Risk Analysis, Vol. 27, No. 5, 2007 DOI: 10.1111/j.1539-6924.2007.00960.x

Reducing the Risk of Human Extinction

Jason G. Matheny[*]

In this century a number of events could extinguish humanity. The probability of these events may be very low, but the expected value of preventing them could be high, as it represents the value of all future human lives. We review the challenges to studying human extinction risks and, by way of example, estimate the cost effectiveness of preventing extinction-level asteroid impacts.

KEY WORDS: Asteroids; catastrophic risk; cost-effectiveness analysis; discounting; existential risk; human extinction

1. INTRODUCTION

Projections of climate change and influenza pandemics, coupled with the damage caused by recent tsunamis, hurricanes, and terrorist attacks, have increased interest in low-probability, high-consequence "catastrophic risks." Richard Posner (2004) has reviewed a number of these risks and the failures of policy and traditional risk assessment to address them. Richard Horton (2005), editor of *The Lancet*, has recommended creating an international body to rationally address catastrophic risks. The World Economic Forum (2006) recently convened a panel to catalog global catastrophic risks. The OECD (2003) has completed a similar exercise. And national research centers have emerged to study responses to catastrophe—the U.S. Department of Homeland Security recently funded a Center for the Study of High Consequence Event Preparedness and Response that involves 21 institutions.

In this article, I discuss a subset of catastrophic events—those that could extinguish humanity.[1] It is only in the last century, with the invention of nuclear weapons, that some of these events can be both caused and prevented by human action. While extinction events may be very improbable, their consequences are so grave that it could be cost effective to prevent them.

A search of EconLit and the Social Sciences Citation Index suggests that virtually nothing has been written about the cost effectiveness of reducing human extinction risks.[2] Maybe this is because human extinction seems impossible, inevitable, or, in either case, beyond our control; maybe human extinction seems inconsequential compared to the other social issues to which cost-effectiveness analysis has been applied; or maybe the methodological and philosophical problems involved seem insuperable.

Certainly, the problems are intimidating. Because human extinction is unprecedented, speculations about how and when it could occur are highly subjective. To efficiently spend resources in reducing extinction risks, one needs to estimate the probabilities of particular extinction events, the expected duration of humanity in an event's absence, the costs of extinction countermeasures, and the relative value of current and future human lives. Here, I outline how one might begin to address these problems.

---

[*] Department of Health Policy and Management, Bloomberg School of Public Health, Johns Hopkins University, 624 N. Broadway, Room 493, Baltimore, MD 21205, USA; tel: 202-486-1306; jmatheny@jhsph.edu.

[1] Because of the large timeframes discussed below, I use "humanity" and "humans" to mean our species and/or its descendents.

[2] The single exception found was Richard Posner's *Catastrophe*, and reviews of it.

0272-4332/07/0100-1335$22.00/1

2. HUMANITY'S LIFE EXPECTANCY

We have some influence over how long we can delay human extinction. Cosmology dictates the upper limit but leaves a large field of play. At its lower limit, humanity could be extinguished as soon as this century by succumbing to near-term extinction risks: nuclear detonations, asteroid or comet impacts, or volcanic eruptions could generate enough atmospheric debris to terminate food production; a nearby supernova or gamma ray burst could sterilize Earth with deadly radiation; greenhouse gas emissions could trigger a positive feedback loop, causing a radical change in climate; a genetically engineered microbe could be unleashed, causing a global plague; or a high-energy physics experiment could go awry, creating a "true vacuum" or strangelets that destroy the planet (Bostrom, 2002; Bostrom & Cirkovic, 2007; Leslie, 1996; Posner, 2004; Rees, 2003).

Farther out in time are risks from technologies that remain theoretical but might be developed in the next century or centuries. For instance, self-replicating nanotechnologies could destroy the ecosystem; and cognitive enhancements or recursively self-improving computers could exceed normal human ingenuity to create uniquely powerful weapons (Bostrom, 2002; Bostrom & Cirkovic, 2007; Ikle, 2006; Joy, 2000; Leslie, 1996; Posner, 2004; Rees, 2003).

Farthest out in time are astronomical risks. In one billion years, the sun will begin its red giant stage, increasing terrestrial temperatures above 1,000 degrees, boiling off our atmosphere, and eventually forming a planetary nebula, making Earth inhospitable to life (Sackmann, Boothroyd, & Kraemer, 1993; Ward & Brownlee, 2002). If we colonize other solar systems, we could survive longer than our sun, perhaps another 100 trillion years, when all stars begin burning out (Adams & Laughlin, 1997). We might survive even longer if we exploit nonstellar energy sources. But it is hard to imagine how humanity will survive beyond the decay of nuclear matter expected in 10^{32} to 10^{41} years (Adams & Laughlin, 1997).[3] Physics seems to support Kafka's remark that "[t]here is infinite hope, but not for us."

While it may be physically possible for humanity or its descendents to flourish for 10^{41} years, it seems unlikely that humanity will live so long. *Homo sapiens* have existed for 200,000 years. Our closest relative, *homo erectus*, existed for around 1.8 million years (Anton, 2003). The median duration of mammalian species is around 2.2 million years (Avise *et al.*, 1998).

A controversial approach to estimating humanity's life expectancy is to use observation selection theory. The number of *homo sapiens* who have ever lived is around 100 billion (Haub, 2002). Suppose the number of people who have ever or will ever live is 10 trillion. If I think of myself as a random sample drawn from the set of all human beings who have ever or will ever live, then the probability of my being among the first 100 billion of 10 trillion lives is only 1%. It is more probable that I am randomly drawn from a smaller number of lives. For instance, if only 200 billion people have ever or will ever live, the probability of my being among the first 100 billion lives is 50%. The reasoning behind this line of argument is controversial but has survived a number of theoretical challenges (Leslie, 1996). Using observation selection theory, Gott (1993) estimated that humanity would survive an additional 5,000 to 8 million years, with 95% confidence.

