STEPHEN DAVIES

APOCALYPSE NEXT

Why another pandemic – or worse – is inevitable. And what we can do about it.

About IEA publications

The IEA publishes scores of books, papers, blogs and more each and every year – covering a variety of stimulating and varied topics.

To that end, our publications will, in the main, be categorised under the following headings, from May 2023 onwards.

Hobart Editions
Papers or books likely to make a lasting contribution to free market debate, either on policy matters or on more academic grounds

IEA Foundations
Primers or introductions to free market concepts, ideas and thinkers

IEA Perspectives
Majoring on current policy-focused issues

IEA Briefings
Short, punchy reports on contemporary issues

IEA Special Editions
Spotlighting lecture transcripts, celebrated writers and more

Much of our work is freely available on the IEA website. To access this vast resource, just scan the QR code below – it will take you directly to the IEA's Research home page.

IEA Perspectives – Edition 1

APOCALYPSE NEXT

The Economics of Global Catastrophic Risks

STEPHEN DAVIES

iea
Institute of
Economic Affairs

First published in Great Britain in 2024 by
The Institute of Economic Affairs
2 Lord North Street
Westminster
London SW1P 3LB
in association with London Publishing Partnership Ltd
www.londonpublishingpartnership.co.uk

The mission of the Institute of Economic Affairs is to improve understanding of the fundamental institutions of a free society by analysing and expounding the role of markets in solving economic and social problems.

A CIP catalogue record for this book is available from the British Library.

ISBN 978-0-255-36822-3 (interactive PDF)

Typeset in Kepler by T&T Productions Ltd
www.tandtproductions.com

CONTENTS

ABOUT THE AUTHOR

Stephen Davies is Senior Education Fellow at the Institute of Economic Affairs. From 1979 to 2009 he was Senior Lecturer in the Department of History and Economic History at Manchester Metropolitan University. He has also been a Visiting Scholar at the Social Philosophy and Policy Center at Bowling Green State University, Ohio, and a Program Officer for the Institute for Humane Studies in Arlington, Virginia. His many publications include *Empiricism and History* (Palgrave Macmillan, 2003) and *The Wealth Explosion: The Nature and Origins of Modernity* (Edward Everett Root, 2019).

TABLES

INTRODUCTION

The popular mind today is haunted by the images of apocalypse; narratives of catastrophe and post-apocalyptic survival are prominent features of popular fiction, and television has seen several successful series with those themes, such as *Walking Dead* and *Revolution*. A thriving 'prepper' subculture of people is getting ready for an anticipated collapse of civilisation. Fears and anticipations of this kind are not found only in science fiction, technothriller novels or television and film, though. In the last two decades, a galaxy of stars of science and technology, from a former President of the Royal Society Sir Martin Rees to the late Stephen Hawking, have warned of the growing chances of a disaster that could destroy civilisation all over the world or even bring about human extinction.[1] Latterly, some politicians have taken up this theme, and not all of them do so from a 'green' position. They are joined by a growing number of important figures from the world of business, such as Bill Gates (2021).

Recently, everyone has had a taste of the kind of global disaster they had previously experienced only vicariously in fictional books and television shows. The Covid-19

1 Apocalypse soon: the scientists preparing for the end times. *New Statesman*, 25 September 2014 (https://www.newstatesman.com/sci-tech/2014/09/apocalypse-soon-scientists-preparing-end-times).

pandemic led to a global crisis that certainly counts as a disaster, in terms of the economic impact and lives lost, and one that massively disrupted everyday life for billions of people. More significantly, it affected the patterns and workings of world civilisation in a number of destructive ways. The sobering reflection is that this could easily have been much worse – instead of thinking about how we have experienced a disaster (true though that is), we should recognise that we have dodged a bullet or had a serious warning.

All of these dangers – real and imagined – are about the same matter: global catastrophic risks (GCRs). A GCR is a possible event that will have a global impact which is utterly disastrous and long-lasting (Schneier 2015). A simple way of understanding GCR is that, as in the fictional portrayals, we are talking about a truly world-changing event. It might destroy civilisation and usher in a new, possibly permanent, dark age. Or it might dramatically reduce global human populations, maybe even to the point of human extinction. As we shall see, there are many such dangers.

The idea of such a disaster is an old one, going back to the deluge narratives of the Bible and many other mythological traditions (Carlin 2019). In the genre of science fiction, stories about such topics can be found from the early nineteenth century onwards, and it has often been a device of speculative fiction (Wagar 1982). There has also been a long tradition of authors of non-fictional jeremiads arguing that if certain trends are allowed to continue, the result will be the end of everything. The many followers of

Thomas Malthus over the last two hundred years are an example of this (although not Malthus himself, as he was not making a prophecy).

