Nuclear Security Science
& Policy Institute

A Division of the
Texas Engineering Experiment Station

NSSPI-09-024

NUCLEAR DISARMAMENT:

CAN RISK ANALYSIS INFORM THE DEBATE?

Paul Nelson
Texas A&M University
College Station, TX 77843

December 14, 2009

This report is a preprint of a manuscript that has been submitted for possible journal publication. It should not be quoted or referenced without contacting the author to ensure it remains current. The author may be contacted at penelson@ne.tamu.edu, or +1-979-845-4132.

EXECUTIVE SUMMARY

The possibility of adapting risk analysis to events deemed to have infinite consequences is illustrated in the context of a simple dynamic two-event model. The results are employed in an effort to illuminate the divergent opinions regarding the desirability of nuclear disarmament. Risk functions (disutility functions to be minimized) considered include discounted expected risk (discount rate = γ), maximum risk rate encountered at any time, and long-term (asymptotic) risk rate. In the case of finite consequences (i.e., consequences of conventional war are considered commensurate with catastrophes occasioned by nuclear weapons), the conclusion that nuclear disarmament is preferable to the status quo rests partially on the *conventional nondeterrence* proposition. This proposition is expressed mathematically as

$$C_1\tilde{\lambda}_1^{(a)} \geq C_1\tilde{\lambda}_1^{(d)},$$

where C_1 is the consequence of conventional war, and $\tilde{\lambda}_1^{(a)}, \tilde{\lambda}_1^{(d)}$ are respectively the rates of occurrence of conventional conflicts under the status quo and a regime of nuclear disarmament. This condition is a mathematical representation of the belief that nuclear arms do not deter conventional warfare sufficiently to offset the risk, as defined by the specified risk (disutility) function, posed by the nuclear arms *per se*.

Under the same conditions a preference for nuclear disarmament is associated with some form of a benign disarmament proposition, which represents mathematically a belief that the two distinct risks associated with first achieving and second maintaining a global condition of nuclear disarmament are not sufficient, per the chosen risk function, to counter that associated with maintaining the nuclear-armed status quo. Under reasonable modification of the objective, a preference for nuclear disarmament under the assignment of infinite consequences to catastrophes associated with nuclear weapons depends *only* on some form of a benign disarmament principle.

However, several mathematically different forms of benign disarmament appear. All can be written in the form

$$\tilde{\lambda}_2^{(a)} > \tilde{\lambda}_2^{(d)} + w\Lambda_0,$$

where $\tilde{\lambda}_2^{(a)}, \tilde{\lambda}_2^{(d)}$ are respectively the rates of occurrence of nuclear-weapon catastrophes under the status quo and a regime of nuclear disarmament, Λ_0 is the probability of a nuclear-weapon catastrophe at the time of (idealized as instantaneous) disarmament, and w is the weight assigned to the threat of nuclear-weapons catastrophe at disarmament. The different forms of the benign disarmament proposition correspond as follows to different values of this weight: impossible ($\lim_{\Lambda_0 \to 0+} \lim_{w\to\infty}$), strong ($\lim_{w\to\infty} \lim_{\Lambda_0 \to 0+}$), gamma ($w = \gamma$), lambda ($w = \lambda_2^{(a)}$) and weak ($w = 0$).

The table given below displays the form of the benign disarmament proposition appropriate to the various disutility functions and associated parametric values considered. As the impossible form implies the strong form implies either of the gamma or lambda forms, and either of the latter two imply the weak form, it follows that:

- *The strongest case for the status quo arises from either discounted expected risk in the limit of arbitrarily large discount rate* (always) *or the maximum risk rate disutility function* (strong form of benign disarmament)*, according respectively as the consequences of conventional war are or are not considered commensurate to those of nuclear-weapons catastrophes.*

Table: Versions of the benign disarmament proposition relevant to different circumstances

	finite C_i, i=1,2	$C_2=\infty$
discounted expected risk, $0<\gamma<\infty$ (γ=discount rate)	gamma	lambda
discounted expected risk, $\gamma\to\infty$	impossible	lambda
discounted expected risk, $\gamma\to 0$	weak	lambda
maximum risk rate	strong	strong
long term	weak form	weak form

- *The strongest case for nuclear disarmament arises from the long-term disutility function* (only the weak form of benign disarmament is required, regardless of whether the consequences of a nuclear-weapons catastrophe are taken as finite or infinite).

Thus conclusions favoring the status quo or disarmament correlate respectively with a bias toward the current generation, or a decision to give generations however far into the future the same weight as that now alive.

TABLE OF CONTENTS

1. INTRODUCTION

Risk analysis is an instance of "science-based" methods[1] for deciding between various possible alternatives. In the realm of public policy this class of approaches is not universally accepted, with rules based on moral intuition often being suggested as an alternative approach to decision making.[2] This note seeks to explore the possibility of finding some common ground between these views by providing for the possibility of consequences having infinite value. The considerations are illustrated primarily through a simple two-alternative model, initially with time-discounted expected risk as the associated disutility function. In order to provide a concrete example this model is employed as a prism through which to view and understand various widely divergent viewpoints of nuclear disarmament.[3] Extensions to alternative forms of disutility function are considered, as is (briefly) multiple-alternative situations. It is concluded that providing for events having infinite consequences has some potential for employing risk analysis to better understand the origins of differing viewpoints on public policy.

Consider two alternative types of untoward events, indexed as $i = 1,2$. Suppose the occurrence of each is governed by a Markov model with (possibly time-dependent) transition rate $\lambda_i(t)$, and with instantaneous recovery. Let C_i be the consequence associated to events of type i. Then the corresponding expected risk, discounted from present time $t = 0$, is

$$R(\lambda_1,\lambda_2) = \int_0^\infty e^{-\gamma t}\left[C_1\lambda_1(t) + C_2\lambda_2(t)\right] dt, \tag{1}$$

where γ is the discount rate.[4]

Given two different "rate profiles," say $\left(\lambda_1^{(a)},\lambda_2^{(a)}\right)$ and $\left(\lambda_1^{(d)},\lambda_2^{(d)}\right)$ one can in principle select the more desirable as that having the smaller value of the associated time-discounted expected risks,[5]

$$R_a = R\left(\lambda_1^{(a)},\lambda_2^{(a)}\right) \text{ vs. } R_d = R\left(\lambda_1^{(d)},\lambda_2^{(d)}\right).$$

---

[1] For example, Byrd and Cothern (2000). See Short (1984) for a general critique of science-based methodologies.

