98) Since Iraq has a large military (37), regional political aspirations (32), and petroleum reserves that give economic control (16). Because Iraq is a threat to the security of the region and the world. WARRANT (51) Because the U.S. has an obligation to intervene. WARRANT (16) Immediately and without hesitation. BACKING (16-37) Since the U.S. is a member of the U.N., committed to Israel and moderate Arab states, and opposed to human rights violations. BACKING (16) QUALIFIER (30) But the authorization is unwarranted. It is an artificial creation of the President and no vital interests are at stake. OBJECTION (37) But sanctions are working. GNP is cut in half and imports and exports have been reduced by 90–97%. OBJECTION (47) But available diplomatic options have not been pursued. OBJECTION (33) However, sanctions have not produced a withdrawal. They will cause economic problems for allies and cause the suffering of innocent persons. REBUTTAL (34) However, diplomacy has not worked in the past and will not work now. REBUTTAL (9) FIGURE C8.2 Senate arguments supporting and opposing U.S. involvement under security council resolution 678 authorizing military intervention to force the withdrawal of Iraq from Kuwait (November 29, 1990) security is only a Pan Am 103 away at any moment.” Senator Hatch (R-Utah) used similar language:

“We have a major political interest in preventing Hussein from radicalizing the Arab world.” Among the supporters of the Mitchell-Nunn resolution were Senator Joseph Biden (D-Delaware), who argued: “Yes, we have interests in the Middle East. We wish to support the free flow of oil. We wish to promote stability, including the securing of Israel. But we have not heard one cogent argument that any vital American interest is at stake in a way that impels us to war.” As for military capability, Senator Patrick Moynihan (D-New York) argued that “The Iraqis do not (even) have the technology to print their own paper money.” In support of the continued use of economic sanctions, Senator Frank Lautenberg (D-New Jersey) argued:“Historical analysis of the use of economic sanctions suggests that they can be effective over time in forcing the withdrawal of Iraqi troops from Kuwait.” Figure C8.1 maps the arguments supporting the resolution sponsored by Bob Dole (R-Kansas) and John Warner (R-Virginia).The figure also maps the objections of senators favoring the counter-resolution sponsored by George Mitchell (D-Maine) and Sam Nunn (D-Georgia). Rebuttals to objections are also displayed. The numbers in parentheses are the numbers of votes for the different parts of the argument supporting military intervention. The votes were along party lines, except for two Democrats who voted for the Dole-Warner Resolution. Because two senators were not present for the vote, the total number is 98 votes.

The Senate debate includes arguments that

appear to be the consequence of at least three

images of Iraq. Arguments supporting the Dole

Warner Resolution authorizing a military attack

appear to be a consequence of images of Iraq as a

radical-terrorist state, or a degenerate state.

In turn, arguments offered in support of the continuation of economic sanctions appear to originate in an image of Iraq as a conventional enemy state such as Germany during World War II. The main point is that images drive problem representations, which in turn drive processes of argumentation surrounding critical issues of foreign and domestic policy.

9 Chapter 9 Communicating Policy Analysis

CHAPTER SUMMARY knowledge and skills in developing policy documents and giving oral briefings. To be effective, analysts need to master and apply a broad range of communications skills, thereby narrowing the present divide between intellectual and political dimensions of policy making.

This concluding chapter has provided an

overview of the process of policy communi

cation and its importance to the use of analy

sis by policy makers. Policy analysis is the

beginning, not the end, of efforts to improve

policy making. Knowledge and skills in con

An important skill is the translation of policy

arguments into structured written narratives.These

written narratives go by different names: white

papers, issue papers, memos, press releases, and

letters. All are used to communicate the results of

analysis in hopes of maximizing the likelihood that

they will actually be used to make and influence

decisions.

CASE 9.1 TRANSLATING POLICY

ARGUMENTS INTO ISSUE PAPERS, POSITION

PAPERS,AND LETTERS TO THE EDITOR—THE

CASE OF THE SECOND PERSIAN GULF WAR

INFORMATION (98)

The Dole-Warner resolution/November 29,

1990 supports UN Security Council

Resolution 678 authorizing U.S. military

intervention to force an Iraqi withdrawal

from Kuwait. WARRANT (51) Because Iraq is a threat to the security of the region and the world. BACKING (16-37) Since Iraq has a large military (37), regional political aspirations (32), and petroleum reserves that give economic control (16). OBJECTION (37) But the authorization is unwarranted. It is an artificial creation of the president and no vital interests are at stake. WARRANT (16) Because the U.S. has an obligation to intervene. BACKING (16) Since the U.S. is a member of the UN, committed to Israel and moderate Arab states, and opposed to human rights violations. OBJECTION (47) But sanctions are working. GNP is cut in half and imports and exports have been reduced by 90% to 97%. REBUTTAL (34) However, sanctions have not produced a withdrawal. They will cause economic problems for allies and cause the suffering of innocent persons. OBJECTION (33) But available diplomatic options have not been pursued. REBUTTAL (9) However, diplomacy has not worked in the past and will not work now. QUALIFIER (30) Immediately and without hesitation. CLAIM (52) The Senate should authorize the use of force under the DoleWarner resolution.

This case draws on the subprogram in Rationale for creating structured narratives from argument maps.

400 opposing points of view, objections, or rebuttals are included. Position papers state one side of an argument as persuasively as possible. The policy brief. The policy brief is similar to a legal brief. It states the reasons for a claim, followed by objections, which are rebutted. Issue papers and memos also state opposing arguments and then show they are not plausible or valid. The policy issue paper or policy memo. The policy issue paper or policy memo (a shortened version of the policy issue paper) presents both sides of an issue equally, without taking sides. Reasons and objections are provided to dispute issues. The letter to the editor. The letter to the editor begins by considering an objection to the claim, rebutting it, and then providing multiple reasons to support the claim.The letter to the editor is an example of a fortiori analysis described in Chapter 5. In each of the following written documents the argument map on the Dole-Warner resolution has been translated into a draft version of a structured written narrative.The statements in bold letters are taken directly from the argument map.The documents would require

editing before release. Argument Map for a Position Paper Position Paper on the U.S. Invasion of IRAQ There are three reasons to believe that the Senate should authorize the use of force under the Dole-Warner resolution. First, the Dole-Warner resolution is justified by UN Security Council Resolution 678 of November 29, 1990. This is evident given that Resolution 678 authorizes U.S. military intervention in Kuwait. Further evidence is that Resolution 678 will force an Iraqi withdrawal from Kuwait. Second, Iraq is a security threat to the region and the world.This is supported by the reason that Iraq has a large military and regional political aspirations. Furthermore, Iraq has petroleum reserves that give it economic control. Finally, the U.S. has an obligation to intervene. Evidence for this is that the U.S. is a member of the UN and governed by the UN Charter. In addition, The U.S. is committed to the defense of Israel and moderate Arab states, which also shows that the U.S. has an obligation to intervene. Given these considerations, it is clear that the Senate should authorize the use of force under the Dole-Warner resolution. because because because because because because because because because because CLAIM The Senate should authorize the use of force under the DoleWarner resolution. WARRANT The Dole-Warner Resolution is justified by UN Security Council Resolution 678 of November 29, 1990. WARRANT The U.S. has an obligation to intervene. WARRANT Iraq is a security threat to the region and the world. BACKING Resolution 678 authorizes U.S. military intervention in Kuwait. BACKING Resolution 678 will force an Iraqi withdrawal from Kuwait. BACKING Iraq has a large military and regional political aspirations. BACKING Iraq has petroleum reserves that give it economic control. BACKING The U.S. is a member of the UN and governed by the UN Charter. BACKING The U.S. is committed to the defense of Israel and moderate Arab states.

Argument Map for a Policy Brief

Issue Brief on the U.S. Invasion of IRAQ

The Senate should authorize the use of force under the Dole

Warner resolution.There are two main reasons

supporting this contention, and the main objection to

it can be rebutted.

The first reason is that the Dole-Warner resolution is

justified by UN Security Council Resolution 678 of November 29,

1990. This reason is based on the claim that Resolution

678 authorizes U.S. military intervention in Kuwait.

The second reason that the Senate should authorize the

use of force under the Dole-Warner resolution is that the U.S.

has an obligation to intervene.This is because the U.S. is

committed to the defense of Israel and moderate Arab states.

On the other hand, a consideration against the

idea that the Senate should authorize the use of force under the

Dole-Warner tesolution is that fiplomatic options are available

and have not been pursued. This objection is not

convincing, however, because diplomacy has not worked in

the past with Iraq and will not work now.

