cemmap

centre for microdata methods and practice

Adaptive partial policy innovation: coping with ambiguity through diversification

Charles F. Manski

The Institute for Fiscal Studies
Department of Economics, UCL

cemmap working paper CWP10/08

E·S·R·C ECONOMIC & SOCIAL RESEARCH COUNCIL

An ESRC Research Centre

ADAPTIVE PARTIAL POLICY INNOVATION:
Coping with Ambiguity through Diversification

Charles F. Manski
Department of Economics and Institute for Policy Research, Northwestern University

March 2008

Abstract

This paper develops a broad theme about policy choice under ambiguity through study of a particular decision criterion. The broad theme is that, where feasible, choice between a status quo policy and an innovation is better framed as selection of a treatment allocation than as a binary decision. Study of the static minimax-regret criterion and its adaptive extension substantiate the theme. When the optimal policy is ambiguous, the static minimax-regret allocation always is fractional absent large fixed costs or deontological considerations. In dynamic choice problems, the adaptive minimax-regret criterion treats each cohort as well as possible, given the knowledge available at the time, and maximizes intertemporal learning about treatment response.

This research was supported in part by National Science Foundation grant SES-0549544. This paper substantially revises and supercedes an earlier working paper "Fractional Treatment Rules for Social Diversification of Indivisible Private Risks," NBER Working Paper W11675, September 2005. I have benefitted from the opportunity to present this work in a seminar at Universitat Pompeu Fabra.

1. Introduction

Problems of choice between a status quo policy and an innovation occur often. In medicine, the status quo may be the prevalent treatment for a disease and the innovation a new treatment proposed by researchers. In criminal justice, the status quo may be existing guidelines for sentencing convicted offenders and the innovation a new sentencing proposal. In education, the status quo may be the present system for evaluating teachers and the innovation an alternative. In tax policy, the status quo may be the present personal income tax schedule and the innovation a different schedule.

In these and many other settings, it is common to have only partial knowledge of policy impacts, particularly concerning the innovation. The better policy choice may then be ambiguous. Formally, ambiguity occurs when there are multiple feasible states of nature and both treatments are undominated. That is, one policy is superior in some states of nature and the other is superior in other states.

There are myriad sources of ambiguity, many deriving from identification problems that are prevalent in empirical research; see Manski (2007) for exposition. Perhaps the most fundamental identification problem arises from the unobervability of counterfactual policy 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.

Suppose that a planner must act with partial knowledge of the welfare achieved by the status quo and the innovation. How should he cope with ambiguity? The Bayesian prescription is to assert a subjective probability distribution over the feasible states of nature and choose an action that maximizes subjective expected welfare. However, a subjective probability distribution is itself a form of knowledge, and the planner may have no credible basis for asserting one. My research program on planning under ambiguity has studied problems of this type; see Manski (2005; 2006; 2007, Chapter 11) and the references contained within. In particular, I have explored application of the minimax-regret (MR) criterion to problems of treatment choice. This paper builds on my earlier work.

To begin, observe that choice between a status quo and an innovation is commonly framed as a binary decision. Either the status quo will continue in force or the innovation will replace it, becoming the new status quo. When the decision is made with partial knowledge, two types of errors may occur. A Type I error occur when an innovation that actually is worse than the status quo is judged superior with the available information. A Type II error occurs when an innovation that is actually better than the status quo is judged inferior with the available information.

This paper argues that, where feasible, choice between a status quo policy and an innovation should be framed as selection of a treatment allocation rather than as a binary decision. Selection of a treatment allocation is feasible when a planner chooses treatments for each member of a population and can treat different members differentially. In these settings, the planner need not make a *singleton* allocation, assigning all persons to the same treatment. He can instead choose a *fractional* allocation, assigning positive fractions of the population to both the status quo treatment and the innovation.

Fractional allocations cope with ambiguity through diversification. Whereas singleton allocations offer a stark choice between possible commission of a Type I or Type II error, fractional allocations make both types of errors but reduce their magnitudes. Depending on the criterion used to make decisions under ambiguity, a planner may find an interior solution preferable to a corner solution. In particular, this occurs when a planner uses the minimax-regret criterion to choose a treatment allocation. The MR criterion places equal weight on Type I and Type II errors and chooses an allocation that balances their potential welfare effects.

When considering fractional treatment allocations, it is important to distinguish differential treatment of persons who vary in observable respects from differential treatment of persons who are observationally identical. It is well known that enabling treatment choice to vary systematically with observed covariates of population members can improve utilitarian welfare if treatment response varies with these covariates; see, for example, Manski (2005, Sec. 1.2). My concern here is with differential treatment of persons who are observationally identical. Differential treatment of this type necessarily is random, not systematic.

For example, a physician could in principle assign some observationally identical patients to the status quo treatment and others to the innovation. A judge could apply the existing sentencing guidelines to some convicted offenders and a new sentencing proposal to others. A school district could use the status quo system to evaluate some teachers and apply an alternative system to others. The federal government could apply the present personal income tax schedule to some persons and a different schedule to others.

Why might fractional treatment allocations be beneficial? I develop three reasons in Sections 2 through 4 respectively. First, fractional allocations enlarge the set of feasible policy choices by convexifying the singleton allocations. Second, such allocations are advantageous for learning because they generate randomized experiments that yield informative outcome data on both treatments. Third, fractional allocations enable better results when policy is determined by non-cooperative decision processes. I elaborate below and then call attention to a possible ethical objection to fractional allocations.

Convexifying the Set of Policy Choices

Binary choice between the status quo and the innovation is an all-or-nothing decision. Fractional allocations convexify the singleton allocations, thereby greatly enlarging the set of feasible policy choices. The relative desirability of fractional and singleton allocations solutions depends on the criterion used for decision making under ambiguity. In previous work studying static planning problems, I have shown that the minimax-regret criterion always yields a fractional allocation when there are two undominated treatments, outcomes are bounded, and welfare increases linearly with the population mean outcome (Manski, 2007, Complement 11A). Moreover, the MR allocation has a simple explicit form.

Section 2 reviews this finding and uses the sentencing of convicted offenders to illustrate. I then extend the analysis to a broader class of welfare functions than I have considered previously. In particular, I permit monotonic transformations of the welfare function, address planning problems with non-additive cost of treatment, and consider deontological welfare functions.

Learning Treatment Outcomes

The convexification argument for differential treatment of observationally identical persons applies equally to static and dynamic planning problems. The learning argument pertains specifically to dynamic problems. Suppose that, in each period, a planner chooses treatments for the current cohort of a population. Then learning may be possible, with observation of treatment outcomes in earlier periods informing treatment choice in later periods. Fractional treatment allocations randomize persons into treatment and, hence, are particularly informative.

Section 3 considers dynamic planning under ambiguity from the minimax-regret perspective. I suggest use of the nicely tractable *adaptive minimax-regret* (*AMR*) criterion, which treats each cohort as well as possible in the static minimax-regret sense, using the information available at the time. The result is a fractional treatment allocation whenever the available knowledge does not suffice to determine which treatment is better. The criterion is adaptive because knowledge of treatment response accumulates over time, so successive cohorts may receive different fractional allocations. I use medical treatment to illustrate application of the AMR criterion. I explain how the AMR criterion differs from the current practice of randomized clinical trials in medicine.

