regulation2point0.org

J O I N T C E N T E R

AEI-BROOKINGS JOINT CENTER FOR REGULATORY STUDIES

Why De Minimis?*

Matthew D. Adler†

Related Publication 07-17

June 2007

† Leon Meltzer Professor of Law, University of Pennsylvania Law School.

Executive Summary

De minimis cutoffs are a familiar feature of risk regulation. This includes the quantitative "individual risk" thresholds for fatality risks employed in many contexts by EPA, FDA, and other agencies, such as the 1-in-1 million lifetime cancer risk cutoff; extreme event cutoffs for addressing natural hazards, such as the 100-year-flood or 475-year-earthquake; de minimis failure probabilities for built structures; the exclusion of low-probability causal models; and other policymaking criteria. All these tests have a common structure, as I show in the Article. A de minimis test, broadly defined, tells the decisionmaker to determine whether the probability of some outcome is above a low threshold and makes this determination relevant, in some way, to her choice.

De minimis cutoffs are deeply problematic, and have been generally misunderstood by scholars. First, they are warranted -- if at all -- by virtue of policymakers' bounded rationality. If policymakers were fully rational, de minimis cutoffs would have no justification. Second, although it seems plausible that some de minimis tests are justified once bounded rationality is brought into the picture, it is not clear which those are, or even how we should go about identifying them.

Why De Minimis?

Matthew D. Adler

I. Introduction

De minimis criteria are a widespread feature of U.S. risk regulation. As I explain below, a governmental decision procedure incorporates a de minimis criterion if it instructs the decisionmaker to determine whether the probability of some outcome is above or below a low-probability threshold and makes this determination relevant, in some way, to her choice.[1] This definition includes the classic example of a de minimis test: the quantitative cut-offs for incremental individual cancer risk, such as 1×10^{-6}, that structure regulation of carcinogenic toxins by agencies such as EPA, OSHA, and FDA. But it subsumes considerably more, including: the use of "safety factors" in regulating noncarcinogens; procedures that ignore low probability causal models; extreme event cutoffs in natural hazards policymaking; and de minimis failure probabilities for built structures.

Are de minimis criteria morally justifiable? Obviously, they have a legal basis -- in the provisions of statutes, regulations, or guidance documents that legally require the criteria -- but are these provisions themselves normatively supportable? Should we keep them in place, or instruct risk regulators to choose policies without using de minimis tests?

This Article addresses these questions and reaches skeptical conclusions. First, de minimis tests have no basis in *ideal moral theory*. Ideal moral theory provides norms for idealized decisionmakers, who are fully rational and are motivated to comply with the norms. In other words, ideal moral theory ignores problems of bounded rationality and imperfect compliance. I consider a range of moral views, both "consequentialist" and "nonconsequentialist," and argue that -- absent bounded rationality or imperfect compliance -- de minimis criteria are difficult to justify.

---

[1] I focus in this Article on the use of de minimis criteria in the decision procedures employed by governmental agencies, such as EPA, FDA, or OSHA, but the analysis naturally extends to the use of de minimis criteria by non-governmental or quasi-governmental actors -- for example, by institutional review boards in approving medical research. (Kopelman, 2004).

This is an important conclusion, because virtually all of the existing scholarly literature on de minimis tests discusses why they might be justified, and how they might be set, without reference to bounded rationality or imperfect compliance. In other words, this literature ignores the crucial point that de minimis tests are morally justifiable -- if at all -- as a matter of non-ideal, not ideal, moral theory.

Second, the Article claims that the justifiability of de minimis tests as a matter of *non-ideal theory* is unsettled. Boundedly rational decisionmakers are justified in economizing on decision costs, to some extent, and certain de minimis tests help them do that. De minimis tests might be defended, in other words, as kind of heuristic. But no one has yet provided a well-specified normative account to identify *appropriate* heuristics -- to determine which mechanisms for reducing decision costs are sufficiently accurate, and produce sufficient decision-cost economies, to be justified. In particular, value-of-information theory does not provide such an account. We therefore are currently in a position of knowing that *some* de minimis tests may be morally justified given the bounded rationality of government decisionmakers, but lacking any good methodology for sorting between justified and unjustified tests.

As for imperfect compliance: because government decisionmakers cannot be counted on to follow moral rules correctly, it is certainly possible for any given moral theory, consisting of a set of moral norms *M*, to justify the enactment of a set of legal norms *L*, where *L* and *M* are not identical. For example, the moral norms comprising *M* might be "standards," which are vague and fuzzy. The set of laws *L* that they justify might consist of bright line "rules," breaches of which are easier to detect and sanction than the fuzzy norms in *M*. The potential argument for de minimis tests, building on the *M*/*L* distinction, would be that some of the laws in *L* incorporate de minimis cutoffs even though none of the moral norms in *M* do. The trouble with *this* argument is that decisionmaker compliance with de minimis tests can itself be difficult to monitor -- either because the tests employ fuzzy qualitative probabilistic language, or because they require a quantitative risk assessment, the correctness of which will often be contestable.

Scholarship about the justifiability and calibration of de minimis levels was especially active a generation or so ago. (Adler 1978; Cross 1991; Shrader-Frechette 1985; Travis et al. 1987; Whipple 1987) Since then, scholarly interest in the topic has died down (Sandin 2005) -- but de minimis levels remain a pervasive and, I suggest, problematic feature of the regulatory landscape. Scholars need to return to the problem and, when they do, to shift focus from arguments sounding in ideal theory (which are unpersuasive) to arguments sounding in non-ideal theory (which offer more promise, particularly with respect to bounded rationality, but require much more development).

II. What Are De Minimis Criteria?

What, exactly, is a de minimis test? The scholarly literature on risk regulation sometimes equates de minimis tests with "individual risk" tests, which ask whether some individual's risk of dying from some risk source exceeds a threshold such as 1×10^{-6}. I offer a broader definition.

> *De Minimis Test: A Broad Definition*
>
> A "risk management framework" or "decision procedure" is a set of instructions for choosing policies to address some risk source. A "de minimis" criterion is a component of a risk management framework which instructs the decisionmaker to determine whether the probability of some outcome is above or below a low-probability threshold and makes this determination relevant, in some way, to her choice.

As I shall discuss in a moment, this broad definition includes both "individual risk" tests and a variety of other criteria for choosing risk regulation policies. All such tests are problematic, for the reasons presented in Parts III and IV below.

First, though, let me clarify the definition. By "outcome," I mean any possible state of affairs. A given individual dying of cancer is an outcome. So is the fact that a particular causal model is true, such as a particular dose-response or exposure model. So is the occurrence of a particular kind of natural hazard, such as an earthquake, flood, or hurricane of a certain magnitude. To be a little more precise, we might distinguish between "possible worlds" and "outcomes." A possible world is a fully specified history of the universe. An outcome is a set of possible worlds. (Lewis, 1981, 5; Loux 1998, 178-79)

A set of instructions for policy choice includes a de minimis test if, at some point, the decisionmaker is asked to determine whether the probability of some stipulated outcome is above or below a low threshold, and if the ultimate policy choice hinges -- in some way -- on that determination. The relevant probabilities can be *epistemic or frequentist*, *quantitative or qualitative, simple or incremental*, and *periodic or nonperiodic.*

Epistemic probabilities measure someone's degree of belief. There is no conceptual difficulty in assigning epistemic probabilities to outcomes, since outcomes are the object of beliefs. Frequentist probabilities measure the frequency of a property in a class of things, such as the frequency of bald men or the frequency of exposures with the property of causing cancer. (Adler 2005, 1142-46, 1206-20). Although there *are* conceptual difficulties in assigning frequentist probabilities to outcomes, these can in some cases be overcome.[2] I therefore include, under the rubric of "de minimis" criteria, both (1) epistemic de minimis tests that tell the decisionmaker to determine whether (her or someone else's) degree-of-belief in some stipulated outcome is low, and (2) frequentist de minimis tests that tell the decisionmaker to determine whether the frequentist probability of some stipulated outcome is low.

Quantitative probabilities are numbers ranging between zero and one. Qualitative probabilities are expressed by terms such as "beyond a reasonable doubt," "highly certain," "more likely than not," or "unlikely." These terms might be understood as referring to ranges of quantitative probabilities, where the boundaries of the range corresponding to a given qualitative probability are possibly vague rather than precise.
Qualitative probabilities have a central role in civil and criminal litigation, where they are used to express the burden of proof, and can and do figure in risk regulation as well - inter alia, in structuring de minimis test. For example, a decisionmaker told to ignore some potential risk source if he is "reasonably certain" that the source will not cause harm has been given a de minimis test defined in qualitative rather than quantitative terms.

[2] For example, if the outcome consists in some particular individual thing having some property, a frequentist probability might be assigned to the outcome by subsuming the individual thing in a reference class of things, and determining the frequency of the property in that class. The quantitative "individual risk" de minimis tests that governmental agencies currently employ involve frequentist, not epistemic, probabilities. (Adler 2005)

Simple probabilities are the sorts of probabilities we have been discussing thus far: a quantitative or qualitative expression of someone's degree in belief in some outcome, or of the frequency of that outcome. An incremental probability is the difference between two simple probabilities. Imagine that the probability of the outcome "Jim gets cancer," conditional on Jim's not being exposed to some toxin, is .01; and that the probability of that outcome, conditional on his being exposed to the toxin, is .0101. Then the difference between these two probabilities is .0001 = 1×10^{-4}. As this example suggests, one way to structure a de minimis test is to ask whether the incremental probability of an outcome, conditional on taking versus not taking steps to mitigate a risk source, is less than a low threshold.

Periodic probabilities involve the occurrence of some event within a time period -- for example, the annual probability of cancer or flooding. Periodic probabilities can be assimilated to the "outcome" framework I am using here by understanding them as the (epistemic or frequentist) probabilities of an outcome that is characterized, in part, in temporal terms. The nonperiodic probability of Jim getting cancer is the probability of the outcome in which Jim incurs cancer at some point in his lifetime. The annual probability of Jim getting cancer is the probability of the outcome in which Jim gets cancer during some stipulated one-year period.

A final clarification concerns the heterogeneous decisional role of de minimis tests. As we shall see, these tests can drive policy choice in a variety of ways. The decision procedure might be very abbreviated: permit a risk source if the probability of some outcome (presumably involving harm to some individual) is below a de minimis level, otherwise ban it. Alternatively, the decisionmaker might be told to allow the risk source if the probability of some outcome is below a de minimis level, and otherwise to proceed to further consideration of possible regulatory responses (using cost-benefit analysis, technology-based risk management criteria, or other criteria). Yet a different possibility is a decision procedure that tells the decisionmaker to apply one or another criterion for policy choice, but in doing so to ignore low-probability outcomes (for example, outcomes in which low-probability dose-response models obtain, or outcomes in which natural hazards with an extreme, low-probability magnitude occur).

With these clarifications in mind, it is easy to see that de minimis tests play an important role in risk regulation, at least in the U.S. What follows are some salient examples. There may well be others.