Based on this reasoning, it is clear that the Senate

should authorize the use of force under the Dole-Warner

resolution. Argument Map for a Policy Issue Paper or Policy Memo Policy Memo on the U.S. Invasion of IRAQ The question has been raised whether the Senate should authorize the use of force under the Dole-Warner resolution.This analysis considers a number of perspectives to determine whether to support or reject this claim. One reason to accept that the Senate should authorize the use of force under the Dole-Warner resolution is the suggestion that the Dole-Warner resolution is justified by UN Security Council Resolution 678 of November 29, 1990.This reason is supported by the fact that Resolution 678 authorizes U.S. military intervention in Kuwait. An objection to the claim that the Dole-Warner resolution is justified by UN Security Council Resolution 678 of November 29, 1990, is that the resolution is an artificial creation of the president and no vital interests are at stake. Another reason to believe

that the Senate should authorize the use of force under the Dole-Warner resolution is that Iraq is a security threat to the region and the world. This is based on the idea that Iraq has a large military and regional political aspirations. The Senate should authorize the use of force under the Dole-Warner resolution. The Dole-Warner resolution is justified by UN Security Council resolution 678/November 29, 1990. But diplomatic options are available and have not been pursued. The U.S. has an obligation to intervene. The U.S. is committed to the defense of Israel and moderate Arab states. However, diplomacy has not worked in the past with Iraq and will not work now. CLAIM WARRANT OBJECTION
WARRANT BACKING REBUTTAL Resolution 678 authorizes U.S. military intervention in Kuwait. BACKING Nonetheless, the opposing view that sanctions are working and military action is not necessary provides an interesting perspective. On balance, however, it seems more reasonable to reject this argument. On the other hand, a consideration against the claim that the Senate should authorize the use of force under the Dole-Warner resolution is that the authorization is unwarranted. This objection is supported by the claim that the U.S. is a sovereign nation and should not be subservient to the UN. Nevertheless, the rebuttal that the resolution is supported by a 51-47 vote of the U.S. Senate must be considered to determine whether the objection is indeed acceptable. Finally, there is another significant objection to the claim that the Senate should authorize the use of force under the Dole-Warner resolution. Diplomatic options are available and have not been pursued. To support this view is the reason that diplomacy should be given a chance. In rebuttal, it is suggested that diplomacy has not worked in the past with Iraq and will not work now. In summary, it would appear more reasonable to accept the claim that the Senate should authorize the use of force under the Dole-Warner resolution.The strongest line of reasoning for this view is that Iraq is a security threat to the region and the world. Iraq has a large military and regional political aspirations. Argument Map for a Letter to the Editor Letter to the Editor on the U.S. Invasion of IRAQ Dear Sir / Madam, I am writing to contend that the Senate should authorize the use of force under the Dole-Warner resolution. This is contrary to the view voiced by the Senate minority, which has argued that sanctions are working and military action is not necessary. I reject this claim because diplomacy has not worked in the past with Iraq and it will not work now. Furthermore, there are three strong reasons to support my view that the Senate should authorize the use of force under the Dole-Warner

resolution. The first reason is that the Dole-Warner resolution is justified by UN Security Council Resolution 678 of November 29, 1990, since Resolution 678 authorizes U.S. military intervention in Kuwait. In addition, this reason is supported by the fact that Britain and other key allies support Resolution 678. The second reason is that Iraq is a security threat to the region and the world. WARRANT The Dole-Warner Resolution is justified by UN Security Council resolution 678 of November 29, 1990. WARRANT Iraq is a security threat to the region and the world. BACKING Resolution 678 authorizes U.S. military intervention in Kuwait. OBJECTION The resolution is an artificial creation of the president and no vital interests are at stake. BACKING Iraq has a large military and regional political aspirations. OBJECTION Sanctions are working and military action is not necessary. WARRANT The U.S. is a sovereign nation and should not be subservient to the UN. REBUTTAL The resolution is supported by a 51--47 vote of the U.S. Senate. WARRANT Diplomacy should be given a chance. REBUTTAL However, diplomacy has not worked in the past with Iraq and will not work now. OBJECTION The authorization is unwarranted. WARRANT Diplomatic options are available and have not been pursued. CLAIM The Senate should authorize the use of force under the DoleWarner resolution.

Finally, it is clear that the Senate should authorize the

use of force under the Dole-Warner resolution, given that Iraq

has large military and regional political aspirations. This reason

is supported by the contention that Saddam Hussein has

threatened to remain in Kuwait. I restate my contention that the Senate should authorize the use of force under the Dole-Warner resolution. I urge you to also adopt this view. Yours sincerely, [Your name]

Policy analysts must deal with many kinds

of empirical evidence. Often the constraints

of time and resources, as well as the nature

of the policy problem under consideration, force

analysts to rely on qualitative and fragmentary

data. Sometimes, however, analysts can find data

that enable them to estimate the magnitudes because because because because because because because because CLAIM The Senate should authorize the use of force under the DoleWarner resolution. WARRANT The Dole-Warner Resolution is justified by UN Security Council Resolution 678 of November 29, 1990. but OBJECTION Sanctions are working and military action is not necessary. because REBUTTAL Diplomacy has not worked in the past with Iraq and it will not work now. WARRANT Iraq has a large military and regional political aspirations. WARRANT Iraq is a security threat to the region and the world. BACKING Resolution 678 authorizes U.S. military intervention in Kuwait. BACKING Britain and other allies support resolution 678. BACKING Saddam Hussein has threatened to remain in Kuwait.

CASE 9.2 COMMUNICATING STATISTICAL

ANALYSES TO MULTIPLE AUDIENCES:

REGULATING LEADED GASOLINE

When Statistics Count: Revising the EPA Lead Standard 20 of the effects of policy interventions on the social, economic, or political conditions.The estimates may permit the analysts to calculate the likely net social benefits of proposed policies or even apply formal optimization techniques to find better alternatives. 20 From David Weimer and Aidan Vining, Policy Analysis: Concepts and Practice, 2d ed. (Englewood Cliffs, NJ: Prentice Hall 1992), Ch. 13, pp. 382–406. Footnotes 21 through 56 below are from Chapter 13 of Weimer and Vining. The effective use of quantitative data requires an understanding of the basic issues of research design and a facility with the techniques of statistical inference. Even when analysts do not have the resources available to do primary data analysis, they often must confront quantitative evidence produced by other participants in the policy process or extracted from the academic literature. If they lack the requisite skills for critical evaluation, they risk losing their influence in the face of quantitative evidence that may appear more objective and scientific to decision makers.Well-trained analysts, therefore, need some facility with the basic concepts of statistical inference for reasons of self-defense, if for no others. The basics of research design and statistical inference cannot be adequately covered in an introductory course in policy analysis. Similarly, we cannot provide adequate coverage in

this book.Yet we can provide an example of a policy change where quantitative analysis was instrumental: the decision by the U.S. Environmental Protection Agency in 1985 to reduce dramatically the amount of lead permitted in gasoline.The story of the new lead standard has a number of elements frequently encountered when doing quantitative analysis in organizational settings: replacement of an initial “quick and dirty” analysis with more sophisticated versions as more time and data become available; repeated reanalysis to rule out alternative explanations (hypotheses) offered by opponents of the proposed policy; and serendipitous events that influence the analytical strategy. Although our primary purpose is to tell what we believe to be an intrinsically interesting and instructive story about the practice of policy analysis, we hope to offer a few basic lessons on quantitative analysis along the way. Background:The EPA Lead Standards The Clean Air Amendments of 1970 give the administrator of the Environmental Protection Agency authority to regulate any motor fuel component or additive that produces emission products dangerous to public health or welfare. 21 Before exercising this authority, however, the administrator must consider “all relevant medical and scientific evidence available to him” as well as alternative methods that set standards for emissions rather than for fuel components. 22 In 1971 the EPA administrator announced that he was considering possible controls on lead additives in gasoline. 23 One reason given was the possible adverse health effects of emissions from engines that burned leaded gasoline, the standard fuel at the time. Available evidence suggested that lead is toxic in the human body, that lead can be absorbed into the body from ambient air, and that gasoline engines account for a large fraction of airborne lead.The other reason for controlling the lead content of gasoline was the incompatibility of leaded fuel with the catalytic converter, a device seen as having potential for reducing hydrocarbon emissions from automobiles.The first reason suggests considering whether reductions in the lead content of all gasoline might be desirable; the second argues for the total removal of lead from gasoline for use by new automobiles equipped with catalytic converters. Without going any further one might ask:Why should the government be concerned with the level of lead in gasoline? And why had the federal government already decided to require the installation of catalytic converters? The following analysis does not explicitly deal with the desirability of catalytic converters. As we have argued, however, it is always important to make sure that there is a convincing rationale for public action.The interventions at hand, the catalytic converter and the related lead

restrictions, deal with market failures— viewed as problems of either negative externalities or public goods (ambient public goods). In addition, another market failure may come into play with respect to lead—imperfect consumer information about the impacts of lead on health and the maintenance cost of vehicles. Because the health and maintenance impacts may not manifest themselves 21 The Clean Air Act Amendments of 1970, Public Law 91-604, December 31, 1970. 22 Section 211(c)(2)(A), 42 U.S.C. @ 1857f-6c(c)(2)(A). 23 36 Fed. Reg. 1468 (January 31, 1971).

for many years (leaded gasoline is a post-experience

good), markets may be inefficient because of

information asymmetry.These apparent market

failures make a case for considering government

intervention.

But intervention itself may be costly. It is quite

conceivable that the costs of intervention could

exceed the benefits. Several questions must be

answered before the comparison of costs and benefits

can be made:What are the impacts in the Economic

Analysis Division [that] had worked on the 1982

regulations. Schwartz used his accumulated

knowledge from previous analyses to complete in two

or three days a quick and dirty benefit-cost analysis

of a total ban on lead additives.

Because a total ban was one of the options

considered in 1982, it was fairly easy for Schwartz

to come up with a reasonable estimate of cost. As is

often the case in evaluating health and safety

regulations, the more difficult problem was

estimating the benefits of a total ban. Schwartz looked at two benefit categories: increased IQ scores of children due to lower levels of blood lead and the avoided damage to catalytic converters.

Schwartz used estimates from a variety of sources to piece together a relationship between lead emissions and the present value of lifetime earnings of children.The first step involved using estimates from the 1982 analyses of the relationship between lead emissions and blood lead levels in children. He next turned to epidemiological studies that reported a relationship between blood lead levels and IQs. Finally, he found econometric studies that estimated the contribution of IQ points to the present value of future earnings.

