Northwestern | INSTITUTE FOR POLICY RESEARCH

IPR Working Paper Series

WP-23-27

Credible Social Planning Under Uncertainty

Charles F. Manski
Northwestern University and IPR

Version: July 17, 2023

DRAFT

Abstract

Economists have long studied policy choice by a social planner who aims to maximize welfare in democracies or other political systems where, in some sense, welfare is intended to express the wellbeing of a society rather than the personal preferences of a dictator. The motivation for studying planning is most transparent when actual planners face specific decision problems. Welfare economics has also sought to shed light on noncooperative societal decision processes, where no actual planner exists. Researchers have generally assumed that the actual or hypothetical planner knows enough about the choice environment to be able to determine an optimal action. However, the consequences of decisions are often highly uncertain. Addressing the failure of research to come to grips with uncertainty has motivated Manski's program of study of credible social planning under uncertainty. This paper describes the main themes and summarizes several applications. He first discusses practices that have promoted planning with incredible certitude, using specific cases to illustrate. He next describes and contrasts the conceptions of uncertainty in consequentialist and axiomatic decision theory. Manski then summarizes his studies of five problems of planning under uncertainty.

Prepared for delivery as the Marshall Lectures at the University of Cambridge, November 2023. This research received no external funding. The author has no relevant financial relationships to disclose.

1. Introduction

A foundational objective of the Constitution of the United States is to "promote the general Welfare." The Preamble states:

> "**We the People** of the United States, in Order to form a more perfect Union, establish Justice, insure domestic Tranquility, provide for the common defence, promote the general Welfare, and secure the Blessings of Liberty to ourselves and our Posterity, do ordain and establish this Constitution for the United States of America."

The Constitution does not define general welfare. A report on clinical practice guidelines by a committee of the U.S. Institute of Medicine (IOM) states (Institute of Medicine, 2010, p. 4):

> "Clinical practice guidelines are statements that include recommendations intended to optimize patient care that are informed by a systematic review of evidence and an assessment of the benefits and harms of alternative care options."

The report does not specify what it means to optimize patient care.

The Constitutional premise that the United States should "promote the general welfare" and the IOM premise that clinicians should "optimize patient care" exemplify broad assertions that entities making societal decisions should aim to maximize social welfare. Such assertions may have rhetorical appeal, but they lack substance per se. They become meaningful only when several questions are answered: What constitutes social welfare? What are the feasible actions? What is known about the welfare consequences of alternative choices?

Maximization of welfare is a well-defined objective if enough is known about the welfare consequences of alternative choices to determine an unambiguous best action. Maximization is ill-defined if the consequences are sufficiently uncertain that no action is clearly best. The concern of this paper is reasonable societal decision making in such settings.

1.1. The Prevalent Study of Planning with As-If Certainty

Economists have long studied policy choice by an actual or hypothetical social planner who aims to maximize welfare in democracies or other political systems where, in some sense, welfare is intended to express the well-being of a society rather than the personal preferences of a dictator. The public at large may not be familiar with the formal structure of welfare economics, but basic ideas are familiar through the widespread use of the term *benefit-cost analysis*. Economists often study planning with utilitarian welfare functions. They sometimes specify ones that express a form of paternalism or principles of fairness.

The motivation for studying planning is most transparent when actual planners face specific decision problems. For example, a national government must design an income tax structure and develop a system for national defense. Local governments choose how to maintain roads, perform policing, and organize public education. Planners need not be governmental. Clinicians make medical choices on behalf of patients. Parents act as planners for their families. In these settings and many more, the objective of the planner may be to maximize some idea of social welfare.

Welfare economics has also sought to shed light on noncooperative societal decision processes, where no actual planner exists. In the mid-1700s, Adam Smith metaphorically suggested that an “invisible hand” makes decentralized decision making in market economies promote social welfare. Between then and the mid-1900s, economists gradually formalized this notion to develop what have become known as the “fundamental theorems of welfare economics.” These give idealized conditions under which equilibrium outcomes in markets have the desirable welfare property of Pareto efficiency, which would be sought by a utilitarian planner.

A central concern of research in public economics, often using the term *mechanism design*, has been to study planning in situations of market failure, where the idealized conditions of the fundamental theorems of welfare economics do not hold. When studying market failure, the Pareto efficiency that a utilitarian planner would achieve serves as a benchmark for measuring the inefficiency of market outcomes. The social welfare achieved by a hypothetical planner has also served as a benchmark in social-choice theory, which studies the outcomes produced by voting and other decentralized mechanisms that attempt to aggregate individual preferences.

I wrote above that welfare economics has studied maximization of welfare. Whether performing abstract theoretical studies or applied benefit-cost analyses, researchers have generally assumed that the planner knows enough about the choice environment to be able to determine an optimal action. However, the consequences of decisions are often highly uncertain. Aiming to circumvent this difficulty, researchers commonly invoke strong unsubstantiated assumptions and use them to study solvable optimization problems. I have referred to this practice as policy analysis with *incredible certitude* (Manski, 2011, 2013).

Planning with incredible certitude can harm society in multiple ways. Most obviously, it seeks to maximize the social welfare that would prevail if untenable assumptions were to hold rather than actual social welfare. If planners incorrectly believe that existing analysis provides an errorless description of the current state of society and accurate predictions of policy outcomes, they may make substantively poor decisions. Moreover, they will not recognize the potential value of new research aiming to improve knowledge. Nor will they appreciate the potential usefulness of decision strategies that may help society cope with uncertainty and learn.

The dearth of study of planning under uncertainty is apparent in the comprehensive textbook on public economics of Atkinson and Stiglitz (1980), which mentions uncertainty only a few times and then only in passing. Mongin and Privato (2016) began their axiomatically oriented review article with this sentence (p. 711): “PERHAPS surprisingly, uncertainty plays no role whatsoever in the classical works on social welfare.”

Addressing the failure of research in welfare economics to come to grips with uncertainty has motivated my research program on credible social planning under uncertainty, which has developed over the past twenty years. The word “credible” is inevitably subjective and often difficult to pin down, but I will use it nonetheless. This paper describes the main themes of my work and uses several applications to illustrate.

As far as I am aware, only a small body of other research engages any of the themes that will be discussed here. In the late 1970s, Johansen (1978) made an early call for research on macroeconomic planning under uncertainty, stating (p. 263-264):

> "Uncertainty is not something which should be considered as a theoretically interesting refinement or extension of standard theory and methodology, but a central factor of eminently practical importance. Sometimes uncertainty is itself the heart of the matter when decisions are to be taken."

In the early 2000s, Hansen and Sargent initiated a program of work on robust macroeconomic policy, considering possible local deviations of reality from the assumptions maintained in conventional macroeconomic models; see Hansen and Sargent (2008). Barlevy (2011) reviews work on macroeconomics policy under ambiguity.

1.2. Uncertainty in Decision Theory

A fundamental difficulty with welfare maximization under uncertainty is apparent even in a simple setting with two feasible actions, say A and B, and two possible choice environments, say s_1 and s_2. Suppose that action A yields higher welfare in environment s_1 and action B yields higher welfare in s_2. If it is not known whether s_1 or s_2 is the actual choice environment, it is not known which action is better. Thus, maximization of welfare is logically impossible. At most one can seek a reasonable way to make a choice. A basic issue is how to interpret and justify the word "reasonable."

Research in decision theory has posed and characterized various principles for reasonable decision making under uncertainty. Decision theory is not specifically concerned with societal decisions. It presumes the existence of an abstract decision maker who must choice among a specified set of actions. The decision maker could be an individual, a firm, or another institution. When the decision maker is an entity making societal decisions, it is a social planner. Thus, decision theory provides the formal basis for the study of social planning under uncertainty.

The description of uncertainty in decision theory is abstract. One supposes that outcomes are determined by the chosen action and by some feature of the environment, called the *state of nature*. The decision maker is assumed able to list all states of nature that could possibly occur. This list, called the *state space*, is a primitive concept which provides the most basic expression of uncertainty. The larger the state space, the less the decision maker knows about the consequences of each action. Decision theorists usually describe the state space mathematically, without reference to an actual choice problem. For example, they might describe it as a finite or a convex set.

Much of decision theory adds a secondary expression of uncertainty in the form of a probability distribution over the state space. Some studies view the probability distribution as a cognitive concept, expressing how decision makers might actually perceive uncertainty. Others view it as a purely mathematical construct, whose existence might be inferred from analysis of choice behavior.

Two conceptually distinct but mathematically related approaches have been used to develop criteria for reasonable decision making. One poses choice axioms as primitives. The other focuses on the substantive consequences of choices.

1.2.1. Axiomatic Decision Theory

Axiomatic decision theory poses principles, called axioms, for consistency of hypothetical behavior across a class of potential choice problems. Researchers sometimes assert it to be self-evident that a decision maker should adhere to these choice axioms. The central research activity of axiomatic decision theorists has been to pose and prove representation theorems establishing that adherence to a specified set of axioms is equivalent to acting as if one wants to maximize some welfare function, coping with uncertainty in some manner. Perhaps the most famous such theorems are those of Von Neumann and Morgenstern (1944) and Savage (1954). Both theorems establish that adherence to certain axioms is equivalent to maximization of expected utility/welfare. They differ mainly in that the probability distribution on the state space used to form expected utility is pre-specified in the former work and determined within the theory in the latter. Von Neumann and Morgenstern viewed the probability distribution as a primitive concept. Savage viewed the distribution as a construct that may in principle be inferred from analysis of choice behavior. In neither theorem does the distribution have any necessary connection to an objective reality,

When studying consistency axioms of the types posed by Von Neumann-Morgenstern and Savage, decision theorists do not differentiate between private entities and social planners. The presumption is that all decision makers should behave consistently in the same manner. However, some theorists have proposed that social planners should adhere to additional ethical axioms that require them, in some sense, to respect the preferences of their populations and/or behave fairly. Recent review articles include Fleurbaey (2018) and Mongin and Privato (2016).

Axiomatic decision theory provides no description of substantively good decisions. Standard choice axioms only aim to characterize procedural reasonableness (often called rationality), in the sense of consistency of hypothetical behavior across potential choice problems. Ethical axioms proposed for social planning only consider some broad properties that a welfare function should presumably satisfy.

1.2.2. Consequentialist Decision Theory

Consequentialist decision theory specifies a welfare function and an expression of uncertainty as primitives. It then seeks reasonable criteria to make decisions. The most prevalent recommendation has been maximization of expected utility. One places a probability distribution on the state space and chooses an action that maximizes the expected value of welfare with respect to this distribution. To assist decision makers who do not find it credible to express uncertainty through a probability distribution, decision theorists have mainly studied criteria that, in some sense, works uniformly well over all of the state space. Two prominent interpretations of this broad idea are the maximin and minimax-regret criteria. I will formalize and apply these criteria later in this paper.

The decision theory used in my research is consequentialist rather than axiomatic. I suppose throughout that the objective of social planning is to make substantively good decisions in particular settings. To accomplish this, I suppose that a planner specifies a relevant welfare function, expresses uncertainty in a credible manner, and uses these primitives to make a decision. The relevance of a welfare function and the credibility of an expression of uncertainty are context specific.

As I see it, a fundamental problem with axiomatic decision theory is that it does not engage the fundamental substantive matter of the credibility of a decision maker's expression of uncertainty. Theory in the tradition of Von Neumann and Morgenstern often assumes that a decision maker holds accurate probabilistic expectations, commonly called rational expectations, without explaining how this may be accomplished in practice. The accuracy of probabilistic expectations is not a concern to theory in the Savage tradition. However, t he realism of expectations should matter to any decision maker. It should matter particularly to a planner that represents a population.

1.3. Characterizing Uncertainty Regarding Objective Probability Distributions

To characterize uncertainty with enough concreteness to be useful to the study of social planning, I draw primarily on my own econometric research on partial identification. Perhaps the most fundamental difficulty when studying planning is the identification problem arising from the unobservability of counterfactual outcomes. At most one can observe the outcomes that occur under realized policies. The outcomes of unrealized policies are logically unobservable. Yet determination of an optimal policy requires comparison of all feasible policies. For this and many other reasons, planners usually have only partial knowledge of the welfare achieved by alternative policies.

I first connected identification analysis with decision making under uncertainty in Manski (2000), writing (page 416):

> "This paper connects decisions under ambiguity with identification problems in econometrics. Considered abstractly, it is natural to make this connection. Ambiguity occurs when lack of knowledge of an objective probability distribution prevents a decision maker from solving an optimization problem. Empirical research seeks to draw conclusions about objective probability distributions by combining assumptions with observations. An identification problem occurs when a specified set of

> assumptions combined with unlimited observations drawn by a specified sampling process does not reveal a distribution of interest. Thus, identification problems generate ambiguity in decision making."

Here and elsewhere, I have followed Ellsberg (1961) in using the word "ambiguity" to signify uncertainty when one specifies a set of feasible states of nature but does not place a probability distribution on the state space. Synonyms for ambiguity include "deep uncertainty" and "Knightian uncertainty."