Improving Non-Cooperative Decisions

Sections 2 and 3 are written from the perspective of a planner with the power to dictate policy. Section 4 considers situations in which polices are determined by non-cooperative decision processes. Just as convexification of the set of policy choices can be beneficial to a planner, it can also improve non-cooperative decisions. I examine a two-agent setting where the agents may have different welfare functions and beliefs about the feasible states of nature. If both agents face ambiguity and use the AMR criterion to compare allocations, then there exist fractional allocations that both prefer to the singleton allocations. I use an educational policy choice to illustrate.

Equal Treatment of Equals

A possible ethical objection to fractional treatment allocations is that they violate one interpretation of the normative principle calling for "equal treatment of equals." Fractional allocations are consistent with this principle in the *ex ante* sense that all observationally identical people have the same probability of receiving a particular treatment. They violate the principle in the *ex post* sense that observationally identical persons ultimately receive different treatments.

The ex post sense of equal treatment expresses a deontological consideration that is absent from the consequentialist welfare functions usually asssumed in economic analysis of planning. I formalize this consideration in Section 2.4 and show how it affects the treatment allocation of a planner who uses the minimax-regret criterion.

2. Static Planning Problems

Section 2.1 reviews relevant elements of my previous analysis of planning under ambiguity. Sections 2.2 through 2.4 extend the analysis in several new directions.

2.1. Treatment Allocation with Linear Welfare

Basic Concepts and Notation

There are two treatments, labeled a and b; the set of feasible treatments is $T \equiv \{a, b\}$. Treatment a is the status quo and b is the innovation. The semantic distinction between the status quo and the innovation plays no role in the general analysis described here, which broadly concerns choice between two treatments. The distinction becomes meaningful when the analysis is applied. In particular, more may be known about the status quo treatment than about the innovation.

Each member j of a population denoted J has a response function $y_j(\cdot)$: $T \rightarrow Y$ that maps treatments $t \in T$ into outcomes $y_j(t) \in Y$. The subscript j in $y_j(\cdot)$ indicates that treatment response may vary across the population. Let $u_j(t) \equiv u_j[y(t), t]$ denote the net contribution to social welfare that occurs if person j receives treatment t and realizes outcome $y_j(t)$. For example, $u_j(t)$ may have the "benefit-cost" form $u_j(t) = y_{j1}(t) - y_{j2}(t)$, where $y_{j1}(t)$ is the benefit of treatment t and $y_{j2}(t)$ is its cost.

I assume for simplicity that all members of the population are observationally identical. In practice, persons may have observable covariates, and a planner may be able to differentially treat persons with different covariates. In such cases, the present analysis can be applied separately to each sub-population of persons who share the same covariates.

Let $P[y(\cdot)]$ denote the population distribution of treatment response. I suppose that the population is large in the formal sense of being atomless; that is, $P(j) = 0$ for all $j \in J$. This idealization implies that if the planner randomly assigns a positive fraction of the population to a treatment, the sub-population of persons who receive this treatment is infinite. This eliminates sampling variation as an issue when comparing alternative treatment allocations and analyzing treatment response.

The planner's task is to allocate the population between the two treatments. A treatment allocation is a number $\delta \in [0, 1]$ that randomly assigns a fraction δ of the population to treatment b and the remaining $1 - \delta$ to treatment a. I assume that the planner wants to choose a treatment allocation that maximizes mean welfare in the population. Let $\alpha \equiv E[u(a)]$ and $\beta \equiv E[u(b)]$ be the mean welfare that would result if a randomly drawn person were to receive treatment a or b respectively. Social welfare with allocation δ is

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

$W(\cdot)$ is a consequentialist social welfare function that additively aggregates individual contributions to welfare. If the function $u(\cdot)$ expresses private preferences, then $W(\cdot)$ is the utilitarian social welfare function usually assumed in research on welfare economics.

Treatment Choice Under Ambiguity

The optimal treatment allocation is obvious if (α, β) are known. The planner should choose $\delta = 1$ if $\beta > \alpha$ and $\delta = 0$ if $\beta < \alpha$. All allocations yield the same welfare if $\beta = \alpha$. The problem of interest is treatment choice when (α, β) is partially known.

To formalize the problem, let S index the feasible states of nature. Thus, the planner knows that (α, β) lies in the set $[(\alpha_s, \beta_s), s \in S]$. I assume that this set is bounded and denote the extreme feasible values of α and β as $\alpha_L \equiv \min_{s \in S} \alpha_s$, $\beta_L \equiv \min_{s \in S} \beta_s$, $\alpha_U \equiv \max_{s \in S} \alpha_s$, and $\beta_U \equiv \max_{s \in S} \beta_s$. Partial knowledge is unproblematic for decision making if $(\alpha_s \geq \beta_s, s \in S)$ or if $(\alpha_s \leq \beta_s, s \in S)$; choosing $\delta = 0$ is optimal in the former case and $\delta = 1$ in the latter. The planner faces ambiguity if both treatments are undominated; that is, if $\alpha_s > \beta_s$ for some values of s and $\alpha_s < \beta_s$ for other values. I assume that the planner faces ambiguity.

There is no optimal treatment allocation under ambiguity. Yet the planner must somehow choose an allocation. To accomplish this, decision theorists have proposed various ways of transforming the original optimization problem, which cannot be solved, into another one that can be solved.

Bayesians recommend that the planner assert a subjective distribution on the states of nature and choose an allocation that maximize subjective mean welfare with respect to this distribution. The maximin and the minimax-regret criteria do not use a subjective distribution. Instead they choose allocations that, in different senses, perform uniformly well over all states of nature.

I briefly discuss the Bayes and maximin criteria and then consider the minimax-regret criterion more fully. It turns out that the minimax-regret criteria yields a qualitatively different allocation than do the Bayes and maximin criteria. Bayesian and maximin treatment allocations are generally singleton, assigning all persons to the same treatment. In contrast, the minimax-regret allocation under ambiguity is always fractional, assigning positive fractions of the population to both treatments.

Bayes Rules

A Bayesian planner places a subjective probability distribution π on the states of nature, computes the subjective mean value of social welfare under each treatment allocation, and chooses an allocation that maximizes this subjective mean. Thus, the planner solves the optimization problem

$$(2) \quad \max_{\delta \in [0, 1]} \quad E_\pi(\alpha) + [E_\pi(\beta) - E_\pi(\alpha)]\delta,$$

where $E_\pi(\alpha) = \int \alpha_s d\pi$ and $E_\pi(\beta) = \int \beta_s d\pi$ are the subjective means of α and β. The Bayes decision assigns everyone to treatment b if $E_\pi(\beta) > E_\pi(\alpha)$ and everyone to treatment a if $E_\pi(\alpha) > E_\pi(\beta)$. All treatment allocations are Bayes decisions if $E_\pi(\beta) = E_\pi(\alpha)$. Thus, a Bayesian planner behaves as would a planner who knows that the population means in (1) have the values in (2).

Although Bayesian planning is conceptually straightforward, it may not be straightforward to form a credible subjective distribution on the states of nature. The allocation chosen by a Bayesian planner depends on the subjective distribution used. Here, as always, the Bayesian paradigm is appealing only when a decision maker is able to form a subjective distribution that really expresses his beliefs.

The Maximin Criterion

To determine the maximin allocation, one first computes the minimum welfare attained by each allocation across all states of nature. One then chooses an allocation that maximizes this minimum welfare.