Quantitative Individual Risk Cutoffs. EPA, FDA, OSHA, and other regulatory bodies frequently employ "individual risk" tests which instruct the decisionmaker to compare the risk of dying incurred by some individual in the exposure distribution to a numerical cutoff such as 1×10^{-6}. (Adler 2005, 1149-79). For example, Section 112 of the Clean Air Act tells EPA to consider lowering the technology-based emissions standards that govern an industrial category's emissions of a carcinogenic air pollutant if the pollutant imposes an incremental lifetime cancer risk on the maximally exposed individual exceeding 1×10^{-6}. EPA's criteria for cleaning up inactive waste sites under CERCLA state that a clean-up will not be required if the incremental lifetime cancer risk consequent upon "reasonable maximum exposure" is less than 1×10^{-6}. OSHA will not regulate a workplace carcinogen if, given existing uses, the incremental lifetime cancer risk to a worker exposed for his entire working lifetime is substantially below 1×10^{-3}. The FDA licenses carcinogenic food additives exempt from the Delaney Clause so as to ensure that the incremental lifetime cancer risk to the 90th percentile food consumer is no greater than 1×10^{-6}. One of the Nuclear Regulatory Commission's fundamental safety goals in licensing and regulating reactors is that "the risk to an average individual in the vicinity of a nuclear power plant of prompt fatalities that might result from reactor accidents should not exceed one-tenth of one percent … of the sum of prompt fatality risks resulting from other accidents to which members of the U.S. population are generally exposed" (NRC 1986), which translates into an annual prompt fatality risk of 1 in 2 million. Numerous other examples of "individual risk" cutoffs in regulatory practice could be furnished.

Quantitative "individual risk" tests identify some individual -- not by name, but by her exposure characteristics. They are typically understood as asking about *incremental* probabilities. To illustrate, consider a test that asks whether the cancer risk that some risk source imposes on the maximally exposed individual exceeds 1×10^{-6}. The relevant outcome, then, is the outcome in which that person dies from cancer; and the test has the structure of asking whether the difference in the probability of that person dying of cancer, conditional on exposure versus conditional on nonexposure, exceeds 1 x 10^{-6}.[3]

A different way to understand quantitative individual risk tests -- less common among the risk assessment community, but perhaps appropriate for certain moral theories -- is as asking about the *simple* (rather than incremental) probability of an outcome characterized by a causal connection between a particular risk source and the death of a particular person. The individual risk imposed by a particular waste dump, a particular kind of food additive, a particular kind of air pollution, etc., on a particular person might be the simple probability that this particular source *causes the death* of that person or *causes cancer* to that person. In other words, John Smith's "individual risk" from source S might be interpreted not as (1) the difference in probability between the outcome "John Smith dies" conditional on his exposure to S, and conditional on nonexposure; but instead as (2) the simple probability of the outcome, "Source S causes John Smith to die."

Whether framed in simple or incremental terms, quantitative "individual risk" tests are a key example of de minimis tests.

NOAEL/Safety Factor Tests. In many of the contexts where agencies employ quantitative "individual risk" tests to regulate carcinogens, they use the NOAEL/safety factor approach to regulate noncarcinogens. (Adler 2005, 1161-65). For example, whereas FDA licenses carcinogenic food additives (exempt from the Delaney Clause) so that the "individual risk" of the 90th percentile food consumer is no greater than 1 x 10^{-6}, it licenses non-carcinogenic food additives so that the 90th percentile exposure is no greater than the NOAEL ("no observed adverse affect level") observed in animal tests, divided by a safety factor (typically 100). There are many other examples.

The use of the NOAEL/safety factor approach for noncarcinogens derives from the traditional view, among toxicologists, that each individual has a physiological threshold below which exposure will, determinately, not cause harm. However, because the regulator is not omniscient, she cannot be absolutely certain that a given (nonzero)

[3] More precisely, because the maximally exposed person from some risk source is not the same particular individual in every possible world, the "individual risk" test here is really best seen as a weighted average of incremental risks. For each particular individual in the population, we ask about the difference between his cancer risk, conditional on being the maximally exposed person versus nonexposure; and then sum those values, discounting each by the probability that the particular person is the maximally exposed person. How we estimate this weighted average will depend on what characteristics we use to estimate "individual risk" other than exposure, e.g., gender or age.

exposure of some individual to a noncarcinogen is below that individual's threshold. The technique of dividing the NOAEL by safety factors -- to account for animal-to-human extrapolation and the heterogeneity in thresholds among humans -- is a way to arrive at an exposure level that the regulator is reasonably certain will not cause harm. As Rodricks has explained:

> [The NOAEL/Safety Factor method] … does not eliminate absolutely the possibility of harm for use of [an additive], but it does provide a reasonable degree of assurance -- a "reasonable certainty" -- that harm will not result. The 100-fold safety factor … is considered prudent and protective, but it does not draw a sharp line between absolute safety and hazard. It is merely an accepted stopping point on a spectrum of possible safety factors. … [A]n even greater assurance of safety … could be achieved by raising the safety factor to 200, 500, 1000, or greater. Only at an infinite safety factor, however, would the assurance of safety be absolute. ….
>
> Thus, while the consensus of scientific opinion is that a 100-fold safety factor usually yields a "reasonable certainty" that harm will not result, scientists also recognize that there remains some residual risk or possibility, however remote, that harm could result, even after applying the protective 100-fold safety factor.

(Rodricks et al. 1991, 530-31; Brock et al. 2003, 439). In short, the NOAEL/safety factor criterion is a kind of de minimis test that involves *qualitative* rather than *quantitative* probabilities. Dividing the NOAEL by safety factors does not bring the probability of harm to any single quantitative level. Rather, it brings the probability of harm to a low qualitative level -- to the level where harm is "very unlikely," "very low," "reasonably certain not to occur," or similar qualitative language used to indicate a very small, but nonzero, probability.

More precisely, there are two possible interpretations of NOAEL/Safety factor tests, simple and incremental, which nicely parallel the two sorts of quantitative "individual risk" tests. As already explained, a simple quantitative "individual risk" test asks whether the probability of a causally-characterized outcome (one in which some particular risk source causes death, cancer, or some other harm to a particular individual) exceeds a numerical cutoff such as 1×10^{-6}. The matching *qualitative* test asks whether the simple probability of an outcome in which some particular risk source causes death, disease (other than cancer), or some other harm to a particular individual is very low. An incremental quantitative "individual risk" test asks whether the difference in the probability of some individual dying, or incurring some other harm, conditional on exposure versus nonexposure, is less than a numerical cutoff such as 1×10^{-6}. The matching qualitative test asks whether that difference in probability is very small (a qualitative incremental threshold).

Model Uncertainty and the Exclusion of Low Probability Models. The problem of model uncertainty is that the (nonomniscient) regulator is confronted by a plurality of possible models of some process relevant to risk regulation, for example, exposure models for some aspect of the environmental dissemination of toxins, or dose-response models, and cannot know for certain which model is true. The "Bayesian" approach to model uncertainty tells the regulator to assign epistemic probabilities to the different possible models, and to determine individual and population risks by integrating over the models. Although leading health and safety agencies in the U.S. have generally resisted this approach, it is an approach which, plausibly, risk regulators *should* employ -- and, in any event, one that is regularly used in risk regulation scholarship. (Adler 2005, 1209).

I suggest that the Bayesian procedure for handling model uncertainty must inevitably include a de minimis cutoff to exclude low probability models. The analyst cannot literally consider every possible model, given her cognitive limitations. For example, there are an infinity of different possible functional forms for dose-response curves, and an infinity or at least a very large number to which the reasonable analyst would ascribe some non-zero epistemic probability. In practice, Bayesian risk regulation scholarship attends to a relatively small number of models.

A low-probability model cutoff involves simple, not incremental, probabilities. The cutoff can be defined quantitatively (ignore all dose-response models with a probability below *t*), or qualitative (ignore all dose-response models that are very unlikely).

Extreme Event Cutoffs and Natural Hazard Mitigation. Structural mitigation measures for natural hazards have often been selected with reference to extreme event cutoffs. (Adler 2006, 27-28, 32-35). These cutoffs are defined in terms of periodic probabilities, either quantitative or qualitative. The "100 year flood" (that flood whose annual exceedance probability is 1%) has a central role under the National Flood Insurance Program, and the Army Corps of Engineers has designed many levees so as to protect settled areas from the 100-year flood. Levees to protect urban areas are typically designed to prevent against a yet more extreme event: the "standard project flood" discharge. "The [standard project flood] discharge in a river represents the flow that can be expected from the most severe combination of meterologic and hydrologic conditions reasonably characteristic of the geographic region involved." (Interagency Floodplain Management Review Committee 1994, 60). Note that it is possible (if improbable) for the standard project flood, thus defined, to be exceeded; indeed, the standard-project-flood level in practice seems to translate into something like the 500 year flood.

Similarly, building codes for earthquakes are typically drafted so as to ensure that buildings do not collapse except in the event of an extreme earthquake -- for example, the 475-year earthquake, a fairly standard cutoff in this context. And governmental bodies may build or design critical infrastructure, such as roads, bridges, pipelines, and so forth, so as to remain functioning in all but extreme earthquakes.

Extreme event cutoffs can be seen as de minimis tests. Their generic structure is this: The decisionmaker is concerned about the occurrence of a natural hazard in a given location, such as a city or region, and is told to identify that magnitude *m* of the natural hazard, such that the simple probability of the outcome "the hazard occurs, in this location, within the next year, with magnitude exceeding *m*" is below a de minimis level. The decisionmaker is then told to aim at some goal (protecting a settlement within the location from flooding, preventing buildings or infrastructure from collapsing), but in so doing to ignore the outcome "the hazard occurs, in this location, with magnitude exceeding *m*." Consider, for example, a procedure which tells engineers working for the Army Corps of Engineers to design a levee so as to protect some settlement, in some location, from the 100-year flood. In effect, the procedure tells the engineers to identify the flood magnitude *m*, such that the probability of the outcome "a flood occurs in this location, in the next year, with magnitude exceeding *m*" equals the de minimis level, 1%. It then says to design a levee so as to be certain (or at least virtually certain) that the settlement is not flooded, *ignoring the possibility of a flood in this location greater than m.*

Failure Probabilities for Built Structures. Structural mitigation measures for natural hazards, such as building levees or earthquake resistant structures, are often selected so as to ensure that no one in a given population (everyone in a given settlement, or inside a given building), suffers harm if the natural hazard ensues, excepting extreme events. However, designers can never be absolutely certain that a structure will contain a non-extreme event. For example, a levee designed to survive the 100-year flood may be overtopped by a smaller flood (because of low-probability conditions in the flood channel), or may just collapse. The Army Corps of Engineers historically dealt with the possibility of overtopping by adding 3 feet of additional height to its levees, thereby creating a de minimis (but not zero) probability that the levees would be overtopped by non-extreme floods. (Adler 2006, 20).

There are various ways to formalize this sort of de minimis test. Most straightforwardly, one might understand it as testing the simple probability of the outcome "the structure fails," conditional on the structure being built and a non-extreme event occurring. The decision procedure bifurcates, following one path if that probability is below a de minimis threshold (e.g., approve the structure), and a different path if that probability is above a de minimis threshold (do not approve the structure).

