



Four New Horsemen of an Apocalypse? Solar Flares, Super-volcanoes, Pandemics, and Artificial Intelligence

Ilan Noy¹ · Tomáš Uher¹

Received: 14 September 2021 / Accepted: 3 January 2022 / Published online: 15 January 2022
© The Author(s) 2022

Abstract

If economists have largely failed to predict or prevent the Global Financial Crisis in 2008, and the more disastrous economic collapse associated with the pandemic of 2020, what else is the profession missing? This is the question that motivates this survey. Specifically, we want to highlight four catastrophic risks – i.e., risks that can potentially result in global catastrophes of a much larger magnitude than either of the 2008 or 2020 events. The four risks we examine here are: Space weather and solar flares, super-volcanic eruptions, high-mortality pandemics, and misaligned artificial intelligence. All four have a non-trivial probability of occurring and all four can lead to a catastrophe, possibly not very different from human extinction. Inevitably, and fortunately, these catastrophic events have not yet occurred, so the literature investigating them is by necessity more speculative and less grounded in empirical observations. Nevertheless, that does not make these risks any less real. This survey is motivated by the belief that economists can and should be thinking about these risks more systematically, so that we can devise the appropriate ways to prevent them or ameliorate their potential impacts.

Keywords Solar · Volcano · Pandemic · Artificial intelligence · Extinction · D81 · O33

JEL D81 · O33

The Existential Risks we Ignore

“So in summary, Your Majesty, the failure to foresee the timing, extent and severity of the crisis [of 2008] and to head it off, while it had many causes, was principally a failure of the collective imagination of many bright people, both in this country and internationally, to understand the risks to the system as a whole.” This excerpt is from a letter written by members of the British Academy to Queen Elizabeth, in response to the question she famously posed at the London School of Economics in the aftermath of the 2008 Global Financial Crisis: Why did nobody see it coming? (Besley and Hennessy 2009).

✉ Ilan Noy
ilan.noy@vuw.ac.nz

¹ Victoria University of Wellington, Wellington, New Zealand

Disasters have been the subject of social science research for some time (O’Keefe et al. 1976). Catastrophic events, ones that have the potential to extinguish the human species are also not inevitable ‘natural’ events. Indeed, the insurance term ‘Acts of God’ to describe acts that are not within the immediate power for us to prevent are mis-labelled. Disasters occur because of our actions or inactions. Epidemics are no different to earthquakes or volcanic eruptions, and their impact is also largely dependent on how we prepare and react to them. And yet, practically every country was unprepared for the COVID-19 pandemic that started in early 2020.

If economists have failed to predict or prevent the Global Financial Crisis in 2008, and the more disastrous pandemic of 2020, what else is the profession missing? This is the question that motivates this survey. Specifically, we want to highlight four catastrophic risks – i.e., risks that can potentially result in global catastrophes of a much larger magnitude than either the 2008 or 2020 events. Inevitably, and fortunately, these catastrophic scenarios have not yet occurred, so the literature investigating them is by necessity more speculative and less grounded in observations. Nevertheless, that does not make these risks any less real, any less dangerous, or any less worthy of our attention.

The four risks we examine here are space weather and solar flares, super-volcanic eruptions, high-mortality pandemics, and misaligned (or malignant) artificial intelligence. All four have a non-trivial probability of occurring and all four can potentially, and disastrously, lead to a catastrophe not very different from human extinction. As Weitzman’s Dismal Theorem suggests, the basic approach to economic assessments of risks fails for these extremely catastrophic ones (Weitzman 2009). “There is a race being run in the extreme tail between how rapidly probabilities are declining and how rapidly damages are increasing. Who wins this race, and by how much, depends on how fat (with probability mass) the extreme tails are.” (Weitzman 2011). As Weitzman points out, it is often difficult to assess how fat these tails are, because they represent sets of events that are far outside the realm of our concrete experiences. Still, this survey is motivated by the belief that economists can and should be thinking about these risks more systematically, so that we can devise appropriate ways to prevent them or ameliorate their potential impacts.

Space Weather and Solar Flares

Active regions on the surface of the sun can produce solar flares, which are jets of solar energy and coronal mass ejections (sudden release of plasma accompanied by a magnetic field) (Kahler 1992). These phenomena are produced when local regions of the sun’s magnetic field suddenly change configuration (Knipp and Biesecker 2015) and are typically interconnected, occurring within a relatively short period of time. The resulting electromagnetic radiation can create disruptions of the Earth’s magnetic field (termed geomagnetic storms), commonly known as “space weather events” (Schwenn 2006).

The most important implication of these events is in terms of disruptions to electricity-powered technology. Due to the global economy’s increasing reliance on such technologies, it has grown increasingly more vulnerable to the impacts of space weather.¹ An extended

¹ For example, while the 1998 five-week power outage in Auckland, New Zealand was not caused by a space weather event and was limited in its scale to inhabitants of a single city, it demonstrated the severity of the consequences that a large-scale collapse of the power grid may induce. The Auckland power outage was caused by a failure of power cables (Rennie et al. 1998).

power outage that lasts for weeks or months, affecting a large population (potentially tens of millions), can easily become a major catastrophe, and even a systemic threat to society.

Societal Impacts of Severe Space Weather and Solar Flares

Space weather events have been associated with a multitude of negative, mainly technological, direct economic consequences such as blocked radio communications, satellite damage and malfunctions, disruption to rail networks and wireless networks, and global navigation systems such as GPS (Cannon et al. 2013; Eroshenko et al. 2010). However, the most consequential possible effect is extensive damage to electricity transformers and therefore potentially long-term disruption to the electricity transmission infrastructure (Khurshid et al. 2020; Kappenman 2012). Such a disruption would cause severe problems with the supply of many basic services, including the access to pumped potable water, and the loss of all perishable foods and medications that depend on refrigeration (National Research Council 2009; Kappenman 2012).

The observed history of space weather in modern economies is very limited. The biggest geomagnetic storm of the last few decades happened in 1989 and caused transformer damage in multiple countries and an approximately 9 h-long power outages for over 6 million inhabitants of Quebec (Lakhina et al. 2004; Barnes and Van Dyke 1990).

The most severe directly observed events were the storms of 1859 (the Carrington Event) and 1921, which were roughly three times as powerful as the 1989 event and are typically considered to be a 1 in a 100-year events (Kappenman 2010). However, since economies at the time were not so dependent on electricity (especially, of course, in 1859), the impact of both these events was much less severe than the impacts of the 1989 event and consisted mostly of disruption to the telegraph service (Boteler 2006).

While these historical events are associated only with relatively mild societal impacts, a similar event today would have severe global consequences. Many of the studies attempting to analyse the future potential impacts of such an event focus on the economic effects associated with major power outages. A study by Kappenman (2010) estimates that a Carrington storm hitting the United States could put more than 200 large power transformers at risk of permanent damage and cause severe damages to the power grid, leading to long-lasting (months or potentially years) blackouts for approximately 130 million people in the US and a full recovery time of 4–10 years. The total cost of a long-term power outage affecting a significant area is estimated at USD 1–2 trillion during the first year (National Research Council 2009).

In a report for Lloyd's, Maynard et al. (2013) points out that the duration of the power outage depends largely on the availability of spare transformers. These authors estimate that, in the worst-case scenario of no spare transformers, a Carrington-like event would lead to a power outage in the US affecting 20–40 million people and lasting 5 months. They estimate an associated economic loss of USD 0.5–2.6 trillion.²

Moran et al. (2014) analyse the economic impacts of space weather from a global perspective and conclude that “a severe space-weather event could be the worst natural disaster in modern history” (p. 8). Assuming a power outage lasting 1 year for an event of a magnitude between the 1989 storm and the Carrington event (with annual probability of occurrence likely to be higher than 1%), the authors conclude a major

² The data and methodology on which this estimate is based are not disclosed.

disruption of global supply chains affecting all industries and estimate a global economic loss of up to USD 3.4 trillion (5.6% of global GDP) in the first year.

Oughton et al. (2016) find that an extended power outage in the US caused by a similar space weather event would lead to global economic losses with respect to global supply chain disruptions valued at USD 0.5–2.7 trillion. Using a globally integrated economic model to account for the post-event dynamic responses of global trade, they estimate a decrease of global GDP by up to USD 1.1 trillion over a five-year period. The economic loss to the US manufacturing industry is estimated to be USD 350 billion with approximately half of the losses being indirect (roughly equally split between the losses associated with upstream and downstream disruptions to the supply chain). The losses to the US insurance industry are estimated to be up to USD 334 billion, with 90% of the losses caused by service interruption within property insurance policies.

In another paper, Oughton et al. (2017) suggest that a severe geomagnetic storm, causing a power outage for 66% of the population in the US, would create a daily economic loss of USD 41.5 billion. Approximately half of the total economic loss is estimated to be inflicted indirectly outside of the blackout zone due to supply chain disruptions; Moran et al. (2014) reach similar conclusion for the global economy.

Oughton et al. (2019) estimate an economic impact of a significant geomagnetically-induced power grid failure in the UK and distinguish between scenarios in terms of the ability to forecast the event. They find that a 1 in a 100-year event with the current level of forecasting would cause a GDP loss of GBP 2.9 billion in the UK, but an enhanced forecasting ability based on further investment in this technology would bring the loss down significantly, to GBP 0.9 billion.

In all these modelling exercises, power outages for a significant portion of the population are assumed to lead to cascading effects over many sectors. National Research Council (2009) emphasised banking and finance, government services and emergency response. Eventually, however, all economic sectors would be adversely affected (Moran et al. 2014; Riley et al. 2018). Apart from the sectoral vulnerability to power outages due to the reliance on electricity, Moran et al. (2014) note that the global economic production system today is made even more vulnerable due to the common use of practices such as just-in-time production, reduced inventories, and increased reliance on long-distance supply chains with many links.