3. ESTIMATING THE NEAR-TERM PROBABILITY OF EXTINCTION

It is possible for humanity (or its descendents) to survive a million years or more, but we could succumb to extinction as soon as this century. During the Cuban Missile Crisis, U.S. President Kennedy estimated the probability of a nuclear holocaust as "somewhere between one out of three and even" (Kennedy, 1969, p. 110). John von Neumann, as Chairman of the U.S. Air Force Strategic Missiles Evaluation Committee, predicted that it was "absolutely certain (1) that there would be a nuclear war; and (2) that everyone would die in it" (Leslie, 1996, p. 26).

More recent predictions of human extinction are little more optimistic. In their catalogs of extinction risks, Britain's Astronomer Royal, Sir Martin Rees (2003), gives humanity 50-50 odds on surviving the 21st century; philosopher Nick Bostrom argues that it would be "misguided" to assume that the probability of extinction is less than 25%; and philosopher John Leslie (1996) assigns a 30% probability to extinction during the next five centuries. The "Stern Review" for the U.K. Treasury (2006) assumes that the probability of human extinction during the next century is 10%. And some explanations of the "Fermi Paradox" imply

---

[3] However, some cosmologies might allow life to exist indefinitely (Freese & Kinney, 2003).

a high probability (close to 100%) of extinction among technological civilizations (Pisani, 2006).[4]

Estimating the probabilities of unprecedented events is subjective, so we should treat these numbers skeptically. Still, even if the probability of extinction is several orders lower, because the stakes are high, it could be wise to invest in extinction countermeasures.

4. REDUCING EXTINCTION RISK

We already invest in some extinction countermeasures. NASA spends $4 million per year monitoring near-Earth asteroids and comets (Leary, 2007) and there has been some research on how to deflect these objects using existing technologies (Gritzner & Kahle, 2004; NASA, 2007). $1.7 billion is spent researching climate change and there are many strategies to reduce carbon emissions (Posner, 2004, p. 181). There are policies to reduce nuclear threats, such as the Non-Proliferation Treaty and the Comprehensive Test Ban Treaty, as well as efforts to secure expertise by employing former nuclear scientists.

Of current extinction risks, the most severe may be bioterrorism. The knowledge needed to engineer a virus is modest compared to that needed to build a nuclear weapon; the necessary equipment and materials are increasingly accessible and because biological agents are self-replicating, a weapon can have an exponential effect on a population (Warrick, 2006; Williams, 2006).[5] Current U.S. biodefense efforts are funded at $5 billion per year to develop and stockpile new drugs and vaccines, monitor biological agents and emerging diseases, and strengthen the capacities of local health systems to respond to pandemics (Lam, Franco, & Shuler, 2006).

There is currently no independent body assessing the risks of high-energy physics experiments. Posner (2004) has recommended withdrawing federal support for such experiments because the benefits do not seem to be worth the risks.

As for astronomical risks, to escape our sun's death, humanity will eventually need to relocate. If we survive the next century, we are likely to build self-sufficient colonies in space. We would be motivated by self-interest to do so, as asteroids, moons, and planets have valuable resources to mine, and the technological requirements for colonization are not beyond imagination (Kargel, 1994; Lewis, 1996).

Colonizing space sooner, rather than later, could reduce extinction risk (Gott, 1999; Hartmann, 1984; Leslie, 1999), as a species' survivability is closely related to the extent of its range (Hecht, 2006). Citing, in particular, the threat of new biological weapons, Stephen Hawking has said, "I don't think the human race will survive the next thousand years, unless we spread into space. There are too many accidents that can befall life on a single planet" (Highfield, 2001). Similarly, NASA Administrator, Michael Griffin (2006), recently remarked: "The history of life on Earth is the history of extinction events, and human expansion into the Solar System is, in the end, fundamentally about the survival of the species."

Perhaps more cost effective than building refuges in space would be building them on Earth. Elaborate bunkers exist for government leaders to occupy during a nuclear war (McCamley, 2007). And remote facilities are planned to protect crop seeds from "nuclear war, asteroid strikes, and climate change" (Hopkin, 2007). But I know of no self-sufficient, remote, permanently occupied refuge meant to protect *humanity* from a range of possible extinction events. Hanson (2007) argues that a refuge permanently housing as few as 100 people would significantly improve the chances of human survival during a range of global catastrophes. The Americas and Polynesia were originally populated by fewer than 100 founders (Hey, 2005; Murray-McIntosh *et al.*, 1998). Although it would take thousands of years for 100 people to repopulate Earth, this would be a small setback compared to extinction.

---

[4] Since 1947, the Bulletin of Atomic Scientists, an organization founded by former Manhattan Project physicists, has maintained a "Doomsday clock" that "conveys how close humanity is to catastrophic destruction—the figurative midnight—and monitors the means humankind could use to obliterate itself. First and foremost, these include nuclear weapons, but they also encompass climate-changing technologies and new developments in the life sciences and nanotechnology that could inflict irrevocable harm." However, the clock hands do not correspond to any specific probability.

[5] One environmental group, the Gaia Liberation Front (1995), recommended using biological weapons to extinguish humanity. Its manifesto stated: "we can ensure Gaia's survival only through the extinction of the Humans as a species . . . we now have the specific technology for doing the job . . . several different [genetically-engineered] viruses could be released (with provision being made for the release of a second round after the generals and the politicians had come out of their shelters)." Asked whether an engineered pathogen could be virulent enough to "wipe out all of humanity," "[National Institute of Allergy and Infectious Diseases Director Anthony] Fauci and other top officials . . . said such an agent was technically feasible but in practice unlikely" (Fiorill, 2005). Below, I argue that even unlikely events may be cost-effectively prevented.