As a first cut at quantifying the benefits resulting from the more effective control of other emissions, Schwartz estimated the cost of catalytic converters being contaminated under the current standards that would not be contaminated under a total ban. He used the number of converters saved multiplied by the price per converter as the benefit measure. Assuming that converters themselves have benefit cost ratios greater than one, this benefit measure would be conservative.

These “back-of-the-envelope” calculations

suggested that the benefits of a total ban on lead additives would be more than twice the costs. Schwartz discussed the results with his branch chief, G. Martin Wagner, who sent word to the office of the administrator that further analysis of a lead ban seemed worthwhile. A few weeks later Schwartz and fellow analyst Jane Leggett began a two-month effort to move from the back of an envelope to a preliminary report. Pulling the Pieces Together The most urgent task facing Schwartz and Leggett was to develop better measures of the benefits of a lead ban.The key to improving the measure of benefits from avoided converter poisonings was a more sophisticated accounting of the future age composition of the U.S. vehicle fleet.The key to improving the measure of benefits from reduced lead emissions was a better quantitative estimate of the relationship between gasoline lead and blood lead. Modeling and statistical analysis would be an important part of their work. The age composition of the vehicle fleet continually changes as old vehicles, some of which have contaminated or partially contaminated converters, are retired and new ones, with fresh converters, are added. A ban on lead would have different effects on different vintage vehicles. For instance, it would be irrelevant for vehicles that already have contaminated converters, but very important for vehicles that would otherwise be contaminated and remain in service for a long time. The analysts developed an inventory model that tracked cohorts of vehicles over time. Each year a new cohort would enter the fleet. Each successive year a fraction of the vehicles would be retired because of accidents and mechanical failures. In addition, a fraction would suffer converter poisoning from misfueling. By following the various cohorts over time, it was possible to predict the total number of converters that would be saved in each future year from specified reductions in the lead content of leaded gasoline today and in the future. Avoided loss of the depreciated value of the catalytic converters (later avoided costs of the health and property damage from pollutants other than lead) could then be calculated for each future year.With appropriate discounting, the yearly benefits could then be summed to yield the present value of the lead reduction schedule being considered. Two important considerations came to light during work on the vehicle fleet model. One was a body of literature suggesting that lead increases the routine maintenance costs of vehicles. Subsequently, a new benefit category, avoided maintenance costs, was estimated using the vehicle fleet model.The

other consideration was the possibility that some engines might suffer premature valve-seat wear if fueled with totally lead-free gasoline. Although the problem was fairly limited (primarily automobile engines manufactured prior to 1971 and some newer trucks, motorcycles, and off-road vehicles), it suggested the need to consider alternatives to a total ban. Efforts to quantify better the link between gasoline lead and blood lead focused on analysis of data from the second National Health and Nutrition Examination Survey (NHANES II).The NHANES II survey was designed by the National Center for Health Statistics to provide a representative national sample of the population between ages 6 months and 74 years. It included 27,801 persons sampled at sixty-four representative sites from 1976 to 1980. Of the 16,563 persons who were asked to provide blood samples, 61 percent complied.The lead concentrations in the blood samples were measured, providing data that could be used to track average blood levels over the four-year survey period.The amount of lead in gasoline sold over the period could then be correlated with the blood lead concentrations. A positive relationship between blood lead concentrations from the NHANES II data and gasoline lead had already been found by researchers. The positive correlation is apparent in [Figure 9.2], which was prepared by James Pirkle of the Centers for Disease Control to show the close tracking of blood lead concentrations and total gasoline lead. Much more than the apparent correlation, however, was needed for the benefit-cost analysis. Schwartz used multiple regression techniques . . . to estimate the increase in average micrograms of lead per deciliter of blood (µg/dl) due to each additional 100 metric tons per day of lead in consumed gasoline. He also employed models that allowed him to 1976 1977 1978 1979 1980 Year 50 60 70 80 90 100 110 T o t a l l e a d u s e d 6 m o n t h p e r i o d (1 , 0 0 0 t o n s) 10 9 11 12 13 14 15 16 A v e r a g e b l o o d l e a d l e v e l s (g / d l) Average blood lead levels Lead used in gasoline FIGURE 9.2 Lead use in gasoline production and average NHANES II blood lead levels Source: J. Schwartz et al., Costs and Benefits of Reducing Lead in Gasoline: Final Regulatory Impact Analysis, Publication No. EPA-230-05-85-006 (Washington, D.C.: EPA, 1985), p. E-5.

estimate the probabilities that children with

specified characteristics will have toxic levels of

blood lead (then defined by the Centers for Disease

Control as greater than 30 µg/dl) and gasoline lead.

These probabilities could then be used to predict the number of children in the population who would avoid lead toxicity if a ban on gasoline lead were imposed.

In early November of 1983, Schwartz and Leggett pulled together the components of their analyses and found a ratio of benefits to costs of a total ban to be greater than Schwartz's original back-of-the-envelope calculation.Together with their branch chief, they presented their results to Deputy Administrator Alm, who found them encouraging. He gave the green light for the preparation of a refined version that could be presented to the administrator as the basis for a new regulation. Alm also wanted the various components of the refined study to be subjected to peer review by experts outside the EPA. At the same time, he urged speed in order to reduce the chances that word would get out to refiners and manufacturers of lead additives, the primary opponents of a ban, before the EPA had an opportunity to review all the evidence.

The branch chief expanded the analytical team in order to hasten the preparation of the refined report that would be sent to the administrator. Joining Schwartz and Leggett were Ronnie Levin; Hugh

Pitcher, an econometrician; and Bart Ostro, an expert on the benefits of ozone reduction. In little more than one month the team was ready to send a draft report out for peer review.

The effort involved several changes in analytical approach. Because of the valve-head problem, the analysis focused on a major reduction in grams of lead per leaded gallon (from 1.1 gplg to 0.1 gplg) as well as a total ban.The available evidence suggested that the 0.1 gplg level would be adequate to avoid excessive valve-head wear in the small number of engines designed to be fueled only with leaded gasoline. At the same time, considerable effort was put into quantifying the maintenance costs that would be avoided by owners of other vehicles if the lead concentration were reduced. It soon appeared that the maintenance benefits consumers would enjoy would more than offset the higher prices they would have to pay for gasoline. Finally, the team decided that the benefits based on the blood lead (through IQ) to future earnings relationship would be too controversial. Instead, they turned their attention to the costs of compensatory education for children who suffered IQ losses from high blood lead levels. In late December, sections of the report were sent to experts outside the EPA for comments.The list included automotive engineers, economists, biostatisticians, toxicologists, clinical researchers, transportation experts, and a psychologist. During January 1984, the team refined [its] analysis and incorporated, or at least responded to, the comments made by the external reviewers. Finally, in early February, the team was ready to present its results to Administrator Ruckelshaus. He agreed that their analysis supported a new standard of 0.1 gplg. He told the team to finalize [its] report without a proposed rule and release it for public comment. Ruckelshaus also directed the Office of the Assistant

Administrator for Air and Radiation to draft a proposed rule. The team's Draft Final Report was printed and eventually released to the public on March 26, 1984. 24 The team continued to refine the analysis in the following months. It also had to devote considerable time to external relations. Executive Order 12291 requires regulatory agencies to submit proposed regulations that would have annual costs of more than $100 million to the Office of Management and Budget for review.The team met with OMB analysts several times before securing their acceptance of the benefit-cost analysis of the tighter standard. The political environment was taking shape pretty much as expected. Opposition would come from refiners and manufacturers of lead additives. The refiners, however, generally seemed resigned to the eventual elimination of lead from gasoline. 24 Joel Schwartz, Jan Leggett, Bart Ostro, Hugh Pitcher, and Ronnie Levin, Costs and Benefits of Reducing Lead in Gasoline: Draft Final Report (Washington, DC: Office of Policy Analysis, EPA, March 26, 1984). Their primary concern was the speed of implementation. Some refiners seemed particularly concerned about the first few years of a tighter standard, when they would have difficulty making the required reductions with their existing configurations of capital equipment. In response, the team began exploring the costs and benefits of less stringent compliance schedules. The manufacturers of lead additives were ready to fight tighter standards. In May Schwartz attended a conference at the Centers for Disease Control in Atlanta, Georgia, on the proposed revision of the blood lead toxicity standard for children from 30 µg/dl to 25 µg/dl. Representatives of the lead manufacturers were also there.They openly talked about their strategy for challenging tighter standards. They planned to argue that refiners would blend more benzene, a suspected carcinogen, into gasoline to boost octane if lead were restricted. Schwartz investigated this possibility following the conference. He found that, even if more benzene were added to gasoline, the total emissions of benzene would decline because of the reduction in the contamination of catalytic converters, which oxidize the benzene if not contaminated.The day that the proposed rule was published Schwartz put a memorandum on the docket covering the benzene issue, thus preempting the manufacturers' main attack. The EPA published the proposed rule on August 2, 1984. 25 It would require that the permitted level of gasoline lead be reduced to 0.1 gplg on January 1, 1986.The proposal stated the EPA assumption that the new standard could be met with existing refining equipment, but indicated that alternative phase-in schedules involving more gradual reductions also were being