In Manski (2000), I also noted that (page 416): "Statistical problems of induction from finite samples to populations generate further ambiguity." Statistical imprecision in empirical analysis is relevant to social planning. However, identification is generally the deeper and more profound source of uncertainty. Hence, I discuss statistical imprecision only in a few places in this paper.

What are the objective uncertainties with which social planning must cope? They are too many and varied to summarize easily. For the moment, I will simply list those that I have studied. There are numerous identification problems in medical risk assessment and prediction of individualistic treatment response; see Manski (2019a) for a broad exposition. There is deep uncertainty in the epidemiological models used to predict the spread of infectious diseases, which inform choice of vaccination policy (Manski, 2010, 2017). There is deep uncertainty in the physical-science climate models used to predict future climate change, which inform choice of climate policy (Manski, Sanstad, and DeCanio, 2021), and in the discount rate used to form a social welfare function (DeCanio, Manski, and Sanstad, 2022).

Identification problems arise when studying the preferences and behavior of human populations. Knowledge of preferences is essential to policy evaluation when welfare is utilitarian. An ability to predict behavior is required to evaluate policy consequences whatever the welfare function may be. Manski (2007a) provides an abstract analysis. I have examined how uncertainties about preferences and behavior complicate evaluation of income tax policies, where a central consideration is the relative preferences of potential workers for consumption goods and for availability of time to enable non-paid activities (Manski, 2014a, b). And I have shown how uncertainty about the deterrent effect of policing on criminal behavior complicates evaluation of proactive policing programs (Manski, 2006).

1.4. Organization of the Paper

As prelude before engagement with uncertainty, Section 2 discusses practices that have promoted planning with incredible certitude, using specific cases to illustrate. Section 3 describes and contrasts the conceptions of uncertainty in consequentialist and axiomatic decision theory. Sections 4 through 8 summarize my studies of fiv problems of planning under uncertainty: diversified medical or other individualistic treatment under ambiguity (Section 4), treatment choice using data from a classical randomized trial (Section 5), choice of a vaccination policy (Section 6), choice of an income tax policy (Section 7), and choice of a climate policy (Section 8). Section 9 makes concluding comments on the institutional separation of policy analysis and decision making.

2. Planning with Incredible Certitude

Analyses of public policy regularly express certitude about the consequences of alternative policy choices. Expressions of uncertainty are rare. Yet predictions often are fragile. Conclusions may rest on critical unsupported assumptions or on leaps of logic. Then the certitude of policy analysis is not credible.

One 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 and policy makers respond to incentives. The scientific community rewards strong novel findings. The public wants unequivocal policy recommendations. These incentives make it tempting to maintain assumptions far stronger than can be persuasively defended, in order to draw strong conclusions. Morgenstern (1963) remarked that federal statistical agencies may perceive a political incentive to express incredible certitude about the state of the economy when they publish official statistics. He wrote (p. 11):

> "All offices must try to impress the public with the quality of their work. Should too many doubts be raised, financial support from Congress or other sources may not be forthcoming. More than once has

> it happened that Congressional appropriations were endangered when it was suspected that government statistics might not be 100 percent accurate. It is natural, therefore, that various offices will defend the quality of their work even to an unreasonable degree."

Expressing certitude 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.

Manski (2011a) introduced a typology of practices that contribute to incredible certitude. I have since elaborated in Manski (2013, 2015, 2019b, 2020a): The typology is

* conventional certitude: A prediction that is generally accepted as true but is not necessarily true.
* dueling certitudes: Contradictory predictions made with alternative assumptions.
* conflating science and advocacy: Specifying assumptions to generate a predetermined conclusion.
* wishful extrapolation: Using untenable assumptions to extrapolate.
* illogical certitude: Drawing an unfounded conclusion based on logical errors.
* media overreach: Premature or exaggerated public reporting of policy analysis.

I have provided illustrative examples and have offered suggestions to improve practices. I provide examples here of the three components of the typology that are most directly relevant to planning.

2.1. Conventional Certitude: CBO Scoring of Legislation

Conventional certitude is exemplified by U.S. Congressional Budget Office (CBO) scoring of 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, unaccompanied by measures of uncertainty. CBO scores exemplify conventional certitude because they have achieved broad acceptance. They are used by both Democratic and Republican members of Congress. Media reports largely take them at face value.

A well-known example is the scoring of the Patient Protection and Affordable Care Act of 2010, commonly known as Obamacare or the ACA. In March of 2010 the CBO and the Joint Committee on

Taxation (JCT) jointly scored the combined consequences of the ACA and the Reconciliation Act of 2010 and reported (Elmendorf, 2010, p.2): “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 score as fact without questioning its validity, the hallmark of conventional certitude.

A simple approach to avoid incredible certitude 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. Or it could present a full probabilistic forecast in a graphical fan chart such as the Bank of England uses to predict GDP growth. If the CBO must provide a point prediction, it can continue to do so, with some convention used to locate the point within the interval forecast.

2.2. Dueling Certitudes: The Deterrent Effect of the Death Penalty

American society has long debated the deterrent effect of the death penalty as a punishment for murder. Disagreement persists because research has not been able to settle the question. Researchers have used data on homicide rates and sanctions across states and years to examine the deterrent effect of the death penalty. The fundamental difficulty is that the outcomes of counterfactual policies are unobservable. Data alone cannot reveal what the homicide rate in a state without (with) a death penalty would have been had the state (not) adopted a death penalty statute. Data must be combined with assumptions to predict homicides under counterfactual deterrence policies.

A large body of work has addressed deterrence and the death penalty, yet the literature has failed to achieve consensus. Researchers studying the question have used much the same data, but they have maintained different assumptions and have consequently reached different conclusions. Rather than acknowledge uncertainty about the realism of its maintained assumptions, each published article touts its findings as accurate. The result is dueling certitudes across articles.

Two committees of the National Research Council have documented the substantial variation in research findings and have investigated in depth the problem of inference on deterrence; see Blumstein, Cohen, and Nagin (1978) and National Research Council (2012). The latter committee, reiterating a basic conclusion of the former one, wrote (p. 2): "The committee concludes that research to date on the effect of capital punishment on homicide is not informative about whether capital punishment decreases, increases, or has no effect on homicide rates."

To illustrate in a simple setting how research that uses the same data but different assumptions can reach very different findings, Manski and Pepper (2013) examined data from the critical 1970s period when the Supreme Court decided the constitutionality of the death penalty. The 1972 Supreme Court case *Furman vs. Georgia* resulted in a multi-year moratorium on the application of the death penalty, while the 1976 case *Gregg vs. Georgia* ruled that the death penalty could be applied subject to certain criteria. We examined the effect of death penalty statutes on homicide rates in two years: 1975, the last full year of the moratorium, and 1977, the first full year after the moratorium was lifted. In 1975 the death penalty was illegal throughout the country. In 1977 thirty-two states had legal death penalty statutes. For each state and year, we observe the homicide rate and whether the death penalty is legal.

We computed three simple estimates of the effect of death penalty statutes on homicide. A "before-and-after" analysis compares homicide rates in the treated states in 1975 and 1977. The 1975 homicide rate in these states, when none had the death penalty, was 10.3 per 100,000. The 1977 rate, when all had the death penalty, was 9.7. The before-and-after estimate is the difference between the 1977 and 1975 homicide rates; that is 9.7 - 10.3 = -0.6. This is interpretable as the average effect of the death penalty on homicide in the treated states if one assumes that nothing germane to homicide occurred in these states between 1975 and 1977 except for legalization of capital punishment.

Alternatively, one might compare the 1977 homicide rates in the treated and untreated states. The 1977 rate in the treated states, which had the death penalty, was 9.7. The 1977 rate in the untreated states, which did not have the death penalty, was 6.9. The estimate is the difference between these homicide rates; that is, 9.7 - 6.9 = 2.8. This is interpretable as the nationwide average effect of the death penalty on homicide in

1977 if one assumes that persons living in the treated and untreated states have the same propensity to commit murder in the absence of the death penalty and respond similarly to enactment of the death penalty. With this assumption, the observed homicide rate in the treated states reveals what the rate would have been in the untreated states if they had enacted the death penalty, and vice versa.

Yet a third way to use the data is to compare the temporal changes in homicide rates in the treated and untreated states. Between 1975 and 1977 the homicide rate in the treated states fell from 10.3 to 9.7, while the rate in the untreated states fell from 8.0 to 6.9. The so-called *difference-in-difference* (*DID*) estimate is the difference between these temporal changes; that is, (9.7 - 10.3) - (6.9 - 8.0) = 0.5. This is interpretable as the nationwide effect of the death penalty on homicide if one assumes that all states experience a common time trend in homicide and that enactment of the death penalty has the same effect in all states.

These three estimates yield different empirical findings regarding the effect of the death penalty on homicide. The before-and-after estimate implies that enactment of a death penalty statute reduces the homicide rate by -0.6 per 100,000. The other two estimates imply that having the death penalty raises the homicide rate by 2.8 or 0.5 per 100,000. The idea that capital punishment may increase the homicide rate is contrary to the traditional view of punishment as a deterrent. However, some researchers have argued that the death penalty shows a lack of concern for life that brutalizes society into greater acceptance of commission of murder.

Which estimate is correct? Given certain assumptions, each appropriately measures the effect of the death penalty on homicide. However, the assumptions that justify this interpretation differ across estimates. One may be correct, or none of them. If three researchers were to each maintain a different one of the assumptions and report one of the three estimates, they would exhibit dueling certitudes.

The antidote to dueling certitudes about the deterrent effect of capital punishment is to recognize uncertainty by generating a set of estimates under alternative assumptions. To formalize this idea in a flexible manner, Manski and Pepper (2013) studied the conclusions implied by relatively weak ‘bounded variation’ assumptions that restrict variation in treatment response across places and time. The results are findings that bound the deterrent effect of capital punishment. By successively adding stronger identifying assumptions, we seek to make transparent how assumptions shape inference. We performed empirical analysis using state-level data in the United States in 1975 and 1977. Under the weakest restrictions, there is substantial ambiguity: we cannot rule out the possibility that having a death penalty statute substantially increases or decreases homicide. This ambiguity is reduced when we impose stronger assumptions, but inferences are sensitive to the maintained restrictions. Combining the data with some assumptions implies that the death penalty increases homicide, but other assumptions imply that the death penalty deters it.

2.3. Wishful Extrapolation: From Medical Research to Patient Care

Extrapolation is essential to policy analysis. A central objective is to inform policy choice by predicting the outcomes that would occur if past policies were to be continued or alternative ones were to be enacted. Researchers often use untenable assumptions to extrapolate. I have called this manifestation of incredible certitude *wishful extrapolation*. To illustrate, I will discuss extrapolation from randomized trials in medicine to inform patient care, drawing on Manski (2019a).

Trials have long enjoyed a favored status within medical research on treatment response and are often called the “gold standard” for such research. The appeal of trials is that, with sufficient sample size and complete observation of outcomes, they deliver credible findings on treatment response in the study population. However, extrapolation of findings from trials to clinical practice can be difficult. Researchers and guideline developers often use untenable assumptions to extrapolate.

2.3.1. From Study Populations to Patient Populations

Study populations in trials often differ from patient populations. It is common to perform trials studying treatment of a specific disease only on subjects who have no co-morbidities. Another source of difference between study and patient populations is that a study population consists of persons with specified demographic attributes who volunteer to participate in a trial. Participation in a trial may be restricted to persons in certain age categories who reside in certain locales. Among such persons, volunteers are those who respond to financial and medical incentives to participate. It may be wishful extrapolation to assume that treatment response in trials performed on volunteers with specified demographic attributes who lack co-morbidities is the same as what would occur in actual patient populations.

To justify trials performed on study populations that may differ substantially from patient populations, researchers often cite Donald Campbell, who distinguished between the internal and external validity of studies of treatment response (Campbell and Stanley, 1963). A study has *internal validity* if it has credible findings for the study population. It has *external validity* if an invariance assumption permits credible extrapolation. The appeal of randomized trials is their internal validity. Wishful extrapolation is an absence of external validity.

Campbell argued that studies should be judged primarily by their internal validity and secondarily by their external validity. This perspective has been used to argue for the primacy of experimental research over observational studies, whatever the study population may be. The Campbell position is well grounded if treatment response is homogeneous. Then researchers can learn about treatment response in easy-to-analyze study populations and clinicians can confidently extrapolate findings to patient populations. However, homogeneity of treatment response seems the exception rather than the rule. Hence, it may be wishful to extrapolate from a study population to a patient population.

2.3.2. From Experimental Treatments to Clinical Treatments

Treatments in trials often differ from those that occur in clinical practice. This is particularly so in trials comparing drug treatments. Drug trials are double-blinded, neither the patient nor the clinician knowing the assigned treatment. A double-blinded drug trial reveals the distribution of response in a setting where patients and clinicians are uncertain what treatment a patient is receiving. It does not reveal what response would be when patients and clinicians know what drug is being administered and can react to this information.