5. DISCOUNTING

An extinction event today could cause the loss of thousands of generations. This matters to the extent we value future lives. Society places some value on future lives when it accepts the costs of long-term environmental policies or hazardous waste storage. Individuals place some value on future lives when they adopt measures, such as screening for genetic diseases, to ensure the health of children who do not yet exist. Disagreement, then, does not center on whether future lives matter, but on how much they matter.[6] Valuing future lives less than current ones ("intergenerational discounting") has been justified by arguments about time preference, growth in consumption, uncertainty about future existence, and opportunity costs. I will argue that none of these justifications applies to the benefits of delaying human extinction.

Under time preference, a good enjoyed in the future is worth less, intrinsically, than a good enjoyed now. The typical justification for time preference is descriptive—most people make decisions that suggest that they value current goods more than future ones. However, it may be that people's time preference applies only to instrumental goods, like money, whose value predictably decreases in time. In fact, it would be difficult to design an experiment in which time preference for an intrinsic good (like happiness), rather than an instrumental good (like money), is separated from the other forms of discounting discussed below. But even supposing individuals exhibit time preference within their own lives, it is not clear how this would ethically justify discounting across different lives and generations (Frederick, 2006; Schelling, 2000).

In practice, discounting the value of future lives would lead to results few of us would accept as being ethical. For instance, if we discounted lives at a 5% annual rate, a life today would have greater intrinsic value than a billion lives 400 years hence (Cowen & Parfit, 1992). Broome (1994) suggests most economists and philosophers recognize that this preference for ourselves over our descendents is unjustifiable and agree that ethical impartiality requires setting the intergenerational discount rate to zero. After all, if we reject spatial discounting and assign equal value to contemporary human lives, whatever their physical distance from us, we have similar reasons to reject temporal discounting, and assign equal value to human lives, whatever their temporal distance from us. I Parfit (1984), Cowen (1992), and Blackorby *et al.* (1995) have similarly argued that time preference across generations is not ethically defensible.[7]

There could still be other reasons to discount future generations. A common justification for discounting economic goods is that their abundance generally increases with time. Because there is diminishing marginal utility from consumption, future generations may gain less satisfaction from a dollar than we will (Schelling, 2000). This principle makes sense for intergenerational transfers of most economic goods but not for intergenerational transfers of existence. There is no diminishing marginal utility from having ever existed. There is no reason to believe existence matters less to a person 1,000 years hence than it does to a person 10 years hence.

Discounting could be justified by our uncertainty about future generations' existence. If we knew for certain that we would all die in 10 years, it would not make sense for us to spend money on asteroid defense. It would make more sense to live it up, until we become extinct. A discount scheme would be justified that devalued (to zero) anything beyond 10 years.

Dasgupta and Heal (1979, pp. 261–262) defend discounting on these grounds—we are uncertain about humanity's long-term survival, so planning too far ahead is imprudent.[8] Discounting is an approximate way to account for our uncertainty about survival (Ponthiere, 2003). But it is unnecessary—an analysis of extinction risk should equate the value of averting extinction at any given time with the expected value of humanity's future from that moment forward, which includes the probabilities of extinction in all subsequent periods (Ng, 2005). If we discounted the expected value of humanity's future, we would count future extinction risks twice—once in the discount rate and once in the undiscounted expected

---

[6] Some philosophers hold that future lives have no value, but this view is at odds with many of our deepest moral intuitions. See, for instance, Broome (2004), Hare (1993), Holtug (2001), Ng (1989), Parfit (1984), and Sikora (1978).

[7] This is made explicit by some policymakers, as well. For instance, the U.K. National Radiological Board's (1992) and the U.S. Environmental Protection Agency's (1985) standards on disposal of radioactive waste do not discount future lives.

[8] Dasgupta and Heal (1979) also argue, without a positive discount rate, we would have an infinite stream of future benefits, around which social policies could not be optimized. This argument requires an infinite duration of human existence, which does not seem possible given our cosmology. But if we attach a nonzero probability to the possibility of an infinite existence, then even with discounting, we are left with infinite expected values. Discounting does not help us avoid the problem unless outcomes beyond a certain horizon are valued at exactly zero. There does not yet seem to be a good way of dealing with infinite values in aggregative ethics.

value—and underestimate the value of reducing current risks.

In any case, Dasgupta and Heal's argument does not justify traditional discounting at a constant rate, as the probability of human extinction is unlikely to be uniform in time.[9] Because of nuclear and biological weapons, the probability of human extinction could be higher today than it was a century ago; and if humanity colonizes other planets, the probability of human extinction could be lower then than it is today.

Even Rees's (2003) pessimistic 50-50 odds on human extinction by 2100 would be equivalent to an annual discount rate under 1% for this century. (If we are 100% certain of a good's existence in 2007 but only 50% certain of a good's existence in 2100, then the expected value of the good decreases by 50% over 94 years, which corresponds to an annual discount rate of 0.75%.) As Ng (1989) has pointed out, a constant annual discount rate of 1% implies that we are more than 99.99% certain of not surviving the next 1,000 years. Such pessimism seems unwarranted.

A last argument for intergenerational discounting is from opportunity costs: without discounting, we would always invest our money rather than spend it now on important projects (Broome, 1994). For instance, if we invest our money now in a stock market with an average 5% real annual return, in a century we will have 130 times more money to spend on extinction countermeasures (assuming we survive the century). This reasoning could be extended indefinitely (as long as we survive). This could be an argument for investing in stocks rather than extinction countermeasures if: the rate of return on capital is exogenous to the rate of social savings, the average rate of return on capital is higher than the rate of technological change in extinction countermeasures, and the marginal cost effectiveness of extinction countermeasures does not decrease at a rate equal to or greater than the return on capital.