considered in case this assumption proved false. Finally, the proposal raised the possibility of a complete ban on gasoline lead by 1995. A Closer Look at the Link Between Gasoline Lead and Blood Lead Calculation of the direct health benefits of tightening the lead standard requires quantitative estimates of the contribution of gasoline lead to blood lead. The NHANES II data, combined with information on gasoline lead levels, enabled the study team to make the necessary estimates.Their efforts provide an excellent illustration of how statistical inference can be used effectively in policy analysis. The Need for Multivariate Analysis A casual inspection of [Figure 9.2] suggests a strong positive relationship between gasoline lead and blood lead.Why is it necessary to go any further? One reason is the difficulty of answering, directly from [Figure 9.2], the central empirical question: How much does the average blood lead level in the United States decline for each 1,000-ton reduction in the total gasoline lead used over the previous month? [Figure 9.2] indicates a positive correlation between gasoline lead and blood lead—changes in blood lead track changes in gasoline lead quite closely. But for the same correlation in the data, we could have very different answers to our central question. Figure 9.3 illustrates the difference between correlation and magnitude of effect with stylized data. If our data were represented by the diamonds, we might “fit” line one as our best guess at the relationship between blood lead and gasoline lead. The effect of reducing gasoline lead from 500 tons per day to 400 tons per day would be to reduce average blood lead from 10.1 µg/dl to 10.0 µg/dl, or 0.1 µg/dl per 100 tons per day. An alternative sample of data is represented by the dots, which are plotted to have approximately the same correlation between gasoline lead and blood lead as the data represented by triangles.The slope of line two, the best fit to the dots, is 1.0 µg/dl per 100 tons per day—ten times greater than the slope of line one.That is, although the two data sets have the same correlation, the second implies a much greater contribution of gasoline lead to blood lead than the first. Even after we display the data in [Figure 9.2] as a plot between blood lead and gasoline lead, our analysis is far from complete because the apparent relationship (the slopes of lines one and two in our example) may be spurious. Blood lead and gasoline lead may not be directly related to each other but to 25 49 Fed. Reg. 31031 (August 2, 1984). 2 3 4 5 Gasoline lead (100 metric tons/day) A v e r a g e b l o o d l e a d (g / d l) 9.0 10.0 11.0 12.0 13.0 14.0 Line two Line one

some third variable that causes them to change

together.The classic illustration of this problem is the correlation sometimes found between the density of stork nests and the human birth rate. If one were to plot birth rates against the density of stork nests for districts in some region, one might very well find a positive relationship and perhaps conclude that there might be something to the myth about storks bringing babies.

Of course, there is a more plausible explanation. A variable measuring the degree to which a district is rural "intervenes" between birth rate and nest density—rural areas have farmers who want a lot of children to help with chores as well as lot of open land that provides nesting grounds for storks; more urbanized areas tend to have people who want smaller families as well as less hospitable nesting grounds. Looking across a sample that included both rural and urban districts could yield the positive correlation between birth rate and nest density. If we were to "control" statistically for the intervening variable by looking at either rural or urban districts separately rather than pooled together, we would expect the correlation between birth rate and nest density to become negligible. Returning to our analysis of lead, we must be concerned that one or more intervening variables might explain the positive correlation between gasoline lead and blood lead. For example, there is some evidence that cigarette smokers have higher blood lead levels than nonsmokers. It may be that over the period that the NHANES

II data were collected the proportion of smokers in the sample declined, so that much of the downward trend in average blood lead levels should be attributed to reductions in smoking rather than reductions in gasoline lead. To determine whether smoking is an intervening, or confounding, variable, we might construct separate diagrams like [Figure 9.3] for smokers and nonsmokers. In this way, we could control for the possibility that changes in the proportion of smokers in the sample over time were responsible for changes in blood lead levels. If we fit lines with similar positive slopes to each of the samples, we would conclude that smoking behavior was not an FIGURE 9.3 Data samples with identical correlations but different regression lines intervening variable for the relationship between gasoline lead and blood lead. 26 With unlimited quantities of data, we could always control for possible intervening variables in this way. Unfortunately, with a fixed sample size, we usually stretch our data too thin if we try to form subsets of data that hold all variables constant (e.g., everyone is a nonsmoker) except the dependent variable (blood lead levels) we are trying to explain and the independent variable (gasoline lead) we are considering as a possible explanation. For example, occupational exposure, alcohol consumption, region, and age are only a few of the other variables that might be causing the relationship in our data between gasoline lead and blood lead. If we selected from our sample only adult males living in small cities in the South who are moderate drinkers and nonsmokers and have no occupational exposure to lead, we might have too few data to reliably fit a line as in [Figure 9.3]. Even if we had adequate data, we would end up with estimates of the relationship between gasoline lead and blood lead from all of our subsets. (The number of subsets would equal the product of the number of categories making up our control variables.) We might then have difficulty combining these subset estimates into an overall estimate. 27 The Basic Linear Regression Model Linear regression provides a manageable way to control statistically for the effects of several independent variables. 28 Its use requires us to assume that the effects of the various independent variables are additive.That is, the marginal effect on the dependent variable of a unit change in any one of the independent variables remains the same no matter what the values of the other independent variables [are]. 29 We can express a linear regression model in mathematical form: y b 0 b 1 x 1 b 2 x 2 b 2 x 2 . . . b k x k e where y is the dependent variable, x 1 , x 2 , . . . , x k are the k independent variables, b 0 , b 1 , . . . , b k are parameters (coefficients) to be estimated, and e is an

error term that incorporates the cumulative effect on y of all the factors not explicitly included in the model. If we were to increase x 1 by one unit while holding the values of the other independent variables constant, y would change by an amount b 1 . Similarly, each of the coefficients measures the marginal effect of a unit change in its variable on the dependent variable. Imagine that we set the values of all the independent variables except x 1 equal to zero.We could then plot y 26 If the two lines coincided, we would conclude that smoking did not contribute to blood lead. If the lines were parallel but not identical, the difference between their intercepts with the vertical axis would represent the average effect of smoking on blood lead. If the lines were not parallel, we might suspect that smoking interacted with exposure to gasoline lead so that their effects are not additive. That is, smokers were either more or less susceptible to exposure to gasoline lead than nonsmokers. 27 If the relationship truly varies across our subsets, then it would generally not make sense to make an overall estimate. As we will explain later, however, we never really observe the true relationship—it always contains some unknown error. Therefore, we may not be sure if it is reasonable to combine our subsample data. Because the variance of the error will be larger, the smaller the size of our subsamples, the more we divide our data, the more difficult it is to determine whether observed differences reflect true differences. 28 For clear introductions, see Eric A. Hanushek and John E. Jackson, Statistical Methods for the Social Sciences (New York: Academic Press, 1977); and Christopher H. Achen, Intrepreting and Using Regression (Beverly Hills, Calif.: Sage Publications, 1982). 29 This assumption is not as restrictive as it might at first seem. We can create new variables that are functions of the original independent variables to capture nonlinearities. For example, if we thought smokers were likely to absorb environmental lead faster than nonsmokers, we might include in our linear regression model a new variable that is the product of the number of cigarettes smoked per day and the level of gasoline lead. This new variable, capturing the interaction of smoking and gasoline lead, would have an effect that would be additive with the other independent variables in the model, including the smoking and gasoline lead variables. The marginal effect of gasoline lead on blood lead would consist of the contribution from the gasoline lead variable and the contribution from the variable representing its interaction with smoking, which will change depending on the particular level of smoking.

against x 1 in a graph like [Figure 9.3].The equation

y b 0 b 1 x 1 would represent the line fit to our sample

of observations.The slope of the line is b 1 , the

magnitude of the change in y that will result from a unit

change in x 1 , other things equal.The actual observations

do not lie exactly on the line, however.Their vertical

distances from the line will be equal to the random

error, represented in our model by e, incorporated in

each of the observations of y. If the values of e are

small, our line will fit the data well in the sense that the

actual observations will lie close to it.

How should we go about fitting the line? The most

commonly used procedure is the method of ordinary least

squares (OLS).When we have only one independent

variable, so we can plot our data on a two-dimensional

graph like [Figure 9.3], the OLS procedure picks the

line for which the sum of squared vertical deviations

from the observed data is smallest.When we have

more than one independent variable, OLS determines

the values of the coefficients (b 0 , b 1 , . . . , b k)
that

minimize the sum of squared prediction errors. 30 As

long as the number of observations in our sample

exceeds the number of coefficients we are trying to

estimate, and none of our independent variables can be

expressed as a linear combination of the other

independent variables, the commonly available regression software packages will enable us to use computers to find the OLS fitted coefficients. 31 The estimates of coefficients that we obtain from OLS will generally have a number of very desirable properties. If the independent variables are all uncorrelated with the error term (e), then our coefficient estimators will be unbiased. 32 (NOTE: an estimator is the formula we use to calculate a particular estimate from our data.) To understand what it means for an estimator to be unbiased, we must keep in mind that our particular estimate depends upon the errors actually realized in our sample of data. If we were to select a new sample, we would realize different errors and hence different coefficient estimates.When an estimator is unbiased, we expect that the average of our estimates across different samples will be very close to the true coefficient value. For example, if gasoline lead had no true effect on blood lead, we would almost certainly estimate its coefficient to be positive or negative, rather than exactly zero. Repeating OLS on a large number of samples and averaging our estimates of the coefficient of gasoline lead, however, would generally yield a result very close to zero. 33 Indeed, by adding more and more samples, we could get the average as close to zero as we wanted. 30 The prediction error is the observed value of the dependent variable minus the value we would predict for the dependent variables based on our parameter estimates and the values of our independent variables. For the ith observation the prediction error is given by: The prediction error is also referred to as the residual of the observation. OLS selects the parameter estimates to minimize the sum of the squares of the residuals. 31 When one independent variable can be written as a linear combination of the others, we have a case of perfect multicollinearity. A related and more common problem, which we can rarely do anything about, occurs when the independent variables in our sample are highly correlated. This condition, called multicollinearity, is not a problem with the specification of our model but with the data we have available to estimate it. If two variables are highly correlated, positively or negatively, OLS has difficulty identifying their independent effects on the dependent variable. As a result, the estimates of the parameters associated with these variables will not be very reliable. That is, they will have large variances, increasing the chances that we will fail to recognize, statistically speaking, their effects on the dependent variable. One way to deal with multicollinearity is to add new observations to our sample that lower the correlation. For example, if we had a high positive correlation between smoking and drinking in our