[2] This debate has been framed in terms of consequentialists who desire to make decisions based on their consequences, and nonconsequentialists who prefer judgments otherwise based, often including a preference for inaction ("omission bias," defined as "people are more willing to cause harmful outcomes through their omissions than through their acts"; cf. Baron, 1993). See Parfit (1984, esp. Section 10) for objections to consequentialism.

[3] There have been many recent discussions of the pros and cons of nuclear disarmament. This note was much motivated by a recent article by Fred C. Iklé (2009), and an ensuing exchange of views (Kampelman, 2009; Blechman, 2009; and Iklé, 2009a).

[4] Jason Matheny (2007, esp. Section 5) argues that many of the traditional justifications for discounting are inapplicable to the "doomsday scenarios" (i.e., to the "benefits of delaying human extinction") that are integral to the present note. The use of discounting here is justified by the rather more technical and pragmatic desire to have some hope of finite time-integrated risks for the apparently reasonable case that the two components of the rate profile are positive constants and the time horizon is infinite. Other alternatives are to impose some arbitrary finite time horizon, or to require all rate-profile components to approach zero asymptotically sufficiently strongly so as to be (finitely) integrable. Among these alternatives the use of discounting seems the least arbitrary, although in practice the choice of a discount rate - which is unnecessary for present purposes - is quite troublesome. Richard Posner (2004, p. 17) seems to argue for the alternative of a finite time horizon that is somewhere between thousands and millions of years. See also Chapter 3 of Posner (2004) for a more detailed discussion of some of the alternatives mentioned above, and extensive references to further discussion in the literature. See Gollier and Weitzmann (2009) for an argument, in the context of climate change, favoring use of a discount rate that varies with time in a particular manner.

[5] Alternatives to the time-discounted expected risk as the disutility function to be minimized are briefly considered in Section 4.

This rather naïve depiction masks a number of well-known and widely discussed difficulties that affect any actual implementation of such an approach, such as obtaining an evaluation of consequences, and uncertainty in the frequencies. These matters are especially acute for the high-consequence low-probability threats that motivate the present work, but here we set them aside in order to focus on a different difficulty.

The difficulty of interest here arises when one of the threats, say the second, is assigned the value infinity,[6] $C_2 = \infty$. In such a case then the associated discounted expected risk is also infinite, if only the second component of the rate profile (i.e., λ_2) is not identically zero.[7] If neither $\lambda_2^{(a)}$ nor $\lambda_2^{(d)}$ is identically zero, then both of the corresponding discounted expected risks are infinite, and therefore provide no basis for discrimination between the two rate profiles.[8]

Even though the discounted expected risk provides no basis for discrimination, there is a rational basis. One can reasonably prefer the rate profile that has the largest expected time to "doomsday,"

$$T(\lambda_2) = \int_0^\infty t\lambda_2(t) e^{-\int_0^t \lambda_2(\tau)d\tau}\, dt. \tag{2}$$

In case $T\left(\lambda_2^{(a)}\right) = T\left(\lambda_2^{(d)}\right)$, one could reasonably "break the tie" by choosing the rate profile having the discounted risk with smallest "finite part,"

$$R_1\left(\lambda_1\right) = C_1 \int_0^\infty e^{-\gamma t} \lambda_1(t)\, dt.$$

Other alternatives for an algorithm that chooses a preferred rate profile on the basis of the values of $T(\lambda_2)$ and $R_1(\lambda_1)$ also are possible.[9]

---

[6] Valuation of consequences always is a subjective (and therefore troublesome) aspect of risk analysis, especially in the cost/benefit version. When the value infinity should be assigned a threat is, if anything, even more troublesome, as will be exemplified in the ensuing discussion of nuclear disarmament. The one class of threats that most indubitably deserve such assignment is the doomsday events that have as their physical consequence the extinction of the human species. See Matheny (2007, esp. p. 1342) for a very explicit argument (attributed to Derek Parfit) that there is a bigger qualitative difference between the doomsday of a nuclear war killing 100% of the world's existing population and a nuclear war killing 99% than there is between the latter ("merely"?) catastrophic event and peace. Consideration of events with infinite consequences is by no means unique to the present work. For example such matters have been explicitly considered by Manson (1999, 2002), under the term "catastrophe argument." He regards this argument as an extreme instance of the "precautionary principle" that has found its way into some legal arguments (especially relating to environmental matters, including concerns related to climate change; *cf.* Gollier, 2001), and sees "a strong similarity between the Catastrophe Argument and one of the most famous arguments of philosophy: Pascal's Wager." See Bostrom and Ćirković (2008) for discussion by various authors on both specific possible doomsday scenarios (under the term "global catastrophic risks"), methodologies for response and a variety of related perspectives. Note also that at sufficiently large consequences the assumption of instantaneous recovery that underlies the present model necessarily becomes invalid. This is certainly the case for doomsday events. Recovery from events having infinite consequences must be considered as impossible, which is the most extreme form of recovery that is not instantaneous.

[7] Technically, is not almost everywhere zero on $[0,\infty)$.

[8] The subtlety associated to the mathematical indeterminacy arising when p_2 is identically zero also will not be considered here.

[9] The temptation to "tradeoff" between T and R_1 is especially strong if both $T^{(a)}$ and $T^{(a)}$ are large compared to a typical human lifespan, their difference is small compared to the uncertainty in either, and the larger of $R_1^{(a)}, R_1^{(d)}$ is associated with a (now time-dependent) consequence that is largest in the near-term. The current discussion of possible actions attendant to the threat of anthropogenic global warming seems such an example. Tonn (2009) suggests criteria for

In the following we employ this simple model in a manner that seems to capture many of the essential considerations underlying the present debate regarding the desirability of global nuclear disarmament. In more detail, in Section 2 we define two rate profiles corresponding respectively to maintenance of the status quo relative to nuclear weapons, and to disarmament at some specified future timed carry out the details of the calculations outlined above, as based on discounted expected risk, and for the case of finite consequences. In Section 3 these results are extended to the case of an infinite consequence for nuclear-weapons catastrophes. Section 4 is devoted to the corresponding developments for two different risk (disutility) functions: The "maximum risk rate" disutility function, and the "long-term risk" disutility function, both of which are related to arguments that have appeared in the literature.