Consider drug treatments for hypertension. Patients react heterogeneously to the various drugs available for prescription. A clinician treating a specific patient may sequentially prescribe alternative drugs, trying each for a period in an effort to find one that performs satisfactorily. Sequential experimentation is not possible in a blinded trial. The standard protocol prohibits the clinician from knowing what drug a subject is receiving and from using judgment to modify the treatment. Blinding is also problematic for interpretation of noncompliance with assigned treatments.

2.3.3. Wishful Meta-Analyses of Disparate Studies

The problems discussed above concern analysis of findings from a single trial. Further difficulties arise when one attempts to combine findings from multiple trials. It is easy to understand the impetus for combination of findings.

Decision makers must somehow interpret the mass of information provided by empirical research. The hard question is how to interpret this information sensibly. Combination of findings is sometimes performed by *systematic review* of a set of studies. This is a subjective process similar to exercise of clinical judgment.

Statisticians have proposed *meta-analysis,* attempting to provide an objective methodology for combining the findings of multiple studies. Meta-analysis was originally developed to address a purely statistical problem. Suppose that multiple trials have been performed on the same population, each drawing an independent random sample. The best way to use the data combines them into one sample.

Suppose that the raw data are unavailable. Instead, multiple parameter estimates are available, each computed with the data from a different sample. Meta-analysis proposes methods to combine the estimates. A common proposal computes a weighted average, weighting estimates by sample size.

The original concept of meta-analysis is uncontroversial, but its applicability is limited. It is common to have multiple disparate studies. The studies may examine distinct patient populations, whose members may have different risk of disease or different distributions of treatment response. Administration of treatments and measurement of outcomes may vary. Gene Glass, who introduced the term *meta-analysis,* wrote (Glass, 1977): "The tough intellectual work in many applied fields is to make incommensurables commensurable, in short, to compare apples and oranges."

Meta-analysis is performed often in such settings, computing weighted averages of estimates for distinct study populations and trial designs. Meta-analyses often use a *random-effects* model (DerSimonian and Laird, 1986). The model considers trials to be drawn at random "from a population of possible studies." Then each trial estimates a parameter drawn at random from a population of possible parameters. A weighted average estimates the mean of these parameters.

The relevance to clinical practice is obscure. DerSimonian and Laird do not explain what is meant by a population of possible studies, nor why published studies should be considered a random sample from this population. They do not explain how a population of possible studies connects to what matters to a clinician−−the distribution of health outcomes across the relevant population of patients.

Manski (2020b) draws on econometric research on partial identification to propose principles for patient-centered meta-analysis. One specifies a prediction of concern and determines what each available study reveals. Given common imperfections in internal and external validity, studies typically yield credible set-valued rather than point predictions. Thus, a study may enable one to conclude that a probability of disease, or mean treatment response, lies within a range of possibilities. Patient-centered meta-analysis would combine the findings of multiple studies by computing the intersection of the set-valued predictions that they yield.

3.Uncertainty in Consequentialist and Axiomatic Decision Theory

3.1. Consequentialist Decision Theory

3.1.1. The Choice Set, State Space, and Welfare Function

As observed in the Introduction, welfare economics has studied maximization of a welfare function that maps actions into welfare. Researchers have commonly assumed that a planner knows enough about the choice environment to be able to determine an optimal action. Maximization of welfare expresses a consequential perspective on decision making. Given a choice setting, the goal is to do as well as possible in achieving a specified objective.

To study decision making under uncertainty, the starting point of consequentialist decision theory has been to suppose that the planner or other decision maker faces a predetermined choice set C and believes that the true state of nature s^* lies in a state space S. The welfare function $w(\cdot, \cdot)$: $C \times S \Rightarrow R^1$ maps actions and states into welfare. The planner wants to maximize $w(\cdot, s^*)$ over C but does not know s^*. Hence, maximization is infeasible except in special cases.

The state space S provides the basic decision theoretic expression of uncertainty. In lay language, S is a list of "known unknowns." States of nature that are not elements of S are presumed impossible to occur. Decision theory supposes that the decision maker does not contemplate the possible existence of unlisted "unknown unknowns."

Discussions of the state space often consider it to express uncertainty purely about the physical or social environment within which choice takes place. However, a state space can also express uncertainty about the welfare function that a planner should maximize. For example, this may occur when the planner is utilitarian and, hence, must know the preferences of the population in order to maximize welfare. The preferences of the population may be uncertain.

Being a primitive of the decision problem, the state space is necessarily subjective. This does not imply, however, that it is an arbitrary construction. Credibility is a fundamental matter in consequential decision theory in general and in the study of social planning specifically. If planning decisions are to enhance societal well-being in the real world, the planner should specify a state space that embodies some reasonable sense of credibility.

Although the concept of credibility is difficult to pin down, scientific research seeks to help by providing at least a partially objective basis for specification of the state space. This basis is obtained by combining plausible theory with empirical analysis. See Section 3.1.4 for further discussion.

3.1.2. Decision Criteria without Sample Data

It is generally accepted that decisions should respect dominance. Action $c \in C$ is weakly dominated if there exists a $d \in C$ such that $w(d, s) \geq w(c, s)$ for all $s \in S$ and $w(d, s) > w(c, s)$ for some $s \in S$. To choose among undominated actions, decision theorists have proposed various ways of using $w(\cdot, \cdot)$ to form functions of actions alone, which can be optimized. In principle, one should only consider undominated actions, but it often is difficult to determine which actions are undominated. Hence, in practice it is common to optimize over the full set of feasible actions. I define decision criteria accordingly.

I initially consider settings without sample data, describing three prominent criteria. I extend these criteria to settings with sample data in the next sub-section. Consequentialist decision theory views the welfare function, state space, and decision criterion as meta-choices made by a decision maker. It views these meta-choices as predetermined rather than matters to be studied within the theory.

A familiar idea is to place a subjective probability distribution π on the state space, average state-dependent welfare with respect to π, and maximize subjective average welfare over C. The criterion solves

$$(1) \qquad \max_{c \in C} \int w(c, s) d\pi.$$

Given a subjective distribution π on S, one need not average over π. Any criterion respecting stochastic dominance has a consequential claim to be reasonable. For example, one might maximize quantile welfare, as studied in Manski (1988).

Another idea seeks an action that, in some sense, works uniformly well over all of S. This yields the maximin and minimax-regret (MMR) criteria. The maximin criterion maximizes the minimum welfare attainable across S, solving the problem

$$(2) \qquad \max_{c \in C} \; \min_{s \in S} \; w(c, s).$$

The MMR criterion solves

$$(3) \quad \min_{c \in C} \; \max_{s \in S} \; [\max_{d \in C} w(d, s) - w(c, s)].$$

Here $\max_{d \in C} w(d, s) - w(c, s)$ is the *regret* of action c in state s. The true state being unknown, one evaluates c by its maximum regret over all states and selects an action that minimizes maximum regret. The maximum regret of an action measures its maximum distance from optimality across states. Hence, maximum regret is uniform nearness to optimality.

The above confines attention to polar cases in which a planner asserts a complete subjective distribution on the state space, or none. A planner might assert a partial subjective distribution, placing lower and upper probabilities on states as in Dempster (1968) or Walley (1991), and then maximize minimum subjective average welfare or minimize maximum average regret. These criteria combine elements of averaging across states and concern with uniform performance across states. Statistical decision theorists refer to them as Γ-maximin and Γ-minimax regret (Berger, 1985). The former has drawn attention from axiomatic decision theorists, who call it maxmin expected utility (Gilboa and Schmeidler, 1989).

3.1.3. Decision Criteria with Sample Data

Abraham Wald, in a series of contributions culminating in Wald (1950), extended consequentialist decision theory to encompass settings where the decision maker observes sample data. Wald's formulation of statistical decision theory supposes that a decision maker observes data generated by a sampling distribution which is a known function of the state of nature. To express this, let the feasible sampling distributions be denoted $(Q_s, s \in S)$. Let Ψ_s denote the sample space in state s; Ψ_s is the set of samples that may be drawn under distribution Q_s. The literature typically assumes that the sample space does not vary with s and is known. I do likewise and denote the sample space as Ψ. A statistical decision function (SDF), $c(\cdot): \Psi \rightarrow C$, maps the sample data into a chosen action.

An SDF is a deterministic function after realization of the sample data, but it is a random function ex ante. Hence, an SDF generically makes a randomized choice of an action. The only exceptions are SDFs that make almost-surely data-invariant choices. An SDF $c(\cdot)$ is almost-surely data-invariant in state s if there exists a $d \in C$ such that $Q_s[c(\psi) = d] = 1$.

Given that SDFs are random functions, welfare using a specified SDF is a random variable ex ante. Wald's theory evaluates the performance of SDF $c(\cdot)$ in state s by $Q_s\{w[c(\psi), s]\}$, the ex-ante distribution of welfare that it yields across realizations ψ of the sampling process. In particular, Wald measured the performance of $c(\cdot)$ in state s by its expected welfare across samples; that is, $E_s\{w[c(\psi), s]\} \equiv \int w[c(\psi), s]dQ_s$. Not knowing the true state, a planner evaluates $c(\cdot)$ by the state-dependent expected welfare vector $(E_s\{w[c(\psi), s]\}, s \in S)$, which is computable.

Statistical decision theory has mainly studied the same decision criteria as has decision theory without sample data. Let Γ denote the set of feasible SDFs, which map $\Psi \rightarrow C$. The statistical versions of criteria (1), (2), and (3) are

$$(4) \qquad \max_{c(\cdot) \in \Gamma} \int E_s\{w[c(\psi), s]\} \, d\pi,$$

$$(5) \qquad \max_{c(\cdot) \in \Gamma} \ \min_{s \in S} \ E_s\{w[c(\psi), s]\},$$

$$(6) \qquad \min_{c(\cdot) \in \Gamma} \ \max_{s \in S} \ \left(\max_{d \in C} w(d, s) - E_s\{w[c(\psi), s]\}\right).$$

Wald's statistical decision theory is frequentist. In settings of choice between two actions, SDFs functions can be viewed as hypothesis tests. However, evaluation of tests in the Wald theory differs fundamentally from the well-known methodology of Neyman-Pearson hypothesis testing. Decision theory does not restrict attention to tests that yield a predetermined upper bound on the probability of a Type I error. Nor does it aim to minimize the maximum value of the probability of a Type II error when more than a specified minimum distance from the null hypothesis. Wald proposed for binary choice, as elsewhere, evaluation of the performance of SDF $c(\cdot)$ in state s by the expected welfare that it yields across realizations of the sampling process. See Section 5 below and Manski (2021) for further discussion.

3.1.4. Learning Objective Probability Distributions

In many planning settings, it has become standard to specify the state space as a set of objective probability distributions that may possibly describe the system under study. Haavelmo (1944) did so for economic systems when he introduced "The Probability Approach in Econometrics." Studies of treatment choice do so when they consider the population to be treated to have a distribution of treatment response functions.

Research seeks to enhance the credibility of planning by constructive combination of theory and empirical analysis to provide information about the possible distributions. The Koopmans (1949) formalization of identification analysis contemplated unlimited data collection that enables one to shrink the state space, eliminating distributions that are inconsistent with theory and with the information revealed by observation. Sample data generally are not informative enough to shrink the state space. Wald's development of statistical decision theory shows how sample data can be informative, nonetheless.

For most of the 20th century, econometricians commonly thought of identification as a binary event – a feature of a probability distribution (a parameter) is either identified or it is not. Empirical researchers applying econometric methods combined available data with assumptions that yield point identification, and they reported point estimates of parameters. Many economists recognized with discomfort that point identification often requires strong assumptions that are difficult to motivate. However, they saw no other way to perform inference.

Yet there is enormous scope for fruitful inference using weaker and more credible assumptions that partially identify population parameters. A parameter is partially identified if the sampling process and maintained assumptions reveal that the parameter lies in a set, its 'identification region' or "identified set', that is smaller than the logical range of the parameter but larger than a single point. Estimates of partially identified parameters generically are set-valued; a natural estimate of an identification region is its sample analog.

These were scattered contributions to analysis of partial identification as early as the 1930s, but the subject remained at the fringes of econometric consciousness and did not spawn systematic study. However, a coherent body of research took shape in the 1990s and has grown rapidly. Reviews of this work include Manski (2003, 2007b), Tamer (2010), and Molinari (2020).

The modern literature on partial identification emerged out of concern with traditional approaches to inference with missing outcome data. Empirical researchers have commonly assumed that missingness is random, in the sense that the observability of an outcome is statistically independent of its value. Yet this and other point-identifying assumptions have regularly been criticized as implausible. So it was natural to ask what random sampling with partial observability of outcomes reveals about outcome distributions if nothing is known about the missingness process or if assumptions weak enough to be widely credible are imposed.