More difficult to model, but maybe not less important to evaluate, is the public response to such an event. Hapgood et al. (2021) suggest that a power grid disruption caused by a severe space weather event would lead to panic buying and stockpiling of essential goods such as petrol, bottled water, non-perishable foods, and toilet paper. There could be various flow-on effects from this panic. Ultimately, the longer-term economic consequences of that are unclear.

It is uncontroversial that an extended power outage for a significant portion of the population would lead to disastrous economic impacts. However, the potential of space weather events of the magnitudes discussed above to cause such extended power outages are questioned by studies such as NERC (2010, 2012) and Cannon et al. (2013). They argue that extensive power grid hardware damage is unlikely, and the more probable consequence is a temporary system collapse due to voltage instability. In their view, this will result in only short-term stoppages in electricity supply and thus modest economic consequences.

Superflares

Research focusing on the risk of even more severe space weather events is scarce, partially because there have so far been no direct observations of “superflares” in our solar system. Some evidence suggests that a superflare 100 times stronger than the Carrington event may have happened in AD 775 (Melott and Thomas 2012; Usoskin et al. 2013; Mekhaldi et al. 2015) and may have led to regional changes in the Earth’s surface temperature (Sukhodolov et al. 2017). The nature of this AD 775 event, however, is inconclusive (Cliver et al. 2014; Stephenson 2015; Neuhäuser and Neuhäuser 2015). Another extremely powerful event possibly happened in AD 993 (Miyake et al. 2013; Mekhaldi et al. 2015). Furthermore, astronomical observations of other sun-like stars in our galaxy suggest that such events are indeed possible (Maehara et al. 2012; Nogami et al. 2014).

Lingam and Loeb (2017a) propose that the most powerful superflares might have been the cause of some of the previous mass extinction events and that a very rare superflare with energy 100,000 times larger than the Carrington event might be able to destroy the ozone layer and lead to widespread destruction of ecosystems with potentially existential consequences. In a follow-up paper, Lingam and Loeb (2017b) suggest that the societal vulnerability to superflares in terms of economic damage is increasing rapidly, due to the growth of technological infrastructure. They propose a mitigation strategy of setting up a protective shield between the Earth and the Sun to avoid these dire consequences.

Probability

In the academic literature, a Carrington-like space weather event is typically referred to as about a 1 in a 100-year event (Riley 2012; Riley and Love 2017). However, Morriña et al. (2019) find this probability of occurrence to be an order of magnitude lower at 0.46%–1.88% over a ten-year period.

For more severe space weather events, Usoskin and Kovaltsov (2012) estimate an annual probability of 0.1–1%, but for an event 100 times more powerful. Maehara et al. (2012) note that flares of such magnitude occur on sun-like stars every 800 years, similarly to Shibayama et al. (2013). While Shibata et al. (2013) further suggest that this frequency applies for the Sun as well, Karoff et al. (2016) argue that the frequency for the Sun might be lower by an order of magnitude due to the relatively lower solar activity compared to other stars. Fortunately, Lingam and Loeb (2017a) propose that an extinction-level superflare with energy levels 100,000 times larger than the Carrington event are much rarer. These might occur with a frequency of once in 20 million years.

Conclusion

Society’s vulnerability to the adverse impacts of severe space weather (such as a Carrington event) has increased dramatically during the twentieth and twenty-first centuries due to our rapidly increasing reliance on electricity, economic globalisation, and other modern economic practices (e.g., just-in-time production with long supply chains). These events now have the potential to create globally destabilizing effects and economic damages worth trillions of dollars due to the extended power outages and other effects they may initiate (such as massive migrations, or local and international conflicts over scarce resources).

Within the regions directly affected by the power outage, all economic sectors could be affected. However, adverse economic effects would also manifest outside of the directly affected areas due to global supply chain disruptions, changing demand patterns, and shifts in the location of people. These can potentially account for up to half of the total economic loss. The economic losses caused by these events may be influenced by factors such as availability of spare transformers or the level of our forecasting ability with respect to their occurrence. However, the potential of these events to create extended power outages across vast regions remain debated and require a lot more research, not only by physicists, but especially by economists and other social scientists.

Super-volcanoes

Super-volcanic eruptions are typically defined as eruptions of magnitude 8.³ These volcanic explosions are tens or hundreds of times larger than eruptions observed in the recent millennia and have potential to create global or even existential catastrophic outcomes (Rampino 2008). These large eruptions are likely to generate many local and global negative consequences. After a large eruption, the surrounding area of possibly hundreds of km² can be covered with falling rock and magma, a thick layer of ash can cover continental amount of land and massive amounts of volcanic ash and gases might be released into the atmosphere (Self 2006). The literature on the climatic and societal impacts of such eruptions is not as sparse as for solar flares, given the richer historical record which includes large volcanic eruptions of magnitude ~7. A pertinent example is the eruption of Tambora in 1815, which led to a global ‘year without a summer’ and had an estimated magnitude of 6.9; but even in this case, most of this literature is not interested in the super-eruptions that are the focus here (Choumert-Nkolo et al. 2021).

The main, global, and potentially existential threat of a major eruption lies in the adverse climatic effects it can induce. Large volcanic eruptions of magnitudes ~7 and higher can cause significant atmospheric cooling, lasting for years or potentially decades. These are often referred to as “volcanic winters” (Newhall et al. 2018; Self 2015, 2006; Jones et al. 2005; Gathorne-Hardy and Harcourt-Smith 2003; Handler 1989; Timmreck and Graf 2006). The global drop in temperatures is caused by the injection of volcanic ash and sulphate aerosols into the stratosphere which block the incoming sunlight. The duration of the climatic cooling may be extended due to feedback loops of ocean cooling and ocean and land ice formations further cooling the atmosphere by reflecting sunlight (Rampino and Self 1992, 1993).⁴ Super-volcanic eruptions also appear to have a negative effect on the ozone layer, increasing the amount of life-endangering ultraviolet radiation that reaches the Earth’s surface (Xu et al. 2019; Kutterolf et al. 2013).

Studies of large historical eruptions and their climatic impacts provide relevant insights. The eruption of Mount Tambora in Indonesia in 1815 was one of the biggest known eruptions in the last two thousand years, releasing approximately 30 km³ of magma (Self et al. 2004). While typically not considered a super-volcanic eruption, it is estimated to have caused severe climatic effects of global atmospheric cooling of 1 °C (Robock 2005;

³ A volcanic explosion’s magnitude is typically measured by the Volcanic Explosivity Index (VEI). It ranges from 0 (a non-explosive eruption) to 8, based mainly on the volume of erupted material (Newhall and Self 1982).

⁴ This is exactly one of the most plausible global warming feedback loops, reversed.

Rampino 2008; Rampino and Self 1982) and the following year has often been referred to as “the year without summer”. Major climatic impacts were also identified for possibly the strongest volcanic eruption of the last two millennia in Rinjani, Indonesia in 1257 (estimated magnitude 7) (Stothers 2000; Oppenheimer 2011; Newhall et al. 2018).

The super-volcanic eruption of Mount Toba in Indonesia around 74,000 years ago (magnitude 8) is considered one of the greatest known volcanic events in geological history. It is estimated to have produced severe and long-lasting global cooling of 3–5 °C (Rampino and Self, 1992, 1993). However, these findings about the longevity and the severity of the climatic impacts are disputed by Timmreck et al. (2010) and Lane et al. (2013).

Economic Impacts of Super-volcanic Eruptions

Decreased temperature, sunlight, and precipitation have a negative impact on the photosynthetic ability of plants and therefore vast and cascading agricultural, ecological, environmental, and other implications (Robock 2000). Furthermore, associated cooling of the oceans can lead to the weakening of ocean circulation and produce further weather anomalies (Rampino 2008). The resulting disruption to agriculture, crop yields and global food production will be quite sudden, and is therefore considered to be the most consequential effect of large volcanic eruptions.

Severe agricultural impacts were shown to have followed the 1815 Tambora eruption. The evidence suggests that the climatic effects of a temperature drop and decreased precipitation in monsoon regions led to agricultural productivity decline and crop failures on multiple continents (Raible et al. 2016; Brönnimann and Krämer 2016; Oppenheimer 2003). This resulted in an increase in food prices and in combination with political turmoil led to famine and food scarcity along with epidemic disease spread in Europe, the US and China (Brönnimann and Krämer 2016; Oppenheimer 2003). Some observers suggest that the famine crisis led to modernizations of the agricultural sector that consequently increased long-term agricultural productivity (Brönnimann and Krämer 2016). The same observers suggest that, in China, these effects may have also led to a shift from subsistence crops to opium poppy farming because of its higher tolerance for dryer conditions and its higher cash value (Brönnimann and Krämer 2016).

Others suggest that the disruptions caused by the Tambora eruption possibly contributed to societal and economic changes characterized by social unrest, the Cholera pandemic, and the economic crisis in the US, and finally led to the emergence of the early welfare state (Wood 2014). In contrast, other researchers propose that the social and economic upheavals caused a political shift to the right instead (e.g., Post 1977). Nevertheless, these conflicting accounts suggest that the eruption was not the primary driver, but a contributing factor to social change.

For the 1257 Rinjani eruption, the evidence is obviously more tentative. Some, however, do suggest that the eruption’s climatic effects and the associated agricultural decline contributed to famine and high mortality in England and Japan and to disease spread in Europe and the Middle East (Stothers 2000; Guilett et al. 2017; Campbell 2017). Furthermore, it has been argued that this event may have contributed to the decline of the Mongol and Byzantine empires (Oppenheimer 2011; and Xoplaki et al. 2016; respectively).

Today, in contrast with the 19th or thirteenth centuries, society is more exposed to a major eruption since the world is more densely populated. We are also more vulnerable because we increasingly rely more on technology, and because of the interconnectedness of the global economy (Newhall et al. 2018). A super-volcanic eruption would likely cause

global food supply disruptions due to a loss of one or more growing seasons (Puma et al. 2015; Toon et al. 1997). Such a decline in food production could easily surpass the amount of stored food worldwide and lead to large-scale famine, unlike local food system shortages due to a hazard event, that could almost always be mitigated by imports from elsewhere, when such trade is permitted.⁵

Rampino (2002) notes that the Asian rice crop could be destroyed by a single night of freezing temperatures during the growing season and a 3–4 °C temperature drop would destroy the Canadian grain production. A similar destruction can also be wrought on cold-sensitive tropical vegetation (Williams 2012). In any case, the overall effect on plant life would definitely create a major ecological and environmental crisis (Self 2015). Besides famines, these dynamics can also lead to the spread of specific infectious diseases, social unrest, and violent conflict over resources (Rampino 2008; Stothers 2000).

