



The far future argument for confronting catastrophic threats to humanity: Practical significance and alternatives



Seth D. Baum *

Global Catastrophic Risk Institute, United States

ARTICLE INFO

Article history:

Available online 12 March 2015

Keywords:

Catastrophic threats
Global catastrophic risk
Existential risk
Far future
Co-benefits
Mainstreaming

ABSTRACT

Sufficiently large catastrophes can affect human civilization into the far future: thousands, millions, or billions of years from now, or even longer. The far future argument says that people should confront catastrophic threats to humanity in order to improve the far future trajectory of human civilization. However, many people are not motivated to help the far future. They are concerned only with the near future, or only with themselves and their communities. This paper assesses the extent to which practical actions to confront catastrophic threats require support for the far future argument and proposes two alternative means of motivating actions. First, many catastrophes could occur in the near future; actions to confront them have near-future benefits. Second, many actions have co-benefits unrelated to catastrophes, and can be mainstreamed into established activities. Most actions, covering most of the total threat, can be motivated with one or both of these alternatives. However, some catastrophe-confronting actions can only be justified with reference to the far future. Attention to the far future can also sometimes inspire additional action. Confronting catastrophic threats best succeeds when it considers the specific practical actions to confront the threats and the various motivations people may have to take these actions.

© 2015 Elsevier Ltd. All rights reserved.

1. Introduction

Over several decades, scholars from a variety of fields have advanced an argument for confronting catastrophic threats to humanity, rooted in the far future benefits of doing so.¹ In this context, the far future can loosely be defined as anything beyond the next several millennia, but will often emphasize timescales of millions or billions of years, or even longer.² Likewise, the catastrophic threats in question—also known as global catastrophic risks (GCRs) and existential risks, among other things—are those that would affect the trajectory of human civilization over these timescales. The simplest case is catastrophes resulting in human extinction, which is a permanent result and thus affects the trajectory of human civilization into the far future. More subtle but comparably relevant cases include catastrophes resulting in the permanent collapse of human civilization, preventing humanity from ever achieving certain very great things, and catastrophes resulting in delays

* Tel.: +1 5677045804.

E-mail address: seth@gcrinstitute.org

¹ See Asimov (1979), Sagan (1983), Parfit (1984), Ng (1991), Tonn (1999), Ćirković (2002), Bostrom (2003), Matheny (2007), and Beckstead (2013) among others.

² This definition of the far future is most explicitly stated in Beckstead (2013). In contrast, psychology and cognitive science research commonly defines “far future” in timescales of years (e.g., D’Argembeau, Xue, Lu, Van der Linden, & Bechara, 2008; Ebert & Prelec, 2007).

in the subsequent rise of civilization toward these achievements. The scholarship argues that people should care about human civilization into the far future, and thus, to achieve far future benefits, should seek to confront these catastrophic threats. Call this the *far future argument* for confronting catastrophic threats to humanity.

In this paper, I will not dispute the basic validity of the far future argument. Indeed, I agree with it, and have advanced it repeatedly in my own work (Baum, 2009, 2010; Maher and Baum, 2013). Instead, I assess the extent to which the far future argument is necessary or helpful for actually confronting the threats. In other words, what is the practical significance of the far future argument? I also propose and assess two alternative approaches to confronting the threats. One alternative emphasizes near future benefits of avoiding near future catastrophes. The other alternative emphasizes other (unrelated) benefits of actions that also help confront the threats, creating opportunities even for people who have zero care about the threats.