Thus, the criterion is

$$(3) \quad \max_{\delta \in [0, 1]} \; \min_{s \in S} \; \alpha_s + (\beta_s - \alpha_s)\delta.$$

The solution has a simple form if (α_L, β_L) is a feasible value of (α, β). Then the maximin allocation is $\delta = 0$ if $\alpha_L > \beta_L$, $\delta = 1$ if $\alpha_L < \beta_L$, and all $\delta \in [0, 1]$ if $\alpha_L = \beta_L$.

The Minimax-Regret Criterion

By definition, the regret of treatment allocation δ in state of nature s is the difference between the maximum achievable welfare and the welfare achieved with this allocation. The maximum welfare achievable in state of nature s is $\max(\alpha_s, \beta_s)$. Hence, allocation δ has regret $\max(\alpha_s, \beta_s) - [\alpha_s + (\beta_s - \alpha_s)\delta]$. The minimax-regret rule computes the maximum regret of each allocation over all states of nature and chooses an allocation to minimize maximum regret. Thus, the criterion is

$$(4) \quad \min_{\delta \in [0, 1]} \; \max_{s \in S} \; \max(\alpha_s, \beta_s) - [\alpha_s + (\beta_s - \alpha_s)\delta].$$

Let S(a) and S(b) be the subsets of S on which treatments a and b are superior. That is, let $S(a) \equiv \{s \in S: \alpha_s > \beta_s\}$ and $S(b) \equiv \{s \in S: \beta_s > \alpha_s\}$. Let $M(a) \equiv \max_{s \in S(a)} (\alpha_s - \beta_s)$ and $M(b) \equiv \max_{s \in S(b)} (\beta_s - \alpha_s)$. Define $M(a) = 0$ if S(a) is empty and $M(b) = 0$ if S(b) is empty. Manski (2007, Complement 11A) proves that the MR criterion always makes a fractional treatment allocation when both treatments are undominated. The result is

$$(5) \quad \delta_{MR} = \frac{M(b)}{M(a) + M(b)}.$$

The proof is short and straightforward, so I reproduce it here.

Proof: The maximum regret of rule δ on all of S is max [R(δ, a), R(δ, b)], where

$$(6a) \quad R(\delta, a) \equiv \max_{s \in S(a)} \alpha_s - [(1 - \delta)\alpha_s + \delta\beta_s] = \max_{s \in S(a)} \delta(\alpha_s - \beta_s) = \delta M(a),$$

$$(6b) \quad R(\delta, b) \equiv \max_{s \in S(b)} \beta_s - [(1 - \delta)\alpha_s + \delta\beta_s] = \max_{s \in S(b)} (1 - \delta)(\beta_s - \alpha_s) = (1 - \delta)M(b),$$

are maximum regret on S(a) and S(b). Both treatments are undominated, so R(1, a) = M(a) > 0 and R(0, b) = M(b) > 0. As δ increases from 0 to 1, R(·, a) increases linearly from 0 to M(a) and R(·, b) decreases linearly from M(b) to 0. Hence, the MR rule is the unique δ ∈ (0, 1) such that R(δ, a) = R(δ, b). This yields (5). □

Expressions M(a) and M(b) simplify when (α_L, β_U) and (α_U, β_L) are feasible values of (α, β). Then $M(a) = \alpha_U - \beta_L$ and $M(b) = \beta_U - \alpha_L$. Hence,

$$(7) \quad \delta_{MR} = \frac{\beta_U - \alpha_L}{(\alpha_U - \beta_L) + (\beta_U - \alpha_L)}.$$

Result (7) simplifies further if α or β is fully known. Full knowledge for the innovation is rarely realistic, but one may have full knowledge for the status quo from observation of past experience. Thus, suppose that $\alpha_L = \alpha_U = \alpha$. Then (7) becomes

$$(8) \quad \delta_{MR} = \frac{\beta_U - \alpha}{\beta_U - \beta_L}.$$

Choosing Sentences for Convicted Juvenile Offenders

To illustrate, consider the problem of choosing sentences for a population of convicted offenders. I apply findings reported in Manski and Nagin (1998), who studied the sentencing and recidivism of male youth in the state of Utah who were convicted of offenses before they reached age 16.

In this illustration, the planner is the state of Utah and the population are males under age 16 who are convicted of an offence. The status quo policy is a decentralized system where judges have discretion to choose between residential confinement and a sentence that does not involve confinement. I take the innovation to be a policy of mandatory confinement for all convicted offenders. I take the outcome of interest to be a binary measure of recidivism. Specifically, $y(t) = 1$ if an offender who receives treatment t is not convicted of a subsequent crime in the two-year period following sentencing, and $y(t) = 0$ if the offender is convicted of a subsequent crime. Let $u(t) = y(t)$. Then $\alpha = P[y(a) = 1]$ and $\beta = P[y(b) = 1]$.

Analyzing data on outcomes under the status quo policy, Manski and Nagin (1998) find that $\alpha = 0.61$. The data do not fully identify β. In the absence of knowledge of how judges choose sentences or how juveniles respond to their sentences, the data reveal only that $\beta \in [0.03, 0.92]$. Thus, the innovation may be much better or worse than the status quo. Manski and Nagin (1998) argue that this "worst-case" bound on β is germane to policy making because criminologists have found it difficult to learn how sentencing affects recidivism. Researchers have long debated the counterfactual outcomes that offenders would experience if they were to receive other sentences.

Consider policy choose when the state of Utah knows that $\alpha = 0.61$ and $\beta \in [0.03, 0.92]$. If the state applies the Bayesian paradigm, it fully adopts the innovation of mandatory confinement if $E_\pi(\beta) > 0.61$ and leaves the status quo of judicial discretion in place if $E_\pi(\beta) < 0.61$. If the state applies the maximin criterion, it leaves the status quo in place because $\beta_L = 0.03 < 0.61$. If the state applies the minimax-regret criterion, it randomly sentences to confinement $(\beta_U - \alpha)/(\beta_U - \beta_L) = (0.92 - 0.61)/(0.92 - 0.03) = 0.35$ of the offenders and leaves judicial discretion in place for the remaining fraction 0.65.

2.2. Monotone Transformations of the Welfare Function

The planning problem described in Section 2.1 has many extensions that warrant analysis. Section 3 will study its extension from static to dynamic settings. Before that, Sections 2.2 through 2.4 perform three extensions within the static context.

A monotone generalization of welfare function (1) is

$$(9) \qquad W(\delta) = f[\alpha + (\beta - \alpha)\delta],$$

where $f(\cdot)$ is strictly increasing in its argument. The specific shape of $f(\cdot)$ is immaterial to treatment choice when one treatment is superior in all states of nature. Whatever monotone function $f(\cdot)$ may be, $\delta = 0$ is optimal if $(\alpha_s \geq \beta_s, s \in S)$ and $\delta = 1$ if $(\alpha_s \leq \beta_s, s \in S)$. However, shape matters when a planner faces ambiguity.

It is tempting to say that the shape of $f(\cdot)$ expresses social risk preferences, with linear $f(\cdot)$ conveying risk neutrality and concave $f(\cdot)$ implying risk aversion. This language has a clear interpretation in Bayesian planning, where linear $f(\cdot)$ implies indifference between mean-preserving spreads of a gamble and concave $f(\cdot)$ implies a preference for gambles with smaller spreads. However, the Bayesian definition of risk preferences does not carry over to maximin and minimax-regret planning, which do not use a subjective probability distribution. Hence, I do not associate the shape of $f(\cdot)$ with risk preferences here.