The deposited ash layer, potentially covering a continent-sized landmass, would also negatively affect infrastructure, transportation, electricity transmission, and leave large areas with contaminated water sources (Donovan and Oppenheimer 2014; Self 2015). Atmospheric ash would severely disrupt aviation, GPS systems and other satellite services and telecommunication (Newhall et al. 2018; Self 2015). Furthermore, the disruptions to telecommunications and aviation would hinder the emergency response (Rymer et al. 2005).

Probability

Assessments of return periods for large volcanic and super-volcanic eruptions are very difficult due to the incomplete geological records and consequently the potential to underestimate their probability (Deligne et al. 2010). With respect to magnitude 7 eruptions, there have been at least three such eruptions in the last two millennia (Plag et al. 2013). It has been estimated that on average these eruptions happen once in ~780 years (Deligne et al. 2010) or ~1,200 years (Rougier et al. 2018).

Super-volcanic eruptions of magnitude 8 or higher are fortunately much rarer. Previously, it was estimated that these events happen on average once in 100,000–200,000 years (Mason et al. 2004). However, a recent study by Rougier et al. (2018), to better account for the under-recording bias, estimates a return period of these events to be significantly lower at 17,000 years. Ord (2020) suggests that the likelihood of an existential catastrophe caused by a super-volcanic eruption within the next 100 years may be on the order of 0.01%.

Conclusion

A large volcanic eruption of magnitude 7 would likely cause global catastrophic outcomes due to its climatic effects and consequent disruptions of global food production, infrastructure, and communication. A super-volcanic eruption of magnitude 8 or higher could potentially lead to a temporary or (in the worst case) an irreversible collapse of civilization or human extinction (Plag et al. 2013; Ord 2020). Apart from these direct effects, the volcanic-induced societal disruptions could lead to wars and conflicts and thus exacerbate or even create other existential risks.

⁵ In the past, food shortages affecting multiple regions were also sometimes accompanied by prohibitions on exports (and sometimes even by stockpiling by exporting countries).

While some scientists are exploring ways to prevent super-volcanic eruptions (Wilcox et al. 2015), currently the best approach to minimize their potential impacts seems to be monitoring, prediction and preparation of food reserves and the development of alternative food production methods (Plag et al. 2013; Denkenberger and Pearce 2016). It appears that, in the case of already identified volcanoes, super-eruptions may be predictable some time in advance (Newhall and Dzurisin 1988). However, it is likely that some super-volcanoes remain unidentified and are thus unmonitored. Plag et al. (2013) calculate that a global volcano monitoring system (with an estimated cost of USD 370 million per year) would have an expected economic benefit of at least ten times the total cost, but potentially hundreds or thousands of times. Such an estimate seems to be of even greater importance in light of the relatively recent higher frequency estimate of super-eruptions by Rougier et al. (2018).

Compared to the large historical volcanic eruptions discussed here, today's scientists have developed a much improved ability for early detection of these events. Early warning systems are now regularly leading to pre-emptive evacuations in various volcanic hot spots (as in Indonesia or the Philippines). However, more investment is necessary, especially in areas in which large populations make emergency evacuations very challenging (e.g., for Yogyakarta or Manila).

Despite the severity of the potential outcomes, the overall risk posed by these events may still be underestimated. Large uncertainties remain. More research is needed to determine the specific climatic effects of super-volcanic eruptions and establishing which magnitudes or threshold may lead to catastrophic risk. With respect to the economic effects of large eruptions, further research of their implications for global and local economies, economic output, supply chains and the impacts on socio-economic outcomes would be valuable. Equally, research on any possible ways to ameliorate impacts (such as with food storage) should be pursued more rigorously.

High-Mortality Pandemics

A naturally occurring pandemic (i.e., not from an engineered pathogen) that would threaten human extinction is a very small probability event. However, historical accounts point to several instances where disease spread played an important role in causing very significant decline of specific populations. For example, the introduction of novel diseases to the Native American population during the European colonization of the Americas had deadly consequences. It is difficult to distinguish the effects of the diseases that came with the Europeans from the war and conflict they also brought with them. Nevertheless, during the first hundred years of the colonization period, the American population may have been reduced by as much as 90% (Ord 2020).

Moreover, two major pandemic events, the Justinian Plague in the sixth century and the Black Death in the fourteenth century appear to have been severe enough to cause a significant population decline of tens of percent in the populations they affected. Both events are believed to have been caused by plague, an infectious disease caused by the bacteria *Yersinia Pestis* (Christakos et al. 2005; Allen 1979). While there is a certain degree of uncertainty involved in studying these events' societal impacts, historical accounts in combination with modern scientific methods provide us with some valuable insights into the effects they may have had on the societies of the time.

With respect to the possibility of a future catastrophic global pandemic, it appears that this risk is increasing significantly along with the advances in the field of synthetic

biology and the rising possibility of an accidental or intentional release of an engineered pathogen. While some of the scientific efforts in the field of synthetic biology are directed towards increasing our understanding and our ability to prevent future catastrophic epidemic threats, the risk stemming from these activities is non-trivial, and may outweigh their benefits.

The Justinian Plague

The Justinian Plague severely affected the people of Europe and East Asia, though estimates of its overall mortality vary. Focusing exclusively on the first wave of the pandemic (AD 541–544), Muehlhauser (2017) suggests the pandemic was associated with a 20% mortality in the Byzantine empire. This estimate is based on the mortality rate estimated for the empire's capital, Constantinople, by Stathakopoulos (2007) to produce a death toll of roughly 5.6 million. For a longer time span, AD 541 to 600, which included subsequent waves of the plague, scholars estimate a higher mortality rate of 33–50% (Allen 1979; Meier 2016).

The demographic changes associated with this high mortality led to a significant disruption of economic activity in the Byzantine empire (Gârdan 2020). A decline in the labour force caused a decline in agricultural production which led to food shortages and famine (Meier 2016). Trade also collapsed. Decreased tax revenues caused by the population decline initiated a major fiscal contraction and consequently a military crisis for the empire (Sarris 2002; Meier 2016). In the longer run, however, the massive reduction of the labour force appears to have had a positive economic effect for the surviving laborers, as the increased marginal value of labour caused a rise in real wages and per capita incomes. These beneficial effects for the survivors were also observed after the Black Death (Pamuk and Shatzmiller 2014; Findlay and Lundahl 2017).

The mortality and the disruption of activity the plague caused in the Byzantine empire also led to further direct and indirect cultural and religious consequences. Meier (2016) particularly highlights the plague's indirect effect of an increase in liturgification (a process of religious permeation and internalization throughout society as defined by Meier 2020), the rise of the Marian cult, and the sacralization of the emperor.

The direct and indirect effects of the plague also appear to have had far-reaching and long-term political repercussions. The societal disruptions caused by the plague are believed to have significantly weakened the position of the Byzantine empire and arguably led to the decline of the Sasanian empire (Sabbatani et al. 2012). Interestingly, the pandemic indirectly favoured the nomadic Arab tribes who were less vulnerable to the contagion while traveling through desert and semi-desert environments during the initial expansion of Islam (Sabbatani et al. 2012).

Of note is the absence of a scientific consensus on the severity of the Justinian Plague's impacts. For example, Mordechai and Eisenberg (2019) and Mordechai et al. (2019) argue against the maximalist interpretation of the historical evidence described above. They suggest that the estimated mortality rate of the plague is exaggerated, and that the pandemic was not a primary cause of the transformational demographic, political and economic changes in the Mediterranean region between the sixth and eighth century. Recently, White and Mordechai (2020) highlighted the high likelihood of the plague having different impacts in the urban areas of the Mediterranean outside of Constantinople.

The Black Death

The Black Death which ravaged Europe, North Africa, and parts of Asia in the middle of the fourteenth century is considered the deadliest pandemic in human history and potentially the most severe global catastrophe to have ever struck mankind. With respect to its mortality, Ord (2020) argues that the best estimate of its global mortality rate is 5–14% of the global population, largely based on Muehlhauser (2017).

The plague created a large demographic shock in the affected regions. It reduced the European population by approximately 30–50% during the 6 years of its initial outbreak (Ord 2020). It took approximately two centuries for the population levels to recover (Livi-Bacci 2017; Jedwab et al. 2019b). As the mortality rates appear to have been the highest among the working-age population, the effects on the labour force were acute (Pamuk 2007).

The plague's mortality, morbidity and the associated societal disruption led to a major decline in economic output both in Europe (Pamuk 2007) and the Middle East (Dols 2019). In Europe, however, this decline in economic output was smaller than the decline in population; output per capita began to increase within a few years of the initial outbreak (Pamuk 2007).

The large demographic shock caused by the plague led to a shift in the relative price of labour which, similarly to the Justinian Plague, had a positive impact on wages. With a reduced labour force, real wages and per capita incomes in many European countries increased and were sustained at higher levels for several centuries (Voigtländer and Voth 2013a; Jedwab et al. 2020; Pamuk and Shatzmiller 2014). Scott and Duncan (2001) point out that real wages approximately doubled in most countries of Europe in the century following the plague.

An additional insight into the long-run relationship between the Black Death's mortality and per capita incomes in Europe is offered by Voigtländer and Voth (2013a). Using a Malthusian model, they suggest that over time, the rise in income caused by the plague's mortality led to an increase in urbanization and trade. Furthermore, the increased tax burden (per capita), combined with the contemporary political climate, increased the frequency of wars. Consequently, higher urbanization and trade led to an increase in disease spread which along with a more frequent war occurrence caused a long-term increase in mortality and a further positive effect on per capita incomes. In this way, the Black Death appears to have created a long-lasting environment of high-mortality and high-income specifically in Western Europe, functioning as an important contributing factor to its economic growth in the next centuries (Alfani 2020). However, while in Western Europe incomes remained elevated over the next centuries, in Southern Europe they began to decline as the Southern European population started recovering after AD 1500 (Jedwab et al. 2020).

