Public Policy in an Uncertain World: Analysis and Decisions

A lecture by Charles F. Manski, Northwestern University

27 March 2013, The British Academy

Public Policy in an Uncertain World: Analysis and Decisions, Harvard University Press, 2013

Broad Themes

- Society should face up to the uncertainties that attend policy formation.

- The current practice of policy analysis hides uncertainty.

- Researchers use untenable assumptions to make exact predictions of policy outcomes.

- Credible policy analysis would explicitly express the limits to knowledge.

- I consider how policy makers can reasonably make decisions in an uncertain world.

Motivation for the Book

- I typically write "in math" for an audience of academic economists and policy analysts.

- I have written this book "in English" to reach policy makers, civil servants, journalists, and policy students.

- Modern societies institutionally separate policy analysis and decision, with analysts reporting findings to policy makers.

- However, the current practice of policy analysis does not serve the public well.

- Consumers of policy analysis cannot trust the producers.

- This book explains how policy analysis is performed and suggests how it can better inform policy making.

Book Contents

Introduction

I. Policy Analysis

1. Policy Analysis with Incredible Certitude

2. Predicting Policy Outcomes

3. Predicting Behavior

II. Policy Decisions

4. Planning with Partial Knowledge

5. Diversified Treatment Choice

6. Policy Analysis and Decisions

POLICY ANALYSIS WITH INCREDIBLE CERTITUDE

The Logic and Credibility of Empirical Research

The logic of inference is summarized by the relationship:

assumptions + data => Conclusions.

The Unobservability of Counterfactual Policy Outcomes

- A researcher may observe the outcomes realized under a status quo policy. These observations are data.

- Policy analysis seeks to use such data to predict outcomes under alternative policies. These counterfactual outcomes are unobserved.

- Prediction of counterfactual outcomes requires assumptions.

- Weak assumptions yield interval predictions (bounds).

- Stronger assumptions tighten the intervals.

- Sufficiently strong assumptions yield exact predictions.

- There is a tension between the strength of assumptions and their credibility.

- *The Law of Decreasing Credibility:* The credibility of inference decreases with the strength of the assumptions maintained.

- *Credibility* is a primitive concept that defies deep definition.

- The Oxford English Dictionary (OED) defines *credibility* as "the quality of being credible."

- The OED defines *credible* as "capable of being believed; believable."

- It defines *believable* as "able to be believed; credible."

- And so we come full circle.

- Whatever credibility may be, it is subjective.

- Analysts should agree on the logic of inference, but they often disagree about the credibility of assumptions.

- Disagreement can persist when multiple contradictory assumptions are consistent with the available data.

- Such assumptions are *nonrefutable (untestable)*.

- An analyst can pose a nonrefutable assumption and displace the burden of proof, stating:

  "I will maintain this assumption until it is proved wrong."

Incentives for Certitude

- A researcher can resolve the tension between the credibility and power of assumptions by posing assumptions of varying strength and determining the conclusions that follow.

- In practice, policy analysis tends to sacrifice credibility in return for strong conclusions. Why so?

- Analysts respond to incentives.

* The scientific community rewards strong novel findings.

* The public wants unequivocal policy recommendations.

- These incentives make it tempting for researchers to maintain assumptions far stronger than they can persuasively defend, in order to draw strong conclusions.

- A story circulates about an economist's attempt to describe his uncertainty about a forecast to U. S. President Lyndon B. Johnson.

- The economist presented the forecast as a likely range of values for the quantity under discussion. Johnson is said to have replied:

  "Ranges are for cattle. Give me a number. "

- An econometrics colleague who frequently acts as a consultant stated the incentive argument this way:

  "You can't give the client a bound. The client needs a point."

- I have found a common perception that policy makers are either psychologically unwilling or cognitively unable to cope with uncertainty.

- Consultants argue that pragmatism dictates point prediction, even though it may not be credible.

- Making point predictions also has been advocated in philosophy of science.

- When there are multiple explanations for available data, philosophers recommend using a criterion such as "simplicity" to choose one of them. (*Occam's razor*)

Some Manifestations of Incredible Certitude

conventional certitudes

- Predictions that are generally accepted as true, but that are not necessarily true.

  (Examples: CBO scoring, reporting of official statistics)

dueling certitudes

- Contradictory predictions based on alternative non-refutable assumptions.

  (Example: RAND and IDA studies of drug policy)

conflating science and advocacy

- Reversal of the direction of the logical relationship

  assumptions + data => conclusions,

- seeking assumptions that imply specified conclusions.

  (Example: Friedman advocacy of school vouchers)

wishful extrapolation

- The use of untenable assumptions to extrapolate.

  (Example: FDA drug approval)

illogical certitudes

- Deductive errors, particularly non sequiturs.

  (Example: heritability research)

media overreach

- Premature or exaggerated reporting of policy analysis.