Studying inference with missing outcome data led naturally to consideration of conditional prediction and analysis of treatment response. A common objective of empirical research is to predict an outcome conditional on given covariates, using data from a random sample of the population. Analysis of treatment response must contend with the fundamental problem that counterfactual outcomes are not observable; hence, findings on partial identification with missing outcome data are directly applicable. Yet analysis of treatment response poses much more than a generic missing-data problem. One reason is that observations of realized outcomes, when combined with suitable assumptions, can provide information about counterfactual ones. Another is that practical problems of treatment choice motivate research on treatment response and thereby determine what population parameters are of interest. For these reasons, it has been productive to study partial identification of treatment response as a subject in its own right.

Whatever the specific subject under study, a common theme runs through the literature on partial identification. One first asks what the sampling process alone reveals about the population of interest and then studies the identifying power of assumptions that aim to be credible in practice. This conservative approach to inference makes clear the conclusions one can draw in empirical research without imposing untenable assumptions. It establishes a domain of consensus among analysts who may hold disparate beliefs about what assumptions are appropriate. It also makes plain the limitations of the available data. When credible identification regions turn out to be large, we should face up to the fact that the available data do not support inferences as tight as we might like to achieve.

From the perspective of planning, findings on partial identification imply that empirical research may shrink the state space for decision making but not reduce it to a single state of nature. Let S be the state space without observation of the unlimited data assumed in an identification study. Let $S_0 \subset S$ be the shrunken state space with these data. Then decision criteria (1) through (3) posed in Section 3.1.1 have the same forms, but with S_0 replacing S. In (1), the conditional subjective distribution $\pi(s|s \subset S_0)$ replaces $\pi(s)$.

3.2. Axiomatic Decision Theory

Axiomatic decision theory studies consistency of behavior across hypothetical choice settings. Consistency is expressed through adherence to sets of choice axioms that researchers consider worthy of analysis. Axiomatic theorists have long debated which specific axioms have normative appeal. Appraisal of normative appeal rests on introspection, so there should be no expectation that consensus will emerge. Indeed, decision theorists exhibit considerable difference in opinion. Binmore (2009) catalogues and assesses a wide spectrum of consistency axioms.

3.2.1. Representation Theorems

The staple formalism of axiomatic decision theory is a representation theorem that considers a collection of hypothetical choice settings and proposes axioms that mandate specific forms of consistency of behavior across settings. The theorem proves that adherence to the axioms is necessary and sufficient for behavior across settings to be representable as solution of a particular optimization problem.

Consider the famous Von Neumann and Morgenstern (1944) and Savage (1954) representation theorems. Both begin with a basic axiom stipulating that an agent has a complete binary preference ordering over a class A of actions. They then propose further axioms mandating certain consistency properties for the preference ordering. The theorems prove that adherence to the axioms is necessary and sufficient for representation of behavior when facing any hypothetical choice set $D \subset A$ as maximization of expected utility.

Consequentialist decision theory takes the utility function to be a primitive specified by the decision maker to express what he wants to achieve. In contrast, the representation theorems of axiomatic theory view the utility function as a mathematical construct implied by hypothetical choice behavior. In neither the Von Neumann-Morgenstern nor the Savage theorem does the distribution on the state space have any necessary connection to an objective reality. Considering this distribution, Berger (1985) cautioned that (p. 121): “a Bayesian analysis may be ‘rational’ in the weak axiomatic sense, yet be terrible in a practical sense if an inappropriate prior distribution is used.” Berger’s comment expresses the consequentialist perspective that a decision maker should express uncertainty in a realistic manner.

Although the Von Neumann-Morgenstern and Savage theorems both represent behavior as maximization of expected utility, they differ in how they view uncertainty. A central primitive of Von Neumann and Morgenstern is an externally specified probability distribution on the state space. Discussions of the theorem often presume this to be a credible objective distribution, but it could be a subjective distribution formed by a cognitive process. The Savage theorem does not pre-specify a distribution on the state space. Instead, it proves that a decision maker who adheres to the axioms behaves as if he maximizes expected utility using a (utility function, state-space distribution) pair implied by hypothetical choice behavior. Thus, the utility function and the probability distribution of the Savage theorem are both constructs determined with his representation theorem.

Axiomatic decision theorists often use language that obscures the distinction between hypothetical and actual choice behavior. They often describe axiomatic theory as revealed preference analysis. Consider, for example, this passage in Savage (1954) concerning two actions labeled f and g (p. 17): “I think it of great importance that preference, and indifference, between f and g be determined, at least in principle, by decisions between acts and not by response to introspective questions.” The critical phrase in this sentence is “at least in principle.” The enormously rich choice data contemplated in the Savage axioms are essentially never available in practice. This has been pointed out repeatedly over the years, at least as early as Sen (1973). Nevertheless, some theorists continue to describe their subject as revealed preference analysis; see Gul and Pesendorfer (2008).

3.2.2. Is Axiomatic Theory Relevant to Planning?

In Manski (2011), I argued from a consequentialist perspective that a decision maker facing an actual choice problem is not necessarily concerned with the consistency of his behavior across hypothetical choice scenarios. The decision maker wants to make a substantively reasonable choice in the setting that he actually faces. I called this idea *actualist rationality,* stating (p. 196): "Prescriptions for decision making should promote welfare maximization in the choice problem the agent actually faces."

From the perspective of actualist rationality, one need not introspect regarding the normative appeal of choice axioms. Axiomatic theory might become relevant if researchers were to show that adherence to certain axioms promotes substantively good decision making. However, this has not been the objective of axiomatic theory. The representation theorems of axiomatic theory are interpretative rather than prescriptive. Indeed, the decision maker contemplated in axiomatic theory is assumed to know how he would behave when facing any choice set. Hence, he has no need for prescriptions.

3.3. Minimax-Regret Planning

Among the decision criteria posed in Section 3.1, maximization of subjective average welfare places a specified subjective distribution on the state space, whereas maximin and MMR do not. Concern with the basis for specification of a subjective distribution motivated Wald (1950) to study the minimax criterion (maximin in my description), writing (p. 18): "a minimax solution seems, in general, to be a reasonable solution of the decision problem when an a priori distribution . . . . does not exist or is unknown."

I am similarly concerned with decision making with no subjective distribution on states. However, I have mainly measured performance of decisions by maximum regret rather than by minimum welfare. The maximin and MMR criteria both provide ex ante evaluations of the worst result that a decision maker may experience ex post. However, the criteria are equivalent only in special cases, particularly when optimal welfare is invariant across states. They differ more generally. Whereas maximin considers the worst absolute outcome that an action may yield across states, MMR considers the worst outcome relative to what is achievable in a given state.

A conceptual appeal of using maximum regret to measure performance is that it quantifies how lack of knowledge of the true state of nature diminishes the quality of decisions. The term “maximum regret” is shorthand for the maximum sub-optimality of a decision criterion across the feasible states of nature. A decision with small maximum regret is uniformly near optimal across all states. This is a desirable property.

MMR has drawn diverse reactions from axiomatic decision theorists. In a famous early critique, Chernoff (1954) observed that MMR decisions are not always consistent with the choice axiom known as independence of irrelevant alternatives (IIA). He considered this a serious deficiency. Chernoff’s view has been endorsed by some modern decision theorists, such as Binmore (2009). However, Sen (1993) argued that adherence to axioms such as IIA does not per se provide a sound basis for evaluation of decision criteria. He asserted that consideration of the context of decision making is essential.

Manski (2011) argued that adherence to the IIA axiom is not a virtue per se. What matters is how violation of the axiom affects welfare. I observed that the MMR violation of the IIA axiom does not yield choice of a dominated decision. The MMR decision is always undominated when it is unique. There generically exists an undominated MMR decision when the criterion has multiple solutions. Hence, I concluded that violation of the IIA axiom is not a sound rationale to dismiss minimax regret.

4. Diversified Treatment under Ambiguity

I summarize here the study of diversified treatment under ambiguity initiated in Manski (2007b, Chapter 11) and expanded in Manski (2009). In these works, I considered settings in which a planner can treat persons differentially. Examples include medical treatment, sentencing of offenders, and active labor-market programs. The planner may make a *singleton* allocation, assigning all observationally identical persons to the same treatment. Or he could choose a *fractional* allocation, randomly assigning positive fractions of these persons to different treatments.

Portfolio choice in finance has long been framed as a choice among fractional allocations, but social planning has commonly been viewed as a choice between singleton allocations. Fractional allocations cope with ambiguity through diversification. Suppose that there are two feasible treatments, labeled A and B. Diversification enables a decision maker to balance two types of potential error. A Type A error occur when treatment A is chosen but is actually inferior to B, and a Type B error occurs when B is chosen but is inferior to A. The singleton allocation assigning everyone to treatment A entirely avoids type B errors but may yield Type A errors, and vice versa for singleton assignment to treatment B. Fractional allocations make both types of errors but reduce their potential magnitudes.

4.1. One-Period Problems with Individualistic Treatment and Linear Welfare

I begin with a simple leading case. Each member j of a population J of observationally identical persons has a response function $y_j(\cdot)$: mapping treatments t into outcomes $y_j(t)$. $P[y(\cdot)]$ is the population distribution of treatment response. The population is large, with $P(j) = 0$ for all $j \in J$.

The task is to allocate the population to treatments A and B. An allocation $\delta \in [0, 1]$ randomly assigns a fraction δ of the population to treatment B and the remaining $1 - \delta$ to treatment A. A utilitarian or paternalistic planner wants to maximize mean individual welfare. Let $u_j(t) \equiv u_j[y(t), t]$ be the welfare of person j when this person receives treatment t and realizes outcome $y_j(t)$. Assuming that this welfare is cardinal and interpersonally comparable, let $\alpha \equiv E[u(A)]$ and $\beta \equiv E[u(B)]$ be mean welfare if all members of the population receive treatment A or B. Then mean welfare with allocation δ is

$$(7) \qquad W(\delta) = \alpha(1 - \delta) + \beta\delta = \alpha + (\beta - \alpha)\delta.$$

$\delta = 1$ is optimal if $\beta \geq \alpha$ and $\delta = 0$ if $\beta \leq \alpha$.

The problem is treatment choice when the planner has partial knowledge of the population distribution of welfare, rendering (α, β) partially known. Let S index the feasible states of nature. Let the planner know that (α, β) lies in a bounded set $[(\alpha_s, \beta_s), s \in S]$. Let $\alpha_L \equiv \min_{s \in S} \alpha_s$, $\beta_L \equiv \min_{s \in S} \beta_s$, $\alpha_U \equiv \max_{s \in S} \alpha_s$, and $\beta_U \equiv \max_{s \in S} \beta_s$. Then the planner faces ambiguity if $\alpha_s > \beta_s$ for some values of s and $\alpha_s < \beta_s$ for other values. I suppose below that the planner faces ambiguity.

4.1.1. Bayesian and Maximin Planning

A Bayesian planner places a subjective distribution π on S and solves

$$(8) \qquad \max_{\delta \in [0, 1]} \; E_\pi(\alpha) + [E_\pi(\beta) - E_\pi(\alpha)]\delta,$$

where $E_\pi(\alpha) = \int \alpha_s d\pi$ and $E_\pi(\beta) = \int \beta_s d\pi$. The planner chooses $\delta = 0$ if $E_\pi(\beta) < E_\pi(\alpha)$ and $\delta = 1$ if $E_\pi(\beta) > E_\pi(\alpha)$. He is indifferent among all δ if $E_\pi(\beta) = E_\pi(\alpha)$. Thus, a Bayesian planner makes a singleton choice if $E_\pi(\beta) \neq E_\pi(\alpha)$.

A maximin planner solves

$$(9) \qquad \max_{\delta \in [0, 1]} \; \min_{s \in S} \; \alpha_s + (\beta_s - \alpha_s)\delta.$$

If (α_L, β_L) is feasible, the decision is $\delta = 0$ if $\beta_L < \alpha_L$, $\delta = 1$ if $\beta_L > \alpha_L$, and all δ if $\beta_L = \alpha_L$. Thus, a maximin planner makes a singleton choice if (α_L, β_L) is feasible and $\beta_L \neq \alpha_L$.

4.1.2. MMR Planning

The regret of allocation δ in state of nature s is the difference between the maximum achievable welfare and the welfare achieved with allocation δ. Maximum welfare in state of nature s is max (α_s, β_s). The minimax-regret criterion is

$$(10) \qquad \min_{\delta \in [0, 1]} \quad \max_{s \in S} \quad \max(\alpha_s, \beta_s) - [\alpha_s + (\beta_s - \alpha_s)\delta].$$

Manski (2007b, Chapter 11) derived the MMR treatment allocation. Let $S(A) \equiv \{s \in S: \alpha_s > \beta_s\}$ and $S(B) \equiv \{s \in S: \beta_s > \alpha_s\}$. Let $M(A) \equiv \max_{s \in S(A)} (\alpha_s - \beta_s)$; $M(B) \equiv \max_{s \in S(B)} (\beta_s - \alpha_s)$. The MMR allocation to treatment B is the fraction $\delta_{MR} = M(B)/[M(A) + M(B)]$. If (α_L, β_U) and (α_U, β_L) are feasible, this expression reduces to $\delta_{MR} = (\beta_U - \alpha_L)/[(\alpha_U - \beta_L) + (\beta_U - \alpha_L)]$. In contrast to the situation with Bayesian and maximin planning, the MMR allocation of two treatments is always fractional under ambiguity.