Apart from the positive effects on wages, the increased marginal value of labour combined with other factors had further economic and social implications. A decreased relative value of land and the lack of workforce to use it effectively caused land prices and land rents to decrease (Jedwab et al. 2020; Pamuk 2007). A decreased marginal value of capital assets in general led to a lapse in the enforcement of property rights (Haddock and Kiesling 2002). Interest rates and real rates of return on assets also decreased (Pamuk 2007; Jedwab et al. 2020; Pamuk and Shatzmiller 2014; Jordà et al. 2021; Clark 2016).

Higher wages in combination with a relative abundance of land increased people's access to land/home ownership, likely reducing social inequality (Alfani 2020). On the

other end of the income distribution, decreased incomes for landowners led to an overall decrease in income inequality (Jedwab et al. 2020; Alfani and Murphy 2017).

With respect to the effects on agriculture, the structure of agricultural output moved away from cereals to other crops following the plague. Furthermore, the workforce shortages and the incentives to increase the labour supply are believed to have caused a shift from male-labour intensive arable farming towards pastoral farming, consequently raising the demand for female labour (Voigtländer and Voth 2013b). However, while the Black Death appears to have caused certain structural agricultural changes, Clark (2016) finds no effect of the plague on agricultural productivity in the long run.

In terms of other social consequences, the evidence suggests that the plague's mortality reduced labour coercion, particularly throughout Western Europe (Jedwab et al. 2020; Haddock and Kiesling 2002; Gingerich and Vogler 2021). The increased bargaining power of labour caused by the plague's demographic shock contributed to and accelerated the decline in serfdom and development of a free labour regime. Gingerich and Voler (2021) further argue that these effects may have had long-lasting political implications and that a decline of repressive labour practices (such as serfdom) permitted the development of more inclusive political institutions. They find that the regions with the highest mortality were more likely to develop participatory political institutions and more equitable land ownership systems. They find that centuries later, in Germany, the populations in these high-mortality regions were less likely to vote for Hitler's National Socialist (Nazi) Party in the 1930 and 1932 elections in Germany.

However, the positive effects on the emergence of freer labour did not take place in Eastern Europe, where serfdom was sustained and even intensified. Robinson and Torvik (2011) attempt to explain this asymmetry arguing that these differential outcomes may have been caused by the varying power and quality of institutions. The authors suggest that opportunities generated by the increased bargaining power of labour, in an environment of weak institutions, were less likely to lead to a positive effect than in the case of regions with stronger institutions (with more robust rule-of-law or less corrupt or predatory practices).

Apart from causing a negative demographic shock to the affected populations, the Black Death appears to have caused further indirect demographic changes, particularly in Western Europe. The increased employment opportunities for females caused by worker shortages and a higher female labour demand led to a decline in fertility rates and an increased age of marriage (Voigtländer and Voth 2013b). This demographic transition to a population characterized by lower birth rates likely helped to preserve the high levels of per capita incomes and contributed to further economic development of certain parts of Europe, enabling it to escape the "Malthusian trap" in the following centuries (Pamuk 2007). Siuda and Sunde (2021) confirm the pandemic's effect on the accelerated demographic transition empirically, as they find that greater pandemic mortality was associated with an earlier onset of the demographic transition across the various regions of Germany.

Unfortunately, the Black Death also led to an increase in the persecution of Jews (Finley and Koyama 2018; Jedwab et al. 2019a). Interestingly, Jedwab et al. (2019a) were able to estimate that in the case of regions with the highest mortality rates, the probability of persecution decreased if the Jewish minority was believed to benefit the local economy.

It is important to highlight that the long-term repercussions of the Black Death were highly asymmetrical. While in Western Europe the pandemic appears to have led to some long-term dynamic shifts associated with increased wages, decreased inequality and a decrease in labour coercion, this was not the case for other regions. A decrease in wages was observed for example in Spain (Alfani 2020) and Egypt. In Spain, the plague's

demographic impact on an already scarce population caused a long-lasting negative disruption to the local trade-oriented economy. The workforce disruption in Egypt led to a collapse of the labour-intensive irrigation system for growing crops in the Nile valley, with consequent disastrous effects on the rural economy (Alfani 2020). Borsch (2005) argues that the economic decline in Egypt caused by the Black Death “put an end to the power in the heartland of the Arab world” (p. 114) and to the impressive scientific and technological developments that came out of this region.

A consensus for an explanation of the Black Death’s varied impacts across regions, and their determinants, does not appear to exist. However, several researchers attempt to provide partial insights. For example, Alfani (2020) considers the differential outcomes to be broadly dependent upon the initial conditions in each region. More specifically, both Robinson and Torvik (2011) and Pamuk (2007) propose that the asymmetry of impacts can largely be explained by the differences in the institutional environments of the affected societies.

It is argued that the Black Death defined the threshold between the medieval and the modern ages, similarly to the way the Justinian Plague did for antiquity and the Middle Ages (Horden 2021). Furthermore, the differential long-term outcomes of the Black Death likely provided a significant contribution to the so-called “Great Divergence” between Europe and the rest of the world and the “Little Divergence” between North-western and Southern and Eastern Europe (Jedwab et al. 2020; Pamuk 2007).

From this perspective, it would seem rational to conclude that apart from causing substantial and long-term demographic, economic, political, and cultural changes, both the Justinian Plague and the Black Death likely significantly altered the course of human history.

Considering the above, it is not unreasonable to expect that a pandemic of a similar magnitude to these past catastrophes would do the same in the present day. However, what societal impacts a pandemic of similar or higher mortality would inflict in the twenty-first century has not really been the subject of any study, as far as we were able to identify. A possibility exists, given the newly developed capacity of humanity to create new pathogens, that the outcomes of a future catastrophic pandemic will be even more adverse than those of the Justinian Plague and the Black Death.

Probability

In terms of the probability of naturally occurring pandemics, an informal survey of participants of the Global Catastrophic Risk Conference in Oxford in 2008 shows that the median estimate for a probability of a natural pandemic killing more than 1 billion people before the year 2100 was surveyed to be 5%, and the probability of such pandemic to cause human extinction was 0.05%. Ord (2020) uses a slightly broader definition of existential risk, which apart from human extinction also includes a permanent reduction of human potential. He estimates the probability of an existential risk stemming from a natural pandemic in the next 100 years to be 0.01%.

Probability estimates for a pandemic caused by an engineered pathogen are significantly higher. These can occur due to an accidental leak during gain-of-function research creating a deadly pathogen, or even because of an intentional release of an engineered pathogen by a malicious individual or group. In Sandberg and Bostrom (2008), the perceived risk of human extinction caused by an engineered pandemic is believed to be 2% and the risk of an engineered pandemic causing more than 1 billion deaths to be 10% before 2100. Ord

(2020) estimates the probability of an existential risk in the form of an engineered pandemic in the next 100 years to be slightly higher at 3% (300 times higher than his estimate for a natural pandemic). Klotz and Sylvester (2014) find that the general probability of a pandemic caused by a laboratory safety failure could be as high as 27% for a period of 10 years, suggesting there is a significant risk associated with this kind of research.

Conclusion

Severe historical pandemic events may have significantly altered the course of history and were associated with a multitude of consequential and long-lasting societal effects. The mortality and morbidity associated with these events led to demographic, economic, political, and cultural changes and these effects manifested differently over affected regions. Of note is the impact of the Black Death's mortality on the long-run per capita incomes and its likely relevance for the relatively higher long-run economic development of Western Europe compared to other regions. If true, the further implications of this development for the historical events of the next centuries and the present-day structure of our civilization may be hard to over-estimate.

We note the absence of research on societal effects of a plausible present-day high-mortality pandemic event (both natural and engineered) that could be associated with global catastrophic or existential risks. The dynamics of these effects would almost certainly be considerably different from the ones reported for the historical pandemics of the medieval era, due to the qualitatively different state of our technology, medicine, and social organisation. The developments in the past several hundred years have arguably both increased our ability to treat pandemic risks (with modern medicine, disease surveillance, better coordination ability, and capacity to develop new vaccines and treatments) and decreased it (increased population density, more international trade and travel, and more intensive farming). However, the combined effect of these factors is most likely, on balance, favourable.

The dramatic impacts of the COVID-19 pandemic in 2020–2021, caused by a relatively low-risk-of-mortality pathogen (when compared to these past events), suggest that a high-mortality pandemic would still likely be trajectory-changing for humanity, even after accounting for the improvements mentioned above. Indeed, considering the relatively high probability estimates (especially in the case of engineered pandemics) described in the previous section, it would be of great value if future research efforts increased our level of understanding of these risks and the way we can ameliorate them.

Misaligned Artificial Intelligence

Artificial intelligence (AI) refers to computer programs or machines which exhibit behaviour humans would perceive to be intelligent (Kaplan 2016). The significant scientific progress of the field in the last decades has caused many to wonder about the future of this technology and its societal implications. AI may create significant risks, depending on what kinds of systems we create and how we deploy them.

With respect to advances in the field of AI, considerable improvements have been made within specific narrow domains of intelligence (narrow AI) such as computer vision or natural language processing. Notable progress has also been made towards achieving a more general kind of intelligence, such that may eventually be able to employ multiple cognitive abilities, abstraction, and reasoning, as is typical for humans (Strogatz 2018; Badia et al.

2020; Silver et al. 2018). A system with this general reasoning ability is termed artificial general intelligence (AGI).

Broadly speaking, the use of AI can lead to harmful outcomes either if the AI is programmed to achieve a harmful goal, or if the AI is programmed to achieve a beneficial goal but employs a harmful method for achieving it (Future of Life Institute, n.d.; Turchin and Denkenberger 2018a). The latter case is especially relevant for AGI, as it is argued that application of such systems could lead to catastrophic outcomes without any bad intentions or development of harmful methods by its creators (Omohundro 2008).