III. De Minimis Criteria And Ideal Moral Theory

Moral theories are diverse. (Kagan 1998). We can distinguish, first, between "consequentialist" theories that see the primary aim of morality as promoting good outcomes; and "nonconsequentialist" theories, which deny this premise. Within consequentialism, we can further distinguish between welfare consequentialism, which characterizes the goodness of outcomes solely in terms of human well-being, and non-welfarist variants of consequentialism. Within nonconsequentialism, too, there are a range of possibilities. Some moral theorists are "deontologists," who propose that "rights" or "side constraints" -- for example, an absolute or very strong duty not to murder or torture -- constrain the pursuit of good consequences. Others are "contractarians," who analyze moral questions with reference to a hypothetical social contract.

I shall argue that none of these theories, in its ideal form, supports the use of de minimis criteria. That is to say: none of these theories would direct an *idealized* government decisionmaker, who is fully rational and conscientious in complying with the demands of the theory, to employ a de minimis test. To be sure, actual decisionmakers are not idealized in this sense. But, as an analytic matter, it is very helpful to start with ideal theory. In seeing that idealized decisionmakers would *not* incorporate de minimis tests in their choice procedures, we see that it is the possibility of *departures* from the ideal case -- from fully rational and morally conscientious governmental decisionmaking -- that warrant de minimis criteria. This critical point has been insufficiently appreciated in the existing literature.

A. *Welfare Consequentialism*

Welfare consequentialism is not a single moral view, but rather a family of specific moral views, which differ in the specific social welfare functions they generate but share important features. This general approach to moral theory finds substantial support in the philosophical literature and provides the bedrock for normative economics. (Adler 2005, 1183-86; Adler & Sanchirico 2006, 291-304). Here, I focus on welfare consequentialism in its ideal-theory version: as a family of views applicable to a fully rational decisionmaker. The fully rational decisionmaker lacks full information, but her mental abilities are "unbounded." She can deduce any logical or mathematical truth, has a coherent set of probabilities over the entire space of possible outcomes, can perform mental operations instantly and at zero cost, has infinite mental "storage space," and so forth. She is not constrained by the cognitive limitations that Herbert Simon famously characterized as "bounded rationality." (Simon 1982, 1997).

Any welfare consequentialist view (in its ideal-theory version) begins with a complete and transitive ordering of possible worlds. A possible world is a complete possible history of the universe. Equivalently, it is a fully specified outcome, one that fully describes what individuals and other objects exist, what their characteristics are, what events occur, and so forth. (Loux 1998, 167-200). Further, the ordering of possible worlds provided by a given welfare-consequentialist view is sensitive only to facts about well-being. This is one characteristic feature of *welfare* consequentialism, as contrasted with nonwelfarist moral views.

If certain measurability axioms are satisfied, each world w can be represented as a vector of utilities (one for each possible individual, representing that individual's lifetime well-being in that outcome), and the ordering of worlds provided by a given welfare-consequentialist view can be represented by a social welfare function *s* defined on these utility vectors. The social welfare function *s* maps each world w_i onto a scalar number, such that $s(w_i) > s(w_j)$ if and only if w_i is ranked higher than w_j in the ordering.[4] The utilitarian social function, which simply adds up individual utilities, is one kind of social welfare function. But there are also social welfare functions that are sensitive to distributive concerns -- a point which is important to note, here, because the plausibility of welfare consequentialism as a moral framework is enhanced by its potential distributional sensitivity. (Adler & Sanchirico 2006, 291-304).

What guidance does welfare consequentialism provide for the (fully rational) governmental decisionmaker facing a choice situation $\mathbf{A} = \{A_1 \ldots A_I\}$, where each A_i represents a different possible governmental choice? (A given A_i might be the choice to issue a particular regulation, or to undertake some particular spending program to address a hazard, or to remain inactive.) Here, drawing on expected utility theory -- our best normative account of choice under uncertainty -- welfare consequentialism tells the decisionmaker to *maximize expected social welfare*.[5] (Eells 1982; Resnik 1987, 81-120; Adler & Sanchirico 2006). The decisionmaker, because fully rational, is able to ascribe probabilities to every possible world. These are epistemic probabilities: P(*w*) is the decisionmaker's degree of belief that the possible world *w* is the actual world. Integrating the decisionmaker's probability assignment to worlds, and the social welfare function *s*, the maximize-expected-social-welfare ("MESW") decision procedure instructs the decisionmaker to pick the alternative A_i that maximizes: $\sum_{w \in W} P(w|A_i)s(w)$,[6] where $P(w|A_i)$ is the conditional probability of *w* given the choice of A_i, and *W* is the set of all possible worlds.[7]

[4] Technically, there is a utility function *u* that maps *w* onto a utility vector, and *s* is a function of the utility vector. But, to keep the formalism simple, I will have *s* operate directly on *w*.

[5] I do not suggest that expected utility theory provides an accurate descriptive or predictive account of actual decisionmaking. Rather, the claim is that expected utility theory is the most attractive normative account, at least for consequentialists. Indeed, within the normative literature that accepts consequentialism, no real competitor has yet emerged.

[6] This formula assumes a countable set of outcomes. The analysis generalizes to the case of an uncountable infinity of outcomes, but, for simplicity, I will focus on the countable case here.

[7] This formulation is meant to be agnostic in the debate between "evidential" and "causal" decision theorists. (Joyce & Gibbard 1998). This debate is critical for decision theory and, by extension,

What is the connection between "outcomes" (a concept I have used in defining de minimis criteria) and "possible worlds" (a crucial element of ideal-theory welfare consequentialism and the MESW procedure)? The answer, as already explained, is that an outcome is a set of possible worlds. Any kind of outcome relevant to risk regulation (whether it describes the harm that befalls a particular person, the occurrence of an extreme event, the collapse of a structure, the truth of a causal model, etc.) corresponds to some set of possible worlds. For example, the outcome "Matt gets cancer" is the set of possible worlds in which Matt gets cancer. The outcome corresponding to a given model of toxicity (say, a linear model) is the set of worlds in which that model holds true. The outcome "the levees protecting New Orleans are overtopped by flood waters" is the set of possible worlds in which those levees are overtopped.

It is straightforward to see that there is no outcome for which the MESW procedure incorporates a de minimis cutoff. Consider any outcome, *O*. The MESW formula, $\sum_{w \in W} P(w|A_i)s(w)$, is equivalent to $\sum_{w \in O} P(w|A_i)s(w) + \sum_{w \notin O} P(w|A_i)s(w)$. That, in turn, is equivalent to: $P(O|A_i) \times \sum_{w \in O} P(w|A_i \cap O) \times s(w) + P(O^C|A_i) \times \sum_{w \notin O} P(w|A_i \cap O^C) \times s(w)$.

Let us abbreviate $\sum_{w \in O} P(w|A_i \cap O) \times s(w)$ as ESW(O, A_i), and $\sum_{w \notin O} P(w|A_i \cap O^C) \times s(w)$ as ESW(O^C, A_i). ESW(O, A_i) is expected social welfare, conditional on A_i being chosen and *O* occurring. The social welfare value for each possible world within *O* is discounted by the probability of that possible world (conditional on A_i being chosen and *O*

consequentialism (which incorporates decision theory as its account of choice under uncertainty) but does not, I believe, affect the analysis of de minimis rules. I therefore do not take a position in the debate here. For "evidential" decision theorists, $P(w|A_i)$ and $P(O|A_i)$ are ordinary conditional probabilities. By contrast, "causal" decision theorists assign probabilities to worlds and outcomes by using a "state space": a set of mutually exclusive and collectively exhaustive "states of the world," where each state is causally independent of every action in the choice set and where each state, together with a given action, fully determines what world occurs. ("States" can be thought of as outcomes consisting of facts about the world prior to the time of choice, plus causal laws.) For those who accept causal rather than evidential decision theory, $P(w|A_i)$ should be read as the aggregate probability of those states that would lead to *w*, if A_i were performed. $P(O|A_i)$, in turn, is $\sum_{w \in O} P(w|A_i)$.

occurring), and these discounted values are then summed. Similarly, $\text{ESW}(O^C, A_i)$ is expected social welfare, conditional on A_i being chosen and *O not* occurring. The social welfare value for each possible world within the complement of *O* is discounted by the probability of that possible world (conditional on A_i being chosen and *O* not occurring), and *these* discounted values are summed. Observe also that $P(O^C|A_i) = 1 - P(O|A_i)$

In short, for any given outcome *O*, the MESW procedure assigns each possible choice A_i a number equaling a probabilistic mixture of two expected social welfare terms: $P(O|A_i) \times \text{ESW}(O, A_i) + (1 - P(O|A_i)) \times \text{ESW}(O^C, A_i)$. This formulation makes clear that the MESW procedure does not incorporate a de minimis test. Note, first, that this is a single formula. (The MESW procedure is not a bifurcated procedure, which uses one formula to assign a value to A_i where the probability of *O* is low, and another formula where the probability of *O* is high.) Further, note that the single formula used by MESW is a function of the simple probability term $P(O|A_i)$, together with the ESW terms. Incremental probabilities (let alone any cutoff defined with reference to incremental probabilities) are no part of the formula.

As for the terms that *do* enter the formula, the ESW terms themselves are not logically dependent on the probability terms. $\text{ESW}(O, A_i)$ might increase, decrease, or stay constant as $P(O|A_i)$ becomes smaller and $1 - P(O|A_i)$ larger, and the same is true of $\text{ESW}(O^C, A_i)$.[8] To be sure, the formula is a *weighted* average of the two ESW terms. If $P(O|A_i)$ is lower, then $\text{ESW}(O, A_i)$-- the expected social welfare that would be produced, were A_i to be chosen and *O* to occur -- has less weight in determining the value of A_i, and $\text{ESW}(O^C, A_i)$ has more weight. But there is no categorical change in the weighting of

[8] To see this, take a given choice situation **A** and an initial set of probability assignments $P(w|A_i)$ for all the worlds within O and all the worlds within O^C, for each action A_i in **A.** $P(O|A_i)$ initially equals *k*. Decrease $P(O|A_i)$ by some factor $r < 1$ by uniformly decreasing the probability $P(w|A_i)$ of each world within O -- that is, multiplying each such probability by *r* -- and uniformly increasing the probability of every world in O^C -- that is, multiplying each such probability by $(1-rk)/(1-k)$. Then $\text{ESW}(O, A_i)$ and $\text{ESW}(O^C, A_i)$ don't change. Alternatively, if the change in $P(O|A_i)$ is spread nonuniformly among the worlds in *O*, the ESW terms might increase or decrease.

ESW(O, A_i) and ESW(O^C, A_i), depending on whether $P(O|A_i)$ is above or below a probability threshold. There is no cutoff value for $P(O|A_i)$, such ESW(O, A_i) has categorically less weight below that cutoff, and categorically more weight above the cutoff.

It might be countered that, although the MESW procedure itself does not incorporate a de minimis test, there are decisional rules which *do* incorporate a de minimis test and are good approximations for the MESW procedure. Consider, for example, a rule which identifies some outcome O and tells the decisionmaker to ignore O in calculating the expected social welfare of a given policy choice if the probability of O is sufficiently low. In other words, the rule says this: if $P(O|A_i)$ is below a threshold, the decisionmaker should proceed as if $P(O|A_i)$ is zero, and $P(O^C|A_i)$ is one. Rather than assigning A_i the value $P(O|A_i)\times$ ESW(O, A_i) + $P(O^C|A_i)\times$ ESW(O^C, A_i), the decisionmaker should assign it the value ESW(O^C, A_i). Depending on how ESW(O, A_i), ESW(O^C, A_i), and $P(O|A_i)$ vary across choice situations, it may be that the numbers assigned by this procedure to actions are quite close to the MESW numbers where $P(O|A_i)$ is below a threshold.