sample, we should try to add observations on individuals who smoke but do not drink and who drink but do not smoke. Unfortunately, we often have no choice but to work with the data that are already available. 32 Strictly speaking, we must also assume that our independent variables are fixed in the sense that we could construct a new sample with exactly the same observations on the independent variables. Of course, even if the independent variables are fixed, we would observe different values of the dependent variable because of the random error term. In addition, for complete generality, we must assume that the expected value of the error term is constant for all observations. y i (b N 0 + b N 1 x 1i + Á + b N k x ki) Unfortunately, we usually only have a single sample for estimating coefficients. How do we decide if an estimate deviates enough from zero for us to conclude that the true value of the parameter is not zero? Making the fairly reasonable assumption that the error term for each observation can be treated as a draw from a Normal distribution with constant variance, the OLS estimators will be distributed according to the Student-t distribution. 34 That is, we can interpret the particular numerical estimate of a coefficient as a draw from a random variable distributed as a Student-t distribution centered [on] the true value of the coefficient. (The OLS estimator is the random variable; the actual estimate based on our data is a realization of that random variable.) Knowing the distribution of the OLS estimator enables us to interpret the statistical significance of our coefficient estimate.We determine statistical significance by asking the following question: How likely is it that we would observe a coefficient estimate as large as we did if the true value of the coefficient were zero? We answer this question by first assuming that the true value of the coefficient is zero (the null hypothesis) so that the distribution of our estimator is centered [on] zero.We then standardize our distribution to have a variance of one by dividing our coefficient estimate by an estimate of its standard error (a by-product of the OLS procedure).The resulting number, called the t-ratio, can then be compared to critical values in tabulations of the standardized Student-t distribution found in the appendix of almost any statistics text. For example, we might decide that we will reject the null hypothesis that the true value of the coefficient is zero if there is less than a 5 percent probability of observing a t-ratio (in absolute value sense) as large as we did if the null hypothesis is true. (The probability we choose puts an upward bound on the probability of falsely rejecting the null hypothesis.) 35 To carry out the test, we look in the standardized tabulations of the Student-t distribution for the critical value corresponding to 5

percent. 36 If the absolute value of our estimated t-ratio exceeds the critical value, then we reject the null hypothesis and say that our estimated coefficient is statistically significantly different from zero. Fortunately, most regression software saves us the trouble of looking up critical values in tables 33 Our average will not be close to zero if gasoline lead is correlated with a variable excluded from our model that does have an effect on blood lead. In this case, gasoline lead stands as a proxy for the excluded variable. Other things equal, the stronger the true effect of the excluded variable on blood lead and the higher the absolute value of the correlation between gasoline lead and the excluded variable, the greater will be the bias of the coefficient of gasoline lead. We might not worry that much about the bias if we knew that it would approach zero as we increase sample size. (If the variance of the estimator also approached zero as we increased sample size, we would say that the estimator is consistent.) Although OLS estimators are consistent for correctly specified models, correlation with an important excluded variable makes an estimator inconsistent. 34 The Central Limit Theorem tells us that the distribution of the sum of independent random variables approaches the Normal distribution as the number in the sum becomes large. The theorem applies for almost any starting distributions—the existence of a finite variance is sufficient. If we think of the error term as the sum of all the many factors excluded from our model, and further, we believe that they are not systematically related to each other or the included variables, then the Central Limit Theorem suggests that the distribution of the error terms will be at least approximately Normal. 35 Falsely rejecting the null hypothesis is referred to as Type I error. Failing to reject the null hypothesis when in fact the alternative hypothesis is true is referred to as Type II error. We usually set the probability of Type I error at some low level like 5 percent. Holding sample size constant, the lower we set the probability of Type I error, the greater the probability of Type II error. 36 The Student-t distribution is tabulated by degrees of freedom. In the basic OLS framework, the degrees of freedom is the total number of observations minus the number of coefficients being estimated. As the degrees of freedom becomes larger, the Student-t distribution looks more like a standardized Normal distribution. You should also note the difference between a one-tailed and two-tailed test. Because the standardized Student-t is a symmetric distribution centered on zero, a 5 percent test usually involves setting critical values so that 2.5 percent of area lies under each of the tails (positive and negative). A one-tailed test,

appropriate when the null hypothesis is that the true coefficient value is zero or less than zero, puts the entire 5 percent in the positive tail.

by directly calculating the probability under the null hypothesis of observing a t-ratio as large as that estimated.To do a classical test of hypothesis on the coefficient, we simply see if the reported probability is less than the maximum probability of falsely rejecting the null hypothesis that we are willing to accept. If it is smaller, then we reject the null hypothesis.

Consider the regression results presented in [Table 9.3].They are based on data from 6,534 whites in the NHANES II survey for whom blood lead measurements were made. 37 The dependent variable is the individual’s blood lead level measured in µg/dl.The independent variables are listed under the heading “Effect.”The independent variable of primary interest is the national consumption of gasoline lead (in hundreds of metric tons per day) in the month prior to the individual’s blood lead measurement.The other independent variables were included in an effort to control statistically for other factors that might be expected to affect the individual’s blood lead level.With the exception of the number of cigarettes smoked per day and the

dietary factors (vitamin C, riboflavin, and so on),

these other statistical controls are indicator, or

“dummy,” variables that take on the value one if

some condition is met and zero otherwise. So, for

example, if the individual is male, the variable

“male” will equal one; if the individual is female,

it will equal zero.The variables “vitamin C,”

“phosphorus,”“riboflavin,” and “vitamin A,” which

are included as proxy measures for dietary intake of

lead, each measure dietary intake in milligrams. Otherwise, the other variables are intended to capture demographic, income, occupational exposure, drinking habit, and locational effects.The reported R 2 indicates that, taken together, the independent variables explain about 33 percent of the total variation in blood lead levels. 38 The estimated coefficient for gasoline lead is 2.14 µg/dl of blood per 100 metric tons per day of national gasoline lead consumption. Dividing the coefficient estimate by its estimated standard error of 0.192 yields a t-ratio of about 11.The probability of observing a t-ratio this large or larger if the true value of the coefficient were actually zero is less than one chance in 10,000 (the 0.0000 entry in [Table 9.3] under “P-value”).We would thus reject the null hypothesis in favor of the alternative hypothesis that gasoline lead does contribute to blood lead. In other words, we would say that gasoline lead has a statistically significant effect on blood lead. After finding a statistically significant effect, the next question to ask is whether the size of the coefficient is substantively significant. That is, does the variable in question have an effect that is worth considering? 39 One approach to answering this question is to multiply the estimated coefficient by the plausible change in the independent variable that might occur. For example, by the end of the NHANES II survey, gasoline lead was being consumed at a rate of about 250 metric tons per day nationally. A strict policy might reduce the level to, say, 25 metric tons per day. Using the estimated coefficient of gasoline lead, we would expect a reduction of this magnitude to reduce blood lead levels on average by about 4.8 µg/dl (the reduction of 225 metric tons 37 The analysts estimated similar models

for blacks and for blacks and whites together. Their estimates of the coefficient for gasoline lead never deviated by more than 10 percent across the different samples. To conserve space, they reported in detail only their regression results for whites. They chose whites because it was the largest subgroup and because it preempted the assertion that the relationship between blood lead and gasoline lead was due to changes in the racial composition of the sample over time. 38 R 2 is a measure of the goodness of fit of the model to the particular sample of data. It is the square of the correlation between the values of the dependent variable predicted by the model and the values actually observed. An R 2 of one would mean that the model perfectly predicted the independent variable for the sample; an R 2 of zero would mean that the model made no contribution to prediction. 39 The standard errors of the coefficient estimates decrease as sample size increases. Thus, very large samples may yield large t-ratios even when the estimated coefficient (and its true value) [is] small. We refer to the power of a statistical test as 1 minus the probability of failing to reject the null hypothesis in favor of the alternative hypothesis. Other things equal, larger sample sizes have greater power, increasing the chances that we will reject the null hypothesis in favor of alternatives very close to 0. TABLE 9.3 Basic Regression Model for Estimating the Effects of Gasoline Lead on Blood Lead a Effect Coefficient Standard Error P-value Intercept 6.15 Gasoline 2.14 .142 0.0000 Low income 0.79 .243 0.0025 Moderate income 0.32 .184 0.0897 Child (under 8) 3.47 .354 0.0000 Number of cigarettes 0.08 .012 0.0000 Occupationally exposed 1.74 .251 0.0000 Vitamin C −0.04 .000 0.0010 Teenager −0.30 .224 0.1841 Male 0.50 .436 0.2538 Male teenager 1.67 .510 0.0026 Male adult 3.40 .510 0.0000 Small city −0.91 .292 0.0039 Rural −1.29 .316 0.0003 Phosphorus −0.001 .000 0.0009 Drinker 0.67 .173 0.0007 Heavy drinker 1.53 .316 0.0000 Northeast −1.09 .332 0.0028 South −1.44 .374 0.0005 Midwest −1.35 .500 0.0115 Educational level −0.60 .140 0.0000 Riboflavin 0.188 .071 0.0186 Vitamin A 0.018 .008 0.0355 a Dependent variable: Blood lead (μg/dl) of whites in NHANES II Survey. Source: Joel Schwartz et al., Costs and Benefits of Reducing Lead in Gasoline: Final Regulatory Impact Analysis (Washington, DC: EPA, 1985), p. III-15.The original document reported incorrect standard errors.The standard errors reported here were provided by Joel Schwartz. per day times the estimated coefficient of 2.14 μg/dl per 100 metric tons per day). To get a better sense of whether a 4.8 μg/dl reduction is substantively important, we can look at the blood lead levels for representative groups at the 250 and 25 metric ton levels. For example, at the 250 metric ton level a