The development of machines that may potentially become smarter and more powerful than humans could mark the end of an era characterized by humanity's control of its future (Russell 2019). If such a powerful agent does not share our values, the result could be catastrophic (Bostrom 2014; Ord 2020). To prevent the potential disastrous outcomes of future AI, researchers argue it is crucial to align the value and motivation systems of AI systems with human values, a task that is referred to as the alignment problem (Bostrom 2014; Yudkowsky 2016; Critch and Krueger 2020). However, objectively formulating and programming human values into a computer is a complicated task. At present, we do not seem to know how to do it.

Predicting the potential catastrophic impacts of AI is made difficult by several factors. Firstly, the risk posed by AI is unprecedented and cannot be reliably assessed using historical data and extrapolation, unlike the other types of existential risks explored in this review (space weather, super-volcanoes, and pandemics). Secondly, with respect to general and super-intelligence, it may be practically or even inherently impossible for us to predict how a system more intelligent than us will act (Yampolskiy 2020). Considering the difficulty of predicting catastrophic societal impacts associated with AI, we are limited here to hypothetical scenarios of basic pathways describing how the disastrous outcomes could be manifested.

When considering potential global catastrophic or existential risks stemming from AI, it is useful to distinguish between narrow AI and AGI, as the speculated possible outcomes associated with each type can differ greatly. For narrow AI systems to cause catastrophic outcomes, the potential scenarios include events such as software viruses affecting hardware or critical infrastructure globally, AI systems serving as weapons of mass destruction (such as slaughter-bots), or AI-caused biotechnological or nuclear catastrophe (Turchin and Denkenberger 2018a, b; Tegmark 2017; Freitas 2000). Interestingly, Turchin and Denkenberger (2018a) argue that the catastrophic risks stemming from narrow AI are relatively neglected despite their potential to materialize sooner than the risks from AGI. Still, the probability of narrow AI to cause an existential catastrophe appears to be relatively lower than in the case of AGI (Ord 2020).

With respect to the global catastrophic and existential risk of misaligned AGI, much of the expected risk seems to lie in an AGI system's extraordinary ability to pursue its goals. According to Bostrom's instrumental convergence thesis, instrumental aims such as developing more resources and/or power for gaining control over humans would be beneficial for achieving almost any final goal the AGI system might have (Bostrom 2012). Hence, it can be argued that almost any misaligned AGI system would be motivated to gain control over humans (often described in the literature as a 'decisive strategic advantage') to eliminate the possibility of human interference with the system's pursuit of its goals (Bostrom 2014; Russell 2019). Once humans have been controlled, the system would be free to pursue its main goal, whatever that might be.

In line with this instrumental convergence thesis, an AGI system whose values are not perfectly aligned with human values would be likely to pursue harmful instrumental goals,

including seizing control and thus potentially creating catastrophic outcomes for humanity (Russell and Norvig 2016; Bostrom 2002, 2003a; Taylor et al. 2016; Urban 2015; Ord 2020; Muehlhauser 2014). A popular example of such a scenario is the paperclip maximizer, which firstly appeared in a mailing list of AI researchers in the early 2000's (Harris 2018); a later version is included in Bostrom (2003a). Most versions of this scenario involve an AGI system with an arbitrary goal of manufacturing paperclips. In pursuit of this goal, it will inevitably transform Earth into a giant paperclip factory and therefore destroy all life on it. There are other scenarios that end up with potential extinction. Ord (2020) presents one in which the system increases its computational resources by hacking other systems, which enables it to gain financial and human resources to further increase its power in pursuit of its defined goal.⁶

Another possible scenario with respect to the future of AI technology is AI merging with humans and creating artificially enhanced beings. This may indeed be a likely future outcome, considering that the development of these emerging technologies is already ongoing. The potential for this to create globally catastrophic outcomes is relatively less explored. We have not been able to identify any studies focusing on it. Could this be a path that may eventually lead us to great harm?

The hypothetical scenarios discussed here demonstrate the possible pathways through which development and use of AI technology could lead to global or existential catastrophes. However, they do not inform us about concrete societal effects associated with such outcomes, or much on how one can invest in prevention. The catastrophic risk of AI appears to be so speculative and unpredictable that we were not able to identify any estimates of more specific socio-economic impacts of this technology, or any associated risk reduction costs and benefits.

Estimating the likelihood of a catastrophic AI scenario is equally fraught (Armstrong and Sotala 2015). At present, estimates are predominantly based on expert opinion surveys. In these, experts estimate that the probability that an AI system will be capable of performing most human tasks by 2040–2050 to be around 50% (Müller and Bostrom 2016). Similarly, a survey of Grace et al. (2018) suggests a 50% probability of AI systems outperforming humans in all tasks by 2063. The respondents in Müller and Bostrom (2016) predict that such systems would become super-intelligent less than 30 years after their development, with an estimated chance of one in three that machine super-intelligence will be bad or extremely bad for humanity. A survey described in Turchin (2019) suggests that there is an 10% chance of AGI appearing relatively early in the next 10–15 years. By combining this polling with extrapolations of current technological progress, Turchin (2019) estimates that AI capable of causing global catastrophic outcomes (narrow AI included) could be possible as early as 2030–2040.

When questioned directly about existential risk from AI, an informal survey of participants in the Global Catastrophic Risk Conference in Oxford in 2008 indicated that the risk of human extinction due to a super-intelligent AI before 2100 may be as high as 5%; the highest probability out of all those considered in that event (Sandberg and Bostrom 2008). Ord (2020) puts the probability of an AI-caused existential catastrophe even higher at 10% within the next 100 years. Sandberg and Bostrom (2008), Turchin and Denkenberger (2018c) and Ord (2020) all suggest that the existential risk probability of AI is the highest out of the risks described herein.

⁶ A different example is offered by Wiblin (2017): A misaligned AGI is tasked with inventing an anti-cancer drug, which ends up killing humans before they grow old enough to develop cancer.

Conclusion

The previous section argued that Artificial Intelligence (AI) systems most likely pose the highest global catastrophic and existential risk to humanity from the four risks we described here, including solar-flares and space weather, engineered and natural pandemics, and super-volcanic eruptions. While the risk assessments are difficult for all these four hazards, the difficulty of predicting the potential disastrous dynamics of AI is most acute, even as the risk is also likely the highest.

AI may also be very relevant for the other existential risks we described here. Counter-intuitively, maybe, AI may potentially assist us in decreasing the risk of other global catastrophic and existential events such as those associated with a nuclear war, an engineered pandemic, or a solar-flare (Liu et al. 2018; Akiyama 2021; Boulanin 2019; Bostrom 2002). There are undeniably tremendous benefits in well-aligned AI systems in general, but also particularly in terms of their capabilities to reduce existential and other risks (Bostrom 2003b; Yudkowsky 2008) or increase our resilience to these shocks.⁷ We should note that AI systems pose other less extreme risks, and some observers view these are more concerning than the extinction risk we focus on here (e.g., Acemoglu 2021; Lu and Zhou 2021).

The potential impacts of the other catastrophic risks explored in this paper may generally be better understood than the future of AI, partially because they have historical precedents. We still need further inquiry and deeper insights into the differential impacts they would all likely have on any modern society or economy. Afterall, our societies today are different in many ways than the societies that were affected, for example, by the 1918 influenza pandemics, and even more when considering earlier pandemic events. Though these differences may both reduce or increase the risks we have analysed here. When we learn from the past, we also need to take these changes into account. But more generally, it is exactly the uncommonness of these risks that make them difficult to study, and difficult to mobilise the resources required to adequately prepare for them (Wiener 2016).

Apart from analysing the economic consequences of catastrophic and existential risks, some economists have attempted to explore other economic dimensions of these risks related to, for example, decision-making regarding mitigation policy (Martin and Pindyck 2015), risk insurance (Pindyck and Wang 2013) or economic evaluation of potentially catastrophic technology (Posner 2004). However, given the potential importance of the economics of global catastrophic risks, it is still the case that the topic remains under-explored.

We also note that there are, possibly, other existential risks that we have not covered in this paper – for example, nuclear war, or a catastrophic asteroid impact that can change the climate in such a way to make much of the earth uninhabitable for a significant period of time. There have been some attempts to quantify this latter risk (e.g., Reinhardt et al. (2016) and Baum (2018) and there have also been explorations of ways to ameliorate it (e.g., Baum 2019). Our choice, ultimately, was to focus on the four risks we deem most significant, but of course this decision should be questioned, and re-evaluated repeatedly, as circumstances change, and new risks potentially emerge.

The aim of this short survey was not necessarily to provide answers to the many questions that seem to arise once we start to contemplate the economic dimensions of these risks. It seems undeniable that we are currently under-investing in thinking about these

⁷ For very useful definitions of the term ‘resilience’ across different social sciences, and specifically for macro- and micro-economic assessments, see Martin and Sunley (2015) and Rose (2017).

risks, not to say planning for them, or developing the systems that might be necessary to prevent some of the catastrophic scenarios described here from transpiring.

The letter of the British Academy to the Queen ends with a hopeful promise. “Given the forecasting failure at the heart of your enquiry, the British Academy is giving some thought to how...[it] might develop a new, shared horizon-scanning capability so that you never need to ask your question again.... The events of the past year have delivered a salutary shock. Whether it will turn out to have been a beneficial one will depend on the candour with which we dissect the lessons and apply them in future.” The COVID-19 pandemic has clearly shown that, collectively, and arguably specifically in Britain, we have not fully learned these lessons. The only aim of this survey is to highlight this gaping holes in our understanding of the economic dimensions of the risks we face, with the hope that we will soon enough start to plug them.

Funding Open Access funding enabled and organized by CAUL and its Member Institutions.

Data Availability Statement Not applicable (no data included).