First, the assumption of exogeneity can be rejected. Funding extinction countermeasures would require spending large sums; if, instead, we invested those sums in the stock market, they would affect the average market rate of return (Cowen & Parfit, 1992). Second, some spending on countermeasures, such as research on biodefense, has its own rate of return, since learning tends to accelerate as a knowledge base expands. This rate could be higher than the average rate of return on capital. Third, if the probability of human extinction significantly decreases after space colonization, there may be a small window of reducible risk: the period of maximum marginal cost effectiveness may be limited to the next few centuries.

Discounting would be a crude way of accounting for opportunity costs, as cost effectiveness is probably not constant. A more precise approach would identify the optimal invest-and-spend path based on estimates of current and future extinction risks, the cost effectiveness of countermeasures, and market returns.

In summary, there are good reasons not to discount the benefits of extinction countermeasures. Time preference is not justifiable in intergenerational problems, there is no diminishing marginal utility from having ever existed, and uncertainties about human existence should be represented by expected values. I thus assume that the value of future lives cannot be discounted. Since this position is controversial, I later show how acceptance of discounting would affect our conclusions.

6. COST EFFECTIVENESS AND UNCERTAINTY

To establish the priority of delaying human extinction among other public projects, we need to know not only the value of future lives but also the costs of extinction countermeasures and how to account for their uncertain success. Cost-effectiveness analysis (CEA) is often used to prioritize public projects (Jamison, 1993). The ethical premise behind CEA is we should deliver the greatest good to the greatest number of people. With finite resources, this implies investing in projects that have the lowest marginal cost per unit of value—life-year saved, case of disease averted, etc. (McKie *et al.*, 1998). Even when CEA employs distributional constraints or weights to account for fairness or equity, cost effectiveness is typically seen as an essential aspect of the fair distribution of finite resources (Williams, 1997).[10]

The effects of public projects are uncertain. Some projects may not work and some may address problems that never emerge. The typical way of dealing with these uncertainties in economics is to use

---

[9] A declining discounting scheme advanced by Weitzman (2001), based on surveys from 2,160 economists, reduces to zero after 300 years. For other declining discounting schemes that better correspond to observed decision making than constant discounting, see Frederick, Loewenstein, and O'Donoghue (2002) and Price (1993).

[10] Equity weights would increase the priority given to delaying extinction, as ensuring the existence of future generations would be the first step in making their welfare equal to that of the present generation (Holtug, 2007).

expected values. The expected value of a project is the sum of the probability of each possible outcome of the project multiplied by each outcome's respective value.

7. EXAMPLE: THE COST EFFECTIVENESS OF REDUCING EXTINCTION RISKS FROM ASTEROIDS

Even if extinction events are improbable, the expected values of countermeasures could be large, as they include the value of all future lives. This introduces a discontinuity between the CEA of extinction and nonextinction risks. Even though the risk to any existing individual of dying in a car crash is much greater than the risk of dying in an asteroid impact, asteroids pose a much greater risk to the existence of future generations (we are not likely to crash all our cars at once) (Chapman, 2004). The "death-toll" of an extinction-level asteroid impact is the population of Earth, plus all the descendents of that population who would otherwise have existed if not for the impact. There is thus a discontinuity between risks that threaten 99% of humanity and those that threaten 100%.

As an example, consider asteroids. Let p be the probability of a preventable extinction event occurring in this century:

$$p = p_a + p_o,$$

where p_a is the probability of an asteroid-related extinction event occurring during the century, and p_o is the probability of any other preventable extinction event occurring. The (reducible) extinction risk is:

$$Lp = L(p_a + p_o),$$

where L is the expected number of future human life-years in the absence of preventable extinction events during the century. The expected value of reducing p_a by 50% is thus:

$$L(p_a + p_o) - L(0.5p_a + p_o) = 0.5Lp_a.$$

Suppose humanity would, in the absence of preventable extinction events during the century, survive as long as our closest relative, *homo erectus*, and could thus survive another 1.6 million years (Avise *et al.*, 1998).[11] Further suppose humanity maintains a population of 10 billion persons.[12] Then,

$$\begin{aligned} L &= 1.6 \text{ million years} \times 10 \text{ billion lives} \\ &= 1.6 \times 10^{16} \text{ life-years.} \end{aligned}$$

Based on the frequency of previous asteroid impacts, the probability of an extinction-level ($\geq$10 km) asteroid impact in this century is around one in 1 million (Chapman, 2004; NASA, 2007). Thus,

$$\begin{aligned} 0.5Lp_a &= 0.5 \times 1.6 \times 10^{16} \text{ life-years} \times 10^{-6} \\ &= 8 \text{ billion life-years.} \end{aligned}$$

A system to detect all large, near-Earth asteroids would cost between \$300 million and \$2 billion (Chapman, 2004; NASA, 2006, pp. 251–254), while a system to deflect large asteroids would cost between \$1 and 20 billion to develop (Gritzner, 1997, p. 156; NASA, 2006, pp. 251–254; Sommer, 2005, p. 121; Urias *et al.*, 1996).[13] Suppose a detect-and-deflect system costing a total of \$20 billion would buy us a century of protection, reducing the probability of an extinction-level impact over the next century by 50%.[14] Further suppose this cost is incurred even if the deflection system is never used, and the system offers no benefit besides mitigating extinction-level asteroid impacts.[15] Then the cost effectiveness of the detect-and-deflect system is

$$\$20 \text{ billion}/8 \text{ billion life-years} = \$2.50 \text{ per life-year.}$$

By comparison, it is common for U.S. health programs to spend, and for U.S. policies and citizens to value, more than \$100,000 per life-year (Kenkel, 2001; Neumann *et al.*, 2000; Viscusi & Aldy, 2003).[16] Even if

---

[11] If we survive this century, we are likely to develop self-sufficient colonies elsewhere in space. By that point, the probability of extinction by asteroid impact approaches zero, as the product of the probabilities of asteroid impacts destroying every human settlement.