It would be important if the threats can be confronted without the far future argument, because many people do not buy the argument. That people do not is suggested by a range of research. An extensive time discounting literature assess how much people value future costs and benefits. Most discounting studies use time scales of days to decades and focus on future benefits to oneself (Frederick, Loewenstein, & O'Donoghue, 2002); these studies are of limited relevance to valuations of the far future of human civilization. One more relevant time discounting study finds that people discount lives saved 20 years later at a 25% annual rate and lives saved 100 years later at an 8% annual rate (Johannesson & Johannesson, 1996); extrapolating this suggests negligible concern for lives saved in the far future. Similarly, Tonn, Conrad, & Hemrick (2006, p. 821) find that people believe humanity should plan mainly for the upcoming 20 years or so and should plan less for time periods over 1000 years. In a study on social discounting, Jones and Rachlin (2006) find that people are willing to forgo more money to help close friends and family than distant acquaintances; they presumably would forgo even less for members of far future generations. Finally, there are considerations rooted in how societies today are structured. Several researchers have argued that current electoral structures favor the short-term (Ekeli, 2005; Tonn, 1996; Wolfe, 2008). Similarly, Karlsson (2005) suggests that the rise of decentralized capitalist/democratic political economies and the fall of authoritarian (notably communist) political economies has diminished major long-term planning. While none of these studies directly assess the extent to which people buy the far future argument, the studies all suggest that many people do not buy the argument to any significant degree.

To the extent that efforts to confront catastrophic threats can be made synergistic with what people already care about, a lot more can be done. This would seem to be an obvious point, but it has gone largely overlooked in prior research on catastrophic threats. One exception is Posner (2004), who argues that some actions to reduce the risk of human extinction can be justified even if only the current generation and its immediate successor are valued. Another is Baum (2015), who proposes to confront the threat of catastrophic nuclear winter in terms that could appeal to nuclear weapon states; Baum calls this “ethics with strategy”. But most of the prior research, including the studies cited above, emphasize the far future argument.

This paper expands Posner's argument to further argue that some actions can be taken even for those who only care about their immediate communities or even just themselves. This paper also makes progress toward assessing the total practical significance of the far future by presenting a relatively comprehensive survey of GCRs and GCR-reducing actions. Such surveys are also scarce; one example is Leggett (2006), who surveys the space of GCRs to identify priorities for action. The present paper also has commonalities with Tonn and Stiefel (2014), who evaluate different levels of sacrifice that society should make in response to GCRs of different magnitude. The present paper also considers levels of sacrifice, but instead argues that, from a practical standpoint, it is better to start with those actions that require less sacrifice or are in other ways more desirable. Indeed, actions requiring large sacrifice may only be justifiable with reference to far future benefits.

The paper is organized as follows. Section 2 briefly reviews the space of GCRs. All actions to reduce the risk must help on one or more of these so as to result in a net risk reduction. The space of GCRs likewise provides an organizing framework for subsequent sections, as summarized in Table 1. Section 3 discusses the timing of GCRs. For catastrophes that could happen earlier, actions to avoid them will include the earlier benefits of catastrophe avoidance. Almost all GCR reduction actions have near-future GCR reduction benefits. Section 4 discusses co-benefits and mainstreaming of GCR reduction actions. Co-benefits are benefits unrelated to GCR reduction. Mainstreaming is integrating GCR reduction into established

Table 1

Summary of global catastrophic risk categories (Section 2), their timing (Section 3), co-benefits and mainstreaming opportunities (Section 4), and high-cost GCR reduction actions that may only be justifiable with reference to far future benefits (Section 5). The co-benefits and mainstreaming opportunities and high-cost actions are illustrative examples, not complete listings.

GCR category	Timing	Co-benefits & mainstreaming	High-cost actions
Environmental change	Near or far	Money, health, happiness	Pollution abatement
Emerging technologies	Near only	Other tech risks	Technology relinquishment
Large-scale violence	Near or far	Money, small-scale violence	Increase in smaller conflicts
Pandemics	Near or far	Other public health benefits	Aggressive quarantine
Natural disasters	Near or far	Other disasters	Advanced refuges
Physics experiments	Near only	Money	N/A
ET encounter	Near or far	Science, entertainment	N/A
Unknowns	Near or far	Other risks	Extraterrrestrial time capsule

Download English Version:

<https://daneshyari.com/en/article/1015451>

Download Persian Version:

<https://daneshyari.com/article/1015451>

[Daneshyari.com](https://daneshyari.com)