The MMR allocation under ambiguity is not always fractional when a planner allocates the population among more than two treatments. Stoye (2007) studied a class of such problems and has found that the MMR allocations are subtle to characterize. They often are fractional, but he gave an example in which there exists a unique singleton allocation.

4.2. Other Welfare Functions

It is of interest to study planning when the social welfare function is other than mean individual welfare. Manski (2009) reports findings on several cases. I summarize two here.

4.2.1. Welfare Monotone in Mean Individual Welfare

Let $W(\delta) = f[\alpha + (\beta - \alpha)\delta]$, where $f(\cdot)$ is strictly increasing. The Bayes decision is generically singleton if $f(\cdot)$ is convex, but it may be fractional if $f(\cdot)$ has concave segments. In finance, this is the well-known finding that a risk-seeking investor, whose utility is convex in income, does not diversify but a risk-averse investor, whose utility is concave in income, may diversify.

The shape of f(·) does not affect the maximin decision. The reason is that the maximin criterion only uses ordinal, not cardinal properties of the welfare function. Manski (2009) shows that the MMR allocation is fractional whenever f(·) is continuous and the planner faces ambiguity. If f(·) = log(·) and $\{(\alpha_L, \beta_U), (\alpha_U, \beta_L)\}$ are feasible, then the MMR allocation is $\delta_{MR} = [\alpha_U(\beta_U - \alpha_L)]/[\alpha_U(\beta_U - \alpha_L) + \beta_U(\alpha_U - \beta_L)]$.

4.2.2. Deontological Welfare Functions

Deontological ethics supposes that choices may have intrinsic value, apart from their consequences. *Equal treatment of equals* is sometimes thought to be an important deontological principle. Fractional allocations adhere to the principle in the ex ante sense that all persons have equal probabilities of receiving particular treatments. Fractional allocations are inconsistent with equal treatment in the ex post sense that all persons do not actually receive the same treatment.

Manski (2009) studied welfare functions that express concern with ex post equal treatment by subtracting a fixed cost from welfare when the treatment allocation is not singleton. It was found that the MMR allocation may be singleton or fractional, depending on the specifics of the case.

4.3. Sequential Planning Problems

Manski (2009) also considered a sequential planning problem where, in each period $n = 0, \ldots, N$, a planner must choose treatments for the current cohort of a population. Now learning is possible, with observation of the outcomes experienced by earlier cohorts informing treatment choice for later cohorts. Fractional treatment allocations generate randomized experiments, yielding outcome data on both treatments. Sampling variation is not an issue when cohorts are large. All fractional allocations yield the same information.

I considered the *adaptive minimax-regret* (*AMR*) criterion, which applies the static minimax-regret criterion each period using the information available at the time. The AMR criterion is an appealing myopic rule. It treats each cohort as well as possible, in the MMR sense, given the available knowledge. It does not ask the members of one cohort to sacrifice for the benefit of future cohorts. Unless fixed costs make the AMR allocation singleton, it maximizes learning about treatment response.

Randomized clinical trials (RCTs) are regularly used to learn about medical innovations, yielding fractional treatment allocations. However, the allocations produced by the AMR criterion differ from the practice of RCTs in many ways, including these.

(a) The AMR allocation can take any value in [0, 1]. In contrast, the treatment group receiving an innovation in RCTs is typically a small fraction of the population, with sample size determined by conventional calculations of statistical power.

(b) Under the AMR criterion, the persons receiving the innovation are randomly drawn from the full patient population. In contrast, RCTs draw subjects from pools of persons who volunteer to participate.

(c) Under the AMR criterion, one observes outcomes as they unfold over time. In contrast, RCTs typically have short durations. Hence, researchers often measure surrogate outcomes rather than outcomes of real interest.

(d) Under the AMR criterion, assigned treatments are known to subjects and investigators. In contrast, blinded treatment assignment has been the norm in RCTs of new drugs and in some other cases.

(e) Choosing a treatment allocation to minimize maximum regret is remote from the way that RCTs have been used in decision making. The conventional approach has been to perform a hypothesis test, the null hypothesis being that the innovation is no better than the status quo. The innovation is chosen in a singleton manner if and only if the hypothesis is rejected.

5. Using Statistical Decision theory to Choose Treatments with Data from Randomized Trials

5.1. General Ideas

Section 4 studied treatment choice with partial knowledge due to identification problems. I now consider treatment choice using the outcome data generated by a classical randomized trial. There are no identification problems in this setting. The only difficulty is statistical imprecision due to random selection of subjects. Whatever criterion one uses to make treatment decisions based on the results of a trial, there is always some probability that random variation will lead to sub-optimal decisions.

It has been prevalent in medical research to use trial data to compare some notion of standard patient care with an innovation. The prevailing statistical practice has been for researchers to conclude that the innovation is better than standard care only if the estimated average treatment effect comparing the innovation with standard care is statistically significant by conventional criteria. Equivalently, a hypothesis test must reject the null hypothesis that the innovation is no better than standard care.

Considering error probabilities alone is insufficient. The same error probability should be less tolerable when the impact of sub-optimal treatment on patient welfare is larger. To evaluate treatment choice based on trial data, I have in a series of papers used the concept of maximum regret presented abstractly in Section 3.1.3; see Manski (2004) and Manski and Tetenov (2016, 2019, 2021). Maximum regret jointly considers the probability of errors and their magnitudes. It has a simple computable form in the context of randomized trials. I explain here.

Consider specified possible values for average patient outcomes under each treatment. Presuming the common medical focus on average patient outcomes, the ideal clinical decision would be to prescribe a treatment that maximizes average outcome. Trial data do not reveal the best treatment with certainty, so one cannot achieve this ideal. Suppose then that one applies some decision criterion to the trial data. The criterion may be a hypothesis test or another one.

For every treatment that is not best, compute the probability that it would be prescribed when the criterion is applied to the results of a trial. Multiply this error probability by the magnitude of the loss from prescribing this treatment, measured by the difference in average patient outcomes compared to the best treatment. This product, which measures the expected loss from prescribing the inferior treatment, is the regret of Section 3.1.3. The sum of these expected losses across all inferior treatments measures the gap between the ideal of prescribing the best treatment and the reality of having to prescribe the treatment based on trial data that is subject to random variation. Thus, regret measures the expected distance of treatment choice from optimality.

The above calculations are made using specified possible values for average patient outcomes with each treatment. However, trial data do not reveal the true values for average patient outcomes; they only enable one to estimate them. The final measurement step is to look across all a priori possible values for average patient outcomes for all treatments to find the values where the expected loss from prescribing inferior treatments is largest. This measures the uniform nearness to optimality of the proposed criterion for clinical decision making. This is the maximum regret of Section 3.1.3.

5.2. Application to Comparison of Treatments for COVID-19

Manski and Tetenov (2021) illustrated with analysis of the data from an early trial comparing standard-care treatment for severe COVID-19 with standard care augmented by prescription of the drug pair lopinavir–ritonavir. At the beginning of the pandemic, Cao *et al.* (2020) reported on a randomized trial in China comparing these alternatives. The trial assigned 99 hospitalized adult patients to the lopinavir–ritonavir group and 100 to the standard-care only group. The pre-declared primary endpoint measured time to clinical improvement. A secondary outcome was mortality within 28 days.

The authors summarized the primary finding by reporting that lopinavir–ritonavir led to a median time to clinical improvement that was shorter by 1 day than that observed with standard care (hazard ratio, 1.39; 95% CI, 1.00 to 1.91). Regarding mortality, 19 of the 99 patients assigned to lopinavir-ritonavir died within

28 days and 25 of the 100 receiving only standard care died. The authors characterized this finding as follows (p. 1): "Mortality at 28 days was similar in the lopinavir–ritonavir group and the standard-care group (19.2% vs. 25.0%; difference, −5.8 percentage points; 95% CI, −17.3 to 5.7)."

A clinician might reasonably view the estimated reductions in median time to clinical improvement and in mortality to be suggestive, albeit not definitive, evidence that treatment with lopinavir–ritonavir is beneficial relative to standard care alone. Yet the study authors concluded (p. 1) "no benefit was observed with lopinavir–ritonavir treatment beyond standard care." This conclusion was reached because the estimated treatment effects were not statistically significant, having confidence intervals that cover zero.

The statistical analysis of the Cao *et al.* study was not unusual. A report on a trial comparing standard care augmented by the drug remdesivir with standard care alone concluded that remdesivir reduces time to recovery, but the report stated no conclusion regarding mortality (Beigel *et al.* 2020). The report stated that remdesivir reduces median time to recovery from 15 days to 11 days and reduces 14-day mortality from 11.9% to 7.1%. A clinician might reasonably consider both findings to be relevant to treatment choice, but the research team stated a conclusion only for time to recovery. The rationale was that the former finding was statistically significant by conventional criteria and the latter was insignificant.

To illustrate how measurement of nearness to optimality works in practice, Manski and Tetenov (2021) applied two different decision criteria to the trial design in Cao *et al.* (2020). We focused on the outcome of 28-day mortality, presuming that this is the most important outcome for patients with severe COVID-19. One criterion was a standard hypothesis test. The other was the *empirical success* rule, previously studied in Manski (2004) and Manski and Tetenov (2016, 2019). This criterion chooses the treatment with the highest observed average patient outcome in the trial, regardless of the statistical significance of the result. Whereas hypothesis testing favors standard care and places the burden of proof on an innovation, the empirical success rule assesses the evidence on each treatment symmetrically.

Good near-optimality properties of the empirical success rule in two-arm trials are well established in the theoretical literature. Given any specified sample size, the empirical success rule has been shown to minimize maximum regret in trials with binary outcomes that assign an equal number of patients to each arm (Stoye, 2009). In the context of the Cao et al. trial, we found that the empirical success rule yields a dramatic improvement in maximum regret relative to hypothesis testing.

6. Vaccination with Partial Knowledge of the Indirect Effect on the Unvaccinated

Researchers studying optimal vaccination against infectious disease have typically assumed the planner knows how vaccination affects illness rates. However, there are at least two sources of partial knowledge. First, the planner may only partially know the direct effect of vaccination in generating an immune response that prevents a vaccinated person from become ill or infectious. Second, the planner may only partially know the indirect effect of vaccination in preventing transmission of disease to members of the population who are unvaccinated or unsuccessfully vaccinated.

The second issue commonly is more of a problem. A standard randomized clinical trial, which vaccinates a small experimental group of individuals, enables evaluation of the direct effect of vaccination. However, the trial does not reveal the indirect effect of applying different vaccination rates to the population. If the experimental group is small, the population vaccination rate is essentially zero. If a trial vaccinating a non-negligible fraction of the population is undertaken, the resulting outcome data only reveal the indirect effect of the chosen vaccination rate. The outcomes with other vaccination rates remain counterfactual, yet choice of a vaccination policy requires comparison of alternative rates.

Attempting to cope with the absence of empirical evidence, researchers have used epidemiological models to forecast the outcomes that would occur with counterfactual vaccination policies. However, authors typically provide little information that would enable one to assess the accuracy of their assumptions about individual behavior, social interactions, and disease transmission. Hence, it is prudent to view their forecasts more as computational experiments predicting outcomes under specific assumptions than as accurate predictions of policy impacts.

Manski (2010, 2017) studied choice of a vaccination policy when a planner has partial knowledge of indirect effectiveness and, hence, is unable to determine the optimal policy. Both articles posed a planning problem whose objective is to minimize the utilitarian social cost of illness and vaccination. Vaccination is socially costly, as is illness. Vaccination is beneficial to the extent that it prevents illness. The indirect effect of vaccination is expressed through a function that describes how the illness rate of unvaccinated or unsuccessfully vaccinated persons varies with the vaccination rate.

6.1. Unconstrained Choice of a Vaccination Rate

To show concretely how basic principles may be applied, Manski (2010) considered a relatively simple scenario in which a planner must choose a vaccination rate for a population of observationally identical persons. I supposed the planner knows the direct effect of vaccination. However, he has no knowledge of the indirect effect except that it is monotone in the vaccination rate. That is, the rate of illness among unvaccinated persons weakly decreases as the vaccination rate rises. This is a highly credible assumption.

I supposed that the planner observes the illness rate of a study population whose vaccination rate has previously been chosen. The study population may, for example, be a past cohort of the population now about to be treated. The planner knows (or finds it credible to assume) that the indirect effect function is the same in the study population and the population of interest. He also knows that the illness rate of unvaccinated or unsuccessfully vaccinated persons weakly decreases as the vaccination rate increases. However, the planner does not know the magnitude of this indirect effect at counterfactual vaccination rates. This scenario is realistic enough to demonstrate key ideas about vaccination under ambiguity, but it is idealized enough to yield simple analytical findings.

I first showed how the planner can eliminate dominated vaccination rates, ones which are inferior whatever the actual indirect-effect function may be. Broadly speaking, low (high) vaccination rates are dominated when the cost of vaccination is low (high). I then showed how the planner can use the minimax or minimax-regret criterion to choose an undominated vaccination rate. These criteria yield different policies.