[12] Although the terrestrial human population may be limited to 10 billion, if we colonize space, we could significantly increase population size. See, for instance, Bostrom (2003).

[13] The expected operating cost of such a system would be negligible, since there is only one chance in 1 million that it would need to be operated during the century.

[14] The studies cited assume detection success of 90% or more, and deflection success of 50% or more, using technologies with a "high level of technology readiness." In cases of failure, deflection attempts can be repeated. Because of the low probability of a system being used, repeated operating costs are negligible in comparison to development costs. Historically, the success rate of novel space systems such as Apollo or Soyuz has been over 50%.

[15] These are pessimistic assumptions, since such a system could also be used to detect and deflect more frequent subextinction level asteroid threats.

[16] This value is likely to increase, as incomes continue to rise, and as the "value of a statistical life" (VSL) rises with income. The

one is less optimistic and believes humanity will certainly die out in 1,000 years, asteroid defense would be cost effective at $4,000 per life-year.

8. DISCOUNTING REVISITED

Although the usual justifications for discounting do not apply to extinction, we might accept discounting and still conclude that delaying human extinction is cost effective. In the tabular display below I estimate the cost effectiveness of asteroid defense under different discounting schemes. As above, these estimates assume asteroid defense will save an expected 8 billion life-years. However, now the value of future life-years is discounted, relative to the value of a life-year lived now. The cost of asteroid detection and deflection is still assumed to be $20 billion, paid in the present.

Discount Rate (Form)	Present Value of Life-Years Saved	Cost per (Present Value) Life-Year Saved
No discounting	8.0×10^9	$2.50
Gamma (Weitzman, 2001)	1.4×10^8	$140.65
1% constant geometric	5.0×10^5	$40,000
3% constant geometric	1.7×10^5	$120,000
5% constant geometric	1.0×10^5	$200,000

The cost per life-year saved is $2.50 in the undiscounted case and $140.65 in the declining discounted case. Under constant discounting, the cost per life-year saved ranges from $40,000 to $200,000. Because the value of future life-years declines rapidly under constant discounting, these costs change by less than $1 if one pessimistically assumes a human duration of 1,000 years. Thus, even with discounting, and even assuming a 1,000-year human duration, asteroid defense could be more cost effective than much existing health spending.

Even if we expected humanity to become extinct within a generation, traditional statistical life valuations would warrant a $16 billion to $32 billion annual investment in asteroid defense (Gerrard & Barber, 1997). Yet the United States spends only $4 million per year on asteroid detection and there is no direct spending on mitigation.[17]

Some extinction risks are probably greater than asteroid impacts, and some risk-reducing projects are probably more cost effective than asteroid defense. A refuge would probably cost less than $20 billion to build and occupy, and would provide a stronger insurance policy against a broader range of extinction risks. Like other forms of catastrophic insurance, the probability of its being needed is low, but its expected value is high.

9. CONCLUSION

We may be poorly equipped to recognize or plan for extinction risks (Yudkowsky, 2007). We may not be good at grasping the significance of very large numbers (catastrophic outcomes) or very small numbers (probabilities) over large timeframes. We struggle with estimating the probabilities of rare or unprecedented events (Kunreuther *et al.*, 2001). Policymakers may not plan far beyond current political administrations and rarely do risk assessments value the existence of future generations.[18] We may unjustifiably discount the value of future lives. Finally, extinction risks are market failures where an individual enjoys no perceptible benefit from his or her investment in risk reduction. Human survival may thus be a good requiring deliberate policies to protect.

It might be feared that consideration of extinction risks would lead to a *reductio ad absurdum*: we ought to invest all our resources in asteroid defense or nuclear disarmament, instead of AIDS, pollution, world hunger, or other problems we face today. On the contrary, programs that create a healthy and content global population are likely to reduce the probability of global war or catastrophic terrorism. They should thus be seen as an essential part of a portfolio of risk-reducing projects.

Discussing the risks of "nuclear winter," Carl Sagan (1983) wrote:

> Some have argued that the difference between the deaths of several hundred million people in a nuclear war (as has been thought until recently to be a reasonable upper limit) and the death of every person on Earth (as now seems possible) is only a matter of one order of magnitude. For me, the difference is considerably greater. Restricting our attention only to those

---

income elasticity of VSL is around 0.5 (Viscusi & Aldy, 2003), and global per capita GDP has increased an average 1 to 3% per year over the last century (Maddison, 2003). If this growth continues, we can estimate 0.5 to 1.5% annual growth in VSL.

[17] Gerrard (2000) has pointed out that federal funding for asteroid detection is many orders of magnitude smaller than funding for hazardous waste sites, per unit of risk.

[18] For an exception, see Kent (2004).

> who die as a consequence of the war conceals its full impact. If we are required to calibrate extinction in numerical terms, I would be sure to include the number of people in future generations who would not be born. A nuclear war imperils all of our descendants, for as long as there will be humans. Even if the population remains static, with an average lifetime of the order of 100 years, over a typical time period for the biological evolution of a successful species (roughly ten million years), we are talking about some 500 trillion people yet to come. By this criterion, the stakes are one million times greater for extinction than for the more modest nuclear wars that kill "only" hundreds of millions of people. There are many other possible measures of the potential loss—including culture and science, the evolutionary history of the planet, and the significance of the lives of all of our ancestors who contributed to the future of their descendants. Extinction is the undoing of the human enterprise.