In all of these cases the mathematical conditions that would imply nuclear disarmament is preferred to the status quo are found to have a degree of commonality. They involve first a "conventional nondeterrence" proposition, which is an algebraic reflection of the belief that nuclear arms do not deter conventional warfare sufficiently to offset the risk, as defined by the specified risk (disutility) function, posed by the nuclear arms *per se*. Second they involve some form of a "benign disarmament" proposition, which is a mathematical representation of the belief that the (two separate discounted expected) risks in achieving and maintaining a global condition of nuclear disarmament are not sufficient to counter that associated with maintaining the nuclear-armed status quo.

The conventional nondeterrence proposition of course occurs only in the cases that the consequences are taken as finite, but it has the same mathematical form in all cases that it does occur. By contrast some form of the benign disarmament proposition appears in all case, but a total of four different (achievable) algebraic forms of that proposition occur: weak form, gamma form, lambda form and strong form. The weak form is implied by all of the others, and is therefore the easiest to believe in. The strong form implies all of the others, and is therefore the most difficult to argue the validity of. (It in fact holds only under very restrictive circumstances.) The gamma and lambda forms are not comparable, except that they are identical if the discount rate is taken as the rate at which nuclear-weapons catastrophes occur.

In Section 5 an extension of these considerations to scenarios involving a continuum of consequences is briefly sketched. Conclusions are briefly summarized in the closing Section 6. These include a table summarizing the results describe in the two preceding paragraphs. This table demonstrates that:

- *The strongest case for the status quo arises from either discounted expected risk in the limit of arbitrarily large discount rate* (always) *or the maximum risk rate disutility function* (strong form of benign disarmament)*, according respectively as the consequences of conventional war are or are not considered commensurate to those of nuclear-weapons catastrophes.*
- *The strongest case for nuclear disarmament arises from the long-term disutility function* (only the weak form of benign disarmament is required, regardless of whether the consequences of a nuclear-weapons catastrophe are taken as finite or infinite).

Thus conclusions favoring the status quo or disarmament correlate respectively with a strong bias toward the current generation, or a decision to give generations however far into the future the same weight as that now alive.

2. USE OF NUCLEAR WEAPONS AS A LIMITED CATASTROPHE

Under the status quo case of a nuclear-armed world assume the corresponding rate profile satisfies

$$\lambda_1(t) = \tilde{\lambda}_1^{(a)} \text{ and } \lambda_2(t) = \tilde{\lambda}_2^{(a)},$$

---

acceptability that might be modifiable into decision criteria, but seem to presuppose existence of some accepted time (year 1 billion) for human extinction.

where $\tilde{\lambda}_1^{(a)}$ and $\tilde{\lambda}_2^{(a)}$ are nonnegative constants that represent respectively the corresponding status quo frequencies of occurrence of conventional war or a nuclear-weapon catastrophe of some type. The associated discounted risk of nuclear armament is

$$R^{(a)} := R\left(\lambda_1^{(a)}, \lambda_2^{(a)}\right) = \gamma^{-1}\left[C_1\tilde{\lambda}_1^{(a)} + C_2\tilde{\lambda}_2^{(a)}\right],$$

where the C_i, $i = 1,2,$ are the (finite) consequences of respectively a conventional war or a nuclear-weapon catastrophe.

We take the rate profiles corresponding to nuclear disarmament as

$$\lambda_1^{(d)}(t) = \begin{cases} \tilde{\lambda}_1^{(a)}, t < t_0, \\ \tilde{\lambda}_1^{(d)}, t > t_0, \end{cases}$$

and

$$\lambda_2^{(d)}(t) = \Lambda_0\delta\left(t - t_0\right) + \begin{cases} \tilde{\lambda}_2^{(a)}, t < t_0, \\ \tilde{\lambda}_2^{(d)}, t > t_0. \end{cases}$$

Here t_0 is some "zero time" at which disarmament is assumed to occur instantaneously,[10] $\tilde{\lambda}_1^{(d)}$ and $\tilde{\lambda}_2^{(d)}$ are nonnegative constants that represent respectively the corresponding frequencies of occurrence of conventional war or a nuclear-weapon catastrophe during a state of worldwide nuclear disarmament, δ is the Dirac delta function, and Λ_0 is the expected number of nuclear-weapon catastrophes that occur "at" the disarmament time t_0. The corresponding time-discounted expected risk of nuclear disarmament is

$$\begin{aligned} &R^{(d)} := R\left(\lambda_1^{(d)}, \lambda_2^{(d)}\right) = \\ &\gamma^{-1}\left\{\left[C_1\tilde{\lambda}_1^{(d)} + C_2\tilde{\lambda}_2^{(d)}\right]\left(1 - e^{-\gamma t_0}\right) + \left[C_1\tilde{\lambda}_1^{(d)} + C_2\tilde{\lambda}_2^{(d)}\right]e^{-\gamma t_0}\right\} + C_2\Lambda_0 e^{-\gamma t_0}. \end{aligned}$$

Risk analysis then dictates that disarmament is the preferred course of action if $R^{(a)} > R^{(d)}$, while retaining nuclear arms is indicated if the opposing inequality holds. A bit of algebra reveals that this is equivalent to rationally choosing disarmament or armament according respectively as

$$\gamma^{-1}\left[C_1\tilde{\lambda}_1^{(a)} + C_2\tilde{\lambda}_2^{(a)}\right] > \gamma^{-1}\left[C_1\tilde{\lambda}_1^{(d)} + C_2\tilde{\lambda}_2^{(d)}\right] + \Lambda_0 C_2 \tag{3}$$

---

[10] No serious thinker about this matter expects nuclear disarmament to occur within a duration that is short on the scale of human lifetimes. Indeed some of the deeper thinkers among the advocates (e.g., Perkovich and Acton, 2008) suggest a great deal of preparatory work will be required to develop the necessary underlying international agreements, and the corollary technical means necessary to enforce those agreement. That is essentially because as the total number of nuclear weapons becomes smaller, the temptation a state faces to use its few remaining weapons to accomplish its purposes is perceived to become larger; therefore when the total number becomes very small it is necessary to be able to detect small numbers of (illicitly held) weapons, and possibly even more critically to respond appropriately. The assumption here of an instantaneously occurring disarmament merely simplifies the final results, and therefore the subsequent discussion, while retaining (via the factor Λ_0) the key possibility that there may be a nonzero risk associated with the *step* of disarming, as opposed to the *state of being* disarmed.

or the reverse inequality holds.

