## 1AC

### ADV---Precarity---1AC

#### Advantage 1 is PRECARITY.

#### The growth of the gig economy is on a collision course with antitrust law. Courts are shutting down organizing as per se anticompetitive conduct, chilling collective bargaining.

Sowry ’24 [David; February 6; J.D., University of Pennsylvania Carey Law School; University of Pennsylvania Carey Law School, “Off to the Races?: The Antitrust Labor Exemption and The Gig Economy after Jinetes,” Sowry, David, Off to the Races?” http://dx.doi.org/10.2139/ssrn.4702014]

Today, two trends reshaping the American economy—the explosive growth of the gig economy and the return of labor unions—are on a collision course due to their interactions with antitrust law.

Uber, Airbnb, DoorDash. Most had never heard of them a decade ago,1 but today these companies and their competitors have changed not only how Americans live, but the structure of the workforce as well. While temporary work and independent contractors have existed for years, the growth of app-based startups has revolutionized temporary working arrangements.2 The barriers to entry for workers have never been lower, and it is now easier than ever to connect the supply and demand for these services.3 This has resulted in widespread adoption of the gig economy as a source of income, as over 16% of Americans have reported earnings through these platforms, 4 which is a greater percentage of the workforce than government workers, health care, or retail.5 Demand is high for these platforms too, with nearly half of all Americans having used the applications in some capacity.6 Use of these platforms has only grown over time,7 and based on the disruptions caused to their traditional counterparts,8 they seem to be here to stay. 9

Unionized labor, after declining for several decades, seems to be returning in a major way to the American workforce. Public support for unions and organized labor is at a recent high.10 Unionization is occurring at rates not seen in decades.11 Several major strikes dominated the news cycle in recent months.12 The Biden administration made enforcing NLRB obligations a priority. 13 The Biden FTC has even stated that it would prefer to spend less resources pursuing unionizing workers for antitrust violations, the very topic of this article.14 And for the first time in history, a sitting President joined a picket line to show his support for labor.15

Despite the strength and power of both the gig economy and collective labor efforts, they now appear to be on a collision course in the courts. Under current antitrust law, it is likely that the independent contractors who make their livelihoods through the gig economy cannot take collective action or form a union. This paper seeks to assess the current state of independent contractors within the antitrust labor exemption, examine the impact a recent decision may have, and analyze the odds of success various groups of workers (including rental property hosts, skilled laborers, and drivers) in the gig economy may have to unionize.

I. The Antitrust Labor Exemption

Labor unions are antitrust violations, or at least violate the spirit of the laws. They are an attempt to coordinate the supply of a service and use an attempted monopoly to drive higher prices and reduce competition.16 This has created an inherent tension between the competition maximizing policies of antitrust law and the cooperative condition-improving policies behind labor law.17 In the years after the passage of the Sherman Act, courts began to enjoin strikes and other collective actions while trying to navigate this conflict.18 Courts navigating this issue sided with management, using the Sherman Act to limit the ability of labor to collectively act.19 In fact, in the years after its passage, the Sherman Act began to be “perceived [as] a powerful union-busting device.”20 Congress, in recognition of the importance of organized labor in American society, passed laws to immunize unions from antitrust claims, under what has become referred to as the antitrust labor exemption. 21

In 1914, Congress passed the Clayton Act.22 In Section 6, Congress laid clear its intentions by stating that “[t]he labor of a human being is not a commodity or article of commerce.”23 The statute went further and immunized from liability both labor organizations and their members, so long as the action being taken is for the “purposes of mutual help” and is “lawfully carrying out legitimate objects.”24 To enact this policy, Section 20 of the Act prohibited injunctions against specific labor activities such as boycotts and strikes.25

Despite Congress’s best intentions,26 the Clayton Act was used primarily to foil organized labor,27 because it created a private right of injunction against labor unions.28 Reading into the Clayton Act “the very beliefs which [it] was designed to remove,” the Supreme Court removed secondary boycotts from the scope of Section 20.29 Congress was dismayed by judicial interpretations of the Clayton Act,30 and took “the extraordinary step of divesting federal courts of equitable jurisdiction.”31