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Acemoglu D (2021) Harms of AI. NBER Working Paper No. 29247
- Akiyama N (2021) AI nuclear winter or AI that saves humanity? AI and nuclear deterrence. In: von Braun J, Archer MS, Reichberg GM, Sánchez Sorondo M (eds) Robotics, AI, and humanity. Springer, Cham, pp 161–170
- Alfani G, Murphy TE (2017) Plague and lethal epidemics in the pre-industrial world. *J Econ Hist* 77(1):314–343
- Alfani G (2020) The economic consequences of plague: lessons for the age of Covid-19. History & Policy Working Paper. Retrieved 30 June 2021 from <https://www.historyandpolicy.org/policy-papers/papers/the-economic-consequences-of-plaguelessons-for-the-age-of-covid-19>. Accessed 01 Oct 2021
- Allen P (1979) The "Justinianic" plague. *Byzantion* 49:5–20
- Armstrong S, Sotala K (2015) How we're predicting AI—or failing to. In: Romportl J, Zackova E (eds) Beyond artificial intelligence. Springer, Cham, pp 11–29
- Badia A, Piot B, Kapturowski S, Sprechmann P, Vitvitskyi A, Guo D, Blundell C (2020) Agent57: Outperforming the human Atari benchmark. Proceedings of the 37th International Conference on Machine Learning, Online, PMLR 119, 2020
- Barnes PR, Van Dyke JW (1990) Economic consequences of geomagnetic storms (a summary). *IEEE Power Eng Rev* 10(11):3–4
- Baum SD (2018) Uncertain human consequences in asteroid risk analysis and the global catastrophe threshold. *Nat Hazards* 94:759–775
- Baum SD (2019) Risk-risk tradeoff analysis of nuclear explosives for asteroid deflection. *Risk Anal* 39:2427–2442
- Besley T, Hennessy P (2009) Letter to Her Majesty the Queen – 22/07/2009. Available at: <https://www.imperial.ac.uk/~bin06/M3A22/queen-lse.pdf>. Accessed 01 Oct 2021
- Bostrom N (2002) Existential risks: analyzing human extinction scenarios and related hazards. *J Evol Technol* 9(1):1–30

- Bostrom N (2003a) Ethical issues in advanced artificial intelligence. In: Schneider S (ed) Science fiction and philosophy: from time travel to superintelligence. Wiley, pp 277–284
- Bostrom N (2003b) Astronomical waste: The opportunity cost of delayed technological development. *Utilitas* 15(3):308–314
- Bostrom N (2012) The Superintelligent Will: Motivation and Instrumental Rationality in Advanced Artificial Agents. *Minds & Machines* 22:71–85. <https://doi.org/10.1007/s11023-012-9281-3>
- Bostrom N (2014) *Superintelligence: paths, dangers, strategies*, 1st edn. Oxford University Press, Oxford
- Boteler DH (2006) The super storms of August/September 1859 and their effects on the telegraph system. *Adv Space Res* 38(2):159–172
- Boulanin V (2019) The impact of artificial intelligence on strategic stability and nuclear risk: euro-atlantic perspectives. SIPRI, Stockholm
- Brönnimann S, Krämer D (2016) Tambora and the "Year Without a Summer" of 1816. A perspective on earth and human systems science, vol 90. *Geographica Bernensia*
- Campbell BM (2017) Global climates, the 1257 mega-eruption of Samalas volcano, Indonesia, and the English food crisis of 1258. *Trans R Hist Soc* 27:87
- Cannon P, Angling M, Barclay L, Curry C, Dyer C, Edwards R, ... , Underwood C (2013) Extreme space weather: impacts on engineered systems and infrastructure. Royal Academy of Engineering
- Choumert-Nkolo J, Lamour A, Phélinas P (2021) The Economics of Volcanoes. *Econ Disaster Clim Chang* 5(2):277–299
- Christakos G, Olea RA, Serre ML, Yu HL, Wang LL (2005) Interdisciplinary public health reasoning and epidemic modelling: the case of Black Death, pp 103–152
- Clark G (2016) Microbes and markets: was the Black Death an economic revolution? *J Demogr Econ* 82(2):139–165
- Cliver EW, Tylka AJ, Dietrich WF, Ling AG (2014) On a solar origin for the cosmogenic nuclide event of 775 AD. *Astrophys J* 781(1):32
- Critch A, Krueger D (2020) AI Research Considerations for Human Existential Safety (ARCHES). arXiv preprint arXiv:2006.04948
- Deligne NI, Coles SG, Sparks RSJ (2010) Recurrence rates of large explosive volcanic eruptions. *J Geophys Res: Solid Earth* 115(B6). <https://doi.org/10.1029/2009JB006554>
- Denkenberger DC, Pearce JM (2016) Cost-effectiveness of interventions for alternate food to address agricultural catastrophes globally. *Int J Disaster Risk Sci* 7(3):205–215
- Dols MW (2019) *The Black Death in the Middle East*. Princeton University Press
- Donovan A, Oppenheimer C (2014) Extreme volcanism: disaster risk and societal implications. In: Ismail-Zadeh A, Fucugauchi JU, Kijko A, Takeuchi K, Zaliapin I (eds) *Extreme natural hazards, disaster risks and societal implications*. Cambridge University Press, pp 29–46
- Eroshenko EA, Belov AV, Boteler D, Gaidash SP, Lobkov SL, Pirjola R, Trichtchenko L (2010) Effects of strong geomagnetic storms on Northern railways in Russia. *Adv Space Res* 46(9):1102–1110
- Findlay R, Lundahl M (2017) Demographic shocks and the factor proportions model: from the plague of Justinian to the Black Death. In: Findlay R, Lundahl M (eds) *The economics of the frontier*. Palgrave Macmillan, London, pp 125–172
- Finley T, Koyama M (2018) Plague, Politics, and Pogroms: The Black Death, the Rule of Law, and the Persecution of Jews in the Holy Roman Empire. *The Journal of Law and Economics* 61(2):253–277
- Freitas RA Jr (2000) Some limits to global ecophagy by biovorous nanoreplicators, with public policy recommendations. Foresight Institute Technical Report
- Future of Life Institute (n.d.) Benefits and risks of artificial intelligence. Retrieved 1 July, 2021 from <https://futureoflife.org/background/benefits-risks-of-artificial-intelligence/?cn-reloaded=1>. Accessed 01 Oct 2021
- Gårdan GV (2020) "The Justinianic plague": the effects of a pandemic in late antiquity and the early Middle Ages. *Rom J Artist Creat* 8(4):3–18
- Gathorne-Hardy FJ, Harcourt-Smith WEH (2003) The super-eruption of Toba, did it cause a human bottleneck? *J Hum Evol* 45(3):227–230
- Gingerich D, Vogler JP (2021) Pandemics and political development: the electoral legacy of the Black Death in Germany. *World Politics*, forthcoming. <https://doi.org/10.1017/S0043887121000034>
- Grace K, Salvatier J, Dafoe A, Zhang B, Evans O (2018) Viewpoint: When Will AI Exceed Human Performance? Evidence from AI Experts. *JAIR* 62. <https://doi.org/10.1613/jair.1.11222>
- Guillet S, Corona C, Stoffel M et al (2017) Climate response to the Samalas volcanic eruption in 1257 revealed by proxy records. *Nature Geosci* 10:123–128. <https://doi.org/10.1038/ngeo2875>
- Haddock DD, Kiesling L (2002) The Black Death and property rights. *J Leg Stud* 31(S2):S545–S587

- Handler P (1989) The effect of volcanic aerosols on global climate. *J Volcanol Geoth Res* 37(3–4):233–249
- Hapgood M, Angling MJ, Attrill G, Bisi M, Cannon PS, Dyer C, ... , Willis M (2021) Development of space weather reasonable worst-case scenarios for the UK national risk assessment. *Space Weather* 19(4), e2020SW002593
- Harris S (2018) AI: racing toward the brink, a conversation with Eliezer Yudkowsky (#116) [Audio podcast episode]. Making Sense Podcast
- Horden P (2021) Plague of Justinian. *Oxford Classical Dictionary*. Retrieved 3 Sep. 2021, from <https://doi.org/10.1093/acrefore/9780199381135.001.0001/acrefore-9780199381135-e-8566>
- Jedwab R, Johnson ND, Koyama M (2019a) Negative shocks and mass persecutions: evidence from the Black Death. *J Econ Growth* 24(4):345–395
- Jedwab R, Johnson ND, Koyama M (2019b) Pandemics, places, and populations: evidence from the Black Death. CESifo Working Paper No. 7524
- Jedwab R, Johnson N, Koyama M (2020) The economic impact of the Black Death. *J Econ Lit*, forthcoming
- Jones GS, Gregory JM, Stott PA, Tett SF, Thorpe RB (2005) An AOGCM simulation of the climate response to a volcanic super-eruption. *Clim Dyn* 25(7–8):725–738
- Jordà Ò, Singh SR, Taylor AM (2021) Longer-run economic consequences of pandemics. *Rev Econ Stat*, forthcoming 104 (1): 166–175. https://doi.org/10.1162/rest_a_01042
- Kahler SW (1992) Solar flares and coronal mass ejections. *Ann Rev Astron Astrophys* 30(1):113–141
- Kaplan J (2016) *Artificial intelligence: What everyone needs to know*. Oxford University Press
- Kappenman J (2010) Geomagnetic storms and their impacts on the US power grid. Metatech, Goleta
- Kappenman JG (2012) Geomagnetic disturbances and impacts upon power system operation In: *Electric power generation, transmission, and distribution 3rd edn*. CRC Press, Boca Raton
- Karoff C, Knudsen M, De Cat P et al (2016) Observational evidence for enhanced magnetic activity of superflare stars. *Nat Commun* 7:11058. <https://doi.org/10.1038/ncomms11058>
- Khurshid ZM, Ab Aziz NF, Ab Kadir MZA, Rhazali ZA (2020) A Review of geomagnetically induced current effects on electrical power system: principles and theory. *IEEE Access*. <https://doi.org/10.1109/ACCESS.2020.3034347>
- Klotz LC, Sylvester EJ (2014) The consequences of a lab escape of a potential pandemic pathogen. *Front Public Health* 2:116
- Knipp DJ, Biesecker DA (2015) Changing of the guard: satellite will warn earth of solar storms. *Earth & Space Science News* 96(7):12–16
- Kutterolf S, Hansteen TH, Appel K, Freundt A, Krüger K, Perez W, Wehrmann H (2013) Combined bromine and chlorine release from large explosive volcanic eruptions: A threat to stratospheric ozone? *Geology* 41(6):707–710
- Lakhina GS, Alex S, Tsurutani BT, Gonzalez WD (2004) Research on historical records of geomagnetic storms. *Proc Int Astron Union* 2004(IAUS226):3–15
- Lane CS, Chorn BT, Johnson TC (2013) Ash from the Toba supereruption in Lake Malawi shows no volcanic winter in East Africa at 75 ka. *Proc Natl Acad Sci* 110(20):8025–8029
- Lingam M, Loeb A (2017b) Impact and mitigation strategy for future solar flares. *arXiv preprint arXiv:1709.05348*
- Lingam M, Loeb A (2017a) Risks for life on habitable planets from superflares of their host stars. *Astrophys J* 848(1):41
- Liu HY, Lauta KC, Maas MM (2018) Governing Boring Apocalypses: A new typology of existential vulnerabilities and exposures for existential risk research. *Futures* 102:6–19
- Livi-Bacci ML (2017) *A concise history of world population*. Wiley
- Lu Y, Zhou Y (2021) A review on the economics of artificial intelligence. *J Econ Surv* 35:1045–1072
- Maehara H, Shibayama T, Notsu S, Notsu Y, Nagao T, Kusaba S, ... , Shibata K (2012) Superflares on solar-type stars. *Nature* 485(7399): 478–481
- Martin R, Sunley P (2015) On the notion of regional economic resilience: conceptualization and explanation. *J Econ Geogr* 15(1):1–42
- Martin IWR, Pindyck S (2015) "Averting Catastrophes: The Strange Economics of Scylla and Charybdis." *American Economic Review* 105(10):2947–85
- Mason BG, Pyle DM, Oppenheimer C (2004) The size and frequency of the largest explosive eruptions on Earth. *Bull Volcanol* 66(8):735–748
- Maynard T, Smith N, Gonzalez S (2013) Solar storm risk to the North American electric grid. *Lloyd's* 1:11
- Meier M (2016) The 'Justinianic Plague': The economic consequences of the pandemic in the eastern Roman empire and its cultural and religious effects. *Early Mediev Eur* 24(3):267–292
- Meier M (2020) The Roman context of early Islam. *Millennium* 17(1):265–302