The various terms in (3) have simple interpretations. The term on the left is the discounted expected risk of the status quo, as calculated at the contemplated time of disarmament (t_0). The first term on the right has the same interpretation, except for maintenance of disarmament rather than the status quo. The last term on the right is the "instantaneous" risk of attaining disarmament, also as evaluated at the proposed time of disarmament.

This simple criterion seems to permit illustration of many of the arguments advanced in favor of nuclear disarmament, and many against. The most straightforward arguments in favor of nuclear disarmament seem more-or-less to involve two distinct propositions:

1. The *conventional nondeterrence* proposition that the risk per unit time of conventional war is no less with nuclear arms than without, so that

$$C_1\tilde{\lambda}_1^{(a)} \geq C_1\tilde{\lambda}_1^{(d)}, \tag{4}$$

and (3) becomes

$$\gamma^{-1}\tilde{\lambda}_2^{(a)} > \gamma^{-1}\tilde{\lambda}_2^{(d)} + \Lambda_0. \tag{5}$$

2. The *benign disarmament* proposition that the rate of nuclear-weapon catastrophes under nuclear disarmament and the expected number of nuclear-weapon catastrophes at disarmament are (or can be made) so small relative to the rate of such catastrophes while some states legally possess nuclear weapons, that (5) certainly holds.

In the following (5) will be termed as the *gamma form* of benign disarmament, whenever necessary to distinguish it from other versions of a benign disarmament proposition, which have different algebraic representations, but similar interpretations.

Both of the immediately preceding propositions can be questioned, and have been. For example the idea has been offered (Brown, 2007-2008) that nuclear disarmament will tend to increase the frequency of conventional wars (i.e., $\lambda_1^{(d)} > \lambda_1^{(a)}$), which negates (4). The question of the impact of nuclear arms on the frequency of conventional wars has been discussed empirically, seminally by Kugler (1984), who concluded as follows "…there is no evidence that nuclear weapons have added stability to the relation between the three nuclear giants. The terror created by nuclear devastation cannot, in sum, be directly linked to the preservation of peace."

This conclusion seems to support conventional nondeterrence, but more recent empirical work seems to challenge that conclusion. The entire April 2009 issue of the *Journal of Conflict Resolution* consists of further work in this vein that provides at least modest further steps toward interjecting fact-based data into this discussion. The preface by Erik Gartzke and Matthew Kroenig (2009, p. 157) states: "The second set of articles considers the consequences of nuclear proliferation. Taken together, they find that nuclear weapons do not affect the frequency of conflict, but they do affect the timing, duration, severity, and outcome of conflict. These articles provide considerable support for the argument that nuclear weapons enhance the security and diplomatic power of their possessors. Nuclear weapon states are neither more nor less conflict prone, but their conflicts are shorter and less intense, and they tend to emerge victorious from them. Furthermore, the authors find that nuclear powers enjoy enhanced international bargaining power."

While this conclusion is not diametrically opposed to that of Kugler quoted above, it is substantially different in tone, and seems not supportive of the general concept of conventional nondeterrence. Taken together the two different conclusions suggest that while nuclear disarmament might not increase the frequency of conventional war, it might increase the consequences, and thus the associated (discounted expected) risk, which is really what is at issue in the proposition of conventional nondeterrence. The conclusion of Gartzke and Kroenig also perhaps indicates why it might be difficult for the current nuclear-armed states to forgo those weapons. Even though these empirical results might not be conclusive regarding validity of conventional nondeterrence, they arguably should be better known among those prone to qualitative discussion of such matters, if only to temper claims that might someday be subject to empirical validation.

If anything the benign disarmament proposition has been the subject of even greater critical scrutiny. It requires both that the expected number of nuclear-weapon catastrophes during nuclear disarmament be less than the expected number occurring over one discount period,

$$\Lambda_0 < \gamma^{-1}\tilde{\lambda}_2^{(a)} \tag{6}$$

and that the rate at which nuclear-weapon catastrophes occur under nuclear disarmament be less than that at which they occur while nuclear armed,

$$\tilde{\lambda}_2^{(d)} < \tilde{\lambda}_2^{(a)}. \tag{7}$$

Both of these inequalities can be questioned, and have been.

The inequality (6) is contrary to the classical theory of Kenneth Waltz (1981) to the (loose) effect that the more widely nuclear weapons are spread among and within states the more effectively they will deter nuclear warfare (i.e., the smaller $\lambda_2^{(a)}$ will be).[11] The recent tendency is to discount this theory,[12] and argue that (6) holds more now than in the past because terrorists are much more likely to use nuclear weapons than states, as the latter are more subject to deterrence by threat of counterattack.

Yet another reason the Waltzian theory often is minimized in the (usually tacit) context of (6) is the perceived large value of $\lambda_2^{(a)}$ associated with the possibility of an unintentional nuclear detonation. Any discussion of "accidental launches," as for many of the issues discussed here, tends to be contentious if for no other reason than lack of availability of empirical data, at least that are accessible to the public. [13] For example, Iklé (2009) offers "the fact that nuclear weapons have not been used for sixty-four years as reason for "hope." This could be taken as a suggestion that $\lambda_2^{(a)}$ is small, and thus as nonsupportive of benign disarmament. Blechman (2009) counters this with (among other things) the observation that "cases of near uses of nuclear weapons during the Cold War are well known."

---

[11] The concern is that in the process of achieving disarmament it will be necessary to go through circumstances in which only a few states - perhaps only one - have nuclear weapons, which are precisely the circumstances in which the Waltz theory predicts the greatest likelihood of use.

[12] For example Barnaby (2009, p. 10) brushes such considerations aside with a reference to "some misplaced faith in the continued value of nuclear deterrence." Somewhat similarly, if a bit more vehemently, Blechman (2009) suggests it is typical that "opponents of elimination: hyperventilate about the potential risks of a world without nuclear weapons, but shut your eyes to the rising risks of the real, proliferating world in which we live today."

[13] But Sagan (1993) is an excellent publicly available discussion of such matters. However it employs the methodology of case studies, so is largely anecdotal and qualitative. Approaches for quantifying such information are not readily apparent.