In 1932 Congress enacted this plan by passing the Norris-LaGuardia Act.32 There, Congress clarified the national policy towards labor, stating that the protections are given to help workers “exercise actual liberty.”33 The Act removed the court-imposed restraints upon labor unions by limiting the circumstances under which courts could grant injunctions in labor disputes.34 Labor disputes are defined in the Act as “any controversy concerning terms or conditions of employment, or concerning the association or representation of persons in negotiating, fixing, maintaining, changing, or seeking to arrange terms or conditions of employment, regardless of whether or not the disputants stand in the proximate relation of employer and employee.”35 There are two noteworthy elements of the Act’s definition of a labor dispute. The first is that it is defined as “any controversy concerning terms or conditions of employment.”36 The second is that the scope of the Act is not limited by the formal relationship between the two groups in the dispute, as they are not required to “stand in the proximate relation of employer and employee.”37 The Act’s definition of labor dispute has been found to be “broad” by the courts, in part to prevent a redux of the anti-labor issues associated with the Clayton Act.38

Although the Norris-LaGuardia Act is not technically about antitrust law,39 the Supreme Court has stated that it covers antitrust claims, because a failure to do so would “would run counter to the plain mandate of the Act and would reverse the declared purpose of Congress.”40

#### The gig economy are making traditional employment obsolete. Low-paying gig jobs exacerbate economic precarity and inequality.

Kiron ’23 [David, Elizabeth J. Altman, and Christoph Riedl; April 13; Editorial Director at MIT Sloan Management Review,PhD in philosophy from the University of Rochester; Nonresident Fellow of Economic Studies at the Center on Regulation and Markets, MS in Management from the Massachusetts Institute of Technology; Associate Professor at Northeastern University, D’Amore-McKim School of Business, PhD and MSc in Information Systems, Technische Universität München; Brookings, “Workforce ecosystems and AI,” https://www.brookings.edu/articles/workforce-ecosystems-and-ai/]

Companies increasingly rely on an extended workforce (e.g., contractors, gig workers, professional service firms, complementor organizations, and technologies such as algorithmic management and artificial intelligence) to achieve strategic goals and objectives.1 When we ask leaders to describe how they define their workforce today, they mention a diverse array of participants, beyond just full- and part-time employees, all contributing in various ways. Many of these leaders observe that their extended workforce now comprises 30-50% of their entire workforce. For example, Novartis has approximately 100,000 employees and counts more than 50,000 other workers as external contributors.2 Businesses are also increasingly using crowdsourcing platforms to engage external participants in the development of products and services.34 Managers are thinking about their workforce in terms of who contributes to outcomes, not just by workers’ employment arrangements.5

Our ongoing research on workforce ecosystems demonstrates that managing work across organizational boundaries with groups of interdependent actors in a variety of employment relationships creates new opportunities and risks for both workers and businesses.6 These are not subtle shifts. We define a workforce ecosystem as:7

A structure that encompasses actors, from within the organization and beyond, working to create value for an organization. Within the ecosystem, actors work toward individual and collective goals with interdependencies and complementarities among the participants.

The emergence of workforce ecosystems has implications for management theory, organizational behavior, social welfare, and policymakers. In particular, issues surrounding work and worker flexibility, equity, and data governance and transparency pose substantial opportunities for policymaking.

At the same time, artificial intelligence (AI)—which we define broadly to include machine learning and algorithmic management—is playing an increasingly large role within the corporate context. The widespread use of AI is already displacing workers through automation, augmenting human performance at work, and creating new job categories.

What’s more, AI is enabling, driving, and accelerating the emergence of workforce ecosystems. Workforce ecosystems are incorporating human-AI collaboration on both physical and cognitive tasks and introducing new dependencies among managers, employees, contingent workers, other service providers, and AI.