Bayes Rules

A Bayesian planner with welfare function (9) solves the optimization problem

$$(10) \qquad \max_{\delta \in [0, 1]} \int f[\alpha_s + (\beta_s - \alpha_s)\delta]d\pi.$$

The solution is generically singleton if $f(\cdot)$ is convex, but it may be fractional if $f(\cdot)$ has concave segments. Manski and Tetenov (2007, Proposition 5) consider the special case where the planner knows α and is uncertain only about β. They show that the Bayes allocation is $\delta = 0$ if $f(\cdot)$ is concave and $E_\pi(\beta) < \alpha$. However, it is fractional if $f(\cdot)$ is continuously differentiable, $E_\pi(\beta) > \alpha$, and $\int f(\beta)d\pi < f(\alpha)$.

The Maximin Criterion

The maximin problem

$$(11) \quad \max_{\delta \in [0, 1]} \ \min_{s \in S} \ f[\alpha_s + (\beta_s - \alpha_s)\delta]$$

has the same solution for all strictly increasing $f(\cdot)$. Thus, the shape of $f(\cdot)$ does not affect the maximin allocation.

The Minimax-Regret Criterion

The shape of $f(\cdot)$ does affect the solution to the minimax-regret problem

$$(12) \quad \min_{\delta \in [0, 1]} \ \max_{s \in S} \ \max\, [f(\alpha_s), f(\beta_s)] - f[(1 - \delta)\alpha_s + \delta\beta_s].$$

Nevertheless, the central qualitative finding of Section 2.2 continues to hold with almost complete generality. I show here that the MR allocation is fractional whenever $f(\cdot)$ is continuous.

Proof: Recall that $S(a) \equiv \{s \in S: \alpha_s > \beta_s\}$ and $S(b) \equiv \{s \in S: \beta_s > \alpha_s\}$. Let

$$(13a) \quad R(\delta, a) \equiv \max_{s \in S(a)} \ f(\alpha_s) - f[(1 - \delta)\alpha_s + \delta\beta_s],$$

$$\text{(13b)} \qquad R(\delta, b) \equiv \max_{s \in S(b)} f(\beta_s) - f[(1-\delta)\alpha_s + \delta\beta_s],$$

be the maximum regret of allocation δ on S(a) and S(b) respectively. The maximum regret of δ on all of S is max [R(δ, a), R(δ, b)]. As δ increases from 0 to 1, R(·, a) strictly increases from 0 to R(1, a) > 0 and R(·, b) strictly decreases from R(0, b) > 0 to 0. Continuity of f(·) and boundedness of $[(\alpha_s, \beta_s), s \in S]$ imply that R(·, a) and R(·, b) are continuous functions of δ. Hence, there exists a unique δ ∈ (0, 1) such that R(δ, a) = R(δ, b). This is the MR allocation. □

Logarithmic Welfare

Section 2.2 showed that the minimax-regret allocation has the simple form (7) when f(·) is linear and $\{\alpha_L, \beta_U), (\alpha_U, \beta_L)\}$ are feasible values of (α, β). The MR allocation typically must be determined numerically when f(·) is nonlinear. However, a simple form emerges when f(·) is the log function and $\{(\alpha_L, \beta_U), (\alpha_U, \beta_L)\}$ are feasible values of (α, β). Then

$$\text{(14a)} \qquad R(\delta, a) = \max_{s \in S(a)} \log\{[\alpha_s/[(1-\delta)\alpha_s + \delta\beta_s]\} = \log\{[\alpha_U/[(1-\delta)\alpha_U + \delta\beta_L]\},$$

$$\text{(14b)} \qquad R(\delta, b) \equiv \max_{s \in S(b)} \log\{\beta_s/[(1-\delta)\alpha_s + \delta\beta_s]\} = \log\{\beta_U/[(1-\delta)\alpha_L + \delta\beta_U]\}.$$

Hence, the MR allocation solves the equation

$$\text{(15)} \quad \alpha_U/[(1-\delta)\alpha_U + \delta\beta_L] = \beta_U/[(1-\delta)\alpha_L + \delta\beta_U].$$

The solution is

$$(16) \quad \delta_{MR} = \frac{\alpha_U(\beta_U - \alpha_L)}{\alpha_U(\beta_U - \alpha_L) + \beta_U(\alpha_U - \beta_L)}.$$

Comparison of (7) and (16) shows that the MR allocations under linear and logarithmic welfare coincide when $\beta_U = \alpha_U$, but they otherwise generally differ from one another. In particular, the two allocations differ when the planner knows α and has partial knowledge of β. Then (16) reduces to

$$(17) \quad \delta_{MR} = \frac{\alpha(\beta_U - \alpha)}{\alpha(\beta_U - \alpha) + \beta_U(\alpha - \beta_L)}.$$

By assumption $\beta_U > \alpha$. Hence, the fraction of the population allocated to the innovation when welfare is logarithmic is smaller than when welfare is linear. For example, in the sentencing illustration of Section 2.1, the MR allocation with logarithmic welfare is 0.26 rather than the 0.35 found with linear welfare.

2.3. Non-Additive Cost of Treatment

Treatment may be costly. The foregoing analysis covers settings where the aggregate cost of a treatment allocation is the sum of individual treatment costs. This was alluded to in Section 2.1, where I observed that $u_j(t)$ may have the benefit-cost form $u_j(t) = y_{j1}(t) - y_{j2}(t)$, where $y_{j1}(t)$ is the benefit when person j receives treatment t and $y_{j2}(t)$ is the cost. There are many ways in which cost might be non-additive. This section considers the polar cases of capacity constraints and fixed costs.

Capacity Constraints

I have thus far assumed that all treatment allocations $\delta \in [0, 1]$ are feasible. Capacity constraints may place an upper bound on the fraction of the population who receive each treatment. A capacity constraint is a cost that equals zero when the fraction of persons who receive a treatment is below the upper bound and infinity thereafter.

Let the maximum fractions of the population who may receive treatments a and b be $\eta(a)$ and $\eta(b)$ respectively. Then the feasible allocations are $\delta \in [1 - \eta(a), \eta(b)]$. A constrained Bayesian, maximin, or minimax-regret allocation solves the relevant extremum problem over the feasible δ. I focus on the MR allocation when welfare has form (9) and $f(\cdot)$ is continuous.

Let δ_{MR} denote the unconstrained MR allocation and δ_{CMR} the constrained MR allocation. As shown in Section 2.3, maximum regret at any allocation δ equals $R(\delta, b)$ for $\delta \leq \delta_{MR}$ and $R(\delta, a)$ for $\delta \geq \delta_{MR}$, where $R(\cdot, b)$ is strictly decreasing in δ and $R(\cdot, a)$ is strictly increasing. It follows that the constrained MR allocation is the feasible allocation closest to the unconstrained MR allocation. That is,

$$(18) \quad \delta_{CMR} = \begin{cases} 1 - \eta(a) & \text{if } \delta_{MR} < 1 - \eta(a), \\ \delta_{MR} & \text{if } \delta_{MR} \in [1 - \eta(a), \eta(b)], \\ \eta(b) & \text{if } \delta_{MR} > \eta(b). \end{cases}$$

Fixed Costs

A fixed cost is a cost component that equals zero when no one receives a treatment and takes a constant positive value when any positive fraction of the population receives the treatment. Fixed costs give singleton allocations an advantage relative to fractional ones. Suppose that treatments a and b have non-negative fixed costs $C(a)$ and $C(b)$ respectively. Then allocations $\delta = 0$ and $\delta = 1$ have fixed costs $C(a)$ and $C(b)$, but any $\delta \in (0, 1)$ bears the larger fixed cost $C(a) + C(b)$. I show here that the MR allocation is fractional if the fixed costs are small but is singleton if they are large.