If a perfect state of nuclear disarmament could be achieved, the one would have $0 = \lambda_2^{(d)} < \lambda_2^{(a)}$, so that the inequality (7) certainly would be attained. The concern associated to attaining that inequality is the difficulty of detecting and countering efforts to cheat on a disarmament regime. The associated difficulty lies in both the technical challenge of detecting small amounts of nuclear materials that might be either created by an entirely clandestine facility or diverted from a legitimate facility, and the enforcement challenge of having the international community formulate an effective response if such an effort were detected. This difficulty is exacerbated by the fact that any such effort at "breakout" would create precisely a situation (small number of nuclear weapons in the hands of a single state) under which the Waltzian theory already referenced above would predict maximum temptation to use employ nuclear weapons. These considerations have been reinforced by recent events, and mandate that one not naively assume $\tilde{\lambda}_2^{(d)} = 0$.

Perkovich and Acton (2008) very forthrightly acknowledge the difficulties inherent in achieving each of (6) and (7), or more generally benign disarmament. Their objective is not to cast doubt on the ability to attain benign disarmament, but rather to focus upon the steps that might be taken in order to attain it. Somewhat interestingly, although they seem to acknowledge that of the two challenges indicated in the preceding paragraph the foremost is the enforcement challenge, the bulk of their discussion seems focused upon the perhaps more tractable - albeit far from easy - technical challenge of detecting illegal attempts to acquire nuclear weapons.

There is not, and cannot be, any significant empirical evidence regarding the benign disarmament proposition, because nuclear disarmament has never really been tried (as opposed to discussed). Thus it may be the case, as Blechman (2009) suggests, nonsensical to believe the 64 years of disuse of nuclear weapons while in a nuclear-armed regime provides evidence that the rate of nuclear-weapon catastrophes under such a regime is small. However, there is *no* empirical evidence the corresponding numbers for achieving or remaining in a disarmed state (while nuclear weapons are known to be feasible) are smaller, because there is no prior experience of either.

In the limits $\gamma \to 0$ and $\gamma \to \infty$, the benign disarmament proposition becomes simply (7) or $\Lambda_0 < 0$, respectively. That is, in the case that one values future lives equally with current ones the benign disarmament proposition requires only the easy belief that nuclear-weapon catastrophes will be less frequent if nuclear arms are abolished than if they are not;[14] i.e., the consequences of "going to" (as opposed to "being at") a state of disarmament are negligible. For this reason we term (7) as the *weak form* of the benign disarmament proposition. On the other hand, in the alternative case that one values current certain lives sufficiently more highly than uncertain future lives (i.e., $\gamma \to \infty$), the benign disarmament proposition can *never* hold, because Λ_0 can never be negative. Thus a belief in benign disarmament seems to tend toward correlating with a higher value accorded to future generations.

If both the conventional nondeterrence and benign disarmament propositions are false, then nuclear disarmament is not indicated (i.e., the inequality (3) does not hold). It is conceivable that nuclear disarmament could be considered as indicated in the presence of only one of these propositions. Such an argument is most likely to be encountered in the context of an assertion, most likely tacit, that the potential consequences of a nuclear-weapon catastrophe are so much greater than those of a conventional war (i.e.,

---

[14] This assumption raises the question of exactly what it means to "abolish nuclear arms." Under some conceivable definitions of disarmament, some states could remain in a position to reconstitute a nuclear arsenal in a very short time. Perkovich and Acton (2008) correctly devote a great deal of attention to this "latency problem," particularly how to assure current non-weapons states that the weapon states have not retained a short latency, without enhancing the weapons-production capability of those non-weapons states. The seriousness of the latency problem is enhanced by the rather large number of non-weapons states that currently are thought to have latencies on the order of months. Some seem to think the objective of the Iranian nuclear program is to enroll Iran in this club. See Juzaitis and McLaughlin (2008) for a more extensive discussion of the latency problem, and references to the extensive related literature.

$C_2 \gg C_1$) that the benign disarmament proposition alone is sufficient to assure desirability of nuclear disarmament, even without the conventional nondeterrence proposition.

We shall not consider this argument in detail, as in the following section we consider it in the even stronger form of this argument that is associated to the evaluation $C_2 = \infty$. Here we simply note there is some room to question the proposition that $C_2 \gg C_1$. Posner (2004, pp. 71-72) estimates that about 150 million deaths could be directly occasioned by 1000 Hiroshima-sized bombs, as compared to the 40-50 million deaths in World war II. As horrific as either of these events certainly would be, the numbers are not indicative of an orders-of-magnitude difference between the consequences of conventional and nuclear warfare. This would be even more the case if the nuclear threat were expanded to include terrorist-caused events of even improvised nuclear explosives, because the probable lack of technological prowess among such parties will tend to make their nuclear devices less destructive.[15] Such considerations notwithstanding, no less a person than former Secretary of Defense William Perry (2007) has stated that "the greatest danger today is that a terror group will detonate a nuclear bomb in one of our cities."

3. NUCLEAR WEAPONS AS A GLOBAL CATASTROPHE

The discussion of the preceding section contains elements of a rational trade-off study between risks of conventional war versus some type of nuclear-weapon catastrophe, and between the zero-time risk of nuclear-weapon catastrophe as opposed to the long-term risk associated with indefinite retention of nuclear arsenals. But there seem to be voices in the discussion of nuclear disarmament that simply do not brook discussion of any possible merit attaching to retaining nuclear arsenals even one second longer.[16] Here we explore the possibility that this viewpoint of nuclear disarmament as a moral imperative that trumps all contrary indications can be understood in the context of expected risk theory by assigning the evaluation $C_2 = \infty$ to the consequence of a nuclear-weapon catastrophe, in the manner discussed above for the abstract two-threat model underlying (1).

Before proceeding to this exploration it is appropriate to discuss whether there is a basis for the evaluation $C_2 = \infty$, on a scale such that this corresponds to a doomsday threat, as previously discussed. Such a discussion is particularly indicated by the preceding indication of estimates that even a nuclear exchange of 1000 Hiroshima-sized weapons would cause about 150 million deaths. While such an event certainly would be horrific - much more so than for any possible threat emanating directly from an act of "nuclear terrorism" - Posner (2004, p. 71) notes that is only about 2.5% of the world's population, and thus is far short of being a doomsday event.