However, the fact that a candidate de minimis rule is a good approximation to the MESW procedure is irrelevant as a matter of ideal theory. Any procedure which is not equivalent to MESW will pick policies different from those chosen by MESW in some possible choice situations -- even, for example, the procedure just described. If welfare consequentialism is true, and if the decisionmaker is morally conscientious and has zero decision costs, then she has no reason to employ a procedure which deviates from MESW at all, whether the deviation occurs frequently or seldom. Given welfare consequentialism, the only grounds for tolerating a procedure that deviates from MESW even a little bit are non-ideal grounds -- namely, that the procedure is cheaper to employ, or that it helps constrain decisionmakers who are not perfectly conscientious.

B. *Health Consequentialism*

Welfare consequentialism does not give special priority to health and longevity over other aspects of well-being. But risk regulation statutes often place special emphasis on health and longevity, and many scholars in the field of risk regulation believe (or seem to believe) that this moral priority is justified. (Adler & Posner 2006, 73-74). One moral theory that incorporates such a position, to be considered shortly, is a nonconsequentialist or "deontological" view that creates "rights" or "side constraints" protecting health and longevity. Yet it is also possible to prioritize health and longevity within the family of consequentialist moral theories.

This possibility has been explored at some length by public health scholars, who have developed the construct of what might be called a *social health function*. (Dolan, 1998; Lindholm & Rosen, 1998; Osterdal, 2005; Williams, 1997). A traditional social welfare function is sensitive to facts about well-being; represents each possible world as a vector of well-being utilities, measuring each individual's well-being in that world; and maps each such vector onto a scalar representing the place of the world in the social ordering. A social health function is sensitive to facts about individual health; represents each possible world as a vector of health utilities, measuring each individual's health in that world; and maps each such vector onto a scalar, representing the place of the world in the social ordering. There is much flexibility, within this framework, in how health utilities are assigned to individuals and in the form of the function that operates on health utilities. The metric of health might be an individual's longevity, or a "QALY" number that incorporates both health and longevity. The function might be straight additive (a utilitarianism of health), or it might be sensitive to the distribution of health.

Health consequentialism is very different from welfare consequentialism, in focusing on health rather than well-being. But it is isomorophic to welfare consequentialism in one respect critical for our purposes here -- namely, in incorporating expected utility theory in providing guidance to unboundedly rational decisionmakers choosing under uncertainty.[9] Health consequentialism says to *maximize expected social*

[9] As already stated, expected utility theory is the leading normative account of choice under uncertainty for consequentialists. One might reject it by moving outside consequentialism, or perhaps by remaining within

health ("MESH"). It says to choose among **A** = {$A_1 \ldots A_I$}by picking the A_i that maximizes $\sum_{w \in W} P(w|A_i)h(w)$, where h is the social health function. Thus, like the MESW approach, MESH does not incorporate a de minimis threshold. In this crucial respect, the two procedures are similar. Remember that, for any given *O*, and regardless of the probability of *O*, the number assigned by MESW to a given action A_i equals a probabilistic mixture of expected social welfare with and without *O*. Similarly, for any given *O*, and regardless of the probability of *O*, the number assigned by MESH to a given action A_i equals a probabilistic mixture of *expected social health* with and without *O* . Paralleling the decomposition of the MESW formula above, it equals: $P(O|A_i) \times \text{ESH}(O, A_i) + (1 - P(O|A_i)) \times \text{ESH}(O^C, A_i)$, where

$$\text{ESH}(O, A_i) = \sum_{w \in O} P(w|A_i \cap O) \times h(w), \text{ and } \text{ESH}(O^C, A_i) = \sum_{w \notin O} P(w|A_i \cap O^C) \times h(w)$$

Incremental probabilities do not figure in this formula at all; and although the simple probability $P(O|A_i)$ does, this probability term is not compared to a qualitative or quantitative cutoff. There is no categorical change in the relevance of ESH(O,A_i) and $\text{ESH}(O^C, A_i)$ as $P(O|A_i)$ gets small.

C. *Rights Based Views*

A "deontological" or "rights based" moral view includes some "side constraints," paradigmatically constraints on actions such as intentional killing, battery, fraud, or trespass. (Kagan 1998, 70-105; Adler 2005, 1223-32). These side constraints are "agent relative." Roughly, the agent is told that *he* ought not to perform a killing, a trespass, etc. – even if, by engaging in the prohibited action, he prevents identical actions by other agents. Because side constraints are agent relative, a moral theory that includes a side constraint cannot be given a consequentialist representation. Consequentialism (in its ideal-theory variant) instructs every agent to choose A_i in a choice situation **A** so as to

consequentialism but adopting some non-standard account of choice under uncertainty. But there is nothing in the shift, within consequentialism, in the *arguments* for the social value function - from individual welfare to individual health -- to warrant accepting it in the welfare case but rejecting it in the health case.

maximize $\sum_{w \in W} P(w | A_i) v(w)$, where *v* is a single function for all agents – be it a social welfare function, a social health function, or some other function of possible worlds. Agent relative constraints, if expressed in terms of a ranking of possible worlds, would have different rankings for different agents. Thus the dictates of a genuinely deontological theory are not coextensive with the maximization of $\sum_{w \in W} P(w | A_i) v(w)$, for any function *v*.

Deontological philosophers continue to debate numerous issues concerning side constraints: which harms are covered; whether the side constraints are absolute or defeasible; whether they take the "ex post" form of prohibiting certain actions that cause harm, or the "ex ante" form of prohibiting certain actions that risk harm; and what the additional conditions are for the violation of a side constraint (for example, regarding the agent's mental states, the act/omission distinction, or the directness of the causal connection), beyond causing or risking harm. (McCarthy, 1997; Nozick 1974, 26-87; Perry; Thomson, 1990).

Whatever their precise structure and content, would side-constraints incorporate de minimis levels? I will focus my analysis on the harm of death -- considering, first, side constraints of the "ex post" form that prohibit certain actions which cause death and, then, side constraints of the "ex ante" form that prohibit certain actions which impose a risk of death.[10] And I will focus on de minimis tests of the "individual risk" form -- namely, those that instruct the decisionmaker to determine whether the "individual risk" of death that his action imposes on some individual is above or below a quantitative or qualitative threshold.[11] It seems at least plausible that the shift from consequentialism to deontology will rehabilitate de minimis tests of this kind. The purpose of side constraints is to block actions that are wrongful to others – to the person who would be killed,

[10] The analysis is virtually the same for non-fatal harms. One difference is that these harms are compensable. So it is more plausible to include a compensation condition in an "ex post" constraint on causing non-fatal harm than it is to include such a condition in an "ex post" constraint on causing death. This difference, however, does not provide an argument for why constraints regarding non-fatal harms should include a de minimis level whereas constraints regarding death would not.

[11] As already suggested, the risk imposed on a particular individual by an activity might be the simple probability of the outcome that the activity causes his death, or the incremental activity of the relevant kind of death (e.g., cancer) conditional on the activity occurring versus not occurring. My argument against de minimis thresholds, here, does not depend on which interpretation of individual risk is adopted. Nor does it depend on whether an epistemic or frequentist interpretation is adopted.

assaulted, trespassed upon, etc. And there is some intuitive plausibility to the notion that A wrongs B only if the risk that A imposes on B is above a threshold. By contrast, it is very hard to see why the shift from consequentialism to deontology would rehabilitate other sorts of de minimis tests.

Let us start, then, with the "ex post" death-relevant side constraint, a constraint on causing death. If such a constraint incorporates a de minimis cutoff, it takes the following form: it says that an action is (absolutely or prima facie) prohibited if (1) it ends up causing some individual's death; (2) the action had a sufficiently high risk of causing that individual's death (the de minimis condition); and perhaps (3) further conditions are satisfied (regarding, for example, the actor's mental states, the act/omission distinction, or the directness of the causal connection). But do deontologists really have good reason to adopt such a constraint? An important difficulty involves deaths that result from actions which are specifically intended to impose a risk of death, but which have a low individual risk. I play Russian Roulette with you, for thrills, using a bomb that has a 1×10^{-7} chance of exploding. The individual fatality risk that my action imposes upon you is very low, but, if the bomb actually explodes and causes your death, haven't I wronged you? The deontologists who have written about killing and risk-imposition tend to have the intuition that, in this sort of case, a violation of a side-constraint has occurred. (McCarthy 1997, 211-13).

Conceivably, a side constraint with respect to causing death could have a hybrid structure, incorporating a de minimis threshold with respect to deaths that are caused by actions not specifically intended to impose risk on the victim, and not incorporating any such threshold with respect to intentional killings. One putative rationale for this hybrid structure would be this: actions that are specifically intended to impose fatality risks have no benefits at all, and should always be flatly prohibited, while other risk-imposing actions can have benefits (for example, the action of my driving to work, which imposes a small fatality risk on pedestrians) and should not be flatly prohibited. But the fact that risk-imposing actions can have benefits hardly justifies a de minimis threshold. Rather, it justifies a defeasible side constraint. (McCarthy 1997, 208-10). The defeasible variant of a side constraint on causing death would have the following structure: an action that (a) causes death and (b) meets certain further conditions (concerning the agent's mental states, the act-omission distinction, etc.) is prohibited, *unless* (c) it causes sufficiently large benefits.

A thought experiment will illustrate the difference between this side constraint, and one that incorporates a de minimis threshold. Imagine a chemical which is totally inert unless the exposed individual has a very rare genetic mutation. The probability of any individual having the mutation is below the de minimis threshold.[12] One form of the chemical causes individuals with this mutation to die prematurely. Another form causes individuals with the mutation to die prematurely, but also dramatically improves their health while alive. Imagine, now, that a polluter exposes some individual to the chemical; the individual has the very rare mutation, and he ultimately dies as a result of the exposure. If the side constraint on causing death includes a de minimis threshold, the polluter has done nothing wrong -- *regardless of which form of the chemical he emitted.* But deontologists who are sensitive to the possible benefits of harmful actions should resist *that* conclusion. Rather, whatever their view about the case in which the chemical has health benefits, they ought to conclude that the polluter has behaved wrongfully if he emitted the form of the chemical that produces no health benefits. As this case illustrates, a de minimis threshold will shield actions that cause death and do not cause countervailing benefits -- and it is hard to see why ex post deontologists would want to exempt such actions from their side constraint.