6.2. Choice Between a Decentralized Vaccination and a Mandate

Manski (2017) considered the decision problem of a planner who observes the illness outcomes that occur when persons make decentralized vaccination choices and who contemplates whether to mandate vaccination. Mandatory vaccination has been a subject of considerable controversy, centered on the tension between personal freedom and public health. The welfare-economic practice of specifying a social welfare (or social cost) function and considering a planner who wants to optimize this function provides a useful normative framework for policy formation.

In the context of vaccination policy, the planner's objective presumably is to minimize the social cost of illness and vaccination. Vaccination is socially costly, as is illness. Vaccination is beneficial to public health to the extent that it prevents illness. Mandatory vaccination improves public health relative to decentralized decision making. However, a mandate increases the cost of production and administration of the vaccine, enlarges the side effects of vaccination, and reduces personal freedom. The specified social cost function expresses quantitatively how society evaluates these advantages and disadvantages of a mandate.

Researchers have previously studied the choice between a mandate and decentralized vaccination as a deterministic planning problem, supposing that the planner knows the outcomes that alternative policies would yield. As in Manski (2010), the analysis in Manski (2017) assumed that the planner (a) wants to minimize the social cost of illness and vaccination, (b) observes a study population whose vaccination rate has been chosen previously, and (c) knows that the indirect effect of vaccination increases with the vaccination rate, but does not know the magnitude of the indirect effect at counterfactual rates. Here, as there, I derived explicit conditions under which the planner can determine that a feasible policy alternative is dominated. Here, as there, I posed several criteria for choice among undominated alternatives and compared the policies they generate.

The analysis in Manski (2017) differed from the earlier one in important respects. The planner of Manski (2010) was empowered to choose any vaccination rate for the population, randomly selecting the persons to vaccinate when the vaccination rate is neither zero nor one. In contrast, the planner of Manski (2017) chooses between two policy alternatives: decentralized and mandatory vaccination. Constraining the planner to two alternatives simplifies the present problem relative to the previous one. However, decentralized decision making is more complex to evaluate because persons may nonrandomly choose to be vaccinated.

One may ask why public health agencies commonly choose between decentralization and mandatory vaccination when they might, in principle, choose any vaccination rate for their populations. A possible answer is that having the government establish a fractional vaccination rate and randomly select persons to be vaccinated would violate a version of the ethical principle of equal treatment of equals. As discussed in Section 4.2.2, randomly vaccinating a fraction of the population is consistent with the equal-treatment principle in the ex ante sense that all observationally identical people have the same probability of vaccination. However, it violates equal treatment in the ex post sense that only some persons ultimately are vaccinated.

7. Identification of Income-Leisure Preferences and Evaluation of Income Tax Policy

Economists have long recognized that the relative merits of alternative income tax policies depend on the preferences of individuals for income and leisure. Among the simplifying assumptions that Mirrlees (1971) made in his seminal study of optimal utilitarian income taxation, he stated (p. 176): "The State is supposed to have perfect information about the individuals in the economy, their utilities and, consequently, their actions." Mirrlees also recognized the difficulty of inference on population preferences. In the conclusion to his article he wrote (p. 207): "The examples discussed confirm, as one would expect, that the shape of the optimum earned-income tax schedule is rather sensitive to the distribution of skills within the population, and to the income-leisure preferences postulated. Neither is easy to estimate for real economies."

Income-leisure preferences play both positive and normative roles in analysis of tax policy. The positive role is that preferences yield labor supply and other decisions that determine tax revenue. The normative role is that social welfare aggregates individual preferences in utilitarian policy evaluation. Thus, choice of an optimal tax policy requires knowledge of preferences both to predict tax revenues and to compute the welfare achieved by alternative policies.

In Manski (2014a), I studied identification of income-leisure preferences using standard revealed-preference analysis of labor supply and reached this pessimistic conclusion (p. 146): "As I see it, we lack the knowledge of preferences necessary to credibly evaluate income tax policies." With this background, Manski (2014b) considered choice of an income tax policy as a problem of planning under ambiguity. I summarize here.

7.1. Taxation and Labor Supply

To begin, standard economic theory does not predict the response of labor supply to income taxation. To the contrary, it shows that a worker may rationally respond in disparate ways. As tax rates increase, a person may rationally decide to work less, work more, or not change his labor supply at all. The silence of theory on labor supply has long been appreciated; see Robbins (1930). Modern labor economics envisions labor supply as a complex sequence of schooling, occupation, and work effort decisions made under uncertainty over the life course, perhaps with only bounded rationality. However, we need only consider a simple static scenario to see that a person may respond to income taxes in disparate ways.

Suppose that a person with a predetermined wage and no unearned income allocates each day between paid work and the various non-paid activities that economists have traditionally called leisure. Let a proportional income tax reduce his wage by the prevailing tax rate, yielding his net wage. Assume that the person allocates time to maximize utility, which is an increasing function of net income and leisure.

Different utility functions imply different relationships between the tax rate and labor supply. The labor supply implied by utility functions in the Constant-Elasticity-of-Substitution (CES) family increases or decreases with the tax rate depending on the elasticity of substitution. Other utility functions imply that labor supply is *backward-bending*. That is, hours worked may initially increase as net wage rises from zero but, above some threshold, decrease as net wage rises further. Still other utility functions yield more complex non-monotone relationships between net wage and labor supply. The review article of Stern (1986) describes a broad spectrum of possibilities.

Given that theory does not predict how income taxation affects labor supply, prediction requires empirical analysis. Robbins (1930) emphasized this, writing (p. 129): "we are left with the conclusion . . . . that any attempt to predict the effect of a change in the terms on which income is earned must proceed by inductive investigation of elasticities." Economists have performed a huge number of empirical studies of labor supply. I briefly summarized the literature in Manski (2014a), referencing multiple review articles.

Models of labor supply differ across studies, but they have generally shared two key restrictive assumptions. First, they suppose that labor supply varies monotonically with net wages. Thus, model specifications do not generally permit backward-bending labor supply functions or other non-monotone relationships. Second, they suppose that the response of labor supply to net wage is homogeneous within broad demographic groups. With occasional exceptions, researchers specify hours-of-work equations that permit hours to vary additively across group members but that assume constant treatment response. That is, they assume that all group members would adjust hours worked in the same way in response to a conjectural change in net wage.

The reality may be that persons have heterogeneous income-leisure preferences and, consequently, heterogeneous labor-supply functions. Some may increase work time with net wage, others may decrease work time, and still others may exhibit a non-monotone relationship. If so, estimates of models that assume monotonicity and homogeneity of labor supply can at most characterize the behavior of an artificial "representative" person. The estimates may not have even this limited interpretation.

7.2. Identification Analysis

In light of the above, Manski (2014a) examined the problem of identification of income-leisure preferences. I studied inference when data on time allocation under status-quo tax policies are interpreted through the lens of standard theory. To illuminate elemental issues, I found it productive to study the classical static model in which persons with separable preferences for private and public goods must allocate one unit of time to work and leisure.

I considered the use of revealed preference analysis to predict labor supply and tax revenue under a proposed policy that would alter persons' status-quo tax schedules. The policies that I had in mind use tax revenue to produce public goods and/or to redistribute income from persons who pay positive income tax to ones who pay negative tax. I did not consider policies that extract positive income tax from persons and then compensate them through provision of lump-sum transfers that yield pre-tax utility levels.

My objective was to shed light on how maintained assumptions affect the conclusions that one may draw about counterfactual labor supply and tax revenue. As in my other research on identification, I found it illuminating to begin with weak assumptions and then to characterize the identifying power of stronger assumptions.

I first assumed only that persons prefer to have more income and leisure. Basic revealed-preference analysis of the type pioneered by Samuelson (1938) shows that observation of a person's time allocation under a status-quo tax policy may bound his allocation under a proposed policy or may have no implications, depending on the tax schedules and the person's status-quo time allocation. Basic analysis assuming only that more-is-better generically does not predict the sign of labor-supply response to change in the tax schedule.

I then explored the identifying power of assumptions restricting the distribution of preferences across persons. I supposed that one observes the time allocation of each person in a population whose members may have heterogeneous preferences, wages, and face various status-quo tax schedules. I found it analytically helpful to suppose that persons choose among a finite set of feasible (income, leisure) values rather than the continuum often assumed in the literature. I used the discrete-choice framework of Manski (2007a) to characterize preferences and to predict aggregate labor supply and tax revenue when various assumptions restrict the distribution of preferences.

I studied the identifying power of two classes of assumptions. The first assumes exogenous variation in choice sets, in the formal sense that groups of persons who face different choice sets are assumed to have the same distribution of preferences. For example, one may assume that groups of persons who have different wages or who face different tax schedules have the same preference distribution. The second restricted the shape of this distribution. For example, one may assume that all persons have preferences in the CES family, with possibly heterogeneous parameters. The generic finding was partial identification of the preference distribution. This implies partial ability to predict tax revenue under proposed policies. I used a computational experiment to illustrate.

The analysis reached highly cautionary findings about present knowledge of income-leisure preferences, carrying implications for evaluation of tax policy. A familiar exercise in normative public economics poses a utilitarian social welfare function and ranks tax policies by the welfare they achieve. This requires knowledge of income-leisure preferences to predict tax revenues and to compute welfare. Partial knowledge of preferences generates two difficulties for policy evaluation. One can only partially predict tax revenue and one can only partially evaluate the utilitarian welfare of policies.

7.3. Choosing Size of Government under Ambiguity: Infrastructure Spending and Income Taxation

The optimal size of government has been a subject of continuing debate. Disagreements may stem in part from the fact that “size of government” is an imprecise term─persons using it may not interpret it the same way. Persons who share a common understanding of the term may disagree on what size is optimal. They may have different normative perspectives on social welfare or different beliefs about the outcomes yielded by alternative policy choices.

Attempting to shed light on the optimal size of government, economists have posed and analyzed planning problems. A standard exercise specifies a set of feasible policies and a social welfare function, typically utilitarian. The planner is assumed to know the welfare achieved by each policy. The analysis characterizes the optimal policy.

One prominent body of research, stimulated by Mirrlees (1971), has considered the use of income taxation to redistribute income, given fixed public spending. Economists have derived optimal tax schedules under the assumption that the planner knows the income-leisure preferences of the population. Another, following Barro (1990), has considered the use of public spending to promote growth. Economists have derived optimal tax-financed spending levels under the assumption that the planner knows the consumption preferences of the population and how public spending affects aggregate output. As discussed throughout this paper, lack of knowledge of the welfare achieved by alternative policies limits the relevance of optimization studies to actual policy choice.

Manski (2014b) examined choice of size of government as a problem of planning under ambiguity. I focused on tax-financed public spending for infrastructure that aims to enhance private productivity. My focus on infrastructure spending was similar to that of the growth literature exemplified by Barro (1990). However, my formalization of the planning problem stemmed from the one used by Mirrlees (1971) to study optimal income taxation. Following Mirrlees, I posed a static setting in which each member of a population allocates time to paid work and to the various non-paid activities that economists have traditionally called leisure. Persons have predetermined heterogeneous wages. An income tax schedule is applied to gross income, yielding net income. Persons allocate time to maximize utility, which increases with net income and leisure. Social welfare is utilitarian.

I departed from the Mirrlees setup in three main ways. First, government chooses how much to spend on infrastructure and on activities that directly affect personal utility. Second, persons may have heterogeneous preferences for income, leisure, and public spending. Third, the planner may have partial knowledge of population preferences and of the productivity of infrastructure spending. In contrast, research on optimal income taxation has commonly studied the use of taxation to redistribute income while holding government spending fixed and has assumed that all persons have the same, known, preferences.

Partial knowledge generates two distinct difficulties. First, the planner may be unable to predict tax revenue with certitude and, thus, may not know if a policy will yield a balanced budget. Second, he may be unable to determine the welfare achieved by a policy. I bypassed the first issue and focused on the second.

The first issue is difficult to address in generality. Satisfactory evaluation of policies that may not yield balanced budgets requires specification and analysis of a dynamic planning problem that permits surpluses and deficits to occur and recognizes their intertemporal welfare implications. This poses a larger challenge than I felt able to confront. To bypass the complexity of dynamic policy evaluation under ambiguity, I considered settings in which the planner can ensure budget balance by choosing components of policies sequentially rather than simultaneously. This idea can be implemented in certain settings.

When budget balance can be ensured by sequential choice of policy components, planning may be studied using established criteria for static decision making under ambiguity. I did so in an illustrative setting that is simple enough to yield easily interpretable closed-form findings. In this setting, the planner only considers tax schedules that make the tax proportional to income. Persons have Cobb-Douglas income-leisure preferences and no non-labor income.