- Mekhaldi F, Muscheler R, Adolphi F, Aldahan A, Beer J, McConnell JR, ... , Woodruff TE (2015) Multiradionuclide evidence for the solar origin of the cosmic-ray events of AD 774/5 and 993/4. *Nat Commun* 6(1): 1–8
- Melott AL, Thomas BC (2012) Causes of an AD 774–775 14 C increase. *Nature* 491(7426):E1–E2
- Miyake F, Masuda K, Nakamura T (2013) Another rapid event in the carbon-14 content of tree rings. *Nat Commun* 4(1):1–6
- Moran D, Lenzen M, Cairns I, Steenge A (2014) How severe space weather can disrupt global supply chains. *Nat Hazard* 14(10):2749–2759
- Mordechai L, Eisenberg M (2019) Rejecting catastrophe: The case of the Justinianic Plague. *Past Present* 244(1):3–50
- Mordechai L, Eisenberg M, Newfield TP, Izdebski A, Kay JE, Poinar H (2019) The Justinianic Plague: an inconsequential pandemic? *Proc Natl Acad Sci* 116(51):25546–25554
- Moriña D, Serra I, Puig P, Corral Á (2019) Probability estimation of a Carrington-like geomagnetic storm. *Sci Rep* 9(1):1–9
- Muehlhauser L (2014) Three misconceptions in Edge.org’s conversation on The Myth of AI. Machine Intelligence Research Institute: Berkeley, CA, USA, 18. Retrieved 2 July, 2021 from <https://intelligence.org/2014/11/18/misconceptions-edge-orgs-conversation-myth-ai/>. Accessed 01 Oct 2021
- Muehlhauser, L. (2017). How Big a Deal Was the Industrial Revolution. Manuscript. Retrieved 30 June 2021 from <https://lukemuehlhauser.com/industrial-revolution/>. Accessed 01 Oct 2021
- Müller VC, Bostrom N (2016) Future Progress in Artificial Intelligence: A Survey of Expert Opinion. In: Müller VC (eds) *Fundamental Issues of Artificial Intelligence*. Synthese Library (Studies in Epistemology, Logic, Methodology, and Philosophy of Science) vol 376 Cham: Springer. https://doi.org/10.1007/978-3-319-26485-1_33
- National Research Council (2009) Severe space weather events: Understanding societal and economic impacts: A workshop report. National Academies Press
- Neuhäuser R, Neuhäuser DL (2015) Solar activity around AD 775 from aurorae and radiocarbon. *Astron Nachr* 336(3):225–248
- Newhall CG, Self S (1982) The volcanic explosivity index (VEI) an estimate of explosive magnitude for historical volcanism. *J Geophys Res: Oceans* 87(C2):1231–1238
- Newhall C, Self S, Robock A (2018) Anticipating future Volcanic Explosivity Index (VEI) 7 eruptions and their chilling impacts. *Geosphere* 14(2):572–603
- Newhall CG, Dzurisin D (1988) Historical unrest at the large calderas of the world (Vol. 2, No. 1855). Department of the Interior, US Geological Survey
- Nogami D, Notsu Y, Honda S, Maehara H, Notsu S, Shibayama T, Shibata K (2014) Two sun-like superflare stars rotating as slow as the Sun. *Publ Astron Soc Jpn* 66(2). <https://doi.org/10.1093/pasj/psu012>
- North American Electric Reliability Corporation (NERC) (2010) High-impact, low-frequency event risk to the North American bulk power system. Atlanta, GA, Tech Rep. Retrieved 15 June from <https://www.energy.gov/sites/default/files/High-Impact%20Low-Frequency%20Event%20Risk%20to%20the%20North%20American%20Bulk%20Power%20System%20-%202010.pdf>. Accessed 01 Oct 2021
- North American Electric Reliability Corporation (NERC) (2012) Effects of geomagnetic disturbances on the bulk power system. Report
- O’Keefe P, Westgate K, Wisner B (1976) Taking the naturalness out of natural disasters. *Nature* 260:566–567
- Omohundro SM (2008) The basic AI drives. In: Yampolskiy RV (ed) *Artificial intelligence safety and security*. CRC Press, pp 47–54
- Oppenheimer C (2003) Climatic, environmental, and human consequences of the largest known historic eruption: Tambora volcano (Indonesia) 1815. *Prog Phys Geogr* 27(2):230–259
- Oppenheimer C (2011) *Eruptions that shook the world*. Cambridge University Press
- Ord T (2020) *The Precipice: Existential Risk and The Future Of Humanity*. Hachette Books
- Oughton EJ, Skelton A, Horne RB, Thomson AW, Gaunt CT (2017) Quantifying the daily economic impact of extreme space weather due to failure in electricity transmission infrastructure. *Space Weather* 15(1):65–83
- Oughton E, Copic J, Skelton A, Kesaite V, Yeo ZY, Ruffle SJ, Tuveson M, Coburn AW, Ralph D (2016) Helios Solar Storm Scenario; Cambridge Risk Framework series; Centre for Risk Studies, University of Cambridge
- Oughton EJ, Hapgood M, Richardson GS, Beggan CD, Thomson AW, Gibbs M, ... , Horne RB (2019) A risk assessment framework for the socioeconomic impacts of electricity transmission infrastructure failure due to space weather: An application to the United Kingdom. *Risk Anal* 39(5): 1022–1043
- Pamuk Ş (2007) The Black Death and the origins of the ‘Great Divergence’ across Europe, 1300–1600. *Eur Rev Econ Hist* 11(3):289–317