Are there other arguments for the incorporation of a de minimis threshold in a side constraint on causing harm? Risk assessment scholars have periodically suggested that the problem of statistical "discernibility" would warrant a de minims level, and that suggestion might seem particularly relevant here. (Cross et al. 1991) If the side constraint on causing death includes no de minimis level, and I cause your death through an action that imposes a very small individual risk of death upon you, how will the supposed violation ever be demonstrated? How will it be shown that the very unlikely

---

[12] This assumes that (1) if the fatality probabilities relevant to the de minimis threshold are epistemic, they do not reflect the beliefs of an omniscient knower (who would know that any given individual has or lacks the mutation), and (2) if the fatality probabilities are frequentist, they are frequencies relative to a reference class that incompletely rather than completely specifies the characteristics of the person upon whom risk is imposed (because a completely specified reference class would include whether the person has the mutation). Both seem reasonable assumptions; but if the reader objects, the hypothetical can be changed to one in which the inert chemical, via an indeterministic process, triggers a mutation in a very small fraction of the population.

indeed occurred -- namely, that my action did cause your death? But problems of statistical discernibility would seem to be more relevant for institutions such as tort law and criminal law, which respond to deaths after they occur and need to demonstrate the causal origins of particular deaths up to some evidentiary threshold ("more likely than not," "beyond a reasonable doubt"), and less relevant to regulators, who intervene to prevent deaths before they occur. On the ex post deontological view, regulators should intervene to prevent actions that are sufficiently likely to violate the side constraint -- that is, sufficiently likely to cause death and meet the additional conditions required for the violation of the constraint -- regardless of whether those deaths, if they do end up occurring, would be discernible. Note, on this score, that risk regulators currently prohibit a range of activities which impose low individual risks; if those activities were permitted and were to cause deaths, the causal connection between the activities and the deaths might well be indiscernible.

In any event, even if deontological regulators *should* be attentive to considerations of "discernibility," those do not map neatly onto any de minimis level. First, there are some low-risk actions that can meet high evidentiary thresholds for proving causation. Imagine that the action, when fatal, leaves some kind of signature. Second, and reciprocally, evidentiary difficulties can arise for actions that impose a risk on some individual well above any level that can be called de minimis. For example, imagine that two or more actors each imposed an incremental cancer risk of 1 in 10 on the victim, or that the individual's background risk of dying from cancer was 2 in 10, and the actor increased it to 3 in 10.

A different argument for de minimis levels involves the application of ex post side constraints under conditions of uncertainty. The ex post deontologist says that a deontological violation occurs when an action actually causes death (and further conditions are satisfied). An omniscient actor would guide her behavior using this constraint by avoiding actions that will cause death (and satisfy the further conditions), and an omniscient deontological regulator would intervene to prevent just those actions.

But what about nonomniscient actors and regulators? Ex post deontologists, to render their side constraint usable at the point of choice, need to include ancillary instructions about what nonomniscient actors and regulators should do about actions that might cause death. (Jackson & Smith, 2006).

Unfortunately, there has been little discussion, and certainly no convergence among deontologists, as to what those ancillary instructions would be. It is conceivable that ex post deontology outfitted for choice under uncertainty takes the following form: actors should refrain from any action which has a non-de minimis probability of causing death (and meeting the additional conditions). But why should the theory be specified in this way? It is certainly true that the ex post deontologist's theory of choice under uncertainty needs to avoid a kind of probabilistic absolutism. Imagine that the theory says this: an actor is prohibited from performing an action which might cause death without causing sufficient countervailing benefits. Such a theory would lead to moral paralysis: any action might cause death without causing countervailing benefits. But injecting a de minimis threshold into the theory may not solve the problem of probabilistic absolutism. If deontology prohibits actions that have a 1 in 10^6 chance of causing death without causing sufficiently large countervailing benefits, a vast array of seemingly innocuous and laudable activity may be shut down. (My driving has at least a 1 in 10^6 chance of killing both a pedestrian and me, thus causing death without causing sufficient benefits to me, or anyone else, to justify the killing.) Nor is it clear what would motivate a particular threshold level.

A final argument for de minimis thresholds in an ex post constraint on causing death involves compensation. "Defeasible" deontologists tend to think that actions which are on balance justifiable, despite infringing constraints, trigger duties to compensate. If I'm a hiker who breaks into your mountain cabin and eats your food because I'm starving, then I have permissibly infringed your property rights-- but must compensate you for the food. (McCarthy 1997, 219; McCarthy 1996). Death, however, is not compensable. Therefore -- or so the argument might go -- an actor who causes a victim's death only does so permissibly if he compensates her ex ante. One way to compensate the victim ex ante is to include her in a reciprocal risk pool. (Fletcher 1972; McCarthy 1996). If a group of individuals engage in a risky but mutually beneficial practice, each imposing a fatality risk of r^* on the others, then if one of the group is killed by another's action, the action will be permissible (if it also causes sufficient benefits);

the victim was compensated by the fact that everyone else accepted a risk of r^* as well. If, on the other hand, the action that caused the victim's death imposed an individual risk of $r^+ > r^*$, then he was not compensated and the action causing his death was impermissible. The upshot is that r^* is a kind of socially based risk threshold, such that actions imposing a fatality risk below r^* are (or can be) permissible, while actions imposing a risk above it are not.

There are multiple difficulties with *this* line of argument. To begin, a defeasible side constraint framed in terms of a prohibition on causing harm should presumably require compensation for the harm, not just the risk of harm. Why would the side constraint have an "ex post" structure in terms of the infringing activity, but an "ex ante" structure qua compensation? The deontological view which (a) adopts a side-constraint prohibiting actions that cause death; (b) allows infringements of that constraint, even for actions that cause substantial benefits, only if the victim is compensated; but (c) counts compensation for the risk of death as sufficient, is a very odd hybrid. In any event, even if the deontologist *does* adopt this odd hybrid view, the "risk pool" argument sketched in the preceding paragraph is fallacious. What compensated the dead individual, ex ante, was *not* the risk imposed on other individuals. It was the benefits that he reaped from the risky practice. Imagine that the actions of other members of the risk pool imposed a uniquely high fatality risk $r^+ > r^*$, on the victim, while everyone else incurred level r^*, but the ex ante benefits for the victim and the other participants were much larger than both an r^* and an r^+ risk of death. Then the action causing the victim's death may well have been permissible, despite the fact that the action causing his death imposed an above-r^* risk. On the other hand, imagine that everyone in the group incurred the same fatality risk r^*, but none received an ex ante benefit sufficient to justify their participation (which was irrational). Then the action causing the victim's death should be counted impermissible.

To sum up, we have considered four possible arguments for the inclusion of a de minimis threshold in an ex post constraint on causing death: the need for a constraint to be sensitive to the benefits of actions that may cause death; discernibility; outfitting the constraint for choice under uncertainty; and compensation through reciprocal risk pools. None of the arguments seems to succeed.

What if we turn to the "ex ante" variant of the deontological side constraint, namely a constraint on actions that impose a risk of death (plus meet other conditions involving the act-omission distinction, the actor's mental state, etc.)? The shift from ex post to ex ante deontology involves a shift in the timing and prevalence of deontological infringements. On the ex ante view, the infringement occurs at the moment of risking, independent of whether the action actually causes death. This shift is philosophically interesting, and may have important institutional implications -- for example, justifying tort liability for risk imposition per se. It does not, however, strengthen the argument for de minimis levels.

Indeed, David McCarthy, the leading philosophical defender of the view that deontological constraints have an ex ante rather than ex post structure, explicitly considers and rejects the suggestion that the constraint on risk applies only to "high" risks. (McCarthy 1997). Rather, on McCarthy's view, *any* action imposing a non-zero fatality risk on some individual, however small, infringes the constraint. The concern for absolutism, within an ex ante view, can be handled by making the constraint defeasible by sufficiently large expected benefits. Assume (for simplicity) that the expected benefits of violating a constraint must be greater than some multiple $m^* > 1$ of the expected fatality costs. Consider, now, a polluting process that is expected to cause 10 more annual deaths than a more expensive, cleaner process and imposes individual fatality risks as high as 10^{-4}, but achieves cost savings of $70 million annually. If the expected annual fatality costs of the polluting process are measured, using a "value of statistical life" of $6 million, at $60 million, then the use of the process is deontologically prohibited. Surely $m^* > 7/6$.[13] And, if we were to hold constant the benefits of the process, and reduce the risks by a factor of one hundred -- driving down the individual risks to 10^{-6} at the highest, and expected annual deaths to 0.1 -- then the deontological verdict might change. Plausibly, $m^* < 700/6$. On the other hand, if we were to reduce both the risks *and* the benefits of the process by a factor of one hundred, then the very low risk imposition would -- quite appropriately -- be counted deontologically wrongful, because its benefits are also now small. The point is that the wrongfulness of the process

[13] Any deontologist will insist that m^* be substantially greater than 1. If it is 1, then the so-called "constraint" disappears and the actor is simply required to maximize expected net benefits.

is a matter of comparing its expected benefits with some multiple of its expected fatality costs, and that the process can therefore be wrongful even if the individual risks it imposes on any individual are very low.

Like most other defeasible deontologists, McCarthy insists that an actor who justifiably infringes a defeasible constraint must provide compensation. This might seem to lead to a socially defined de minimis level, via the notion of reciprocal risk pools. It might seem that regulators implementing an ex ante deontological constraint would (1) prohibit all actions that impose a fatality risk on anyone (and that meet further conditions concerning the actor's mental state, act versus omission, etc.), if the expected benefits of the action are less than m^* times the expected fatality costs; and (2) permit actions that impose a fatality risk (and meet the further conditions), if the expected benefits of the action are greater than or equal to m^* times the expected fatality costs, *conditional on* (a) payment of monetary compensation to the individual bearing the risk; or (b) the individual being compensated by participating in a risk pool, as evidenced by the risk level lying below a "social" background; or (c) compensated in some other way. A de minimis level has crept into this complicated deontological rule, under prong (2)(b). But, for reasons already discussed, prong (2)(b) of the rule is unwarranted. The fact that the risk an individual incurs in some social practice is above or below the level of risk he imposes on others in the practice, or that the others impose on each other, does not determine whether he has been sufficiently compensated for that risk. What determines that is his benefits from the practice.

D. *Contractarianism*

John Rawls revived the contractarian approach to moral thinking, one that uses the device of the hypothetical contract to justify moral requirements, and it represents a major school of thought within contemporary moral philosophy. (Rawls, 1971). However, the implications of contractarianism for risk regulation are unsettled. First, there are very basic, unsettled questions about the view itself. Are the participants to the hypothetical contract self-interested (Rawls' position) or more altruistic (the position adopted by a leading living contractarian, Tim Scanlon)? (Scanlon, 1998) Do the participants conform to expected utility theory (John Harsanyi's position), or do they choose under uncertainty by adopting a "maximin" rule (Rawls' position)? (Harsanyi 1982, 44-48). Is the hypothetical contract a device for evaluating particular choices, or a second-order device for evaluating general moral rules, which will then provide guidance in particular choice situations? (Adler 2005, 1232-27) Second, little scholarship by contractarians has specifically examined risk regulation. (Oberdiek, 2003)

Still, it is possible to make some progress on the subject of de minimis contractarian criteria. Consider first the case in which the hypothetical contract is used to evaluate particular governmental choice situations, such as **A** = {A_1 … A_I}, and the contractors {C_1 ….C_K} are assumed to be self-interested. Assume, further, that we are dealing with ideal-theory contractarianism: the governmental official and contractors are unboundedly rational. If, as per Harsanyi, the hypothetical contractors obey the axioms of expected utility theory, then each contractor C_k would choose among **A** so as to maximize $\sum_{w \in W} P_k(w|A_i)u_k(w)$, where $u_k(.)$ is a utility function representing C_k's self-interested preferences and P_k is the probability measure for contractor C_k.[14] The government official's choice among **A** is, in turn, a function of the expected utility numbers assigned to each A_i by each C_k. The nature of *that* function is not clear -- at a minimum, however, if there is some A^* that each C_k ranks highest, the government official must rank A^* highest as well. Call this the "Pareto constraint."