There is something novel, though, about the last two decades or so. Concerns expressed today are not tropes of fiction or arguments of the 'if this goes on' variety (although we have plenty of those still). Rather, they are well-informed concerns of a GCR happening in the foreseeable future (Bostrom and Cirkovic 2012). As we shall see, the mathematics suggest that in many cases such possible events are bound to happen sooner or later, the only question is *when*.

This partly reflects social and technological change. Developments in the media have made people more aware of disasters happening in other parts of the world, while simultaneously making it easier for ideas and concerns about possible GCRs to spread more easily. In addition, as more people move out of subsistence living and become wealthier, so they have more time to worry about other issues, including ones that might previously not have been worth worrying about.

But there are also objective reasons for increased concern, which reflect scientific investigation. First, the range of possible GCRs has increased. The reason for this is simple: previously, the main GCRs were natural phenomena of various kinds, many of them with low inherent probability. Increasingly, there is serious concern about GCRs that arise as a product of human activity. In particular, current technological developments, combined with the way the modern social and economic infrastructure has developed,

raise a number of dangers. This means there are now new threats, produced by new technology and social development (Moynihan 2020).

The second reason for increased anxiety is the realisation that the mathematics of GCRs has frequently been misunderstood. In simple language, the chances of certain GCRs happening are higher than previously thought and probably increasing. Furthermore, a GCR only needs to happen once for there to be a disaster, and perhaps for human history to come to the proverbial full stop.

This point about probability should matter to everyone, not just mathematicians. Evidence is clear that the way policymakers think about risk in general and GCRs in particular is based on incorrect statistical assumptions. This was illustrated in a non-catastrophic episode, the financial crisis of 2008. Even more alarming is the evidence that institutions such as markets do not price risk correctly and so do not perform their central information gathering and signalling function. In chapter 6, we will consider the important question of why these market outcomes occur.

On a positive note, serious thought is taking place: a number of professional research organisations and think tanks now focus on studying GCR (see the list of links at the end of this work for details). Many other organisations give the topic some attention, and most think tanks and research institutes concerned with technology and environmental or climate policy list GCR as one of their main concerns. In addition, some organisations, such as Cathedral Thinking and The Long Now Foundation, are concerned with addressing various institutional problems.

More negatively, this has not yet had a wider impact on policymaking. It remains to be seen whether the experience of the Covid-19 pandemic may change this.

How should we react to these dangers? One argument is that we should ignore them, thinking that the prognosticators of possible doom are like Chicken Little, who worried that the sky was going to fall. Proponents of this argument state not only that the chances of a GCR are so low as to not be worth worrying about, but also that taking these fears seriously will waste resources and take effort and attention away from more acute and soluble problems. There are two responses to this view: first, it underestimates the actual probability of some kind of GCR happening soon, because the way probability is thought of is faulty; second, the idea that responding to the threat posed by GCRs must compete with other useful activity is incorrect.

If we do take the advice of people such as Martin Rees and Stephen Hawking seriously, however, what should we do? It is perfectly possible and, in some cases, reasonable to conclude that the risk is real but that it does not make sense to do anything about it, at least not anything that consumes resources. This is where economics comes into play because what we are talking about here is the value of resources and the costs of different uses of those resources in terms of forgone alternative uses (opportunity costs). Questions also exist about the relative costs to put right the damage caused by a GCR and the probability of a GCR occurring compared with the cost of action taken to head off or mitigate such an event. This kind of cost–benefit calculation lies at the heart of economics. Just as

we need good science and probabilistic reasoning and mathematics, so we also need sound economic thinking to avoid costly mistakes.

Given that, a striking and regrettable feature of the discussion so far around GCR is how little input there has been from economists, and how much of that has been of very poor quality. There are some notable exceptions, however. The most prominent coverage has come from Richard Posner, who has written an excellent book on the public policy response to catastrophic risks (Posner 2004). We need more of this if we are to make informed decisions as to what kind of GCR, if any, we should take seriously and what actions, if any, it makes sense to then take. Much of the reasoning required is well known to students of insurance and risk management, and there is a range of tools available to help us understand the options available.

If we conclude that the possibility of a specific GCR happening is sufficiently high that it is worth taking seriously (i.e. incurring costs to prevent or mitigate), and even more if we conclude that the possibility is underestimated or increasing, what kinds of action might we consider? It might be that some actions could reduce the *chances* of that GCR happening or even head it off entirely. In other cases, it might be that the best course of action is one of mitigation, of actions that will reduce the *impact* of a global catastrophe or shorten its duration. This could be because the probability of the event occurring is so high or because, although being unlikely to happen in any given year or short-duration timespan, we can be sure it will happen sooner or later and there is no practical way of stopping it.