- Pamuk Ş, Shatzmiller M (2014) Plagues, wages, and economic change in the Islamic Middle East, 700–1500. *Journal of Economic History* 74(1):196–229
- Pindyck RS, Wang N (2013) "The Economic and Policy Consequences of Catastrophes." *American Economic Journal: Economic Policy* 5 (4):306–39
- Plag HP, Stein S, Brocklebank S, Jules-Plag S, Marsh S, Campus P (2013) Extreme geohazards: Reducing the disaster risk and increasing resilience. In EGU General Assembly Conference Abstracts (pp. EGU2013–8118)
- Posner (2004) *Catastrophe: Risk and Response*. Oxford: Oxford University Press
- Post JD (1977) *The last great subsistence crisis in the Western World*. The Johns Hopkins University Press, Baltimore, p 240
- Puma MJ, Bose S, Chon SY, Cook BI (2015) Assessing the evolving fragility of the global food system. *Environ Res Lett* 10(2):024007
- Raible CC, Brönnimann S, Auchmann R, Brohan P, Frölicher TL, Graf HF, ... , Wegmann M (2016) Tambora 1815 as a test case for high impact volcanic eruptions: Earth system effects. *Wiley Interdiscip Rev Clim Chang* 7(4): 569–589
- Rampino MR (2002) Supereruptions as a threat to civilizations on Earth-like planets. *Icarus* 156(2):562–569
- Rampino MR (2008) Super-volcanism and other geophysical processes of catastrophic import. *Global Catastrophic Risks* 1:203
- Rampino MR, Self S (1982) Historic eruptions of Tambora (1815), Krakatau (1883), and Agung (1963), their stratospheric aerosols, and climatic impact. *Quatern Res* 18(2):127–143
- Rampino M, Self S (1992) Volcanic winter and accelerated glaciation following the Toba super-eruption. *Nature* 359:50–52. <https://doi.org/10.1038/359050a0>
- Rampino MR, Self S (1993) Climate-volcanism feedback and the Toba eruption of ~ 74,000 years ago. *Quatern Res* 40(3):269–280
- Reinhardt JC, Chen X, Liu W, Manchev P, Paté-Cornell ME (2016) Asteroid risk assessment: a probabilistic approach. *Risk Anal* 36:244–261
- Rennie H, Turner K, Sollitt D (1998) Auckland power supply failure 1998, the report of the ministerial inquiry into the Auckland power supply failure. Ministry of Economic Development, New Zealand, Auckland 30
- Riley P (2012) On the probability of occurrence of extreme space weather events. *Space Weather* 10:S02012. <https://doi.org/10.1029/2011SW000734>
- Riley P, Love JJ (2017) Extreme geomagnetic storms: Probabilistic forecasts and their uncertainties. *Space Weather* 15(1):53–64
- Riley P, Baker D, Liu YD, Verronen P, Singer H, Güdel M (2018) Extreme space weather events: From cradle to grave. *Space Sci Rev* 214(1):1–24
- Robinson JA, Torvik R (2011) Institutional comparative statics (No. w17106). National Bureau of Economic Research
- Robock A (2000) Volcanic eruptions and climate. *Rev Geophys* 38(2):191–219
- Robock A (2005) Cooling following large volcanic eruptions corrected for the effect of diffuse radiation on tree rings. *Geophys Res Lett* 32(6)
- Rose A (2017) *Defining and measuring economic resilience from a societal, environmental and security perspective*. Springer, Singapore
- Rougier J, Sparks RSJ, Cashman KV, Brown SK (2018) The global magnitude–frequency relationship for large explosive volcanic eruptions. *Earth Planet Sci Lett* 482:621–629
- Russell S (2019) *Human compatible: artificial intelligence and the problem of control*. Viking, New York
- Russell S, Norvig P (2016) *Artificial intelligence: a modern approach*, 3rd edn. Pearson, Harlow
- Rymer H, Sparks S, Self S, Grattan J, Oppenheimer C, Pyle D (2005) *Super-eruptions: global effects and future threats*. Geological Society of London
- Sabbatani S, Manfredi R, Fiorino S (2012) The Justinian plague (part two). Influence of the epidemic on the rise of the Islamic empire. *Le Infezioni In Medicina* 20(3):217–232
- Sandberg A, Bostrom N (2008) Global catastrophic risks survey. *Civil Wars* 98(30):4
- Sarris P (2002) The Justinianic plague: origins and effects. *Contin Chang* 17(2):169–182
- Schwenn R (2006) Space weather: The solar perspective. *Living Rev Sol Phys* 3(1):1–72
- Scott S, Duncan CJ (2001) *Biology of plagues: evidence from historical populations*. Cambridge University Press
- Self S (2006) The effects and consequences of very large explosive volcanic eruptions. *Phil Trans R Soc A Math Phys Eng Sciences* 364(1845):2073–2097
- Self S (2015) Explosive super-eruptions and potential global impacts. In: Schroder JF, Papale P (eds) *Volcanic hazards, risks and disasters*. Elsevier, pp 399–418

- Self S, Gertisser R, Thordarson T, Rampino MR, Wolff JA (2004) Magma volume, volatile emissions, and stratospheric aerosols from the 1815 eruption of Tambora. *Geophys Res Lett* 31(20):L20608. <https://doi.org/10.1029/2004GL020925>
- Shibata K, Isobe H, Hillier A, Choudhuri AR, Maehara H, Ishii TT, ... , Nogami D (2013) Can superflares occur on our Sun?. *Publ Astron Soc Jpn* 65(3): 49
- Shibayama T, Maehara H, Notsu S, Notsu Y, Nagao T, Honda S, ... , Shibata K (2013) Superflares on solar-type stars observed with Kepler. I. Statistical properties of superflares. *Astrophys J Suppl Ser* 209(1): 5
- Silver D, Hubert T, Schrittwieser J, Antonoglou I, Lai M, Guez A, ... , Hassabis D (2018) A general reinforcement learning algorithm that masters chess, shogi, and Go through self-play. *Science* 362(6419), pp 1140–1144
- Siuda F, Sunde U (2021) Disease and demographic development: the legacy of the plague. *J Econ Growth* 26(1):1–30
- Stathakopoulos D (2007) Crime and punishment: the plague in the Byzantine empire 541–749. In Hays J (ed) *Plague and the End of Antiquity: The Pandemic of 541–750*. Cambridge: Cambridge University Press pp 99–118
- Stephenson FR (2015) Astronomical evidence relating to the observed 14C increases in AD 774–5 and 993–4 as determined from tree rings. *Adv Space Res* 55(6):1537–1545
- Stothers RB (2000) Climatic and demographic consequences of the massive volcanic eruption of 1258. *Clim Change* 45(2):361–374
- Strogatz S (2018) One giant step for a chess-playing machine. *New York Times*
- Sukhodolov T, Usoskin I, Rozanov E, Asvestari E, Ball WT, Curran MA, ... , Traversi R (2017) Atmospheric impacts of the strongest known solar particle storm of 775 AD. *Sci Rep* 7(1): 1–9
- Taylor J, Yudkowsky E, LaVictoire P, Critch A (2016) Alignment for advanced machine learning systems. In: Liao SM (ed) *Ethics of Artificial Intelligence*. Oxford University Press, pp 342–382
- Tegmark M (2017) *Life 3.0: Being human in the age of artificial intelligence*. Knopf
- Timmreck C, Graf HF, Lorenz SJ, Niemeier U, Zanchettin D, Matei D, ... , Crowley TJ (2010) Aerosol size confines climate response to volcanic super-eruptions. *Geophys Res Lett* 37(24). L24705. <https://doi.org/10.1029/2010GL045464>
- Timmreck C, Graf HF (2006) The initial dispersal and radiative forcing of a Northern Hemisphere mid-latitude super volcano: a model study. *Atmos Chem Phys* 6(1):35–49
- Toon OB, Zahnle K, Morrison D, Turco RP, Covey C (1997) Environmental perturbations caused by the impacts of asteroids and comets. *Rev Geophys* 35(1):41–78
- Turchin A (2019) Assessing the future plausibility of catastrophically dangerous AI. *Futures* 107:45–58
- Turchin A, Denkenberger D (2018a) Classification of global catastrophic risks connected with artificial intelligence. *AI Soc* 35(1):147–163
- Turchin A, Denkenberger D (2018) Military AI as a convergent goal of self-improving AI. In: Yampolskiy RV (ed) *Artificial intelligence safety and security*. CRC Press, pp 375–374
- Turchin A, Denkenberger D (2018c) Global catastrophic and existential risks communication scale. *Futures* 102:27–38
- Urban T (2015) *The AI revolution: Our immortality or extinction*. Wait But Why, e-book, 51. Retrieved 30 June, 2021, from <https://waitbutwhy.com/2015/01/artificial-intelligence-revolution-2.html>. Accessed 01 Oct 2021
- Usoskin IG, Kromer B, Ludlow F, Beer J, Friedrich M, Kovaltsov GA, ... , Wacker L (2013) The AD775 cosmic event revisited: the Sun is to blame. *Astron Astrophys* 552: L3
- Usoskin IG, Kovaltsov GA (2012) Occurrence of extreme solar particle events: assessment from historical proxy data. *Astrophys J* 757(1):92
- Voigtländer N, Voth HJ (2013a) The three horsemen of riches: Plague, war, and urbanization in early modern Europe. *Rev Econ Stud* 80(2):774–811
- Voigtländer N, Voth HJ (2013b) How the West “Invented” fertility restriction. *Am Econ Rev* 103(6):2227–2264
- Weitzman M (2009) On modeling and interpreting the economics of catastrophic climate change. *Rev Econ Stat* 91:1–19
- Weitzman M (2011) Fat-tailed uncertainty in the economics of catastrophic climate change. *Rev Environ Econ Policy* 5(2):275–292
- White LA, Mordechai L (2020) Modeling the Justinianic Plague: comparing hypothesized transmission routes. *PloS one* 15(4):e0231256
- Wiblin R (2017) Positively shaping the development of artificial intelligence. *80,000 Hours, 1*. Retrieved 30 June, 2021, from <https://80000hours.org/problem-profiles/positively-shaping-artificial-intelligence/>. Accessed 01 Oct 2021

- Wiener JB (2016) The tragedy of the commons: on the politics of apocalypse. *Global Pol* 7:67–80
- Wilcox BH, Mitchell KL, Schwandner FM, Lopes RM (2015) Defending human civilization from super-volcanic eruptions. Jet Propulsion Laboratory, California Institute of Technology, Pasadena
- Williams M (2012) Did the 73 ka Toba super-eruption have an enduring effect? Insights from genetics, prehistoric archaeology, pollen analysis, stable isotope geochemistry, geomorphology, ice cores, and climate models. *Quatern Int* 269:87–93
- Wood GDA (2014) *Tambora: the eruption that changed the World*. Princeton University Press
- Xoplaki E, Fleitmann D, Luterbacher J, Wagner S, Haldon JF, Zorita E, ... , Izdebski A (2016) The medieval climate anomaly and byzantium: A review of the evidence on climatic fluctuations, economic performance and societal change. *Quatern Sci Revi* 136: 229–252
- Xu L, Wei K, Wu X, Smyshlyaev SP, Chen W, Galin VY (2019) The effect of super volcanic eruptions on ozone depletion in a chemistry-climate model. *Adv Atmos Sci* 36(8):823–836
- Yampolskiy RV (2020) Unpredictability of AI: On the impossibility of accurately predicting all actions of a smarter agent. *J Artif Intell Conscious* 7(01):109–118
- Yudkowsky E (2008) Artificial intelligence as a positive and negative factor in global risk. *Global Catastrophic Risks* 1(303):184
- ampinoYudkowsky E (2016) The AI alignment problem: why it is hard, and where to start. *Symbolic Systems Distinguished Speaker*. <https://intelligence.org/2016/12/28/ai-alignment-why-its-hard-and-where-to-start/>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.