Clearly, policy needs to consider how AI-based automation will affect workers and the labor market more broadly. However, focusing only on the effects of automation without considering the impact of AI on organizational and governance structures understates the extent to which AI is already influencing work, workers, and the practice of management. Policy discussions also need to consider the implications of human-AI collaborations and AI that enhances human performance (such as generative AI tools). Policymakers require a much more nuanced and comprehensive view of the dynamic relationship between workforce ecosystems and AI. To that end, this policy brief presents a framework that addresses the convergence of AI and workforce ecosystems.

Within workforce ecosystems, the use of AI is changing the design of work, the supply of labor, the conduct of work, and the measurement of work and workers. Examining AI-related shifts in four categories—Designing Work, Supplying Workers, Conducting Work, and Measuring Work and Workers—reveals a variety of policy implications. We explore these policy considerations, highlighting themes of flexibility, equity, and data governance and transparency. Furthermore, we offer a broad view of how a shift toward workforce ecosystems and the increasing use of AI is influencing the future of work.

AI and Workforce Ecosystems: A Framework

Workforce ecosystems consist of workforce participants inside and outside organizations crossing all organizational levels and functions and spanning all product and service development and delivery phases. Strikingly, AI usage within workforce ecosystems is increasing and simultaneously accelerating their emergence and growth. The increasing shift toward workforce ecosystems creates new opportunities to leverage AI, and the increased use of AI further amplifies the move toward workforce ecosystems.

In this brief, we present a typology to better understand the interaction between the continuing emergence of AI and the ongoing evolution of workforce ecosystems. With this framework, we aim to assist policymakers in making sense of changes accompanying AI’s growth. The typology includes four categories highlighting four areas in which AI is impacting workforce ecosystems: Designing Work, Supplying Workers, Conducting Work, and Measuring Work and Workers. Each of the four categories suggests distinct (if related) policy implications.

One overarching implication of this discussion is that policy for work-related AI applications is not limited to addressing automation. Despite the clear need for policy to consider implications arising from the use of AI to automate jobs and displace workers, it is insufficient to focus policy discussions only on automation and not fully consider changes in which human work is augmented by AI and in which humans and AI collaborate. Discussions omitting these factors run the risk of understating the current and future influence of AI on work, workers, and the practice of management.

Policy related to AI in workforce ecosystems should balance workers’ interests in sustainable and decent jobs with employers’ interests in productivity and economic growth. If done properly, there is tremendous potential to leverage AI to improve working conditions, worker safety, and worker mobility/flexibility, and to work more collectively and intelligently.8 The goal of these policy refinements should be to allow businesses to meet competitive challenges while limiting the risk of dehumanizing workers, discrimination, and inequality. Policy can offer incentives to limit the use of AI in low value-added contexts, such as for automation of work with small efficiency gains, while promoting higher value-added uses of AI that increase economic productivity and employment growth.9

Designing Work

The growing use of AI has a profound effect on work design in workforce ecosystems. A greater supply of AI affects how organizations design work while changes in work design drive greater demand for AI. For example, modern food delivery platforms like GrubHub and DoorDash use AI for sophisticated scheduling, matching, rating, and routing, which has essentially redesigned work within the food delivery industry. Without AI, such crowd-based work designs would not be possible. These technologies and their impact on work design reach beyond food delivery into other supply chains wherever complex delivery systems exist. Similarly, AI-driven tools enable larger, flatter, more integrated teams because entities can coordinate and collaborate more effectively. For workforce ecosystems, this means organizations can more seamlessly integrate external workers, partner organizations, and employees as they strive to meet strategic goals.

On the flip side, changes in work design drive increasing demand for AI. For example, as jobs are disaggregated into tasks and work becomes more modular and/or project-based, algorithms can help humans become more effective.10 As companies refine their approach to designing work, they gain access to more data (e.g., in medical research and marketing analytics) and AI becomes even more valuable.