In a similar vein, the philosopher Derek Parfit (1984) wrote:

> I believe that if we destroy mankind, as we now can, this outcome will be *much* worse than most people think. Compare three outcomes:
>
> 1. Peace
> 2. A nuclear war that kills 99% of the world's existing population
> 3. A nuclear war that kills 100%
>
> 2 would be worse than 1, and 3 would be worse than 2. Which is the greater of these two differences? Most people believe that the greater difference is between 1 and 2. I believe that the difference between 2 and 3 is *very much* greater . . . . The Earth will remain habitable for at least another billion years. Civilization began only a few thousand years ago. If we do not destroy mankind, these thousand years may be only a tiny fraction of the whole of civilized human history. The difference between 2 and 3 may thus be the difference between this tiny fraction and all of the rest of this history. If we compare this possible history to a day, what has occurred so far is only a fraction of a second.

Human extinction in the next few centuries could reduce the number of future generations by thousands or more. We take extraordinary measures to protect some endangered species from extinction. It might be reasonable to take extraordinary measures to protect humanity from the same.[19] To decide whether this is so requires more discussion of the methodological problems mentioned here, as well as research on the extinction risks we face and the costs of mitigating them.[20]

ACKNOWLEDGMENTS

The author is grateful to John Broome, John Leslie, Richard Posner, Yew-Kwang Ng, Nick Bostrom, Robin Hanson, Anjula Razdan, Peter Singer, Jonathan Wiener, David Bishai, Matthew Ball, Les Knight, Lee Clarke, Tyler Cowen, Seth Baum, Rafaela Hillerbrand, Carl Shulman, and two anonymous reviewers for comments on an earlier draft.

REFERENCES

Adams, F. C., & Laughlin, G. (1997). A dying universe: The long-term fate and evolution of astrophysical objects. *Reviews of Modern Physics*, *69*(2), 337–370.

Anton, S. C. (2003). Natural history of Homo erectus. *American Journal of Physical Anthropology*, *122*, 126–170.

Avise, J. C. D., Walker, D., & Johns, G. C. (1998). Speciation durations and Pleistocene effects on vertebrate phylogeography. *Proceedings of the Royal Society of London Series B-Biological Sciences*, *265*, 1707–1712.

Blackorby, C., Bossert, W., & Donaldson, D. (1995). Intertemporal population ethics: Critical-level utilitarian principles. *Econometrica*, *63*, 1303–1320.

Bostrom, N. (2002). Existential risks: Analyzing human extinction scenarios and related hazards. *Journal of Evolution and Technology*, *9*.

Bostrom, N. (2003). Astronomical waste: The opportunity cost of delayed technological development. *Utilitas*, *15*(3), 308–314.

Bostrom, N. & Cirkovic, M. (Eds.). (2007). *Global Catastrophic Risk*. New York: Oxford University Press, forthcoming.

Broome, J. (1994). Discounting the future. *Philosophy and Public Affairs*, *23*(2), 128–156.

Broome, J. (2004). *Weighing Lives*. New York: Oxford University Press.

Chapman, C. R. (2004). The hazard of near-Earth asteroid impacts on earth. *Earth and Planetary Science Letters*, *222*, 1–15.

Chapman, R. H., Stone, P. W., Sandberg, E. A., Bell, C., & Neumann, P. J. (2000). A comprehensive league table of cost-utility ratios and a sub-table of "panel-worthy" studies. *Medical Decision Making*, *20*, 451–467.

Cowen, T. (1992). Consequentialism implies a zero rate of intergenerational discount. In P. Laslett & J. Fishkin (Eds.), *Justice Between Age Groups and Generations* (pp. 162–168). New Haven: Yale University Press.

Cowen, T., & Parfit, D. (1992). Against the social discount rate. In P. Laslett & J. Fishkin (Eds.), *Justice Between Age Groups and Generations* (pp. 144–161). New Haven: Yale University Press.

Dasgupta, P. S., & Heal, G. M. (1979). *Economic Theory and Exhaustible Resources*. Cambridge: Cambridge University Press.

Fiorill, J. (2005). Top U.S. disease fighters warn of new engineered pathogens but call bioweapons doomsday unlikely. *Global Security Newswire*, 29 July 2005. Available at http://www.nti.org/d_newswire/issues/2005/7/29/5c4288ca-3041-47e3-8b28-1359fc8670ee.html

Frederick, S. (2006). Valuing future life and future lives: A framework for understanding discounting. *Journal of Economic Psychology*, *27*, 667–680.

Frederick, S., Loewenstein, G., & O'Donoghue, T. (2002). Time discounting and time preference: A critical review. *Journal of Economic Literature*, *40*(2), 351–401.

---

[19] Human extinction would also likely condemn all nonhuman terrestrial life to extinction, as a planet-sterilizing asteroid or solar event is probable within the next billion years (Matheny, 2007).

[20] A good start would be to inaugurate a center for catastrophic risk assessment, composed of scientists, economists, and policy experts, as Posner (2004) has suggested. At Oxford University, the Future of Humanity Institute includes among its missions research on extinction risks.

Freese, K., & Kinney, W. (2003). The ultimate fate of life in an accelerating universe. *Physics Letters B*, *558*, 1–8.

Gaia Liberation Front. (1995). The final solution. Available at http://www.42inc.com/~estephen/manifesto/aum00063.html.

Gerrard, M. B. (2000). Risks of hazardous waste sites versus asteroid and comet impacts: Accounting for the discrepancies in U.S. resource allocation. *Risk Analysis*, *20*(6), 895–904.

Gerrard, M. B., & Barber, A. W. (1997). Asteroids and comets: U.S. and international law and the lowest probability, highest consequence risk. *Environmental Law Journal*, *6*(1), 1–49.

Gott, R. J. (1993). Implications of the Copernican principle for our future prospects. *Nature*, *363*, 315–319.