There may be some cases where mitigation is not feasible, but, in that event, it may be possible to take precautions to ensure that catastrophe is not total and that there is some way of ensuring the survival of humanity, or the recovery of civilisation, in the much longer term. All of these choices need input from scientists and engineers, but also historians, sociologists and, not least, economists.

The remainder of this book is organised as follows. The first chapter sets out in more detail what GCRs are and how to define them. The second examines the mathematics regarding the probability of GCRs and the question of why so much of the thinking about this is misguided. The third sets out an analytical framework for understanding the different types of GCR and explores the question of why several of these have only recently come to attention; why they are becoming more rather than less probable; and why concern about them has intensified among the scientifically informed. The fourth chapter makes up the core of the work and contains a survey of the different kinds and categories of possible GCR. The fifth chapter addresses the big question of why we should be concerned at all about GCRs, as opposed to simply ignoring them. The sixth, which is the other main part of the work, explains how to think about this economically. It also introduces two related ideas, the comparatively well-known idea of a *precautionary* principle (but modified by economic reasoning) and the less well-known but very important idea of a *proactionary* principle, again with economic reasoning brought to bear. The seventh and final chapter looks at how, if at all, we should respond to the challenge of GCRs. It

starts by unwrapping 'we' to clarify who or what should be acting and in what manner. It then explores the basic principles that should underlie responses – this involves examining the apparent (to some) tension between the ideas looked at here and free-market principles and insights. It continues by looking at some of the concrete measures that can be justified in each of the particular kinds of GCR identified – justified, that is, by a combination of risk assessment and economic reasoning. The concluding chapter examines in general terms the need for institutional reforms to many market and social institutions, in order to bring about a major shift in time preferences, discount rates and investment horizons.

1 THE NATURE AND SIGNIFICANCE OF GLOBAL CATASTROPHIC RISKS

The concept of risk is a familiar one, found in economics, statistics and probability theory, and with applications in areas such as insurance, portfolio management and gambling (Beck and Kewell 2014). It is also something that all human beings everywhere deal with constantly. Despite this ubiquity, the concept is often misunderstood, partly because on examination it proves to be complex. This is true even at the level of experts, all too many of whom, as historical experience shows, either do not fully understand the idea or approach it using a faulty set of intellectual tools. A risk, simply put, is an event that may occur at some point in the future and that, if it does occur, will lead to a loss for the person bearing it. The term also refers to the chance or probability of the event coming to pass, when this can be calculated. This can be expressed in several different ways, but the usual one is as a probability in a given period of time, which can be anything from a single moment (as in the spin of a roulette wheel) to a much longer period.

It is this last that leads to many intuitive misunderstandings, which often prove costly for some but are lucrative for others. As set out here, a risk is a logically possible

event that results in costs if it occurs. Not all risks (possible events) come with risks (calculable probabilities); the probability of some events happening is unknowable, in which case we are talking of uncertainty rather than risk or known probability. Incalculability, however, does not preclude making a judgement of probability.

An even more troublesome category exists, of events that might happen in the future which lead to costs but which are radically unexpected – in these cases, it is not the probability of the occurrence that is unknown but the actual event itself. You might suppose that unknowable and incalculable risks are not worth bothering about, but that is not the case.

So, two of the central components of risk are *probability* and *cost*. Cost here means harm, damage or the loss or destruction of something valued. This can be the loss of something actually in existence (and therefore capable of being possessed) or of something that would have existed or had a high probability of existing, had the event not happened. In the latter case, we are speaking of virtual loss, the loss of wealth or assets of some kind that *would* have existed, but *would not* if the event came to pass. This description makes it sound purely hypothetical, but, in many cases, virtual loss is a serious and important category. Normal procedure is to allow for such a loss (the forgone growth in asset value or profits, for example) but to apply a discount to it, to reflect the uncertainty of the future. The cost of a risk for actuarial or insurance purposes is arrived at by multiplying the probability of the event happening by the estimated size of the loss, which may or may not include

virtual losses. The product of this operation is the value at risk. Having established what risks and values at risk are, the task is to categorise them so that different concrete risks can be compared to one another. Only once that task has been completed can we place risks into any kind of order or hierarchy. This produces an overall measurement of severity or gravity.