These assumptions imply that time-allocation choices are invariant to policy and they enable the planner to achieve budget balance by first choosing the level of public spending, then observing the resulting population income, and finally choosing the tax rate to balance the budget. For further simplicity, I assumed that all public spending is on infrastructure and that wages are person-specific positive constants multiplied by an aggregate production function expressing the wage-enhancing effect of spending on infrastructure. Finally, I assumed that the planner has partial knowledge of the aggregate production function, obtained by observing the outcome of a status quo policy and by assuming that public spending enhances wages but with diminishing marginal returns. Then the space of possible states of nature indexes all concave-monotone aggregate production functions that yield the outcome of the status quo policy.

In this setting, I showed that the planner can reasonably choose a wide range of spending levels. Thus, a society can rationalize having a small or large government. The choice made depends on the decision criterion that the planner uses to cope with ambiguity. I considered planning that maximizes subjective expected welfare or that uses one of several criteria−maximin, minimax-regret, or a Hurwicz criterion−that do not place a subjective probability distribution on unknown quantities.

The study drew conclusions that are methodologically constructive and substantively cautionary. The methodologically constructive conclusion was that, when performing normative research on size of government, economists need not study optimization problems whose solution requires far more knowledge than researchers can credibly assert. Decision theory provides a suitable formal framework for study of planning under ambiguity.

The substantively cautionary conclusion was that study of planning with credible assumptions shows that a wide range of policy choices can be rationalized. The only way to achieve credible conclusions about the desirable size of government is to vastly improve current knowledge of population preferences and the productivity of public spending. There is no immediate way to achieve this, but a research program with a suitably long-run perspective may make progress possible.

8. Climate Policy with Uncertainty in Climate Modeling and Intergenerational Discounting

Integrated assessment (IA) models enable quantitative evaluation of the benefits and costs of alternative climate policies. Policy comparisons are performed by considering a planner who seeks to make optimal trade-offs between the costs of carbon abatement and the damages from climate change. The planning problem has been formalized as a control problem with these components:

(1) equations coupling greenhouse gas (GHG) emissions and abatement to the accumulation of GHGs in the atmosphere and resulting temperature increases.

(2) a damage function that quantifies economic effects of climate change in terms of the loss of global economic output as a function of temperature increases.

(3) an abatement cost function that expresses the cost of actions to reduce GHG emissions relative to a stipulated baseline emissions trajectory.

Social welfare has usually been measured by present discounted gross world product. This expression of welfare takes a stand on the importance of intergenerational equity through the specified discount rate and time horizon. It ignores intragenerational impacts of climate change and policy. The control problem is to minimize the reduction in discounted gross world product due to abatement and damages.

Studying climate policy as a control problem presumes that a planner knows enough to make optimization feasible, but physical and economic uncertainties abound. Lacking a consensus climate model, physical scientists have developed multiple models and performed multi-model ensemble (MME) analysis. To cope with inter-model *structural uncertainty*, they compute simple or weighted averages of the outputs of MMEs. Choosing appropriate weights has been problematic.

Economists have estimated multiple damage functions and abatement cost functions. In general, economists have not performed MME analyses that combine multiple functions by averaging. They have reported disparate findings from separate studies. To the extent that they have recognized uncertainty, they have placed a subjective probability distribution on unknown quantities. For example, Nordhaus (2008) used this approach to express partial knowledge of parameter values in a book on climate policy, writing (page 27): "This book takes the standard economic approach to uncertainty known as the expected utility model, which relies on an assessment with subjective or judgmental probabilities."

Manski, Sanstad, and DeCanio (2021) framed structural uncertainty in climate modeling as a problem of partial identification. We proposed use of the MMR criterion to recognize deep climate uncertainty without weighting climate model forecasts. We developed a theoretical framework for benefit-cost analysis of climate policy based on MMR and we applied it computationally with a simple illustrative IA model.

Perhaps the most contentious economic issue in evaluation of climate policy has been how a planner should assess the costs and benefits of policies across generations. DeCanio, Manski, and Sanstad (2022), henceforth DMS, studied choice of climate policy that minimizes maximum regret with deep uncertainty regarding both the correct climate model and the appropriate intergenerational assessment of policy consequences. I summarize this work here.

8.1. Uncertainty in Climate Models

All climate models are based on a specific set of deterministic nonlinear partial differential equations describing large-scale atmospheric dynamics. Implementation of the equations is subject to numerous practical choices involving discretization, solution methods, and other details. Some components of the system – such as cloud formation and heat transfer between land surfaces and the atmosphere – are not yet fully understood and must be approximated. For these reasons, multiple climate models have been developed and are in use, each reflecting different but credible choices in model design and implementation.

Existing models yield different projections of the global climate. The range of projections produced by different models is a gauge of deep uncertainty about the climate system given the current state-of-the-science. Virtually all methods of MME analysis combine model outputs into single projections of future climate variables. However, climate researchers have recognized persistent methodological problems in combining model projections.

A common technique is to take the simple average across model projections of policy-relevant variables. Researchers may compute weighted average projections when they believe that models can be ranked with respect to relative accuracy. The practice of averaging model forecasts is similar to performance of meta-analysis, discussed in Section 2.3.3. Combining MME outputs into single projected trajectories of the future global climate remains a challenging and unresolved problem.

8.2. Uncertainty and Disagreement Regarding the Discount Rate

The economic losses from climate change are represented by damage functions that give the decreases in world-wide output resulting from increases in mean global temperature. Economists study dynamic optimization by a planner, which entails discounting to quantify the present value of future economic costs and benefits. The appropriate magnitude of the discount rate has been contentious. The choice is consequential. Low rates favor policies that reduce GHG emissions aggressively and rapidly. High rates favor policies that act modestly and slowly. Controversy persists in part because choice of a discount rate is not only an empirical question regarding the future of the economy. It is also a normative question, concerning social preferences for equity across future generations.

A simple version of the famous *Ramsey formula* provides a transparent expression of the interplay of normative and empirical considerations in choosing a discount rate. Let the social welfare function be additively separable in the utility of future generations. Let ρ be the rate at which the planner discounts the utility of future generations. Let the utility of a representative consumer be an increasing and concave function of consumption, with constant elasticity ($-\eta$) of marginal utility. Let g be the annualized growth rate of consumption between time 0 and a future time t. Ramsey (1928) showed that it is optimal to discount future consumption between the present (time 0) and time t at the rate $\delta = \rho + \eta g$.

From the perspective of the present, the empirical value of g may be uncertain. This uncertainty is similar conceptually to the uncertainty that climate modelers face as they attempt to project the future trajectory of climate variables. ρ formalizes how the planner views intergenerational equity, with $\rho = 0$ if the planner gives equal weight to the welfare of all future generations and $\rho > 0$ if the planner weights welfare more heavily in the near future than in the distant future. η formalizes the desirability of intergenerational consumption equity. A planner may feel normative uncertainty about what values of ρ and η to use.

Supposing that the planner aims to represent society, a source of this uncertainty may be normative disagreements within the present population. Such disagreements were evident in a dispute between Nordhaus (2007), who used the value $\rho = 0.03$, and Stern (2006), who used $\rho = 0.001$. Stern concluded that policy should seek to reduce GHG emissions aggressively and rapidly. Nordhaus favored policies that act more modestly and slowly.

DMS argued against any attempt to cope with empirical and normative uncertainty by choosing a single discount rate. Instead, we studied formation of climate policy recognizing a set of possibly appropriate discount rates. To express deep uncertainty, we supposed that the appropriate discount rate lies within an interval that covers the spectrum of rates that have been used in the literature.

8.3. Minimax-Regret Policy Evaluation

The formal analysis by DMS is a straightforward generalization of Manski, Sanstad, and Decanio (2021). There we supposed that the correct climate model is one of six prominent models in the literature on climate science, whereas the correct economic model is known. We supposed that a planner compares six policies, each of which chooses an emissions abatement path that is optimal under one and only one of the six climate models. Regret is the loss in welfare if the model used in policy making is not correct and, consequently, the chosen abatement path is actually sub-optimal. The MMR rule chooses a policy that minimizes the maximum regret, or largest degree of sub-optimality, across all six climate models.

Generalizing this setup, DMS supposed that the correct climate model is one of the six examined earlier and characterized uncertainty about the discount rate by supposing that it takes one of the seven values $\{0.01, 0.02, \ldots, 0.07\}$, a range that covers the rates commonly used. This range reflects both empirical uncertainty about the future of the economy and normative uncertainty, or perhaps disagreement, about how the current population values the welfare of future generations. Given uncertainty about the climate model and the discount rate, we supposed that a planner compares forty-three policies. Forty-two policies entail choosing an emissions abatement path that is optimal under one of the {discount rate, climate model} pairs. The remaining one is a passive policy in which the planner chooses no abatement. The MMR criterion chooses a policy that minimizes maximum regret across all forty-three potential policies.

8.3.1. The Optimal-Control Problem

To begin, we specified the control problem that a planner would solve with no uncertainty. Let B_t represent baseline GHG emissions at time t, A_t be GHG abatement actions at time t under some climate policy, measured in the same units as emissions, $C(A_t)$ be the cost of these actions, and $E_t^{A_t} = B_t - A_t$ be the resulting net emissions. We refer to A_t and $E_t^{A_t}$ as "paths" or "trajectories," and we assume that abatement paths are chosen from some space of feasible paths.

Emissions paths are used as inputs to a climate model M. We focus on the global mean temperatures projected by M as a function of these paths. Let $T(E_t^{A_t}, M)$ be the global mean temperature at time t determined by the GHG trajectory $E_t^{A_t}$ when it is predicted by the climate model M. Then a damage function can be written as $D\left(T\left(E_t^{A_t}, M\right)\right)$.

For abatement path A_t and climate model M, denote the associated total cost (abatement plus damages) at time t as $\mathbb{C}(A_t, M) \equiv C(A_t) + D\left(T\left(E_t^{A_t}, M\right)\right)$ A policymaker seeks to minimize the present value of cost over a planning horizon. As usual in the climate economics literature, we assume an infinite horizon. The control problem given climate model M is to solve

$$\min_{A_t} \int_0^\infty \mathbb{C}(A_t, M) e^{-\delta t} dt$$

where δ is the discount rate. We suppose that the optimal A_t is chosen with commitment at time zero. That is, it is not updated over time as new climate or cost information is obtained. Under certain assumptions, this optimization problem has a unique solution.

8.3.2. The MMR Decision Rule

Let $\boldsymbol{\Delta} = \{\delta_1,\ldots, \delta_K\}$ be a set of discount rates and $\mathbf{M} = \{M_1,\ldots,M_N\}$ be a model ensemble. The planner now faces the problem of minimizing cost over the horizon while recognizing joint {discount rate, model} uncertainty. For rate δ_i and model M_j, let $A^*_{t;\delta_i,M_j}$ be the optimal abatement path defined by

$$A^*_{t;\delta_i,M_j} = \arg \min_{A_t} \int_0^\infty \mathbb{C}(A_t, M_j) e^{-\delta_i t}\, dt$$

Let $\mathbb{C}^*\left(A^*_{t;\delta_i,M_j}, \delta_i, M_j\right)$ be the associated minimum cost:

$$\mathbb{C}^*\left(A^*_{t;\delta_i,M_j}, \delta_i, M_j\right) = \int_0^\infty \mathbb{C}(A^*_t, M_j) e^{-\delta_i t} dt$$

Now consider any feasible abatement trajectory A_t. The *regret* $\mathbb{R}(A_t, \delta_i, M_j)$ associated with A_t, when discount rate δ_i and climate model M_j describe the world, is the difference between the cost of A_t and the cost of the *optimal* policy associated with δ_i and M_j:

$$\mathbb{R}\left(A_t, \delta_i, M_j\right) = \int_0^\infty \mathbb{C}\left(A_t, M_j\right) e^{-\delta_i t} dt - \mathbb{C}^*\left(A^*_{t;\delta_i M_j}, \delta_i, M_j\right)$$

To apply the MMR rule, the planner considers each feasible abatement path A_t and finds the model and discount rate combination that maximizes regret, solving the problem

$$\max_{\delta_i, M_j} \mathbb{R}(A_t, \delta_i, M_i) = \max_{\delta_i, M_j} \left[\int_0^\infty \mathbb{C}\left(A_t, M_j\right) e^{-\delta_i t} dt - \mathbb{C}^*(A^*_{t;\delta_i,M_j}, \delta_i, M_j)\right]$$

The MMR solution is to find A_t to solve the problem

$$\min_{A_t}\left[\max_{\delta_i, M_j} \mathbb{R}(A_t, \delta_i, M_j)\right]$$

8.3.3. Use of **Δ** to Express Empirical and Normative Uncertainty

The term "uncertainty" has usually referred to incomplete knowledge of the empirical environment of a decision maker. This notion of uncertainty applies to incomplete knowledge of the future global temperature, abatement costs, and damages under alternative climate policies. DMS also consider uncertainty about the discount rate. Our use of the set **Δ** to express both empirical and normative uncertainty regarding the discount rate departed from the usual decision-theoretic focus on empirical uncertainty.