Gott, R. J. (1999). The Coppernican principle and human Survivability. In D. M. Haye (Ed.), *Human Survivability in the 21st Century: Proceedings of a Symposium, November 1998 Royal Society of Canada*. Buffalo, NY: University of Toronto Press.

Griffin, M. (2006). *Speech at NASA Goddard Space Flight Center, September 12, 2006*. Available at http://www.spaceref.com/news/viewsr.html?pid = 22033.

Gritzner, C. (1997). *Analysis of Alternative Systems for Orbit Alteration of Near-Earth Asteroids and Comets*. European Space Agency. Available at http://www.tu-dresden.de/mwilr/space/rfs/Dokumente/Publikationen/ESA-TT-1349.pdf.

Gritzner, C., & Kahle, R. (2004). Mitigation technologies and their requirements. In M. J. S. Belton *et al.* (Eds.), *Mitigation of Hazardous Impacts Due to Asteroids and Comets*. Cambridge, UK: Cambridge University Press.

Hanson, R. (2007). Catastrophe, social collapse, and human extinction. In N. Bostrom & M. Cirkovic (Eds.), *Global Catastrophic Risk*. New York: Oxford University Press, forthcoming. Available at http://hanson.gmu.edu/collapse.pdf.

Hare, R. M. (1993). Possible people. In R. M. Hare (Ed.), *Essays on Bioethics* (pp. 67–83). Oxford, UK: Clarendon Press.

Hartmann, W. K. (1984). Space exploration and environmental issues. *Environmental Ethics*, *6*(3), 227–239.

Haub, C. (2002). How many people have ever lived on Earth? *Population Today*, November/December. Available at http://www.prb.org/Template.cfm?Section = PRB&template = /ContentManagement/ContentDisplay.cfm&ContentID = 7421.

Hecht, J. (2006). Lessons in the art of species survival. *New Scientist*, 4 November 2006.

Hey, J. (2005). On the number of new world founders: A population genetic portrait of the peopling of the Americas. *PLoS Biology*, *3*(6), 965–975.

Highfield, R. (2001). Colonies in space may be only hope, says Hawking. *Daily Telegraph (UK)*, 16 October 2001.

Holtug, N. (2001). On the value of coming into existence. *Journal of Ethics*, *5*, 361–384.

Holtug, N. (2007). On giving priority to possible future people. In T. Rønnow-Rasmussen *et al.* (Eds.), *Hommage à Wlodek: Philosophical Papers Dedicated to Wlodek Rabinowicz*. Available at www.fil.lu.se/hommageawlodek.

Hopkin, M. (2007). Norway unveils design of "doomsday" seed bank. *Nature*, *445*, 693.

Horton, R. (2005). Threats to human survival: A WIRE to warn the world. *Lancet*, *365*, 191–193.

Ikle, F. C. (2006). *Annihilation from Within: The Ultimate Threat to Nations*. New York: Columbia University Press.

Jamison, D. (1993). Disease control priorities in developing countries. In D. T. Jamison, W. H. Mosley, A. R. Measham, & J. L. Bobadilla (Eds.), *Disease Control Priorities in Developing Countries*. Oxford, UK: Oxford University Press.

Joy, B. (2000). Why the future doesn't need us. *Wired Magazine*, April 2000.

Kargel, J. (1994). Metalliferous asteroids as potential sources of precious metals. *Journal of Geophysical Research*, *99*(E10), 21129–21141.

Kenkel, D. (2001). Using estimates of the value of a statistical life in evaluating regulatory effects. In F. Kuchler (Ed.), *Valuing Statistical Lives, Valuing the Health Benefits of Food Safety: Proceedings*. U.S. Department of Agriculture Economic Research Service, MP1570, April 2001.

Kennedy, R. (1969). *Thirteen Days: A Memoir of the Cuban Missile Crisis*. London: Macmillan.

Kent, A. (2004). A critical look at risk assessments for global catastrophes. *Risk Analysis*, *24*(1), 157–168.

Kunreuther, H, Novemsky, N, & Kahneman, D. (2001). Making low probabilities useful. *Journal of Risk and Uncertainty*, *23*, 103–120.

Lam, C., Franco, C., & Schuler, A. (2006). Billions for biodefense: Federal agency biodefense funding, FY2006-FY2007. *Biosecurity & Bioterrorism*, *4*(2), 113–127.

Leary, W. (2007). Tracking of killer asteroids runs low on money and short on time. *New York Times*, 12 March 2007 Section A, p. 14.

Leslie, J. (1996). *The End of the World: The Science and Ethics of Human Extinction*. London: Routledge.

Leslie, J. (1999). Risking human extinction. In D. M. Haye (Ed.), *Human Survivability in the 21st Century: Proceedings of a Symposium, November 1998 Royal Society of Canada*. Buffalo, NY: University of Toronto Press.

Lewis, J. S. (1996). *Mining the Sky*. Reading, MA: Addison-Wesley Publishing Co.

Maddison, A. (2003). *The World Economy: Historical Statistics*. OECD Development Centre, p. 263.

Matheny, J. G. (2007). Ought we worry about human extinction? *Environmental Ethics*, under review.

McCamley, N. J. (2007). *Cold War Secret Nuclear Bunkers*. London: Pen & Sword Books.

McKie, J., Richardson, J., Singer, P., & Kuhse, H. (1998). *The Allocation of Health Care Resources: An Ethical Evaluation of the QALY Approach*. Aldershot: Ashgate/Dartmouth.

Morrison, D., Harris, A., Sommer, G., & Chapman, C. (2002). Dealing with the impact hazard. In W. F. Bottke *et al.* (Eds.), *Asteroids III*. Tucson, AZ: University of Arizona.

Murray-McIntosh, R. P., Scrimshaw, B. J., Hatfield, P. J., & Penny, D. (1998). Testing migration patterns and estimating founding population size in Polynesia by using human mtDNA sequences. *Proceedings of National Academy of Science USA*, *95*(15), 9047–9052.