Setting probability to one side, there are three measurement criteria for risks. The first is *extent*, which means the number of people and the geographical area affected by the consequential loss or damage if the event happens. Here exists a straightforward scale. At one end is a risk borne only by a specific individual. If the event happens, they suffer a loss but nobody else does. Clearly, there is a continuous and linear progression from that point with the number of people increasing steadily. We can, if we wish, subdivide that progression using a numerical yardstick such as orders of magnitude of a base number such as ten. The end point might appear to be when the entire global population bears a cost, making the risk universal or global. In fact, that is not the terminus. Beyond the point of a loss suffered by all human beings currently alive is the prospect of a loss suffered by a given number of people not yet born, which can extend to all human beings who could possibly exist in the future, given reasonable assumptions (such as the average span of existence for the typical species). Arguably, the extent can also go so far as to include non-human species or the natural environment, with a similar scale in terms of small numbers up to entire species or, even at the ultimate extent, all species. Extent also

means literal physical extent, as well as numerical. Here the scale runs from purely local, affecting only a very small area, to global, affecting the entire planet or the entire biosphere.

The second criterion is the *scale* of the damage or loss incurred if the event takes place. This ranges from the trivial to the catastrophic. At one end of the scale, the harm or loss amounts to little more than inconvenience. At the other end of the scale, it is the loss of everything that is valuable to the loss-bearer. This is captured in the distinction between loss and ruin (Taleb 2018). In monetary terms, this is like losing some amount of money, either large or small, versus losing your entire net worth. In terms of investment returns, the scale goes from a minor loss to one which wipes out so much there is no possibility of ever recovering the losses and you are left permanently impoverished. At a personal level, the cost of an accident (the risk, in this case) could range from minor injury to serious injury bringing permanent impairment or death. Here again, loss can be virtual as well as actual – to lose anticipated and expected future income is a loss, for example, even if it does not hurt as much as losing physical property or cash in the bank. A common notion is that the worst kind of loss would be one's death, but that is only true if you have no feeling for fellow human beings. For most people, having their children die is worse than their own death. By extension, the death of others around you including everyone in the community of which you are a part and ultimately the extinction of the human species is much worse than one's own death. Part of this is the condemnation to non-existence of the future

people who would have otherwise lived, including one's own descendants (MacAskill 2022).

The third criterion is that of *duration*. Here the question is one of how long-lasting the damage or loss is and of how easily and quickly it can be made good. The scale ranges from brief and easily reparable to permanent and irreversible. This criterion is often combined with the second (the scale of the damage or loss incurred) to give a single criterion of the severity of damage/cost, since serious and large-scale loss will also very likely be long-lasting and hard to reverse. The most obvious example is that of death or permanent impairment where the duration of the harm (permanent) is a major feature of the size of the loss. Similarly, catastrophic loss or ruin almost always means loss that cannot be recovered or made good, at least not within a humanly meaningful timescale. Most models of types of risk, therefore, have a single criterion of severity that combines scale/intensity and duration. However, it does make sense to separate the two, even if that division is fuzzy, because thinking about duration is important in calculating payoffs for various strategies when dealing with risks and is important if we are to take virtual costs into account as well as concrete or crystallised ones. Payoffs in this context mean both positive and negative payoffs. In addition, the two variables are not perfectly aligned – it is possible to have trivial harms that are permanent or severe ones that are short-lived. An example of the first would be a minor scar or perhaps losing part of a finger. An example of the second would be a large monetary loss that is quickly made up.

Putting these three criteria together, whether as two ultimate variables or three, produces a clear structure for the classification of risks and enables us to understand what is meant by risks that are simultaneously global, catastrophic and permanent or long-lasting. The easiest way to think about this is to have two axes and variables: extent and severity (while remembering that severity is actually the product of two separate variables). The vertical axis will measure severity while the horizontal axis will measure extent. This produces a number of categories of risk, depending on how finely the two axes are subdivided. The person most associated with this analytical framework is the Oxford philosopher Nick Bostrom, who used it to arrive at a typology of GCRs (Bostrom 2002). In Bostrom's model, reproduced here, there are three levels of extent: *individual* (affecting a single person), *local* (affecting anything between a small number of people/very local area and a significant but not majority part of the planet's surface and population) and *global* (affecting the entire planet and its inhabitants). These are combined with two levels of severity, which Bostrom calls *endurable* and *terminal*. These two latter categories combine intensity and duration, since an endurable risk is one that gives rise to low levels of loss and is not long-lasting, while a terminal risk is one that brings very costly and permanent loss. As seen in Table 1, this produces six types of risk through the combinations of three values for one variable and two for the other. GCRs are listed in the upper right cell.

Bostrom's categories of endurable and terminal correspond to those of minor and catastrophic, or loss and ruin.