  (Example: *NYT* on "$320,000 Kindergarten Teachers")

CBO Scoring of Legislation

- *Conventional certitude* is exemplified by Congressional Budget Office (CBO) scoring of U. S. federal legislation.

- The CBO was established in the Congressional Budget Act of 1974. The Act has been interpreted as mandating the CBO to provide point predictions (*scores)* of the budgetary impact of legislation.

- CBO scores are conveyed in letters that the Director writes to leaders of Congress.

- They are not accompanied by measures of uncertainty.

- CBO scores have achieved broad acceptance within American society.

- They are used by both Democratic and Republican Members of Congress.

- Media reports largely take them at face value.

The Patient Protection and Affordable Care Act of 2010

- In March 2010 the CBO scored the combined consequences of the Patient Protection and Affordable Care Act and the Reconciliation Act of 2010.

- Director Douglas Elmendorf wrote to Nancy Pelosi:

  "CBO and JCT estimate that enacting both pieces of legislation . . . would produce a net reduction of changes in federal deficits of $138 billion over the 2010–2019 period as a result of changes in direct spending and revenue."

- Media reports largely accepted the CBO scores as fact.

- A rare commentator who rejected the CBO score was Douglas Holtz-Eakin, a former CBO director. He wrote:

  "In reality, if you strip out all the gimmicks and budgetary games and rework the calculus, a wholly different picture emerges: The health care reform legislation would raise, not lower, federal deficits, by $562 billion."

- The CBO and Holtz-Eakin scores differed by $700 billion. Yet they shared the common feature of certitude. Both were presented as exact, with no expression of uncertainty.

- This provides an example of *dueling certitudes.*

Interval Scoring

- The CBO has established an admirable reputation for impartiality.

- Perhaps it is best to leave well enough alone and have the CBO express certitude when it scores legislation, even if the certitude is conventional rather than credible.

- I worry that the existing social contract to take CBO scores at face value will eventually break down.

- I think it better for the CBO to act to protect its reputation than to have some disgruntled group in Congress or the media declare that the emperor has no clothes.

- A simple approach would be to provide interval forecasts of the budgetary impacts of legislation.

- The CBO would produce two scores for a bill, a low score and a high score, and report both.

- If the CBO must provide a point prediction for official purposes, it can continue to do so, with some convention used to locate the point within the interval forecast.

Can Congress Cope with Uncertainty?

- I have received disparate reactions when I have suggested interval scoring to economists and policy analysts.

- Academics react positively, but persons who have worked in the federal government tend to be skeptical.

- Some assert that members of Congress are psychologically or cognitively unable to deal with uncertainty.

- Some assert that Congressional decision making is a noncooperative game in which expression of uncertainty may yield inferior outcomes.

British Norms

- The norms for government forecasting in the United Kingdom differ from those in the United States.

- The Bank of England publishes probabilistic inflation forecasts presented visually as a *fan chart.*

- The government requires an *Impact Assessment* for legislation submitted to Parliament.

- The originating agency must state lower and upper bounds for the net benefits of the proposal, as well as a point estimate.

Credible Interval Estimates for Official Statistics with Survey Nonresponse

- Government agencies commonly report official statistics based on survey data as point estimates, without accompanying measures of error.

- Examples are the unemployment rate and median income.

- The public may misinterpret the information that the statistics provide.

- Some persons may incorrectly believe that errors are small and inconsequential.

- Persons who understand that official statistics are subject to error must fend for themselves and conjecture the error magnitudes.

- Agencies could mitigate misinterpretation of official statistics if they were to measure errors and report them.

- Agencies could report sampling error using established statistical principles.

- It is more challenging to report nonsampling errors because there are many sources and there has been no consensus about how to measure them.

Error Due to Survey Nonresponse

- Nonresponse is common.

- Unit and item nonresponse often make data missing for substantial fractions of the persons sampled.

- Example: The statistics on household income reported by the U. S. Census Bureau are based on an annual supplement to the Current Population Survey.

- During 2002-2012, 7 to 9 percent of the sampled households yielded no income data due to unit nonresponse.

- 41 to 47 percent of the interviewed households yielded incomplete data due to item nonresponse.

- Statistical agencies form point estimates by assuming that nonresponse is random within specified groups of persons.

- These assumptions are implemented as weights for unit nonresponse and imputations for item nonresponse.

- The agencies do not know the consequences.

- A Census Bureau document describing the American Housing Survey is revealing. The document states:

  "Some people refuse the interview or do not know the answers. When the entire interview is missing, other similar interviews represent the missing ones . . . . For most missing answers, an answer from a similar household is copied. The Census Bureau does not know how close the imputed values are to the actual values."

- Methodological research has shown how to form interval estimates that make no assumptions about nonresponse.

- We also know how to form confidence intervals that jointly measure sampling error and nonresponse error.

- Thus, we know how to report credible interval estimates for official statistics with survey nonresponse.

- I urge statistical agencies to report them.