nonsmoking (number of cigarettes equals zero), moderate drinking (drinker equals one; heavy drinker equals zero), nonoccupationally exposed (occupationally exposed equals zero), adult female (child, teenager, male, male teenager, and male adult equal zero), living in a large Northeastern city (Northeast equals one; small city, rural, South, and Midwest equal zero), with moderate income, a college degree, and a highnutrition diet (low income equals zero; moderate income, educational level, vitamin C, phosphorus, riboflavin, and vitamin A equal one) would be expected to have a blood lead level of 10.6 µg/dl. We would expect the same person to have a blood lead level of only 5.8 µg/dl if the gasoline lead level were cut to 25 metric tons per day—a reduction of about 45 percent.Thus, the effect of gasoline lead on blood lead appears substantively as well as statistically significant. The study team was especially interested in estimating the contribution of gasoline lead to blood lead in children. As a first cut, they developed a

logistic regression model 40 for predicting the

probability that a child between the ages of six

months and eight years will have blood levels in

excess of 30 µg/dl, the definition of lead toxicity used

by the Centers for Disease Control at the time. Logistic

regression, which assumes a nonlinear relationship

between the dependent and the independent

variables, 41 is usually more appropriate than linear

regression when the dependent variable is

dichotomous (Y equals one if the condition holds, Y

equals zero if it does not). 42 The study team found a

strong relationship in the NHANES II data between

gasoline lead and the probability that a child has a

toxic level of blood lead. In fact, they estimated that

the elimination of gasoline lead would have reduced

the number of cases of lead toxicity in the sample

by 80 percent for children under eight years of age.

The study team used logistic regression and other probability models to estimate how reductions in gasoline lead would change the number of children having various blood lead levels.These estimates were essential for their subsequent valuation of the effects of gasoline lead reductions on the health of children.

Reconsidering Causality

Finding that an independent variable in a regression model has a statistically significant coefficient does not by itself establish a causal relationship.That is, it does not guarantee that changes in the independent variable cause changes in the dependent variable. Even if the independent variable has no direct effect on the dependent variable, some other variable, not included in the model, may be correlated with both so as to produce an apparent relationship in the data sample (remember the apparent relationship between birth rates and the density of stork nests). Should the strong relationship between gasoline lead and blood lead be interpreted as causal? The study team considered this question in detail. Although its demonstration was not legally necessary, [team members] believed that adoption of the proposed rule would be more likely if a strong case for causality could be made.Their approach was to apply the criteria commonly used by epidemiologists to determine the likelihood of causality. Not all of the criteria are directly applicable outside of the health area. Nonetheless, the way the study team applied the criteria illustrates the sort of questioning that is valuable in empirical research.Therefore, we briefly review the six criteria that

they considered. Is the Model Biologically Plausible? The study team noted that lead can be absorbed through the lung and gut.They pointed out that gasoline lead, the major source of environmental lead, is emitted predominantly as respirable particulates in automobile exhaust.These particulates can be absorbed directly through the lungs.They also contaminate dust that can be inhaled through the lungs and absorbed through the gut.Therefore, they argued, it is biologically plausible that gasoline lead contributes to blood lead. Biological plausibility is the epidemiological statement of a more general criterion: Is the model theoretically plausible? Prior to looking through data for empirical relationships, you should specify a model (your beliefs about how variables are related). If you find that your data are consistent 40 For an introduction to logistic regression, see Hanushek and Jackson, Statistical Methods, pp. 179–216. 41 The logistic regression model is written: where P(Y) is the probability that condition Y holds, e is the natural base, and Z b 0 b 1 x 1 . . . b k x k is the weighted sum of the independent variables x 1 , x 2 , . . . , x k . The coefficients, b 0 , b 1 , . . . , b k , are selected to maximize the probability of observing the data in the sample. Note that the marginal contribution of x i to the value of the dependent variable is not simply b i , as would be the case in a linear regression model. Rather, it is b i [1 − P(Y)]P(Y), which has its greatest absolute value when P(Y) 0.5. 42 The logistic regression model, unlike the linear regression model, always predicts probabilities that lie between zero and one (as they should). P(Y) = e z /(1 = e z) with your model, then you can be more confident that the relationships you estimate are not simply due to chance. 43 Is There Experimental Evidence to Support the Findings? The study team found reports of several investigations specifically designed to measure the contribution of gasoline lead to blood lead. One was an experiment conducted in Turin, Italy, by researchers who monitored changes in the isotopic composition of blood lead as the isotopic composition of gasoline lead varied. 44 They found that at least 25 percent of the lead in the blood of Turin residents originated in gasoline.Thus, the experiment not only confirmed the biological plausibility of the contribution of gasoline lead to blood lead, but also suggested an effect on the order of magnitude of that estimated by the study team. Being able to find such strong and directly relevant experimental support is quite rare in policy research, much of which deals with the behavioral responses of people. Controlled experiments in the social sciences are rare, not only because they are costly and difficult to implement, but also because they often involve

tricky ethical issues concerning the assignment of people to “treatment” and “control” groups. Nevertheless, there have been a number of policy experiments in the United States over the last twenty years. 45 It is unlikely, however, that any of these experiments will be directly applicable to your policy problem.Therefore, you must typically broaden your search for confirmation beyond experiments to other empirical research. Do Other Studies Using Different Data Replicate the Results? The study team reviewed several studies that also found relationships between gasoline lead and blood lead.These studies were based on data collected in conjunction with community-wide lead-screening programs funded by the Centers for Disease Control during the 1970s 46 and on data collected from the umbilical cord blood of over 11,000 babies born in Boston between April 1979 and April 1981. 47 These studies reported statistically significant relationships between gasoline lead and blood lead, and thus supported the study team’s analysis based on the NHANES II data. Does Cause Precede Effect? The study team used information about the half-life of lead in blood to make predictions about the strengths of relationships between lagged levels of gasoline lead and blood lead that would be expected if gasoline lead contributes to blood lead. Lead has a half-life of about thirty days in blood. Noting that the average NHANES II blood test was done at mid-month, they predicted that the previous month’s gasoline lead (which on average represents emissions occurring between fifteen and forty-five days before the test) should have a stronger impact on blood lead than that of either the current month (average exposure of zero to fifteen days) or the month occurring two months prior (average exposure of forty-five to seventy-five days).They tested their 43 Imagine that you regress a variable on twenty other variables. Assume that none of the twenty independent variables has an effect on the dependent variable. (The coefficients in the true model are all zero.) Nevertheless, if you use a statistical test that limits the probability of falsely rejecting the null hypothesis to 5 percent, then you would still have a 0.64 probability [1 – (.95) 20] of rejecting at least one null hypothesis. In other words, if you look through enough data you are bound to find some statistically significant relationships, even when no true relationships exist. By forcing yourself to specify theoretical relationships before you look at the data, you reduce the chances that you will be fooled by the idiosyncrasy of your particular data sample. For a brief review of these issues, see David L. Weimer, “Collective Delusion in the Social Sciences: Publishing Incentives for Empirical Abuse,” Policy Studies Review, Vol. 5, No. 4, May 1986, pp. 705–8. 44 S. Fachetti

and F. Geiss, Isotopic Lead Experiment Status Report, Publication No. EUR8352ZEN (Luxembourg: Commission of the European Communities, 1982). 45 For a review of the major policy experiments conducted in the United States, see David H. Greenberg and Philip K. Robins, “The Changing Role of Social Experiments in Policy Analysis,” Journal of Policy Analysis and Management, Vol. 5, No. 2, Winter 1986, pp. 340–62. 46 Irwin H. Billick et al., Predictions of Pediatric Blood Lead Levels from Gasoline Consumption (Washington, D.C.: U.S. Department of Housing and Urban Development, 1982). 47 Michael Rabinowitz and Herbert L. Needleman, “Petrol Lead Sales and Umbilical Cord Blood Lead Levels in Boston, Massachusetts,” Lancet, January 1/8, 1983, p. 63.

predictions by regressing blood lead levels on current, one-month lagged, and two-month lagged gasoline lead levels. As predicted, one-month lagged gasoline lead was the most significant of the three. Also, consistent with the thirty-day half-life, two-month lagged gasoline lead had a coefficient approximately one-half that of one-month lagged gasoline lead.Thus, cause did appear to precede effect in the expected way.