It is straightforward to see that the expected utility formulae for the K individual contractors, $\sum_{w \in W} P_k(w|A_i)u_k(w)$ for k=1 to K, do not incorporate a de minimis test. The analysis above of the MESW and MESH procedures is equally applicable here. The government official's procedure takes the numbers generated by the K individual formulae as inputs. Whether that procedure incorporates a de minimis test is less obvious. I conjecture (although will not try to rigorously demonstrate) that any procedure which does incorporate a de minimis test will fail the Pareto constraint in some choice situations.[15]

---

[14] The probabilities, because epistemic, might vary among the contractors.

[15] At a minimum this is true of a governmental procedure which tells the official to ignore some outcome *O* if its simple probability falls below a threshold, calculating the individual contractors' expected utilities as if the probability of *O* were zero. For any outcome and any threshold (however small), there is some

What if the hypothetical self-interested contractors follow the Rawlsian prescription to maximin, rather than Harsanyi's prescription to maximize expected utility? Maximin should not be conflated with a de minimis rule. To begin, maximin does *not* instruct the decisionmaker to ignore outcomes whose simple probability is low. Rather, maximin tells the unboundedly rational contractor C_k to proceed as follows: for each possible choice A_i, consider each possible world w such that $P_k(w|A_i)$ is nonzero; among these positive probability worlds, identify the world $w^{min\text{-}Ai}$ that minimizes C_k's well-being, i.e., the "worst case" world associated with A_i; and choose the A_i whose associated worst-case world $w^{min\text{-}Ai}$ produces the greatest well-being for C_k. The crucial point is that, in identifying the "worst case" world associated with each A_i, the maximiner considers each positive probability world and outcome, even very improbable ones. Imagine, instead, that the contractor employs a simple probability threshold t appended to some outcome O -- so that in determining the "worst case" world associated with A_i, the contractor ignores worlds belonging to O, where $P_k(O|A_i) < t$. Clearly, this variation on maximin would deviate from straight maximin in some choice situations where the worst-case world associated with an A_i belongs to an outcome O with probability below t.

It might seem that de minimis tests which are defined in terms of incremental probabilities *are* assimilable to maximin. Maximin is insensitive to *changes* in the probability of outcomes. More precisely, consider two actions A_i and A_j. There is some given outcome O^* such that $P_k(O^*|A_i) > P_k(O^*|A_j)$. However, the choice between A_i and A_j does not alter the worst-case scenario for contractor C_k. The lowest level of well-being that C_k might attain (with nonzero probability) given A_i, and the lowest level of well-being that C_k might attain (with nonzero probability) given A_j, are the same. Then the maximin rule tells C_k to be indifferent between A_i and A_j, notwithstanding the change in the probability of O^*.

Although maximin does incorporate this indifference feature, it has nothing to do with de minimis as opposed to large changes in probability. As long as the worst-case

---

choice situation and some individual utility function for contractor C_k where ignoring the outcome tips the balance. The contractor prefers one action if she maximizes expected utility in the normal way, and a second action if she ignores the outcome. *If* this is true for every contractor in that choice situation (which could arise, e.g., if every contractor has the same utility function), then the official chooses the second action, while the Pareto constraint requires the first.

level of well-being associated with A_i and A_j is the same, even large changes in the probability of any given outcome do not influence C_k's choice if she maximins. Imagine that, if A_i is chosen, C_k has a 1% chance of dying young after a terribly painful disease, and a 99% chance of a better life. If A_j is chosen, C_k has a 51% chance of dying young after that disease, and a 49% chance of a better life. Then the difference in probability of the outcome in which C_k lives the short, painful life is 50% -- hardly de minimis -- but maximin still enjoins C_k to be indifferent between the choices.

If the basic contractarian structure changes -- if we move to altruistic contractors, to a second-order choice of rules, or both -- do de minimis criteria emerge? This issue merits a lengthier discussion than I have space for here, and so I will leave it unresolved. In the case of first-order contracting by altruistic contractors, we would need to know exactly which considerations motivate them. If those are (say) a mix of self-interested considerations, a concern for the interests of friends and family, a concern for general welfare, and a concern for moral rights, de minimis levels would not seem to be in the offing -- since none of these factors, on its own, includes a de minimis level.[16] In the case of a contractarian theory that uses the device of a hypothetical contract for a second-order choice of rules, rather than to determine directly what agents should do in particular choice situations, we would need to understand why the theory takes this two-tiered form. What considerations drive a potential wedge between (a) what the hypothetical contractors would want any particular decisionmaker to do, and (b) what the rules chosen by the hypothetical contractors would direct any particular decisionmaker to do? *If* the wedge arises for non-ideal reasons -- for example, because the rules are designed to guide boundedly rational decisionmakers-- then (for the sorts of reasons to be discussed in a moment) the rules might plausibly incorporate de minimis cutoffs. If the wedge arises for other reasons, the case for de minimis cutoffs is less clear.

---

[16]The contractor advance his own interests by maximizing the expectation of a utility function representing his own well being; he advance the interests of a third party by maximizing the expectation of a utility function representing that person's well-being; he advances general welfare by following MESW; and he takes account of moral rights by respecting deontological constraints. The first, third, and fourth of these factors have already been analyzed and shown to lack a de minimis cutoff; and the analysis of the second is the same as for the first.

IV. <u>Non-Ideal Moral Theory And De Minimis Levels</u>

A. *Decision Costs*

In this Section, I shall argue that bounded rationality provides a *potential* justification for de minimis tests (at least for consequentialists), but that decision theory and moral philosophy in their current state provides no ready basis for identifying which de minimis tests are in fact justified given bounded rationality. I shall focus, here, on welfare consequentialism -- because, at least within welfare consequentialism, it is quite clear that the decision procedures for ideal and non-ideal agents are *different*. Thus the *potential* for a defense of de minimis tests grounded on non-ideal considerations.

As discussed, the ideal-theory variant of welfare consequentialism instructs the decisionmaker to use the MESW procedure. The MESW procedure tells the decisionmaker to choose among **A** = {$A_1 \ldots A_I$} by maximizing $\sum_{w \in W} P(w|A_i)s(w)$, where each *w* is a complete possible world – a maximally specified outcome, which expresses a complete possible history of the universe – and *W* is the set of such histories. It is impossible for an actual human brain to store a single *w*, let alone the set *W*, and (I assume) impossible or at least hugely expensive for current computers to do so. Thus it is impossible or at least hugely expensive for actual humans to actually employ the MESW procedure.

It is critical, at this juncture, to underscore the distinction between bounded rationality and the lack of full information. MESW does *not* assume that decisionmakers possess full information. The very point of expected utility theory, in general, and of MESW in particular -- which represents the merger of expected utility theory and welfare consequentialism -- is to guide choice by uncertain decisionmakers, who do not know which outcomes will result from the choices available to them. However, MESW *does* assume that decisionmakers are unboundedly rational. The MESW procedure is therefore not usable by actual humans.

How might the MESW procedure be reworked so as to be usable by actual human (perhaps aided by current computers)? One possibility is to employ simplified social welfare functions, with a relatively small number of inputs, and simplified causal models for predicting what these inputs will be. Consider, in this regard, the optimal tax literature – a context in which social welfare functions have actually been used quite extensively, by scholars, to evaluate possible policies. (Tuomala, 1990). Typically, *s* is a function of individual consumption and leisure. That is, the argument for *s* is a vector of utilities, one for each member of the population, where each individual's utility is solely a function of his consumption and his leisure (rather than of his consumption, his leisure, his health, his happiness, his social status, his sex life, his relations with friends and family, his professional accomplishments, his spiritual life, his recreational pursuits, and everything else that, along with income, actually affects individual well-being). And the models used to predict individuals' consumption and leisure are simplified representations of the economy.

A social welfare function, thus simplified, can be applied to a set of policy choices $\mathbf{A}=\{A_1 \ldots A_I\}$ without considering the set *W* of complete possible worlds. Rather, the decisionmaker needs (more tractably) to consider the set *Z* of *relevantly specified* possible worlds. Each *z* is an outcome in which the inputs for the simplified *s*, plus the causal determinants of those inputs for all the model(s) under consideration $\mathbf{M}=\{M_1 \ldots M_L\}$, take some value, and one of the models is true. *Z* is the set of all possible *z*.

The decision maker, given a simplified *s*, a simplified set of causal models, and a corresponding *Z*, will assign expected-social-welfare numbers to each A_i by aggregating across *Z*, not *W*. In other words, she will determine $\sum_{z \in Z} P(z|A_i)s(z)$, for each A_i, and pick the A_i with the highest such value. Since *Z* is smaller than *W*, and since each *z* is specified in much less detail than each *w*, this can be a tractable procedure for actual humans.

How do de minimis tests come into play? Like the choice of a simplified *s* and a simplified modeling structure, the use of de minimis tests can economize on decision costs. Consider, first, a de minimis test that instructs decisionmakers to ignore low-probability models. Such an exclusion can reduce the size of *Z*, and the level of complexity of the various *z* contained in *Z* – thus reducing decision costs. An excluded, low-probability causal model *M** may make the arguments for *s* a function (wholly or in part) of new variables which are not relevant to higher-probability models. If *M** is considered, then each *z* will need to be specified with respect to these new variables as well as the variables relevant to the higher-probability models. Further, even if the low-probability model introduces no new variables, its introduction into the analysis increases decision costs, by increasing the size of Z. If higher-probability models $M_1 \ldots M_L$ are already part of the analysis, the decisionmaker must now calculate *s*-values, and assign probabilities, for each world *z* where M_1 is true, each world where M_2 is true, up to M_L, *and* each world where M^* is true.

De minimis tests that instruct the decisionmaker to ignore low-probability outcomes other than causal models can also economize on decision costs. Imagine that **O** is a partition of outcomes $\{O_1 \ldots O_H)$ -- for example, different possible magnitudes of some natural hazard. Consider first a decision procedure that employs some simplified social welfare function *s* and some set *Z* of relevantly specified possible worlds, but does not use a de minimis test. It tells the decisionmaker to determine $\sum_{z \in O_1} P(z|A_i)s(z) + \sum_{z \in O_2} P(z|A_i)s(z) + \ldots + \sum_{z \in O_H} P(z|A_i)s(z)$, for each possible action A_i. Consider, now, this procedure revised to incorporate a de minimis threshold for outcomes within **O**. The decisionmaker is told to calculate expected social welfare on the assumption that below-threshold outcomes do not occur. In other words, she is told to sort the outcomes within *O* into those with above threshold probabilities $\{O_1 \ldots O_G\}$, and those with below-threshold probabilities $\{O_{G+1} \ldots O_H)$, and to calculate the following for each A_i :

$$(1/\sum_{h=1}^{G} P(O_h|A_i)) \times \left[\sum_{z \in O_1} P(z|A_i)s(z) + \ldots + \sum_{z \in O_G} P(z|A_i)s(z) \right].$$

[17] If there are many outcomes with below-threshold probabilities, this calculation may be less expensive than the first calculation.