POLICY MAKING WITH PARTIAL KNOWLEDGE

Policy analysis with incredible certitude can harm policy making in multiple ways. Policy makers may:

1. seek to maximize the social welfare that would prevail if untenable assumptions were to hold, not *actual* welfare.

2. not recognize the value of new research aiming to improve knowledge.

3. not recognize the usefulness of *flexibility* and *diversification* as means to cope with uncertainty and learn.

DIVERSIFIED POLICY CHOICE

Financial Diversification

- Financial diversification is a familiar recommendation for portfolio allocation.

- A portfolio is diversified if an investor allocates positive fractions of wealth to different investments.

- An investor with full knowledge would not diversify. He would invest fully in the investment with the highest return.

- The rationale for diversification arises purely from incompleteness of knowledge.

- Diversification enables someone who is uncertain about the returns to investments to balance potential errors.

Policy Diversification

- Policy choice is diversified if a planner allocates positive fractions of the population to each of various policies.

- Diversification enables a planner who is uncertain about policy response to balance potential errors.

- In principle, any micro policy (applied to a person, household, or locality) is diversifiable.

- Macro policies (such as monetary or climate policy) are not diversifiable.

Policy Diversification Differs from Profiling

- Diversification calls for *randomly* different treatment of persons.

- Profiling calls for *systematically* different treatment of persons who differ in observed attributes.

- Profiling may be good policy when a planner knows how policy response varies across persons.

- Diversification may appeal when a planner does not know how policy response varies across persons.

Adaptive Diversification

- If a planner treats a succession of cohorts, he may be able to learn about policy response.

- Observation of the outcomes realized by earlier cohorts informs policy choice for later cohorts.

- Over time, the planner can *adapt* his policy allocation, treating successive cohorts differently as data accumulates.

- Diversification is advantageous for learning because it generates randomized experiments.

- Adaptive diversification copes with uncertainty in the short run and reduces it in the long run.

Illustration: Centralized Health Care

- Adaptive diversification might be implemented in centralized health care agencies that directly assign medical treatments. (UK National Health Service, US Military Health System).

- Consider choice between a status quo treatment and an innovation.

- The agency knows the effectiveness of the status quo treatment but not that of the innovation.

- The agency may initially diversify treatment in accord with its knowledge and decision criterion.

- Over time, the agency observes learns response to the innovation. As data accumulates, the agency may update the treatment allocation of new patient cohorts.

- Eventually, it may have sufficient knowledge to determine which treatment is best.

- Thus, the agency may gradually decide between the status quo and the innovation as knowledge of treatment response accumulates.

Diversification and Equal Treatment of Equals

- Proposing that an investor may want to choose a diversified portfolio is uncontroversial.

- I have found it controversial to propose diversification of policy to humans.

- The concern is that policy diversification violates the ethical principle calling for *equal treatment of equals.*

- Diversification is consistent with *ex ante* equal treatment. All members of the population have the same probability of receiving a particular policy.

- It violates *ex post* equal treatment as different persons ultimately receive different policies.

- Democratic societies usually seek ex post equal treatment.

- Some policies seek only ex ante equal treatment.

- American examples include random tax audits, drug testing and airport screening, random calls for jury service, and the Green Card and Vietnam draft lotteries.

- Randomized experiments performed by researchers to learn about policy response seek only ex ante equal treatment.

- The practice of randomized experiments differs from policy diversification in several respects.

  1. Randomized experiments aim only to learn about policy response, not to cope with uncertainty when making policy choices.

  2. Democracies do not ordinarily compel participation in experiments.

  3. Randomized experiments often have short duration, requiring measurement of surrogate outcomes.

Laboratories of Democracy

- I have so far contemplated a planner having full power to set policy.

- The U. S. Constitution gives the federal government limited power, reserving much discretion to the states.

- Federalism permits a rough approximation to adaptive diversification.

- Federalism enables the states to implement diverse policies.

- It has become common to refer to the states as the *laboratories of democracy.*

- The federal government can provide incentives to the states to encourage informative diversity of policies.

- Policy variation across states is similar but not identical to diversification.

- Extrapolation of findings from one state to another requires one to assume that states with different policies have similar distributions of policy response.

In Conclusion

- Analysts today tend to provide exact predictions of policy outcomes, sacrificing credibility.

- They could provide credible interval predictions.

- Some think this idea impractical or unwise.

- Policy makers may be psychologically unwilling or cognitively unable to cope with uncertainty.

- Expression of uncertainty may yield inferior decisions in political games.

- I am not certain that policy analysis providing credible interval predictions will yield better decisions than analysis with incredible certitude.

- Hence, I suggest adaptive diversification of policy analysis.

- Exact prediction of policy outcomes is the status quo.

- Provision of credible interval predictions is an innovation.

- The outcome of interest is the quality of policy decisions.

- Society could implement both forms of policy analysis and learn their outcomes.