For simplicity, I suppose that the welfare function is linear and that the fixed costs have known values that do not vary with the state of nature. Thus, the welfare function is

(19) $$W(\delta) = \alpha + (\beta - \alpha)\delta - C(a)\cdot 1[\delta < 1] - C(b)\cdot 1[\delta > 0].$$

Allocation $\delta = 0$ is optimal if $\alpha - C(a) \geq \beta - C(b)$ and $\delta = 1$ is optimal if $\alpha - C(a) \leq \beta - C(b)$.

The problem of interest is treatment choice under ambiguity. Let S(a) and S(b) be the subsets of S on which treatments a and b are superior. That is, $S(a) = \{s \in S: \alpha_s - C(a) > \beta_s - C(b)\}$ and $S(b) = \{s \in S: \beta_s - C(b) > \alpha_s - C(a)\}$. The planner faces ambiguity if S(a) and S(b) are non-empty.

As earlier, let $M(a) \equiv \max_{s \in S(a)} (\alpha_s - \beta_s)$ and $M(b) \equiv \max_{s \in S(b)} (\beta_s - \alpha_s)$. Recall from (5) that the MR allocation in the absence of fixed costs is $\delta_{MR} = M(b)/[M(a) + M(b)]$. Let δ_{FMR} denote the MR allocation in the presence of fixed costs. The result is

(20)

$$\delta_{FMR} = 0 \quad \text{if} \quad M(b) + C(a) - C(b) \leq \min \{M(a) - C(a) + C(b),\ \delta_{MR}M(a) + C(a)[1 - \delta_{MR}] + C(b)\delta_{MR}\},$$

$$\delta_{FMR} = \delta_{MR} + \frac{C(a) - C(b)}{M(a) + M(b)}$$

$$\text{if} \quad \delta_{MR}M(a) + C(a)[1 - \delta_{MR}] + C(b)\delta_{MR} \leq \min \{M(a) - C(a) + C(b),\ M(b) + C(a) - C(b)\},$$

$$\delta_{FMR} = 1 \quad \text{if} \quad M(a) - C(a) + C(b) \leq \min \{M(b) + C(a) - C(b),\ \delta_{MR}M(a) + C(a)[1 - \delta_{MR}] + C(b)\delta_{MR}\}.$$

Proof: For any $\delta \in [0, 1]$, the maximum regret of δ on all of S is $\max [R(\delta, a), R(\delta, b)]$, where

(21a) $$R(\delta, a) \equiv \max_{s \in S(a)} \alpha_s - C(a) - \{\alpha_s + \delta(\beta_s - \alpha_s) - C(a)\cdot 1[\delta < 1] - C(b)\cdot 1[\delta > 0]\}$$

$$= \delta M(a) - C(a)\cdot 1[\delta = 1] + C(b)\cdot 1[\delta > 0]$$

$$(21b)\quad R(\delta, b) \equiv \max_{s \in S(b)} \beta_s - C(b) - \{\alpha_s + \delta(\beta_s - \alpha_s) - C(a)\cdot 1[\delta < 1] - C(b)\cdot 1[\delta > 0]\}$$

$$= (1 - \delta)M(b) + C(a)\cdot 1[\delta < 1] - C(b)\cdot 1[\delta = 0].$$

are maximum regret on S(a) and S(b).

Application of (21) at $\delta = 0$ and $\delta = 1$ gives the maximum regret values

$$\max [R(0, a), R(0, b)] = M(b) + C(a) - C(b),$$

$$\max [R(1, a), R(1, b)] = M(a) - C(a) + C(b).$$

Application of (21) at $\delta \in (0, 1)$ gives

$$\max [R(\delta, a), R(\delta, b)] = \max [\delta M(a) + C(b), (1 - \delta)M(b) + C(a)].$$

The minimum of maximum regret over $\delta \in (0, 1)$ solves the equation

$$\delta M(a) + C(b) = (1 - \delta)M(b) + C(a).$$

Hence, the minimax regret allocation on $\delta \in (0, 1)$ is

$$\frac{M(b) + C(a) - C(b)}{M(a) + M(b)} = \delta_{MR} + \frac{C(a) - C(b)}{M(a) + M(b)}$$

and the minimax regret value on $\delta \in (0, 1)$ is $\delta_{MR}M(a) + C(a)[1 - \delta_{MR}] + C(b)\delta_{MR}$. The final step is to minimize maximum regret over $\delta \in (0, 1)$, $\delta = 0$, and $\delta = 1$. This yields (20). □

Equation (20) simplifies if there is equal fixed cost associated with use of each treatment. Let $C \equiv C(a) = C(b)$. Then (20) reduces to

$$(22)\qquad \delta_{FMR} = 0 \quad \text{if} \quad M(b) \le \min \{M(a),\ \delta_{MR}M(a) + C\},$$

$$\delta_{FMR} = \delta_{MR} \quad \text{if} \quad \delta_{MR}M(a) + C \le \min \{M(a), M(b)\},$$

$$\delta_{FMR} = 1 \quad \text{if} \quad M(a) \le \min \{M(b),\ \delta_{MR}M(a) + C\}.$$

Thus, a common fixed cost smaller than min {M(a), M(b)} − δ_{MR}M(a) has no effect on the minimax-regret allocation. However, a larger fixed cost makes the allocation singleton.

2.4. Deontological Welfare Functions

Sections 2.1 through 2.3 maintained the traditional consequentialist assumption that policy choices matter only for the outcomes they yield. Deontological ethics supposes that choices may have intrinsic value, apart from their consequences.

In Section 2.3, the fixed costs C(a) and C(b) made the treatment allocation affect welfare directly, regardless of the resulting outcomes. Although I then described C(a) and C(b) in ordinary economic language as fixed costs, welfare function (19) can be interpreted as expressing the deontological idea that any use of treatment a or b is normatively bad per se, with C(a) and C(b) expressing the respective welfare losses. A more general class of deontological welfare functions is

$$(23) \qquad W(\delta) = f[\alpha + (\beta - \alpha)\delta + g(\delta)],$$

where g(·) is a specified function of δ.

Equal Treatment of Equals

When considering fractional treatment allocations, a particularly salient deontological idea is the normative principle calling for "equal treatment of equals." Fractional allocations are consistent with this principle in the sense that observationally identical persons have equal probabilities of receiving particular treatments. They are inconsistent with the principle in the sense that observationally identical persons do not actually receive the same treatment. Thus, equal treatment holds ex ante but not ex post.

A dramatic illustration of the difference between the ex ante and ex post senses of equal treatment occurs in this hypothetical problem of treatment choice considered in Manski (2007, Section 11.7).

Choosing Treatments for X-Pox: Suppose that a new viral disease called x-pox is sweeping the world. Medical researchers have proposed two mutually exclusive treatments, t = a and t = b, which reflect alternative hypotheses, say H_a and H_b, about the nature of the virus. If H_t is correct, all persons who receive treatment t survive and all others die. It is known that one of the two hypotheses is correct, but it is not known which one; thus, there are two states of natures, $s = H_a$ and $s = H_b$. Let welfare be the survival rate of the population. If a fraction δ of the population receives treatment b and the remaining 1 − δ receives treatment a, the fraction who survive is $(1 - \delta)\cdot 1[s = H_a] + \delta\cdot 1[s = H_b]$.