Normative uncertainty may have an empirical source, namely incomplete knowledge of the population preferences that a utilitarian planner would seek to maximize. The planner may face the difficult task of representing a population whose members may not be clear about their time preferences or concern with intergenerational inequalities. Using **Δ** to express normative uncertainty is a more radical departure from the decision-theoretic norm if normative disagreements exist within the present population. A segment of the population may strongly value intergenerational equity whereas another segment may be less concerned with the fate of future generations. Then one may think it necessary to abandon the idealization of a planner and replace it with conceptualization of policy making as a non-cooperative political game.

We nonetheless found it attractive to study MMR decision making in this setting. The MMR rule has some appeal as a broadly acceptable mechanism for policy choice. Recall that the regret of a policy in a specified state of nature measures its degree of sub-optimality in that state, and maximum regret measures the maximum degree of sub-optimality across all states. Suppose that the members of a heterogeneous present population disagree on what {discount rate, model} should be considered the true state of nature. Then use of the MMR rule to choose policy minimizes the maximum degree of sub-optimality that will be experienced across the population.

9. The Institutional Separation of Research on Planning and Actual Planning

I will conclude by extending remarks in my expository book Public *Policy in an Uncertain World* (Manski, 2013). I observed there that modern democratic societies have created an institutional separation between policy analysis and decision making, with professional analysts reporting findings to representative governments. Separation of the tasks of analysis and decision making, the former aiming to inform the latter, appears advantageous from the perspective of division of labor. No one can be expert at everything. In principle, having researchers study planning problems and provide their findings to law makers and civil servants enables these planners to focus on the challenging task of policy choice, without having to perform their own research.

I also observed that the current practice of policy analysis with incredible certitude does not serve planners well. The problem is that the consumers of policy analysis cannot trust the producers. I argued that, to improve analysis and to increase trust, research on planning should transparently face up to uncertainty rather than hide it.

Some think this idea to be naïve or impractical. I have repeatedly heard policy analysts assert that policy makers are either psychologically unwilling or cognitively unable to cope with uncertainty. Some economists with experience in the federal government of the United States have suggested to me that concealment of uncertainty is an immutable characteristic of the American policy environment. Hence, they assert that the prevailing practice of policy analysis with incredible certitude will have to continue as is.

A more optimistic possibility is that concealment of uncertainty is a modifiable social norm. My hope is that salutary change will occur if awareness grows that incredible certitude is harmful. Then I anticipate that the scientific community will reward policy research based on credible analysis more than optimization exercises performed with heroic assumptions. Planners and the public will want researchers to provide reasonable policy recommendations that recognize the subtlety of planning under uncertainty, not unequivocal ones that lack foundation.

References

Atkinson, A. and J. Stiglitz (1980), *Lectures on Public Economics*, Princeton: Princeton University Press.

Barlevy, G. (2011), "Robustness and Macroeconomic Policy," *Annual Review of Economics*, 3, 1-24.

Barro, R. (1990), "Government Spending in a Simple Model of Endogenous Growth," *Journal of Political Economy*, 98, S103-S125.

Beigel, J., K. Tomashek, L. Dodd, A. Mehta, B. Zingman, A. Kalil, E. Hohmann, H. Chu, A. Luetkemeyer, S. Kline, D. Lopez de Castilla, R. Finberg, K. Dierberg, V. Tapson, L. Hsieh, T. Patterson, R. Paredes, D. Sweeney, W. Short, G. Touloumi, D. Lye, N. Ohmagari, M. Oh, G. Ruiz-Palacios, T. Benfield, G. Fätkenheuer, M. Kortepeter, R. Atmar, C. Creech, J. Lundgren, A. Babiker, S. Pett, J. Neaton, T. Burgess, T. Bonnett, M. Green, M. Makowski, A. Osinusi, S. Nayak, and H. Lane (2020), "Remdesivir for the Treatment of Covid-19--Preliminary Report," *New England Journal of Medicine*, 383, 1813-1826.

Berger, J. (1985), *Statistical Decision Theory and Bayesian Analysis*, New York: Springer-Verlag.

Binmore, K. (2009), *Rational Decisions*, Princeton: Princeton University Press.

Blumstein, A., J. Cohen J, and D. Nagin (1978), *Deterrence and Incapacitation: Estimating the Effects of Criminal Sanctions on Crime Rates*, Washington, DC: National Academy Press.

Campbell, D., and J. Stanley (1963), *Experimental and Quasi-Experimental Designs for Research*, Chicago: Rand McNally.

Cao, B., Y. Wang, D. Wen, W. Liu, J. Wang, G. Fan, L. Ruan, B. Song, Y. Cai, M. Wei, X. Li, J. Xia, *et al.* (2020), "A Trial of Lopinavir–Ritonavir in Adults Hospitalized with Severe Covid-19," *New England Journal of Medicine*, 382, 1787-1799.

Chernoff, H. (1954), "Rational Selection of Decision Functions," *Econometrica*, 22, 422-443.

DeCanio, S., C. Manski, and A. Sanstad (2022), "Minimax-Regret Climate Policy with Deep Uncertainty in Climate Modeling and Intergenerational Discounting," *Ecological Economics*, 201, https://doi.org/10.1016/j.ecolecon.2022.107552.

Dempster, A. (1968), "A Generalization of Bayesian Inference," *Journal of the Royal Statistical Society, Series B, 30, 205-247.*

DerSimonian, R., and N. Laird (1986), "Meta-Analysis in Clinical Trials." *Controlled Clinical Trials*, 7, 177-188.

Ellsberg, D. (1961), "Risk, Ambiguity, and the Savage Axioms," *Quarterly Journal of Economics*, 75, 643-69.

Elmendorf D (2010) Letter to Honorable Nancy Pelosi, Speaker, US House of Representatives. Congressional Budget Office, housedocs.house.gov/energycommerce/hr4872_CBO.pdf, accessed July 5, 2023.

Fleurbaey, M. (2018), "Welfare Economics, Risk and Uncertainty," *Canadian Journal of Economics*, 51, 5-40.

Gilboa, I. and D. Schmeidler (1989), “Maxmin Expected Utility with Non-Unique Prior,” *Journal of Mathematical Economics*, 18, 141-153.

Glass G. (1977), “Integrating Findings: the Meta-Analysis of Research,” *Review of Research in Education*, 5, 351-379.

Gul, F. and W. Pesendorfer (2008), “The Case for Mindless Economics,” in *The Foundations of Positive and Normative Economics,* A. Caplan and A. Schotter, editors, New York: Oxford University Press.

Haavelmo, T. (1944), “The Probability Approach in Econometrics,” *Econometrica,* 12, Supplement, iii-vi and 1-115.

Hansen, L. and Sargent, T. (2008), *Robustness,* Princeton: Princeton University Press.

Johansen, L. (1978), *Lectures on Macroeconomic Planning, Part 2*, Amsterdam: North-Holland.

Koopmans, T. (1949), “Identification Problems in Economic Model Construction,” *Econometrica*, 17, 125-144.

Manski, C. (1988), “Ordinal Utility Models of Decision Making under Uncertainty,” *Theory and Decision*, 25, 79-104.

Manski, C. (2000), “Identification Problems and Decisions Under Ambiguity: Empirical Analysis of Treatment Response and Normative Analysis of Treatment Choice,” *Journal of Econometrics*, 95, 415-442.

Manski C (2003), *Partial Identification of Probability Distributions,* Springer-Verlag: New York.

Manski, C. (2004), “Statistical Treatment Rules for Heterogeneous Populations,” *Econometrica*, 72, 221-246.

Manski, C. (2006), “Search Profiling with Partial Knowledge of Deterrence,” *The Economic Journal*, 116, F385-F401.

Manski, C. (2007a), “Partial Identification of Counterfactual Choice Probabilities,” *International Economic Review*, 48, 1393-1410.

Manski, C. (2007b), *Identification for Prediction and Decision*, Cambridge, MA: Harvard University Press.

Manski, C. (2009), “Diversified Treatment under Ambiguity,” *International Economic Review*, 50, 1013-1041.

Manski, C. (2010), “Vaccination with Partial Knowledge of External Effectiveness,” *Proceedings of the National Academy of Sciences*, 107, 3953-3960.

Manski C. (2011a), “Policy Analysis with Incredible Certitude,” *The Economic Journal*, 121, F261-F289.

Manski, C. (2011b), “Actualist Rationality,” *Theory and Decision*, 71, 195-210.

Manski C. (2013), *Public Policy in an Uncertain World,* Cambridge, MA: Harvard University Press.

Manski, C. (2014a), “Identification of Income-Leisure Preferences and Evaluation of Income Tax Policy,”, *Quantitative Economics*, 5, 145-174.

Manski, C. (2014b), “Choosing Size of Government under Ambiguity: Infrastructure Spending and Income Taxation,” *The Economic Journal*, 124, 359-376.

Manski C. (2015), “Communicating Uncertainty in Official Economic Statistics: An Appraisal Fifty Years after Morgenstern,” *Journal of Economic Literature,* 53, 631-653.

Manski, C. (2017), “Mandating Vaccination with Unknown Indirect Effects,” *Journal of Public Economics Theory*, 19, 603-619.

Manski, C. (2019a), *Patient Care under Uncertainty*, Princeton: Princeton University Press.

Manski, C. (2019b), “Communicating Uncertainty in Policy Analysis,” *Proceedings of the National Academy of Sciences*, 116, 7634-7641.

Manski, C. (2020a), “The Lure of Incredible Certitude," *Economics and Philosophy*, 36, 216-245.

Manski, C. (2020b), “Towards Credible Patient-Centered Meta-Analysis,” *Epidemiology*, 31, 345-352.

Manski, C. (2021), “Econometrics for Decision Making: Building Foundations Sketched by Haavelmo and Wald,” *Econometrica*, 89, 2827-2853.

Manski, C. and J. Pepper (2013), “Deterrence and the Death Penalty: Partial Identification Analysis Using Repeated Cross Sections,” *Journal of Quantitative Criminology,* 29, 123-141.

Manski, C., A. Sanstad, and S. DeCanio (2021), “ Addressing partial identification in climate modeling and policy analysis,” *Proceedings of the National Academy of Sciences*, 118, doi:10.1073/pnas.2022886118.

Manski, C. and A. Tetenov (2016), “Sufficient Trial Size to Inform Clinical Practice,” *Proceedings of the National Academy of Sciences*, 113, 10518-10523.

Manski, C. and A. Tetenov (2019), “Trial Size for Near-Optimal Treatment: Reconsidering MSLT-II,” *The American Statistician*, 73, 305-311.

Manski, C. and A. Tetenov (2021), “Statistical Decision Properties of Imprecise Trials Assessing Coronavirus 2019 (COVID-19) Drugs,” *Value in Health*, 24, 641-647.

Mirrlees J. (1971), “An Exploration in the Theory of Optimal Income Taxation,” *Review of Economic Studies*, 38, 175-208.

Molinari, F. (2020), “Microeconometrics with Partial Identification,” *Handbook of Econometrics*, Vol. 7A, S. Durlauf, L. Hansen, J. Heckman, and R. Matzkin editors, Amsterdam: Elsevier, 355-486.

Morgenstern, O. (1963), *On the Accuracy of Economic Observations: Second Edition.* Princeton: Princeton University Press.

National Research Council (2012), *Deterrence and the Death Penalty,* Washington, DC: National Academies Press.

Nordhaus, W. (2007), "A Review of the *Stern Review on the Economics of Climate Change*," *Journal of Economic Literature*, 45, 686–702.

Nordhaus, W. (2008), *A Question of Balance: Weighing the Options on Global Warming Policy*, New Haven: Yale University Press.

Ramsey, F. (1928), "A Mathematical Theory of Saving," *Economic Journal*, 38, 543–559.

Robbins, L. (1930), "On the Elasticity of Demand for Income in Terms of Effort," *Economica*, 29, 123-129.

Samuelson, P. (1938), "A Note on the Pure Theory of Consumer Behavior," *Economica*, 5, 61-71.

Savage L. (1954), *The Foundations of Statistics*, New York: Wiley.

Sen, A. (1973), "Behaviour and the Concept of Preference," *Economica*, 40, 241–259.

Sen, A. (1993), "Internal Consistency of Choice," *Econometrica*, 61, 495-521.

Stern, N. (1986), "On the Specification of Labour Supply Functions," in R. Blundell and I. Walker (editors) *Unemployment, Search and Labour Supply*, Cambridge: Cambridge University Press, 143-189.

Stern, N. (2007), *The Economics of Climate Change: The Stern Review*, Cambridge: Cambridge University Press.

Stoye, J. (2007), "Minimax Regret Treatment Choice with Incomplete Data and Many Treatments," *Econometric Theory*, 23, 190-199.

Tamer, E. (2010), "Partial Identification in Econometrics," *Annual Review of Economics*, 2, 167-195.

Von Neumann, J. and O. Morgenstern (1944), *Theory of Games and Economic Behavior*, Princeton: Princeton University Press.

Wald, A. (1950), *Statistical decision functions*, New York: Wiley.

Walley, P. (1991), *Statistical Reasoning with Imprecise Probabilities*, New York: Chapman and Hall.