"Individual risk" tests can also economize on decision costs.[18] In regulatory practice, "individual risk" tests sometimes function as initial screening devices. In other

[17] The first term is a scaling factor, increasing the probabilities of outcomes O_1 through O_G so they add to 1.

[18] An analogy can be drawn to the de minimis rule in tort law. The rule (which involves de minimis harms rather than de minimis probabilities) bars compensation for such harms. The decision cost rationale would be that, whatever the moral reason to provide compensation for harm (the deterrence of risky activity, corrective justice, etc.), that reason is small in the case of de minimis harms and is outweighed by litigation costs.

words, the decisionmaker engages in a full analysis only if the "individual risk" to the relevant individual is non-de minimis, and otherwise undertakes a simpler analysis. This bifurcated procedure can reduce decision costs if the cost of full analysis exceeds the cost of the "individual risk" determination plus the cost of the simpler analysis.

To see how this might occur within welfare consequentialism, consider the exemplary case of a bifurcated procedure that ranks actions with respect to regulatory intensity and tells the decisionmaker: (a) given two actions A_i and A_j, if the probability of any given individual getting cancer, conditional on A_i and conditional on A_j, differs by less than some amount *t*, then pick the action with the lower degree of regulatory intensity; otherwise (b) use the expected social welfare formula, $\sum_{z \in Z} P(z|A_i)s(z)$, to choose between A_i and A_j. It is not hard to imagine rules for ranking actions with respect to regulatory intensity with substantially lower decision costs than the $\sum_{z \in Z} P(z|A_i)s(z)$ formula -- for example, a rule that says inaction is less intense than the choice of issuing a regulation or other governmental directive, and that (as between two directives) the directive with lower expected compliance costs counts as less intense. If the decision cost of the "individual risk" determination plus the decision cost of the regulatory intensity determination is less than the decision cost of applying the expected social welfare formula, $\sum_{z \in Z} P(z|A_i)s(z)$, then the bifurcated approach reduces decision costs in those instances where the "individual risks" are de minimis; and the approach can reduce decision costs overall if those instances are sufficiently frequent and the "regulatory intensity" determination is sufficiently cheaper than simple expected welfare maximization.

The fact that de minimis tests can economize on decision costs provides a *potential* justification for these tests, within the context of a non-ideal version of welfare consequentialism that is addressed to boundedly rational agents. The qualifier "potential" must be stressed. It is surely not the case that *every* de minimis test -- and more generally, every procedure that economizes on decision costs -- is morally justified, given bounded rationality. For example, the decisionmaker might adopt a policy of making each and every choice presented to her by flipping a coin. That procedure requires no characterization at all of the possible outcomes of any choice, or of their probabilities, but presumably goes too far in reducing the expense of decisionmaking. Or -- to use a de minimis example -- the decisionmaker might use an "individual risk" cutoff that is much too high. (Imagine a rule that says to refrain from further consideration of possible policies to mitigate a hazard unless it imposes an individual risk of cancer on some individual exceeding 1 in 10.) A different example: if the rule directs the decisionmaker to calculate *s* for a given policy by ignoring any low-probability outcome within some partition, but the cumulative probability of all such outcomes is non-trivial, and the *s*-value of these outcomes tends to be different from the *s*-value of high-probability outcomes, the upshot will be a substantially different estimate of a policy's expected *s*-value than if the decisionmaker had considered all occurrences. (Imagine that the decisionmaker is told to evaluate a hazard by using only the models with widespread support in the community of experts, and those models are conservative or anti-conservative relative to less widely supported models.)

The trick, then, is to sort between justified and unjustified de minimis tests. Which tests is a boundedly rational decisionmaker morally justified in employing -- within the framework of welfare consequentialism -- given the presence and level of decision costs, and, given the tests' relatively accuracy in mimicking what a fuller social welfare analysis would conclude? Which tests would it be morally incorrect for the decisionmaker to employ? I suggest that welfare consequentialism, in its current state of scholarly development, lacks a workable methodology for answering these sorts of questions. It lacks a workable account for specifying appropriate heuristics, given positive decision costs.

This claim may seem surprising, given the now-voluminous literature in economics and psychology concerning "heuristics" and biases." (Hastie & Dawes, 2001). But this literature is largely positive, not normative -- concerned to show that individual decisionmakers do in fact fail to conform to the tenets of expected utility theory, rather than to develop *norms* for decisionmakers whose cognitive capacities are limited. And the literature on decision theory, which *is* normative, has failed to develop any normative account of boundedly rational decisionmaking that has wide scholarly support. (Rubinstein 1998, 2-3).

First, welfare consequentialism lacks a workable methodology for identifying an *appropriately* simplified social welfare function *s* and appropriately simplified set of causal models. In a given choice situation, the decisionmaker can use some *s* and some set of models, generating a set *Z* containing possible worlds specified with respect to the arguments of *s* and the attributes relevant to those models, and maximizing $\sum_{z \in Z} P(z|A_i)s(z)$; or, she can use *s** and a different set of models, generating a set *Z** containing possible worlds specified with respect to the arguments of *s** and the attributes relevant to *those* models, and maximizing $\sum_{z \in Z^*} P(z|A_i)s^*(z)$. Which structure, *s-Z*, or *s*-Z**, ought she to use? The intuitive answer is that she should use the structure which maximizes expected social welfare. After all, we want a structure which *optimally* economizes on decision costs. But what, exactly, does that mean? It cannot be that the decisionmaker should employ the MESW procedure at the initial stage of picking between *s-Z* and *s*-Z**. MESW is not usable by boundedly rational actors. So it must mean that the decisionmaker should use some simplified social welfare function *g*, and some simplified modeling structure, in making the initial choice between *s-Z* and *s*-Z**. But then the problem arises of justifying *that* choice. Why *g* and those models, rather than *g** and a different set?

Second, welfare consequentialism lacks a workable methodology for identifying appropriate short-cuts given an *s-Z* structure. Assume that a particular *s-Z* structure is justified for an agent, in some choice situation $\mathbf{A}=\{A_1 \ldots A_I\}$. *If* she should engage in a full social welfare calculation, then (let us assume), she should employ the $\sum_{z \in Z} P(z|A_i)s(z)$ formula. *But* the decision costs of using that formula are high, and (it seems) the decisionmaker will sometimes be justified in truncating that calculation in some way -- for example through a de minimis test that ignores low-probability outcomes, or a bifurcated procedure that triggers the full $\sum_{z \in Z} P(z|A_i)s(z)$ formula only if someone's "individual risk" is above a threshold. Welfare consequentialists presumably want to say that some truncated procedure is justified, as compared to the procedure of undertaking a full social calculus of **A** with the $\sum_{z \in Z} P(z|A_i)s(z)$ formula, if the truncated procedure sufficiently economizes on decision costs to increase expected social welfare. But what, exactly, does *that* mean? The decisionmaker faces a higher-level choice between two options: (*I*) "Use the $\sum_{z \in Z} P(z|A_i)s(z)$ formula to choose among {A_1 … A_I}" and (*II*) "Use the truncated procedure to choose among {A_1 … A_I}." It is *possible* for the decisionmaker to make this higher-level choice using the $\sum_{z \in Z} P(z|A_i)s(z)$ formula itself. To do that, she applies the formula to the higher-level choices *I* and *II*. That is, she determines $\sum_{z \in Z} P(z|I)s(z)$ and $\sum_{z \in Z} P(z|II)s(z)$. But the decision costs of these calculations are *greater* than the decision cost of simply applying the $\sum_{z \in Z} P(z|A_i)s(z)$ formula to the choice among {A_1 … A_I}. To see why, note that to determine $\sum_{z \in Z} P(z|I)s(z)$, the decisionmaker needs to (a) determine which choice *A**, out of {A_1 … A_I}, the higher-level choice of *I* would lead to; and then to (b) calculate $\sum_{z \in Z} P(z|I \cap A^*)s(z)$. The formula $\sum_{z \in Z} P(z|I \cap A^*)s(z)$ "builds in" the decision costs of the expected-social-welfare formula. It expresses the expected social welfare of performing action *A** *after choosing from* {A_1 … A_I} *by using the expected-social-welfare formula*, with whatever decision costs that may involve. The difficulty is that the higher-level decisionmaker needs to identify the choice *A** that the higher-level choice *I* would lead to -- and to do *that* she needs to apply the $\sum_{z \in Z} P(z|A_i)s(z)$ formula to {A_1 … A_I}. The procedure of applying the $\sum_{z \in Z} P(z|A_i)s(z)$ formula to the *I-II* choice includes, as one part, the procedure of applying the $\sum_{z \in Z} P(z|A_i)s(z)$ formula to the choice among{A_1 … A_I} -- and thus has greater decision costs than that latter procedure.

To see the point another way, imagine that the boundedly-rational decisionmaker had an unboundedly rational "adviser" who could give her advice on a single issue: whether to choose *I* or *II*. The "adviser" surely *should* make that choice by undertaking a full social welfare calculation. In particular, he should apply MESW to the *I-II* choice.

But the boundedly-rational actor herself would be irrational to use MESW or $\sum_{z \in Z} P(z|A_i)s(z)$ to choose between *I* and *II*. If a full social welfare calculus has positive decision costs for some decisionmaker, it is irrational for her to decide whether to incur those costs, or employ a cheaper, truncated procedure, by engaging in a full social welfare calculus of that higher-order choice. Yet decision theory in its current state provides no clear guidance as to how else the boundedly rational actor should resolve the higher-order choice.

The so-called "value of information" (VOI) framework, although certainly a powerful tool, is not helpful in answering this question. (Hirshleifer & Riley 1992, 167-208; Winkler 2003, 267-350). VOI recognizes that the physical actions required to gather information (such as conducting experiments and observations) may be costly. It says to evaluate possible such measures in the same way that any actions are evaluated -- by determining which action maximizes the expected value of the relevant objective function. In the case of welfare consequentialism, the VOI framework says this: if a decisionmaker is faced with a first-order choice among $\mathbf{A} = \{A_1 \ldots A_I\}$, and is using some *s-Z* structure to evaluate that choice, and is trying to decide whether to undertake some information gathering measure *g*, with possible outputs $G_1 \ldots G_N$, then she should (1) determine the choice A_j* that each possible output G_j would induce; (2) calculate $P(G_1|g)\sum_{z \in Z} P(z|A_1^* \cap G_1 \cap g)s(z) + \ldots + P(G_N|g)\sum_{z \in Z} P(z|A_N^* \cap G_N \cap g)s(z)$[19]; and (3) determine whether that value is greater than the expected value of gathering no information, i.e., the maximum value of $\sum_{z \in Z} P(z|A_i)s(z)$ for A_i in **A.** Doing these calculations requires applying the $\sum_{z \in Z} P(z|A_i)s(z)$ formula to the choice among $\{A_1 \ldots A_I\}$ multiple times. For each possible informational output G_j, the decisionmaker identifies the matching A_j* by determining which choice maximizes $\sum_{z \in Z} P(z|A_i)s(z)$

---

[19] $A_j^* \cap G_j \cap g$ is the proposition that the information gathering measure *g* is undertaken, G_j results, and A_j* is chosen. The costs of information gathering as well as the change in the probabilities of the various possible worlds in *Z* when the decisionmaker updates on G_j are reflected in the social welfare calculus by conditioning on this proposition.

given G_j. Thus, the VOI framework is only helpful advice when the decision costs of applying the $\sum_{z \in Z} P(z|A_i)s(z)$ formula are not too high. It is not helpful when the limitations of bounded rationality come into play -- when the costs of *analyzing the expected value of information* become substantial. VOI, as currently developed in the literature on decision theory, simply does not speak to that case. It does not attempt to furnish a truncated decision procedure, as distinct from full expected value maximization, to evaluate informational or other actions.