The singleton allocations δ = 0 and δ = 1 provide equal treatment in both the ex ante and ex post senses. These allocations also equalize realized outcomes—the entire population either survives or dies. The minimax-regret allocation is δ = ½. Everyone is treated equally ex ante, each person having a 50 percent chance of receiving each treatment, but not ex post. Nor are outcomes equalized—to the contrary, half the population lives and half dies. □

If one is concerned only with the ex ante sense of equal treatment, then all values of δ are deontologically equivalent. In terms of welfare function (23), the function g(·) is constant. If one is concerned with the ex post sense of equal treatment, singleton allocations have an advantage relative to fractional ones. In terms of (23), $g(0) = g(1) > g(\delta)$ for $\delta \in (0, 1)$.

The equal fixed-cost case considered at the end of Section 2.3 has the form $g(0) = g(1) = -C$ and $g(\delta) = -2C$ for $\delta \in (0, 1)$. Thus, placing value C on the deontological consideration of equal ex post treatment does not affect the minimax-regret allocation if $C < \min \{M(a), M(b)\} - \delta_{MR}M(a)$. However, it makes the minimax-regret allocation singleton if C is larger.

3.Dynamic Planning Problems

This section considers dynamic problems. I suppose that, in each period, a planner must choose treatments for the current cohort of a population. The planner wants to maximize the welfare of each cohort.

The essential new feature of dynamic problems is that learning is possible, with observation of the outcomes experienced by earlier cohorts informing treatment choice for later cohorts. Fractional treatment allocations are advantageous for learning because they generate randomized experiments yielding outcome data on both treatments. Sampling variation is not an issue when cohorts are large, so all fractional allocations yield the same information. Hence, the choice among fractional allocations may be based on other grounds.

I suggest use of the *adaptive minimax-regret* (*AMR*) criterion. In each period, the AMR criterion applies the static minimax-regret criterion of Section 2, using the information available at the time. In the absence of large fixed costs or deontological considerations, the result is a fractional allocation whenever both treatments are undominated. The AMR criterion is adaptive because successive cohorts may receive different allocations as knowledge of treatment response accumulates over time.

Section 3.1 formalizes the AMR criterion. Section 3.2 illustrates its application to a hypothetical problem of medical treatment. Section 3.3 discusses how the AMR criterion differs from the current medical practice of randomized clinical trials.

3.1. The Adaptive Minimax-Regret Criterion

To frame the dynamic planning problem we need to extend the concepts and notation used earlier. Let n = 0, 1, ... , N denote the periods in which treatment allocations must be chosen. In each period, the set of feasible treatments is T = {a, b}. The planner's problem is to allocate each cohort between the two treatments. A treatment allocation is a vector $\delta \equiv (\delta_n, n = 0, \ldots, N)$ that randomly assigns a fraction δ_n of cohort n to treatment b and the remaining $1 - \delta_n$ to treatment a.

Let $P_n[y(\cdot)]$ denote the distribution of treatment response across cohort n. I assume that all cohorts are large and have the same distribution of treatment response. Thus, $P_n[y(\cdot)] = P[y(\cdot)]$ for all n, where $P[y(\cdot)]$ is a time-invariant distribution. This assumption enables learning. Observation of the outcomes experienced by earlier cohorts yields information about $P[y(\cdot)]$ that can inform treatment choice for later cohorts.

As earlier, $u(t) \equiv u[y(t), t]$ denotes the net contribution to social welfare that occurs when a person receives treatment t and realizes outcome y(t). Moreover, $\alpha \equiv E[u(a)]$ and $\beta \equiv E[u(b)]$ are the mean welfare values that would result if all members of a cohort were to receive treatment a or b respectively. Hence, the optimal allocation in each period is $\delta_n = 1$ if $\beta \geq \alpha$ and $\delta_n = 0$ if $\beta \leq \alpha$.

With this background, consider planning under ambiguity. Let S_n index the feasible states of nature in period n. The planner chooses an allocation δ_n with knowledge of $(\delta_{n'}, n' < n)$ and $(S_{n'}, n' \leq n)$, but without knowledge of the information $(S_{n'}, n' > n)$ that he will possess later on. It is conceptually subtle and computationally daunting to approach choice of δ_n in a forward-looking manner, considering all logically possible subsequent sequences of information sets and choices. It is much simpler to proceed myopically, choosing δ_n as if n is the sole period of a static choice problem.

The AMR criterion provides an appealing myopic decision rule. The criterion in period n is

$$(24) \qquad \min_{\delta_n \in [0, 1]} \ \max_{s \in S_n} \ \max\,[f(\alpha_s), f(\beta_s)] - f[(1 - \delta_n)\alpha_s + \delta_n\beta_s].$$

When $f(\cdot)$ is linear, the AMR allocation follows immediately from (5). Let $S_n(a) \equiv \{s \in S_n\colon \alpha_s > \beta_s\}$ and $S_n(b) \equiv \{s \in S_n\colon \beta_s > \alpha_s\}$. Let $M_n(a) \equiv \max_{s \in Sn(a)} (\alpha_s - \beta_s)$ and $M_n(b) \equiv \max_{s \in Sn(b)} (\beta_s - \alpha_s)$. Then

$$(25) \quad \delta_{nAMR} = \frac{M_n(b)}{M_n(a) + M_n(b)}.$$

All of the other findings in Section 2 extend in the same way.

The AMR criterion has practical and normative appeal. The practical appeal is its simplicity. The static minimax-regret allocation has a particularly transparent form when welfare is linear. The AMR allocation (25) inherits this transparency.

The normative appeal is that the AMR criterion treats each cohort as well as possible, in the minimax-regret sense, given the available knowledge. It does not ask the members of one cohort to sacrifice its own welfare for the benefit of future cohorts. Nevertheless, the AMR criterion is informationally beneficial to future cohorts in the broad class of settings where it yields a fractional treatment allocation under ambiguity. Unless large fixed costs or deontological considerations make the AMR allocation singleton, application of the criterion maximizes cross-cohort learning about treatment response.

3.2. Treating a Life-Threatening Disease

This section illustrates application of the AMR criterion. I present a hypothetical treatment-choice problem in which the outcome of interest unfolds over multiple periods. As empirical evidence accumulates, the AMR treatment allocation changes accordingly.

Consider treatment of a life-threatening disease. The planner may be an independent physician, a private health maintenance organization (HMO), or a government agency such as the Veterans Health Administration in the United States or the National Health Service in England. The outcome of interest may be the number of years that a patient survives within some time horizon. For this illustration, I take the horizon to be five years and I define y(t) to be the number of years that a patient lives during the five years following receipt of treatment t. Thus, y(t) has the time-additive form

$$(26) \qquad y_j(t) = \sum_{k=1}^{K} y_{jk}(t),$$

where $y_{jk}(t) = 1$ if patient j is alive k years after treatment, $y_{jk}(t) = 0$ otherwise, and K = 5.

The outcome gradually becomes observable as time passes. At the time of treatment, $y_j(t)$ can take any of the values [0, 1, 2, 3, 4, 5]. A year later, one can observe whether patient j is still alive and hence can determine whether $y_j(t) = 0$ or $y_j(t) \geq 1$. And so on until year five, when the outcome is fully observable.