Risks of intermediate extent for Bostrom mean risks that impact only part of the planet's surface and population, such as an economic collapse or war affecting a single country or at most a continent. A global risk obviously means it affects the global population and entire planet. The idea of a terminal risk is worth elaborating upon. Death is the most obvious, and at the intermediate level means something like the genocide of an entire people, or a biological disaster that wipes out all or almost all of the population of a significant portion of the planet. An example of the latter would be the epidemics of diseases such as smallpox and measles that killed most of the indigenous population of the Americas after these diseases were brought there by Europeans.

Table 1

Terminal	Death, death of family, permanent impairment	Genocide, local nuclear war, post-Columbus epidemics	Global nuclear war, Black Death or higher-level pandemic
Endurable	Injury	Local conventional war, natural disaster, local epidemic	Global depression, conventional world war, Covid-19
	Personal	**Local**	**Global**

However, death, whether individual or collective, is not the only kind of terminal risk. For Bostrom and other authors, it also includes any risk that permanently forecloses future possibilities of flourishing (Ord 2020). At the individual level, this could be permanent disability. At both the individual and intermediate level, it would also

include slavery or sustained and indefinite tyranny. Very importantly, it also includes a permanent (irrecoverable) breakdown of the complex systems that we call civilisation, because that would preclude for the indefinite future a whole range of scientific and technological developments, as well as economic ones, and so eliminate an enormous range of future possible human flourishing, some at least of which is predictable (i.e. we can be fairly sure it would have happened but for the event). This means, therefore, that global risks which are terminal or catastrophic are of two kinds: those that would mean human extinction or the destruction of life on Earth, and those that would bring about the complete and irretrievable collapse of civilisation, or a permanent end to progress and development.

Bostrom's work is very useful because it produces a hierarchy of risks and a clear idea of what counts as a GCR. In this schema, a GCR is an event of global extent that would, if it happens, result in permanent and severe costs for the entire population of the planet, including future generations. This could be something that results in human extermination, or something as close to that as to make no difference, or something that permanently forecloses the future development and flourishing of humanity. The philosopher Toby Ord (2021) has been particularly exercised by this last point and has explored it in a number of places. What this means is that a GCR is distinct from something that has a catastrophic effect but does not affect the whole world or radically impair the future prospects of humanity as a whole (e.g. the post-Columbian epidemics that devastated the Americas but left the Old World untouched) or

from something that impacts the whole world but does not cause a loss amounting to ruin or catastrophe – a global economic depression, a worldwide conventional war or the Covid-19 pandemic would all fall into that category. This means that the category of risks which count as GCRs is more limited and defined than we might imagine and we can discuss this in a more precise and focused way than if we were simply talking about 'something big and really bad'. (For an alternative definition to the one offered by Ord and Bostrom, see Avin et al. (2018).)

That said, it is worth refining Bostrom's original model in some ways, as he himself has done (Bostrom and Cirkovic 2012: ch. 1). One is by splitting the category of 'endurable' into two, which we might call 'endurable and minor' (imperceptible or trivial) and 'endurable but severe'. The reason is that risks which are not fully catastrophic or terminal can still have very extensive and long-lasting effects as well as being terminal for large numbers of people and for a considerable range of future possibilities. In addition, the border between terminal and 'endurable but severe' is a matter of judgement in some cases. Most importantly, though, many of the events that do count as GCRs are more severe or larger-scale versions of ones that fall into the 'endurable but severe' category and therefore derive from the same underlying cause or phenomenon. This means that acting to protect against one risk may also protect against another risk. In terms of the table of types of risk, this means three rows rather than two.

We can also add another category to severity by bringing in the concept of permanent or at least multi-generational

effects. This incorporates the third criterion of duration and brings virtual losses, the non-existence of things that would have existed, or might have existed, into play. This means four categories on the vertical axis, from endurable and minor to permanent. The result, as Table 2 shows, is twelve categories. The four shaded boxes are those we need to consider, in particular the two boxes at the top of the global column: global catastrophic risks, as defined earlier, and global existential risks (GERs), which are an especially severe subset of the wider set of GCRs. So, that brings us back to the question of the probability of an event happening and in particular the probability of an event happening that is both global and terminal – clear thinking about statistics and probability is required.

Table 2

Multi-generational/ permanent	Death of family	Genocide, part of world made permanently uninhabitable	Human extinction, GERs
Terminal	Death	Post-Columbian epidemics, civilisational collapses	End of all civilisations, GCRs
Endurable and severe	Severe injury, permanent disablement	Civil war, economic crisis	World war, severe global depression
Endurable and minor	Minor injury	Local recession, violence	Global recession
	Personal	**Local**	**Global**