Does a Stable Dose-Response Relationship Exist? The regression model used by the study team assumed a linear relationship between gasoline lead and blood lead. Did this relationship remain stable as the level of gasoline lead changed?To answer this question, the study team took advantage of the fact that, on average, gasoline lead levels were about 50 percent lower in the second half of the NHANES II survey than in the first half. If the relationship between gasoline lead and blood lead is stable and linear, then reestimating the

regression model using only data from the second half of the survey should yield a coefficient for gasoline lead comparable to that for the entire sample. [The team] found that the coefficients were indeed essentially the same. In addition, estimation of regression models that directly allowed for the possibility of nonlinear effects supported the initial findings of a linear relationship between gasoline lead and blood lead.

Is It Likely That Factors Not Included in the Analysis Could Account for the Observed Relationship? The study team considered several factors that might confound the apparent relationship between gasoline lead and blood lead: dietary lead intake, exposure to lead paint, seasonality, and sampling patterns.

The basic regression model included nutrient and demographic variables as proxy measures for the intake of lead in the diet.Yet these variables may not adequately control for a possible downward trend in dietary lead that could be causing the estimated relationship between gasoline lead and blood lead. Market basket studies conducted by the Food and Drug Administration over the survey period, however, showed no downward trend in dietary lead intake. Also, lead intake from drinking water is largely a function of acidity, which did not change systematically over the survey period.

Evidence did suggest that changes in solder reduced the content of lead in canned foods over the period. But the study team was able to rule out the lead content in canned foods as a confounding factor when [it] added the lead content in solder as an independent variable, reestimated the basic regression model, and found that the coefficient of gasoline lead remained essentially unchanged. The study team recognized changing exposure to lead paint as another potential confounding factor. [it] dismissed the possibility on three grounds. First, paint lead is a major source in blood lead for children (who eat paint chips) but not for adults. If declining exposure to lead paint were responsible for the estimated relationship between gasoline lead and blood lead, we would expect the reduction in blood lead to be much greater for children than for adults. In fact, the average reduction over the survey period for adults was only slightly smaller than that for children (37 percent versus 42 percent). Second, the ingestion of paint lead usually results in large increases in blood lead levels. If reduced exposure to paint lead were responsible for declining blood lead levels, then we would expect to observe the improvement primarily in terms of a reduction in the number of people with very high blood lead levels. In fact, blood lead levels declined over the survey period even for groups with low initial levels. 48 Third, declining exposure to paint lead should be a more important factor in central cities than in suburbs because the latter tend to have newer housing stocks with lower frequencies of peeling lead paint.Yet the gasoline lead coefficient was essentially the same for separate estimations based on central city and suburban subsamples. Blood lead levels in the United States are on average higher in the summer than in the winter.To rule out the possibility that seasonal variation confounds the relationship between gasoline lead and blood lead, the study team reestimated the basic model with indicator variables included to allow for 48 The study team also used data from the lead-screening program in Chicago to estimate the probability of toxicity as a function of gasoline lead for children exposed and not exposed to lead paint. They found that gasoline lead had statistically significant positive coefficients for both groups. the possibility of independent seasonal effects.The coefficients of the seasonal variables were not statistically significant when gasoline lead was kept in the model.Thus, it appeared that changes in gasoline lead could adequately explain seasonal as well as long-term changes in blood lead levels. As already mentioned, the study team estimated the basic model on a variety of demographic subsamples and found no more than a 10 percent difference across any two estimates of the gasoline lead

coefficient. [Team members were] also concerned, however, that changes in NHANES II sampling locations over the survey period might have confounded the estimation of the gasoline lead coefficient.Therefore, they reestimated the basic model with indicator variables for forty-nine locations and found that the gasoline lead coefficient changed by only about 5 percent. Further, they found that, even when including variables to allow for different gasoline lead coefficients across locations, the coefficient representing the nationwide effect of gasoline lead was statistically and substantively significant. Together, these tests led the study team to dismiss the possibility of serious sampling bias. The Weight of the Evidence The study team [members] produced a very strong case in support of an important causal relationship between gasoline lead and blood lead. In many ways their efforts were exemplary.They drew relevant evidence from a wide variety of sources to supplement their primary data analysis.They gave serious attention to possible confounding factors, considering both internal tests (such as subsample analyses and model respecifications) and external evidence to see if they could be ruled out. As a consequence, opponents of the proposed policy were left with few openings for attacking its empirical underpinnings. Finalizing the Rule The primary task facing the analytical team after publication of the proposed rule was to respond to comments made by interested parties.Team members participated in public hearings held in August and spent much of the fall of 1984 responding to comments placed in the public docket, which closed on October 1. During this process they became more confident that the proposed rule would produce the large net benefits they predicated. At the same time they discovered another benefit category—reductions in adult blood pressure levels—that could potentially swamp their earlier estimates of benefits. In 1983 Schwartz chanced upon a research article reporting a correlation between blood lead and hypertension. 49 He began work with researchers at the Centers for Disease Control and the University of Michigan to determine if a relationship existed between blood lead and blood pressure levels. By the summer of 1984 their analysis of the NHANES II data suggested a strong link. 50 Because high blood pressure contributes to hypertension, myocardial infarctions, and strokes, the potential benefits from blood lead reductions were enormous. Although the final rule was ultimately issued without reference to quantitative estimates of the benefits of lower adult blood lead levels, the team provided estimates in the supporting documents. The one remaining issue was the compliance schedule.The costs of various lead standards were

estimated, using a model of the U.S. refining sector originally developed for the Department of Energy.The model represents the various types of refining capabilities that are available to convert crude oils to final petroleum products. It employs an optimization procedure for finding the allocations of crude oils and intermediate petroleum products among refining units that maximizes social surplus, the sum of consumer and producer surpluses.This allocation corresponds to that which would result from a perfectly competitive market operating without constraints on the utilization of available units. Cost was estimated by looking at the decline in social surplus resulting when the lead constraint was tightened—for instance, from 1.1 gplg to 0.1 gplg.The manufacturers of lead additives challenged these results on the grounds that the model 49 V. Batuman, E. Landy, J. K. Maesaka, and R. P. Wedeen, "Contribution of Lead to Hypertension with Renal Impairment," New England Journal of Medicine, No. 309, 1983, pp. 17-21. 50 The results of their research were later published in J. L. Pirkle, J. Schwartz, J. R. Landes, and W. R. Harlan, "The Relationship between Blood Lead Levels and Blood Pressure and Its Cardiovascular Risk Implications," American Journal of Epidemiology, Vol. 121, No. 2, 1985, pp. 246-58.

assumed more flexibility in capacity utilization across

different refineries than was realistic.

The analytical team held meetings with staffers

from other EPA offices to consider alternative

compliance schedules. Although a tentative decision

was reached to set an interim standard of 0.5 gplg,

to be effective July 1, 1985, and a final standard of

0.1 gplg, to be effective January 1, 1986, several

staffers feared that some refiners would be unable to

comply with their existing equipment. If these fears

were realized, the economic costs of the new rule

would be higher than estimated and perhaps raise

political problems.

A consultant to the project,William Johnson of Sobotka and Company, suggested a solution. If the physical distribution of equipment among refineries interfered with the flexibility in petroleum transfers assumed in the model, he reasoned, perhaps a secondary market in lead rights could be created to facilitate trading to get around specific bottlenecks. Taking the total permitted lead content from July 1, 1985 to January 1, 1988 as a constraint, the key was to create an incentive for refiners who could make the least costly reductions in lead additives below the interim 0.5 gplg standard to do so.Their additional reductions could then be used to offset excess lead in gasoline produced by refiners who could not easily meet the basic standards with the equipment they had in place. Because current reductions below the standard create a right to produce above the standard some time in the future, the trading process was called “banking of lead rights.” Refiners would be free to buy and sell lead rights at prices that were mutually beneficial. As a result, the aggregate cost of meeting the new standard would be reduced.

Representatives from the various EPA offices involved with the lead rule agreed that banking

seemed to be a good way to deal with concerns about

the compliance schedule. Because it had not been discussed in the proposed rule published in August, banking could not be part of the final rule. Nevertheless, by moving quickly to propose banking in a supplemental notice, it would be available shortly after the new standard became final. 51 The remaining task for the team was to prepare the Final Regulatory Impact Analysis, which would be published in support of the final rule. 52 The resulting document began by discussing the misfueling and health problems associated with lead additives along with alternatives (public education and stepped-up local enforcement to deal specifically with misfueling, pollution charges to deal generally with lead as a negative externality, and other regulatory standards) to the final rule. It then detailed the methods used to estimate the costs of tighter lead standards, the link between gasoline lead and blood lead, the health benefits of reducing the exposure of children and adults to lead, the benefits of reducing pollutants other than lead, and the benefits from reduced vehicle maintenance costs and increased fuel economy. The present value of the net benefits of the final rule was presented with various assumptions about misfueling (the use of leaded gasoline in vehicles with catalytic converters).The lower level of permitted lead would reduce the price differential between leaded and unleaded gasoline, thereby reducing the economic incentive for misfueling. It was not possible to predict with confidence, however, how much misfueling would actually decline.Therefore, the reasonable approach was to consider net benefits over the range of possibilities. [Table 9.4] presents the results of this sensitivity analysis. Note that net benefits were given both including and excluding the adult blood pressure benefits. Although the blood pressure benefits appeared huge, they were the last of the benefit measures considered and hence had the least-developed supporting evidence. Nevertheless, even assuming that the standard would produce no reduction in misfueling and no health benefits for 51 Fed. Reg. 718 (January 4, 1985); and 50 Fed. Reg. 13116 (April 2, 1985). 52 Joel Schwartz, Hugh Pitcher, Ronnie Levin, Bart Ostro, and Albert L. Nichols, Costs and Benefits of Reducing Lead in Gasoline: Final Regulatory Impact Analysis, Publication No. EPA-230-05-85-006 (Washington, DC: Office of Policy Analysis, EPA, February, 1985). By the time the final report was prepared, Jane Leggett left the project team. In the meantime, Albert Nichols, a Harvard University professor who was visiting the EPA as the acting director of the Economic Analysis