Table 1 presents hypothetical data on annual death rates following treatment by the status quo and the innovation. The entries show that 20 (10) percent of the patients who receive the status quo (innovation) die within the first year after treatment. In each of the later years, the death rates are 5 and 2 percent respectively. Overall, the mean numbers of years lived after treatment are α = 3.5 and β = 4.3. The former value is known at the outset from historical experience. The latter gradually becomes observable.

Table 1: Treating a Life-Threatening Disease

cohort or year (n or k)	death rate in k^{th} year after treatment		bound on β for cohort n	AMR allocation for cohort n	maximum regret of AMR allocation for cohort n	mean life span achieved by cohort n
	Status Quo	Innovation				
0			[0, 5]	0.30	1.05	3.74
1	0.20	0.10	[0.90, 4.50]	0.28	0.72	3.72
2	0.05	0.02	[1.78, 4.42]	0.35	0.60	3.78
3	0.05	0.02	[2.64, 4.36]	0.50	0.43	3.90
4	0.05	0.02	[3.48, 4.32]	0.98	0.02	4.28
5	0.05	0.02	[4.30, 4.30]	1	0	4.30

Assume that the planner measures welfare by a patient's length of life; thus, u(t) = y(t). Also assume that the planner has no initial knowledge of β. That is, he does not know whether the innovation will be disastrous, with all patients dying in the first year following treatment, or entirely successful, with all patients living five years or more. Then the initial bound on β is $[\beta_{L0}, \beta_{U0}] = [0, 5]$. Applying (8), the initial AMR treatment allocation is $\delta_0 = 0.30$.

In year 1 the planner observes that, of the patients in cohort 0 assigned to the innovation, 10 percent died in the first year following treatment. This enables him to deduce that $P[y(b) \geq 1] = 0.90$. The planner uses this information to tighten the bound on β to $[\beta_{L1}, \beta_{U1}] = [0.90, 4.50]$. It follows that $\delta_1 = 0.28$.

In each subsequent year the planner observes another annual death rate, tightens the bound on β, and computes the treatment allocation accordingly. The result is that $\delta_2 = 0.35$, $\delta_3 = 0.50$, and $\delta_4 = 0.98$. In year 5 he knows that the innovation is better than the status quo, and so sets $\delta_5 = 1$. The final two columns of Table 1 give the maximum regret and mean life span of each cohort, both computed using the AMR treatment allocation.

3.3. The AMR Criterion and the Current Practice of Randomized Clinical Trials

The illustration of Section 3.2 exemplifies a host of settings in which a medical planner must choose between a well-understood status quo treatment and an innovation whose properties are only partially known. When facing situations of this kind, it has been common to commission randomized clinical trials (RCTs) to learn about the innovation. The fractional allocations produced by the AMR criterion are randomized experiments, so it is natural to ask how application of the AMR criterion differs from the current practice of RCTs. There are many major differences, described below.

Fraction of the Population Receiving the Innovation: The AMR treatment allocation δ_{nAMR} can take any value in the interval [0, 1]. In contrast, the sample receiving the innovation in current RCTs is typically a very small fraction of the relevant population, with sample size determined by conventional calculations of statistical power. For example, in trials conducted to obtain Food and Drug Administration approval of new drugs, the sample receiving the innovation typically comprises two to three thousand persons, whereas the relevant patient population may contain hundreds of thousands or millions of persons. Thus, the value of δ in an RCT is generally less than 0.01 and often less than 0.001.

Group Subject to Randomization: Under the AMR criterion, the persons receiving the innovation are randomly drawn from the full patient population. In contrast, present clinical trials randomly draw subjects from pools of persons who volunteer to participate. Hence, a trial at most reveals the distribution of treatment response within the sub-population of volunteers, not within the full patient population.

Measurement of Outcomes: Under the AMR criterion, one observes the health outcomes of real interest as they unfold over time and one uses these data to inform subsequent treatment decisions. In contrast, current RCTs typically have short durations of two to three years at most. For example, a three-year trial on the disease described in Table 1would only reveal that $\beta \in [2.64, 4.36]$. Attempting to learn from trials of short duration, researchers often measure surrogate outcomes rather than health outcomes of real interest. For example, treatments for heart disease may be evaluated using data on patient cholesterol levels and blood pressure rather than heart attacks and life span. Extrapolation from surrogate outcomes to outcomes of interest can be difficult; see Fleming and Demets (1996).

Blinding of Treatment Assignment: When the AMR criterion is applied, assigned treatments are known to patients and their physicians. In contrast, blinded treatment assignment has been the norm in clinical trials of new drugs. Hence, a trial at most reveals the distribution of response in a setting where patients and physicians are uncertain what treatment has been assigned. It does not reveal the distribution of response in a real clinical setting where patients and physicians would know the assigned treatment.

Use of Empirical Evidence in Decision Making: Choosing a treatment allocation to minimize maximum regret is remote from the way that the findings of RCTs are now used in decision making. The conventional approach is to perform a statistical hypothesis test, the null hypothesis being that the innovation is no better than the status quo treatment and the alternative being that it is better. If the null hypothesis is not rejected, the status quo treatment continues in force and no one subsequently receives the innovation. If the null is rejected, the innovation replaces the status quo as the treatment of choice. A decision mechanism of this type is institutionalized in the drug approval process of the U. S. Food and Drug Administration; see Fisher and Moyé (1999).

Adaptive Clinical Trials

The AMR criterion shares a broad familial relationship with the idea of *adaptive clinical trials*, but differs in important respects. Adaptive trials sequentially draw subjects into traditional clinical trials and use a frequentist or Bayesian statistical criterion to make the allocation of new subjects across treatments a function of the outcomes observed to date for subjects drawn earlier. The objective, as stated in Tamura *et al.* (1994, p. 768), is to "use the observed response data to adapt the allocation probabilities, so that more patients will hopefully receive the better treatment."

The AMR criterion shares with adaptive trials the broad objective of using observed treatment responses to inform subsequent treatment choices. However, these ideas differ in two ways. First, the AMR criterion proposes fractional allocation of the entire patient population, not a sample of volunteers. Second, the AMR criterion is intended to cope with ambiguity rather than the statistical imprecision that motivates adaptive trials. Indeed, the large-population assumption maintained in this paper render statistical imprecision a negligible concern.

The idealization of a large population approximates well the actual environment for treatment of widespread conditions such as diabetes, heart disease, and various cancers. However, statistical imprecision in empirical findings on treatment response may be a serious cause of errors when the patient population is small. Prescription of how a planner might reasonably behave in a dynamic choice setting when facing both ambiguity and statistical imprecision is an open and difficult question.

4. Non-Cooperative Decision Processes

Sections 2 and 3 studied decision making by a planner who can dictate the treatment allocation. Planners possessing close to unilateral decision power exist in some important settings. Consider, for example, centralized health care systems where governmental agencies or private HMOs choose medical treatments for their patents. These planners can more or less unilaterally implement the AMR criterion.

Treatment allocations often result from non-cooperative decision processes. Manski (2008) suggests application of a second-best version of the AMR criterion to the institutionally complex matter of medical drug treatment in the United States, where the Food and Drug Administration has the power to approve new drugs but not to determine usage following approval. Here I consider a much simpler problem of non-cooperative choice between a status quo treatment and an innovation. I presume a two-agent setting where the agents may have different welfare functions and assessments of the feasible states of nature. I assume that any departure from the status quo policy requires the agreement of the two players.