On the other hand, the total number of nuclear weapons in nuclear arsenals across the world today probably significantly exceeds 1000. Further, some of those are thermonuclear weapons, and thus have destructive capacity far exceeding that employed so destructively at Hiroshima. Posner (2004, p. 74) indicates that it "is less certain" … "the human race would survive" … "an all-out war with hydrogen bombs," which he suggests "could produce consequences similar to that of a major asteroid collision."[17] Further one can certainly envision even a nuclear incident that is relatively minor, on the scale being discussed here, as breaking the "nuclear taboo" that has been theorized,[18] and thereby unleashing an arms race that could lead

---

[15] Cf. Mueller, 2009, esp. Chap. 15, for elaboration on these thoughts.

[16] This reflects what Jervis (2009) has termed "a general stance that put(s) nuclear weapons into a uniquely dangerous and indeed immoral category. The preceding (Footnote 12) quotations from Blechman (2009) and Barnaby (2009) seem to suggest they are among the many informed by this viewpoint.

[17] One scenario widely discussed in the 1980's was that of "nuclear winter," under which an extensive nuclear exchange would cloud the atmosphere with "debris that would shut down photosynthesis, maybe for years" (Posner, 2004, p. 72). See Turco *et al.* (1983) for a summary of a once widely noted study of this possibility.

[18] Cf. Tannenwald (2007).

to stockpiles of nuclear weapons in multiple sovereign states that dwarf those existing today. These possibilities certainly suggest it is not totally irrational to assign the value infinity to the consequences of even a minimally destructive nuclear-weapon catastrophe. Nonetheless there remains strong dissent to the danger, even the significance, of nuclear weapons, as exemplified by the recent work of John Mueller (2009).

Now consider nuclear-armed and nuclear-disarmed rate profiles exactly as above. For the status quo nuclear-armed case the expected time to first use is then

$$T^{(a)} = T(\lambda_2^{(a)}) = \int_0^\infty t\tilde{\lambda}_2^{(a)} \exp\left(-\tilde{\lambda}_2^{(a)} t\right)\, dt = \frac{1}{\tilde{\lambda}_2^{(a)}}.$$

Similarly

$$T^{(d)} = T(\lambda_2^{(d)}) = \int_0^{t_0} t\tilde{\lambda}_2^{(a)} e^{-\tilde{\lambda}_2^{(a)} t} dt + \Lambda_0 t_0 e^{-\tilde{\lambda}_2^{(a)} t_0} + \int_{t_0}^\infty t\tilde{\lambda}_2^{(d)} \left(1-\Lambda_0\right) e^{-\left(\tilde{\lambda}_2^{(a)} t_0 + \tilde{\lambda}_2^{(d)}(t-t_0)\right)}\, dt =$$

$$-t_0 e^{-\tilde{\lambda}_2^{(a)} t_0} + \frac{1}{\tilde{\lambda}_2^{(a)}}\left(1-e^{-\tilde{\lambda}_2^{(a)} t_0}\right) + \Lambda_0 t_0 e^{-\tilde{\lambda}_2^{(a)} t_0} + \left(t_0 + \frac{1}{\tilde{\lambda}_2^{(d)}}\right) e^{-\tilde{\lambda}_2^{(a)} t_0}.$$

is the expected time to nuclear-weapon catastrophe under the disarmament profile. Disarmament at time t_0 is therefore indicated if

$$t_0 + \frac{1}{\tilde{\lambda}_2^{(a)}} < \Lambda_0 t_0 + \left(1-\Lambda_0\right)\left(t_0 + \frac{1}{\tilde{\lambda}_2^{(d)}}\right). \tag{8}$$

The criterion (8) for nuclear disarmament (at time t_0) has a simple interpretation. The quantity on the left is the expected time of (first) nuclear-weapon catastrophe, given that disarmament never occurs, and that time t_0 is reached without nuclear-weapon catastrophe. Further the apparent dependence on the specific time of disarmament is illusory. The various occurrences of t_0 cancel algebraically, and the criterion (8) simplifies into

$$\frac{1}{\tilde{\lambda}_2^{(a)}} < \frac{1-\Lambda_0}{\tilde{\lambda}_2^{(d)}}. \tag{9}$$

Thus if the three parameters in (9) were independent of time, then disarmament would either be desirable now or never. Of course this assumption of constancy with time does not comport with reality, because these parameters depend upon both technological capabilities and institutional relations between states with the international system, and both of these are highly dynamic.

The criterion (9) also has a ready interpretation. The left and right hand sides are the amounts by which the expected time of first nuclear-weapons catastrophe exceeds any given disarmament time t_0, under the respective rate profiles of a nuclear-armed world and nuclear disarmament, given that such a catastrophe does not occur prior to t_0. Comparing these obviously is equivalent to comparing the expected times of first use, because the probabilities that such a catastrophe occurs prior to t_0 are the same under the two profiles, and therefore so are the respective probabilities of nonoccurrence.

Qualitative arguments regarding the magnitude of the three parameters in (9) already have been extensively discussed above, and need not be repeated here. This inequality can be considered an alternative form of the (gamma form of the) benign disarmament proposition (5), and even a simpler version in that the ethically troubling discount rate does not appear. Note that if γ in (5) were replaced by $\tilde{\lambda}_2^{(a)}$, the result would be identical to (9); for that reason we term (9) as the *lambda form* of benign disarmament. This form implies the weak form (7).

4. ALTERNATE DISUTILITY FUNCTIONS

One could reasonably seek to minimize disutility functions other than the discounted expected risk (1). One such example is the maximum risk rate disutility function

$$R_{\max}\left(\lambda_1,\lambda_2\right)=\max_{0\le t<\infty}\left\{C_1\lambda_1(t)+C_2\lambda_2(t)\right\}.$$

As applied to the status quo and disarmament profiles considered above, with finite consequences, this criterion always would select the status quo, provided only that C_2 and Λ_0 are positive. (Because under the disarmament profile the maximum risk rate is infinite, at the time of disarmament.) Thus under this criterion, and in this case, it necessary to accept both $\Lambda_0 = 0$ and the weak form, (7), of benign disarmament, in order to conclude that it is not acceptable to maintain the status quo indefinitely. We term this combination as the *strong form* of the benign disarmament proposition, in note of the fact that it implies all forms previously encountered in this note.