Division, began working closely with the team to produce the final document. adults, the present value of benefits appeared to be more than double the present value of costs. Indeed, it appeared that maintenance benefits alone would more than cover higher refining costs. The Final Regulatory Impact Analysis was released in February 1985. On March 7, 1985, the final rule was published in the Federal Register. 53 The 0.1 gplg standard would take effect on January 1, 1986, almost three years after work on the supporting analysis began. Conclusion We have described a case where statistical analysis made an important contribution to changing policy. Is this case typical? Yes and No.You should not expect that such a confluence of skill, time, data, resources, and interest will often arise to produce such definitive empirical findings. At the same time, you should expect to encounter empirical questions that at least can be approached, if not confidently answered, with the sort of statistical methods used by the EPA analysts. Consider a few prominent examples from recent years: Does the death penalty deter homicide? 54 Do higher minimum legal drinking ages and the 55-mile-per-hour speed limit reduce traffic fatalities? 55 Do smaller class sizes improve student performance? 56 Although widely accepted answers to these empirical questions would not necessarily be decisive in resolving policy debates, they would at least move the debates beyond disputes over predictions to explicit considerations of values. Such highly controversial issues aside, you are likely to find that making empirical inferences and critically consuming those of others often contribute in important ways to the quality of your policy analyses. TABLE 9.4 Present Values of Costs and Benefits of Final Rule, 1985–1992 (Millions of 1983 Dollars) No Misfueling Full Misfueling Partial Misfueling Monetized benefits Children’s health effects 2,582 2,506 2,546 Adult blood pressure 27,936 26,743 27,462 Conventional pollutants 1,525 0 1,114 Maintenance 4,331 3,634 4,077 Fuel economy 856 643 788 Total monetized benefits 37,231 33,526 35,987 Total refining costs 2,637 2,678 2,619 Net benefits 34,594 30,847 33,368 Net benefits excluding blood pressure 6,658 4,105 5,906 Source: Joel Schwartz et al., Costs and Benefits of Reducing Lead in Gasoline (Washington, DC: EPA, 1985), table VIII-8, p.VIII-26. 53 50 Fed. Reg. 9386 (March 7, 1985). 54 See Isaac Ehrlich, “The Deterrent Effect of Capital Punishment: A Question of Life and Death,” American Economic Review, Vol. 65, No. 3, June 1975, pp. 397–417; and Alfred Blumstein, Jacqueline Cohen, and Daniel Nagin, eds., Deterrence and Incapacitation: Estimating the Effects of Criminal Sanctions on Crime Rates (Washington, DC: National Academy of Sciences, 1978). 55 See Charles A. Lave,

“Speeding, Coordination, and the 55 MPH Limit,” American Economic Review, Vol. 75, No. 5, December 1985, pp. 1159–64; and Peter Asch and David T. Levy, “Does the Minimum Drinking Age Affect Traffic Fatalities?” Journal of Policy Analysis and Management, Vol. 6, No. 2, Winter 1987, pp. 180–92. 56 See Eric A. Hanushek, “Throwing Money at Schools,” Journal of Policy Analysis and Management, Vol. 1, No. 1, Fall 1981, pp. 19–41.

The Policy Issue Paper

Appendices Element of Policy Issue Paper Example/Note Rating Letter of transmittal 1. Are all relevant stakeholders expected to act on the paper addressed? A paper on local law enforcement might be addressed to the chief of police, the mayor, and the director of public safety. [ ] 2. Does the letter describe all attached materials? “Enclosed please find a policy issue paper on options to increase the productivity of municipal employees. An executive summary of this paper is also included, along with relevant statistical materials appended to the text.” [ ] 3. Does the letter specify who is expected to take action, how, and when? “We look forward to your reply well in advance of the meeting scheduled for June 15.” [ ] Executive summary 4. Are all elements of the issue paper described in the executive summary? Every element of the issue paper (Source and Background of the Problem, the Policy Problem, and so on) should be included in the summary. [ ] 5. Is the summary clear, concise, and specific? “This paper reviews problems of air pollution over the past ten years, showing that the level of industrial pollutants has increased by more than 200 percent in the 2000–2009 period.” [ ] 6. Is the summary understandable to all who will read it? The point here is to avoid unnecessary jargon, long sentences, and complex arguments as much as possible. [ ] 7. Are recommendations appropriately highlighted in the summary? “In conclusion, we recommend that funds in the range of $15,000–$20,000 be allocated to conduct an assessment of recreational needs in the Oakland area.” [ ] I. Background of the problem 8. Are all dimensions of the problem situation described? Increasing crime rates may be analyzed in conjunction with data on unemployment, police recruitment, citizen surveys, migration, and so on. [ ] (continued) Box A1 Checklist for Evaluating a Policy Issue Paper Rating Scale 1 = totally adequate 2 = adequate 3 = inadequate 4 = totally inadequate 0 = not applicable

Element of Policy Issue Paper Example/Note Rating 9. Have outcomes of prior efforts to resolve problems in the area been described? “Similar training programs have been implemented in New York, Florida, and California, at a cost of $4,000–$5,000 per trainee.” [ ]

10. Is there a clear assessment of past policy performance? “Training programs in New York and California resulted in. a 1 percent reduction in hard-core unemployment and were judged to be highly successful.” [ ]

II. Scope and severity of problem

11. Is the scope and severity of the problem situation clearly described? “Problems of industrial pollution are more severe today than at any time in our history and affect the physical well-being of 60 percent of the population. The resolution of these problems is an urgent matter that calls for timely and decisive action.” [ ]

III. Problem statement

12. Is the problem clearly stated? “The problem addressed in this paper is how best to satisfy the need for increased fiscal accountability among local governments in the region.” [ ]

13. Is the issue clearly stated? “The issue addressed in this paper is whether the state’s Department of Community Affairs should increase its technical assistance to local governments in the area of financial management.” [ ]

14. Is the approach to analysis clearly specified? “In addressing this issue we have employed costeffectiveness analysis as a way to determine the levels of investment necessary to train 10,000 local officials in cash management techniques.” [ ]

15. Are all major stakeholders identified and prioritized? The task here is to identify all persons and groups who significantly affect and are significantly affected by the formulation and implementation of policies. Stakeholders who are only marginally influential (or affected) may be dropped from the analysis, but this requires prioritization. [ ]

16. Are goals and objectives clearly specified? “The goal of this policy is to improve the physical security of the community. The objective is to reduce reported crimes by 10 percent or more in the 2005–10 period.” [ ]

17. Are measures of effectiveness clearly specified? Here there are many choices: benefit-cost ratios, net benefits, distributional benefits, and so on. [ ]

18. Are all sets of potential solutions outlined? “Any effort to reduce crime in the area should take into account the effectiveness of law enforcement programs, the unemployment rate, population density, and urbanization.” [ ] Element of Policy Issue Paper Example/Note Rating IV. Policy alternatives 19. Are alternative solutions

specified? “Three policy alternatives are analyzed in this paper: educational programs to alert citizens to the role they can play in crime control, policy training programs, and advanced crime control technology.” [ ] 20. Are alternatives systematically compared in terms of their probable costs and effectiveness? “Program A has a benefit-cost ratio that is twice that of program B.” [ ] 21. Are relevant spillovers and externalities included in the analysis of alternatives? “The early childhood educational program is not only cost-effective. It will also produce spillovers in the form of increased participation of family members in the school. At the same time there are important externalities in the form of increased costs for transportation to and from school.” [ ] 22. Are all relevant constraints taken into account? These may be financial (e.g., fixed costs), legal (e.g., laws proscribing certain actions), or political (e.g., lack of political support). [ ] 23. Have alternatives been systematically compared in terms of political feasibility? The point here is to make informed judgments about the probable support, relevance, and influence of key stakeholders in gaining acceptance for and implementing each alternative. [ ] V. Policy recommendations 24. Are all relevant criteria for prescribing alternatives clearly specified? Criteria may include net welfare improvement, distributional improvement, cost effectiveness, and so on. [ ] 25. Has the preferred alternative been clearly described? “We therefore recommend that the second alternative be adopted and implemented, since it is likely to be more cost-effective and is politically feasible under present conditions.” [ ] 26. Is a strategy for implementation outlined? “The proposed policy should be implemented by the Department of Labor and state employment offices. Implementation should be phased over a period of three years and focus on cities with more than 50,000 persons.” [ ] 27. Are provisions made for monitoring and evaluating policies? “The state employment offices should establish special units to engage in process and outcome evaluations, at intervals of 30 days. Results of monitoring and evaluation should be stored in a management information system, and reports should be filed quarterly.” [ ] (continued)

Element of Policy Issue Paper Example/Note Rating

28. Are limitations and possible unintended consequences taken into account? Here the point is to specify how limitations of the analysis affect the confidence one has in the conclusions. Possible unintended consequences (e.g., abuses of social welfare programs, nuclear meltdowns, and

so on) should also be thought through. [ ]

Appendices

30. Is relevant supporting information (statistics, legislation, documents, correspondence) included? All statistical tables should be properly numbered and headed and the source of data specified at the bottom of the table. Legislation, documents, and correspondence should be identified as to its source, date, author(s), and so on. [ ]