National Aeronautics and Space Administration (NASA). (2006). *Near-Object Survey and Deflection Study*. Washington, DC: NASA Office of Program Analysis and Evaluation, December 28, 2006. Available at http://www.b612foundation.org/papers/NASA-finalrpt.pdf.

National Aeronautics and Space Administration (NASA). (2007). *Near-Earth Object Survey and Deflection Analysis of Alternatives Report to Congress*. Washington, DC: NASA Office of Program Analysis and Evaluation.

Neumann, P., Sandberg, E., Bell, C., Stone, P., & Chapman, R. (2000). Are pharmaceuticals cost effective? A review of the evidence. *Health Affairs*, *19*(2), 92–109.

Ng, Y. K. (1989). What should we do about future generations? *Economics and Philosophy*, *5*, 235–251.

Ng, Y. K. (1991). Should we be very cautious or extremely cautious on measures that may involve our destruction? *Social Choice and Welfare*, *8*, 79–88.

Ng, Y. K. (2005). Intergenerational impartiality: Replacing discounting by probability weighting. *Journal of Agricultural and Environmental Ethics*, *18*, 237–257.

Organisation for Economic Co-operation and Development (OECD). (2003). *Emerging Systemic Risks in the 21st Century: An Agenda for Action*. Paris: OECD. Available at http://www.unisdr.org/eng/library/Literature/7754.pdf.

Parfit, D. (1984). *Reasons and Persons*. Oxford, UK: Clarendon Press, Part 4.

Pisani, R. (2006). *The Fermi Paradox: Three Models*. Department of Statistics, University of California at Berkeley. Manuscript.

Ponthiere, G. (2003). Should we discount future generations' welfare? A survey on the "pure" discounting rate debate, CREPP Working Paper, University of Liege.

Posnerm R. (2004). *Catastrophe: Risk and Response*. New York: Oxford University Press.

Posnerm, R. (2005). Catastrophic risks, resource allocation, and homeland security. *Journal of Homeland Security*, October. Available at http://www.homelandsecurity.org/newjournal/Articles/displayArticle2.asp?article = 133.

Price, C. (1993). *Time, Discounting and Value*. Oxford, UK: Blackwell Publishers.

Rees, M. J. (2003). *Our Final Hour: A Scientist's Warning*. New York: Basic Books.

Sackmann, I. J., Boothroyd, A. I., & Kraemer, K. E. (1993). Our sun III: Present and future. *Astrophysical Journal*, *418*, 457–468.

Sagan, C. (1983). Nuclear war and climatic catastrophe: Some policy implications. *Foreign Affairs*, *62*, 275.

Schelling, T. C. (2000). Intergenerational and international discounting. *Risk Analysis*, *20*(6), 833–837.

Sikora, R. I. (1978). Is it wrong to prevent the existence of future generations? In R. I. Sikora & B. Berry (Eds.), *Obligations to Future Generations*. Philadelphia: Temple University Press.

Sommer, G. S. (2005). *Astronomical Odds: A Policy Framework for the Cosmic Impact Hazard*. Santa Monica, CA: RAND Corporation. Available at: www.rand.org/pubs/rgs_dissertations/2005/RAND_RGSD184.pdf.

U.K. National Radiological Board. (1992). *Board Statement on Radiological Protection Objectives for the Land-Based Disposal of Solid Radioactive Wastes*. Documents of the U.K. National Radiological Protection Board 3, 3. Available at http://www.nrpb.org.uk/Absd3-3.htm.

U.K. Treasury. (2006). *Stern Review on the Economics of Climate Change*, Chapter 2, Technical Appendix, p. 47. Available at http://www.hm-treasury.gov.uk/media/8A3/83/Chapter_2_A_-_Technical_Annex.pdf.

Urias, J. M., DeAngelis, I. M., Ahern, D. A., Caszatt, J. S., Fenimore, G. W., & Wadzinski, M. J. (1996). Planetary defense. *United States Air Force*, pp. 44–51. Available at csat.au.af.mil/2025/volume3/vol3ch16.pdf.

U.S. Environmental Protection Agency. (1985). 40 CFR Part 91. *Federal Register*, *50*(182), 38066–38089.

Viscusi, W. K., & Aldy, J. E. (2003). The value of a statistical life: A critical review of market estimates throughout the world. *Journal of Risk and Uncertainty*, *27*(1), 5–76.

Ward, P. D., & Brownlee, D. (2002). *The Life and Death of Planet Earth*. New York: Henry Holt.

Warrick, J. (2006). Custom-built pathogens raise bioterror fears. *Washington Post*, 31 July 2006, p. A01.

Weiss, R. (2006). The world's agricultural legacy gets a safe home. *Washington Post*, 19 June 2006, p. A1. Available at http://www.washingtonpost.com/wp-dyn/content/article/2006/06/18/AR2006061800950.html.

Weitzman, M. L. (2001). Gamma discounting. *American Economic Review*, *91*(1), 260–272.

Wiener, J. B. (2005). Catastrophe: Risk and response. *Journal of Policy Analysis and Management*, *24*(4), 885–890.

Williams, A. (1997). Intergenerational equity: An exploration of the "Fair Innings" argument. *Health Economics*, *6*, 117–132.

Williams, M. (2006). *The Knowledge, Technology Review (MIT)*, March/April.

World Economic Forum. (2006). *Global Risks 2006*. Geneva: World Economic Forum. Available at http://www.weforum.org/pdf/CSI/Global_Risk_Report.pdf.

Yudkowsky, E. (2007). Cognitive biases potentially affecting judgment of global risks. In N. Bostrom & M. Cirkovic (Eds.), *Global Catastrophic Risk*. New York: Oxford University Press, forthcoming.