B. *Law Versus Morality*

Ideal moral theory attempts to identify the norms that unboundedly rational decisionmakers are morally required to follow. That enterprise is not the same as identifying the rules that actual humans should be instructed to follow. In particular, it is not the same as identifying the morally optimal legal rules for some legal system. Law and morality can diverge. (Adler & Posner 2006, 64-65). A given moral theory, such as welfare consequentialism, health consequentialism, contractarianism, or deontology, might require legislators to enact some set of legal rules, even though those legal rules, in some choice situations, require the subjects to act in ways that conflict with the norms of that very same moral theory. The potential for a wedge between law and morality arises (at least in part) because of humans' *motivational* limitations. It may not be morally optimal to enact into law a set of legal rules that mirror moral norms but that actual humans, given self-interest, moral apathy, and so forth, will evade.

The wedge between law and morality creates a second, potential argument for de minimis tests, distinct from the bounded-rationality argument considered in the previous Section. The bounded-rationality argument says: Morality requires actual government officials, given their cognitive limits, to employ de minimis tests in some cases. The argument now under consideration says: Although actual government officials, given their cognitive limits, may not be morally required to employ de minimis tests, the best legal rules require them to do so (in some cases).

Consider, by way of analogy, the arguments against administrative cost-benefit analysis advanced by some legal scholars. (Adler & Posner 2006, 62-123). While one argument points to decision costs, a different argument points to bureaucratic slippage. Whatever the appropriateness of cost-benefit analysis in the abstract, it is (allegedly) too malleable a standard to constrain bureaucrats who will be motivated to distort cost-benefit analysis to serve other goals. Bureaucracies tend to develop their own conception of the public interest, as a result of bureaucrats' educational background and bureaucratic "culture." Bureaucracies also have a powerful incentive to advance the interests of regulated firms, either because those firms have direct influence at the agency, or because they influence the political actors at the White House and in Congress who oversee agencies. Cost-benefit analysis insufficiently constrains bureaucrats who are motivated in morally problematic ways (be it to advance a misguided conception of the public good, or to help powerful groups). It is better to put in place a clearer and less manipulable rule -- or so the argument goes.

While this is a *potential* line of argument for de minimis tests, I am not sure that it is particularly promising. We are to imagine actual governmental officials making some choice $\mathbf{A}=\{A_1 \ldots A_I\}$. Our moral theory *M*, even taking account of decision costs and the officials' bounded rationality, does not require them to employ a de minimis test in choosing among $\{A_1 \ldots A_I\}$. However, because the actors are not perfectly motivated to comply with *M*, *M* instructs legislators to enact rules that require the officials to employ a de minimis test. The difficulty, here, is understanding why de minimis tests, in particular, are a useful device to constrain bureaucratic slippage. If *M* justifies legal rules that deviate from *M* itself, it presumably does so because violations of those legal rules are easier for oversight bodies (such as courts) to monitor and punish. But monitoring agency compliance with de minimis tests is, itself, no trivial task.

As discussed, de minimis tests can be framed in terms of qualitative or quantitative probabilities. Qualitative de minimis tests would use fuzzy terms like "reasonably certain," "highly unlikely," or "minimal." If one is of the view that qualitative probability terms map onto ranges of quantitative probabilities with precise boundaries, then the fuzziness is a matter of ambiguity --- ordinary English speakers do not have a shared, conscious, sense of the ranges referred to by a given qualitative term. If one is of the view that qualitative probability terms are vague in the technical sense -- they map onto ranges of quantitative probabilities with imprecise boundaries -- then the fuzziness is ineliminable. Speakers with a perfect grasp of the meaning of the qualitative terms, and of the quantitative probabilities of various outcomes, might still be unsure whether to ascribe a qualitative probability term to certain of these outcomes (those on the borderline of the range referred to by the term). In either event, qualitative probability terms would not seem to be powerful tools for monitoring bureaucrats.

As for de minimis tests framed in terms of quantitative probabilities, such as the 1×10^{-6} cutoff: the application of these tests requires a risk assessment, i.e., a mathematical analysis to determine what the quantitative probability (epistemic or frequentist) of the relevant outcome is. The correctness of the risk assessments that governmental agencies undertake is itself a regular topic of controversy. There is no good evidence that risk assessment is less malleable than, say, cost-benefit analysis. Monitoring whether an agency has correctly performed a probability assessment is orders of magnitude more difficult than monitoring whether a car is going 55 miles per hour, or whether the President is 35 -- the classic examples of non-manipulable rules.

V. Conclusion

De minimis tests are problematic as a matter of ideal moral theory. This includes not merely quantitative "individual risk" tests, but a much wider range of decisional criteria that ask whether the probability of some outcome is below a low threshold. And the difficulties for de minimis tests arise within a range of moral views, not merely specific and controversial views such as utilitarianism.

De minimis tests are justified, if at all, on non-ideal grounds -- specifically, as heuristic techniques, responsive to the bounded rationality of government decisionmakers. Unfortunately, how exactly a non-ideal analysis of de minimis tests should proceed is far from clear, since we lack a well-developed normative account of choice under bounded rationality. It seem plausible that some de minimis tests are justified; but it is not apparent which they are, or even how we should go about identifying them.

References

Adler, H., & Weinberg, A. (1978). An approach to setting radiation standards. *Health Physics* 34, 719-20.

Adler, M. (2005). Against "individual risk": A sympathetic critique of risk assessment. *University of Pennsylvania Law Review* 153, 1121-1250.

Adler, M. (2006). Policy analysis for natural hazards: Some cautionary lessons from environmental policy analysis. *Duke Law Journal* 56, 1-50.

Adler, M., & Posner, E. (2006). *New Foundations of Cost-Benefit Analysis*. Cambridge, MA.: Harvard University Press.

Adler, M., & Sanchirico, C. (2006). Inequality and uncertainty: Theory and legal applications. *University of Pennsylvania Law Review* 155, 279-377.

Brock, W., Rodricks, J., Rulis, A., Dellarco, V., Gray, G., & Lane, R. (2003). Food safety: Risk assessment methodology and decision-making criteria. *International Journal of Toxicology* 22, 435-41.

Cross, F., Byrd, D., & Lave, L. (1991). Discernible risk--A proposed standard for significant risk in carcinogen regulation. *Administrative Law Review* 43, 61-88.

Dolan, P. (1998). The measurement of individual utility and social welfare. *Journal of Health Economics* 17, 39.

Eells, E. (1982). *Rational Decision and Causality*. Cambridge: Cambridge University Press.

Fletcher, G. (1972). Fairness and utility in tort theory. *Harvard Law Review* 85, 537-573.

Harsanyi, J. (1982). Morality and the theory of rational behavior. In Sen, A. & Williams, B. eds., *Utilitarianism and Beyond*, 39-62. Cambridge: Cambridge University Press.

Hastie, R. & Dawes, R. (2001). *Rational Choice in an Uncertain World.* Thousand Oaks, CA: Sage Publications.

Hirshleifer, J. & Riley, J. (1992). *The Analytics of Uncertainty and Information*. Cambridge: Cambridge University Press.

Interagency Floodplain Management Review Committee (1994). *Sharing the Challenge: Floodplain Management into the 21st Century*. Washington, D.C.: U.S. Govt. Printing Office.

Jackson, F., & Smith, M. (2006). Absolutist moral theories and uncertainty. *Journal of Philosophy* 103, 267-83.

Joyce, J. & Gibbard, A. (1998) Causal decision theory. In Barbera, S. ed., *Handbook of Utility Theory*, vol. 1, 627-66. Dordrecht: Kluwer.

Kagan, S. (1998). *Normative Ethics.* Boulder, CO: Westview.

Kopelman, L. (2004). Minimal risk as an international ethical standard in research. *Journal of Medicine and Philosophy* 29, 351-78.

Lewis, D. (1981). Causal decision theory. *Australasian Journal of Philosophy* 59, 5-30.

Lindholm, L. & Rosen, M. (1998). On the measurement of the nation's equity-adjusted health. *Health Economics* 7, 621.

Loux, M. (1998). *Metaphysics: A Contemporary Introduction*. London: Routledge.

McCarthy, D. (1996). Liability and risk. *Philosophy and Public Affairs* 25, 238-62.

McCarthy, D. (1997). Rights, explanation, and risks. *Ethics* 107, 205-25.

Nozick, R. (1974). *Anarchy, State, and Utopia*. New York: Basic Books.

Nuclear Regulatory Commission (1986). Safety goals for the operation of nuclear power plants. *Federal Register* 51, 30,028-29.

Oberdiek, J. (2003). *The Morality of Risking: On The Normative Foundations of Risk Regulation*. Ph.D. Dissertation, University of Pennsylvania.

Osterdal, L. (2005). Axioms for health care resource allocation. *Journal of Health Economics* 24, 679.

Perry, S. (undated). Imposing risk. Unpublished paper, on file with author.

Rawls, J. (1971) *A Theory of Justice*. Cambridge, MA: Belknap Press.

Resnik, M. (1987). *Choices: An Introduction to Decision Theory*. Minneapolis: University of Minnesota Press.

Rodricks, J., Starr, T., & Taylor, M. (1991). Evaluating the safety of carcinogens in food--current practices and emerging developments. *Food, Drug, Cosmetic Law Journal* 46, 513-52.

Rubinstein, A. (1998). *Modeling Bounded Rationality*. Cambridge, MA: MIT Press.

Sandin, P. (2005). Naturalness and *de minimis* risk. *Environmental Ethics* 27, 191-200.

Scanlon, T.M. (1998). *What We Owe to Each Other*. Cambridge, MA: Belknap Press.

Shrader-Frechette, F. (1985). Techological risk and small probabilities. *Journal of Business Ethics* 4, 431-445.

Simon, H. (1982). *Models of Bounded Rationality*. Vols. 1-2. Cambridge, MA: MIT Press.

Simon, H. (1997). *Models of Bounded Rationality*. Vol. 3. Cambridge, MA: MIT Press.

Thomson, J. (1990). *The Realm of Rights*. Cambridge, MA: Harvard University Press.

Travis, C., Richter, S., Crouch, E., Wilson, R., & Klema, E. (1987). Cancer risk management. *Environmental Science and Technology* 21, 415-20.

Tuomala, Matti. (1990). *Optimal Income Tax and Redistribution*. Oxford: Clarendon Press.

Whipple, C., ed. (1987). *De Minimis Risk*. New York & London: Plenum.

Williams, A. (1997). Intergenerational equity: An exploration of the "fair innings" argument. *Health Economics* 6, 117 (1997).

Winkler, R. (2003). *An Introduction to Bayesian Inference and Decision*. 2d ed. Gainseville, FL: Probabilistic Publishing.