On the other hand, if C_2 is infinite, and neither of the rates of occurrence of nuclear-weapon catastrophes are almost everywhere zero, then the maximum risk rate disutility function is infinite for both profiles, so fails to discriminate. In this case the philosophically kindred criterion would seem to be minimization of the maximum risk rate for nuclear catastrophe,

$$T_{\max}\left(\lambda_2\right)=\max_{0\le t<\infty}\lambda_2(t).$$

Again this maximum risk rate is infinite for the disarmament rate profile (provided Λ_0 is positive), but finite for the status quo. Similarly one therefore concludes that the status quo either is preferable or equivalent, unless the same strong form of the benign disarmament proposition holds.

Thus the maximum risk rate disutility function seems to favor maintaining indefinitely the status quo circumstances of nuclear armament. This criterion suggests that the societal objective should be to minimize the risk (from nuclear arms or the absence thereof) that any generation will ever face. In that respect it is similar to the "maximizing the worst possible outcome" theory of justice that has been proposed as a basis for equity or morality, in a variety of contexts (e.g., Rawls, 1971). Becker (1982) has applied this "maximin principle"[19] to the type of intergenerational equity issues that necessarily underlie the issue of nuclear disarmament.

Yet another alternative objective that can be considered is that of minimizing the long-term disutility function,

$$R_{\infty}\left(\lambda_1,\lambda_2\right)=\lim_{t\to\infty}\left\{C_1\lambda_1(t)+C_2\lambda_2(t)\right\}, \tag{10}$$

with preference going to profiles that achieve a minimal asymptotic risk rate sooner rather than later. In the case of finite consequences this always leads to the choice of immediate disarmament, provided only that

$$C_1\lambda_1^{(a)}+C_2\lambda_2^{(a)}>C_1\lambda_1^{(d)}+C_2\lambda_2^{(d)}.$$

---

[19] In the present context this would be a "minimax" principle, because disutility is considered rather than utility.

This is the earlier criterion (3), except that under the present long-term disutility function the risk at disarmament is completely discounted. It therefore leads to the same form (4) of the conventional nondeterrence proposition, but the weak form (7) of the benign disarmament proposition.

If the consequence of a nuclear-weapon catastrophe is taken as infinite, then both the status quo and disarmament profiles have infinite long-term disutilities. In this case the logical extension of this criterion seems to be to prefer that having the smaller value of the long-term risk rate,

$$T_\infty\left(\lambda_2\right)=\lim_{t\to\infty}\lambda_2(t).$$

If ties again are broken by a preference for profiles leading soonest to this asymptotic minimal rate, then immediate disarmament is the preferred course of action, provided only that the benign disarmament proposition holds in the weak form (7).

Thus the long-term disutility function (10) seems to favor immediate disarmament. The author is unaware of any prior work explicitly considering such disutility functions. Yet they would seem to be a reasonable consequence of a world view to the effect that whatever problems the current generation created, it is up to that generation to solve, as opposed to indefinite postponement of the difficulties attendant thereunto. There seems decidedly to be more than a little of that world view within the current discussion of matters such as nuclear disarmament and climate change.

Neither of the two alternative disutility functions considered in this section requires consideration of the ethically troublesome discount factor. Nonetheless the maximum risk-rate disutility function allows current generations to defer indefinitely the difficult task of nuclear disarmament, while the long-term disutility function puts the entire burden of shouldering the risk of nuclear disarmament upon the current generation, which of course increases the attendant implicit risk to future generations. Thus not explicitly allowing for a discount factor seems only to camouflage the issue of intergenerational equity, not to eliminate it.

5. MULTIPLE TYPES OF EVENTS

Of course one never really has the luxury of considering only two types of events. Here we briefly illustrate the extension of the preceding considerations to the opposing extreme of a continuum of types of undesirable events, as summarized by their associated consequences and risk rates. The associated expected discounted risk is

$$R(\lambda)=\int_0^\infty\int_0^\infty e^{-\gamma t}C\lambda(C,t)dCdt, \qquad (11)$$

where the consequence C takes on a continuum of values, and $\lambda(C,t)$ now is "risk-rate density," denominated as events per unit time and consequence "at" time t and consequence C.

Infinite time-integrated risks can arise in a variety of ways in this setting. However the appropriate generalization of the matters discussed above seems to be consideration of a risk of the form

$$R(\lambda)=\int_0^\infty\int_0^\infty e^{-\gamma t}C\lambda(C,t)dCdt+\infty\bullet\int_0^\infty e^{-\gamma t}\lambda_\infty(t)dt, \qquad (12)$$

where $\lambda_\infty(t)$ is the (time-varying) rate of occurrence of events of infinite consequence. If $\lambda_\infty(t)$ is nonnegative and not almost surely zero, then the generalized risk denominated by (12) is infinite, regardless of its "finite part" (11). Nonetheless one can in principle distinguish between rate profiles based on the expected time to first event of infinite consequence, which now will be exactly the same as (2), except with λ_2 replaced by λ_∞.

Traditionally such globally catastrophic events are dealt with by attempting to ensure the rate of occurrence from some possible event due to human actions is associated with a rate no larger than believed associated with natural events (e.g., collision with an asteroid). The author is unaware of any such effort in the context of events that might be caused by nuclear weapons.

6. CONCLUSION

Risk analysis seems unlikely to persuade those who have an established position on nuclear disarmament. Even if one accepts the premise that such "science-based" methods are applicable,[20] there is simply too little evidence regarding rates of occurrence, and too many possible rational choices of disutility function, to permit of persuasive arguments either way. Nonetheless the methodology as applied in this note reveals two key issues that seem to contain the essential underlying propositions: "conventional nondeterrence" and "benign disarmament."

See Table I for a summary of the versions of these propositions that are relevant to the various disutility functions and associated parametric values considered in the present work. Here the various versions of benign disarmament have been placed in the unifying form

$$\tilde{\lambda}_2^{(a)} > \tilde{\lambda}_2^{(d)} + w\Lambda_0,$$

where w is the relative weight associated to the risk of achieving (as opposed to maintaining) a disarmament regime.

Whether we have yet reached the point that either of these propositions is firmly established - in any of their various forms presented here - is a debatable proposition, albeit that debate seems unlikely to be advanced by suggesting those on the opposing side are purveying "nonsense." On the other hand, Perkovich and Acton (2008) appears to be a concession by two proponents of nuclear disarmament that the conditions to validate the benign disarmament proposition do not yet exist, along with a call for concrete steps intended to bring that day closer.[21] Perhaps that notable work should be matched by a serious effort by doubters of nuclear disarmament to delineate the circumstances that would adequately resolve those doubts

---

[20] See Stirling and Scones (2009) for a summary of discussions regarding applicability of science-based risk assessment, in quite a different context from nuclear disarmament.