Consider decision making in period n. When policy choice is framed as a binary decision, the status quo is chosen if either agent prefers allocation $\delta_n = 0$ to $\delta_n = 1$. The innovation replaces the status quo only if both agents prefer $\delta_n = 1$ to $\delta_n = 0$.

When policy choice is framed as selection of a treatment allocation, there may exist fractional allocations that both agents prefer to $\delta_n = 0$ and $\delta_n = 1$. Section 4.1 shows that this is so when both agents face ambiguity and use the AMR criterion to compare allocations. Section 4.2 uses an educational policy decision to illustrate.

4.1. Noncooperative Application of the AMR Criterion

Reconsider the dynamic planning problem of Section 3.1. Rather than a single planner, there now are two agents, denoted m = 1 and m = 2. Let $u_m(t) \equiv u_m[y(t), t]$ denote the net contribution to the welfare of agent m when a person receives treatment t and realizes outcome y(t). Let $\alpha_m \equiv E[u_m(a)]$ and $\beta_m \equiv E[u_m(b)]$. Let S_{mn} index the feasible states of nature in period n, as perceived by agent m..

Suppose that both agents use the AMR criterion to compare treatment allocations. Let δ_{mnAMR} be the allocation that agent m would choose if he were able to dictate policy choice. Without loss of generality, suppose that $\delta_{1nAMR} \leq \delta_{2nAMR}$.

The analysis of Section 2.2 shows that, for each agent m, the maximum static regret of an allocation δ_n strictly decreases on the interval $[0, \delta_{mnAMR}]$ and increases on the interval $[\delta_{mnAMR}, 1]$. Hence, both agents prefer δ_{1nAMR} to all $\delta_n < \delta_{1nAMR}$, both prefer δ_{2nAMR} to all $\delta_n > \delta_{2nAMR}$, and preferences differ for $\delta_n \in [\delta_{1nAMR}, \delta_{2nAMR}]$. Thus, the set of pareto efficient allocations is $[\delta_{1nAMR}, \delta_{2nAMR}]$.

Consider a decision process that calls on each agent to announce his preferred allocation and then, giving deference to the status quo, selects the smaller of the two reported values. This process makes it optimal for each agent to reveal his preferred allocation truthfully, regardless of what the other agent announces. Thus, δ_{1nAMR} is the chosen allocation. This result is pareto efficient. Alternatively, one could give deference to the innovation and select the larger of the two reported values. This decision rule also makes truthtelling optimal and yields the pareto efficient allocation δ_{2nAMR}.

When there is conflict between the preferences of two agents, society usually defers to the status quo rather than to the innovation. This is especially evident in the American legal system. A longstanding tenet of the legal system is that the plaintiff in a civil proceeding bears the *burden of proof* to show that an action by the defendant (the status quo) is improper. This tenet was recently applied by the U.S. Supreme Court to choice between a status quo treatment and an innovation.

Shaffer v. Weast: The Individuals with Disabilities Education Act is a federal statute requiring that public schools provide to each disabled child "an individualized education program." The language of the statute does not specify who bears the burden of proof when parents believe that a school has not properly complied with the statute. In the case Shaffer v. Weast (Supreme Court of the United States, 2005), the parents of a disabled child challenged the adequacy of the educational services provided by his school (the status quo policy) and proposed an alternative (the innovation). The Court ruled that the parents have the burden of proof of showing the status quo to be inadequate, writing "We hold that the burden lies, as it typically does, on the party seeking relief." □

4.2. Teacher Evaluation in New York City

To illustrate the non-cooperative decision problem, consider an educational setting where the problem is to choose between a status quo policy for teacher evaluation and an innovation. The two agents are a school district and a teacher's union. The status quo is the traditional system basing evaluation on scrutiny of teacher preparation and observation of classroom lesson delivery. The innovation bases teacher evaluation on student performance in standardized tests. The contract between the school district and the union requires that any departure from the status quo be approved by both agents.

A potential instance of this teacher evaluation problem is described in a recent article in the *New York Times* (Medina, 2008):

> "New York City has embarked on an ambitious experiment, yet to be announced, in which some 2,500 teachers are being measured on how much their students improve on annual standardized tests. . . . . . . While officials say it is too early to determine how they will use the data, which is already being collected, they say it could eventually be used to help make decisions on teacher tenure or as a significant element in performance evaluations and bonuses. . . . . Randi Weingarten, the union president, said she had grave reservations about the project, and would fight if the city tried to use the information for tenure or formal evaluations or even publicized it. She and the city disagree over whether such moves would be allowed under the contract."

Thus, New York City is acting unilaterally to collect data that could potentially be used to evaluate teachers. The contemplated change from the status quo differs from a fractional allocation as defined in this paper because the participating schools were not randomly drawn from the population of New York City schools. This difference aside, the allocation that the City has in mind is fractional with δ about equal to 0.10.

New York City appears to see itself as a planner with unilateral power to implement the innovation. However, the teacher's union asserts that any departure from the status quo policy requires their agreement. The *Times* reporter writes that an attempted unilateral decision by the City "would undoubtedly open up a legal battle with the teacher's union."

Suppose that implementation of a new policy requires agreement by the City and the union. As presently framed, the decision problem is a static noncooperative choice between $\delta = 0$ and $\delta = 0.10$. It would be better to frame it as a noncooperative choice of $\delta_n \in [0, 1]$ in a sequence of periods n, with observed outcomes in earlier periods informing treatment choice in later ones.

The fact that the City currently contemplates a fractional allocation suggests that it views itself as facing a problem of policy choice under ambiguity. The union's perception is not yet apparent, because it thus far has no way to voice its preference except to state its opposition to unilateral decision making by the City. The analysis of Section 4.1 suggests that it would be better to have the City and the union each announce their preferred allocation and then select the smaller of the announced allocations.

References

Fisher, L. and L. Moyé (1999), "Carvedilol and the Food and Drug Administration Approval Process: An Introduction," *Controlled Clinical Trials*, 20, 1-15.

Fleming, T. and D. Demets (1996), "Surrogate End Points in Clinical Trials: Are We Being Misled?" *Annals of Internal Medicine,* 125, 605-613.

Manski, C. (2005), *Social Choice with Partial Knowledge of Treatment Response*, Princeton: Princeton University Press.

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

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

Manski, C. (2008), "Adaptive Partial Drug Approval." Department of Economics, Northwestern University.

Manski, C. and D. Nagin (1998), "Bounding Disagreements about Treatment Effects: a Case Study of Sentencing and Recidivism," *Sociological Methodology*, 28, 99–137.

Manski, C. and A. Tetenov (2007), "Admissible Treatment Rules for a Risk-Averse Planner with Experimental Data on an Innovation," *Journal of Statistical Planning and Inference*, 137, 1998-2010.

Medina, J. (2008), "New York Measuring Teachers by Test Score," *The New York Times*, January 21, 2008.

Supreme Court of the United States (2005), "No. 04698, Brian Schaffer, a Minor, by His Parents and next Friends, Jocelyn and Martin Schaffer, et Al., Petitioners V. Jerry Weast, Superinten- Dent, Montgomery County Public Schools, et al.," 546 U. S.__ 2005.

Tamura, R., D. Faries, J. Andersen, and J. Heiligenstein (1994), "A Case Study of an Adaptive Clinical Trial in the Treatment of Out-Patients with Depressive Disorder," *Journal of the American Statistical Association*, 89, 768-776.