[21] In the context of the simple two-alternative model that was the focus of this note, those steps would constitute changes to the probability of nuclear-weapon catastrophe during nuclear disarmament and to the rate of nuclear-weapon catastrophe during nuclear disarmament so as to move toward satisfying the gamma form (5) or the lambda form (9) of the benign disarmament proposition.

Table I: Versions of the conventional nondeterrence and benign disarmament propositions appropriate to various disutility functions and associated parametric values

	finite C_i, i=1,2		C_2=∞
	conventional nondeterrence	benign disarmament	benign disarmament
discounted expected risk, 0<γ<∞ (γ=discount rate)	(4)	gamma form, $w = \gamma$	lambda form, $w = \lambda_2^{(a)}$
discounted expected risk, γ→∞	(4)	impossible form, $\lim\limits_{\Lambda_0 \to 0+} \lim\limits_{w \to \infty}$	lambda form, $w = \lambda_2^{(a)}$
discounted expected risk, γ→0	(4)	weak form, $w = 0$	lambda form, $w = \lambda_2^{(a)}$
maximum risk rate	(4)	strong form, $\lim\limits_{w \to \infty} \lim\limits_{\Lambda_0 \to 0+}$	strong form, $\lim\limits_{w \to \infty} \lim\limits_{\Lambda_0 \to 0+}$
long term	(4)	weak form, $w = 0$	weak form, $w = 0$

7. REFERENCES

Barnaby, F. (2009). The nuclear renaissance: nuclear weapons proliferation and terrorism. London: Institute for Public Policy Research, March.

Becker, R. A. (1982). Intergenerational equity: The capital-environment trade-off. *Journal of Environmental Economics and Management*, *9*(2), 165-185.

Blechman, B. M. (2009). *National Interest online,* November 5, http://www.nationalinterest.org/Article.aspx?id=22436#Kampelmanan, accessed November 8, 2009.

Bostrom, N. and Ćirković, M. (2008). *Global catastrophic risks*. New York: Oxford University Press.

Brown, H. (2007-2008). New Nuclear Realities. *The Washington Quarterly,* Winter.

Byrd, D. M. and Cothern, C. R. (2000). *Introduction to risk analysis : a systematic approach to science-based decision making*. Rockville MD: Government Institutes

Gollier, C. (2001). Should we beware of the Precautionary Principle? *Economic Policy, 16* (33), 301-328.

Gollier, C and Weitzmann, M. L. (2009). How should the distant future be discounted when discount rates are uncertain? Preprint, November 7, http://idei.fr/doc/by/gollier/discounting_long_term.pdf, accessed November 30, 2009.

Iklé, F. C. (2009). Nuclear abolition, a reverie. *National Interest,* Issue 103, September/October, 4-7.

Iklé, F. C. (2009a). *National Interest online,* November 5, http://www.nationalinterest.org/Article.aspx?id=22436#Kampelmanan, accessed November 8, 2009.

Jervis, R. (2009). Or: how I learned to stop worrying. *National Interest*, Issue 104, November/December, 73-85. (Review of Mueller, 2009)

Juzaitis, R. J. and McLaughlin, J. E. (2008). Challenges of verification and compliance within a state of universal latency. In *Reykjavik revisited: steps toward a world free of nuclear weapons*, G.P. Shultz, S.P. Andreasen, S.D. Drell and J.E. Goodby (Eds.), 159-203, Stanford, CA: Hoover Institution Press.

Kampelman, M. M. (2009). *National Interest online,* November 5, http://www.nationalinterest.org/Article.aspx?id=22436#Kampelmanan, accessed November 8, 2009.

Kugler, J. (1984). Terror without deterrence: reassessing the role of nuclear weapons. *Journal of Conflict Resolution*, *28*, 470–506.

Manson, N. A. (2002). Formulating the precautionary principle. *Environmental Ethics*, *24*, (3), 263-274.

Manson, N. A. (1999). The precautionary principle, the catastrophe argument and Pascal's wager. *Ends and Means*, *4*, (1), http://www.abdn.ac.uk/philosophy/endsandmeans/vol4no1/index.shtml, accessed November 30, 2009.

Matheny, J. G. (2007). Reducing the risk of human extinction. *Risk Analysis, 27*, (5), 1335-1344.

Mueller, J. (2009). *Atomic obsession: Nuclear alarmism from Hiroshima to Al-Qaeda.* New York: Oxford University Press.

Parfit, D. (1984). *Reasons and persons.* New York: Oxford University Press.

Perkovich, G. and Acton, J. M. (2008). *Abolishing nuclear weapons.* New York: Routledge.

Perry, W. J. (2007). Testimony before the House Armed Services Committee, Strategic Forces Subcommittee, July 18. http://www.nuclearsecurityproject.org/atf/cf/%7B1FCE2821-C31C-4560-BEC1-BB4BB58B54D9%7D/PERRY_TESTIMONY071807.PDF, accessed November 11, 2009.

Posner, R. A. (2004). *Catastrophe: risk and response*, Oxford: Oxford University Press.

Rawls, J. (1971). *A theory of justice.* Cambridge, MA: Harvard University Press.

Sagan, S. (1993). *The limits of safety: organizations, Accidents, and Nuclear Weapons.* Princeton: University Press.

Short, J. F. (1984). The social fabric at risk: Toward the social transformation of risk analysis. *American Sociological Review, 49*, 711-725.

Stirling, A. C. and Scones, I (2009). From risk assessment to knowledge mapping: science, precaution, and participation in disease ecology. *Ecology and Society, 14*(2): 14. URL: http://www.ecologyandsociety.org/vol14/iss2/art14/, accessed December 5, 2009.

Tannenwald, N. (2007). The nuclear taboo: the United States and the non-use of nuclear weapons since 1945. Cambridge: University Press.

Turco, R. P., *et al.* (1983). Nuclear winter: global consequences of multiple nuclear explosions. *Science*, *222*, 1283-1292.

Waltz, K. N. (1981). *The spread of nuclear weapons: more may be better.* Adelphi Paper, no. 171, London: International Institute for Strategic Studies.