Policy concerns associated with U.S. business’s increasing reliance on contingent labor date back (at least to) the 1994 Dunlop Commission.11 Companies do not want to overcommit to hiring full-time workers with skills that will soon become obsolete and thus prefer to rely on contingent labor in many cases. They design work for maximum flexibility and productivity but not necessarily for maximum economic security for workers.12 The shift in employment away from (full- and part-time) payroll to more flexible categories (e.g., contingent workers such as long-term contractors or short-term gig workers) tends to increase the income and wealth gap between workers in full- and part-time employed positions and those in contracted roles by affecting what leverage and protection is available for various classes of workers.13

Notably, contingent work has a direct relationship with “precarious work.” Precarious work has been defined as work that is “uncertain, unstable, and insecure and in which employees bear the risks of work […] and receive limited social benefits and statutory protections.”14 This is likely to affect workers of different skills in different ways, leading not only to income and wealth inequality but also to human capital inequality as workers with different skill levels have more or less control over their wages. For example, a highly-skilled data scientist may command a premium and may work for more than one client. In the shipping industry, most of the workers who maintain and operate commercial vessels are contractors, but they are less likely to command a premium nor will they be able to offer their services to multiple clients. Flexible, platform-based work arrangements can result in precarious work arrangements for some workers while giving flexibility, higher wages, and the ability to hyper-specialize to others. This creates human capital inequality. The difference may depend on already existing discrepancies like class, race, and gender, and thus further amplify income and wealth inequality.

The growing sophistication of AI makes it easier for managers to source, vet, and hire contingent labor. This new role for AI enables managers to design work in new ways. Instead of focusing on hiring employees and filling in skill gaps with full-time labor, managers are increasingly turning to external talent markets and staffing platforms as a source of shorter-term, skills-based engagements to achieve outcomes. Managers can disaggregate existing jobs into component tasks and then use AI to access external contributors with specific skills to accomplish those tasks.

Policy considerations for designing work

These changes in work design affect policies for tax, labor, and technology. Federal and state governments should consider developing more inclusive and flexible policies that support all kinds of employment models so workers receive equal protection and benefits based on the value they create, not the employment status they hold. If workers are to be afforded protections that ensure sustainable, safe, and healthy work environments, the same protections should be available to all workers regardless of whether they are an employee or a contingent worker. Unemployment insurance should be modernized to expand eligibility to include workers who do not work (or seek work) full-time and to provide flexible, partial unemployment benefits.

Today, firms themselves may be willing to be more flexible and creative with compensation and benefits schemes, but they sometimes only have limited opportunities to do so because of labor regulation constraints. Modernized unemployment and other labor policies would potentially increase contingent workers’ access to reasonable earning opportunities, social safety nets, and benefits. Beyond unemployment insurance, other benefits including retirement savings contributions, health insurance, and medical, family, and parental leaves are similarly restricted to full-time workers for historical reasons (although the restrictions vary across geographic regions). Policies should be updated to allow portability of benefits between employers and improve access to assistance, which would dampen the income volatility faced by many contingent workers.

Supplying Workers

By using AI to increase the supply of workers of more types (e.g., contractors, gig workers) through improved communication, coordination, and matching, workforce ecosystems can grow more easily, effectively, and efficiently. At the same time, the growth of workforce ecosystems increases the demand for all kinds of workers, leading to more demand for AI to help increase and manage worker supply.

Organizations increasingly require a variety of workers to engage in multiple ways (full-time, part-time, as professional service providers, as long- and short-term contractors, etc.). They can use AI to assist in sourcing these workers, for example, by using both internal and external labor platforms and talent marketplaces to find and match workers more effectively.15 Using AI that includes enhanced matching functions, scheduling, recruiting, planning, and evaluations increases access to a diverse corps of workers. Organizations can use AI to more effectively build workforce ecosystems that both align with specific business needs and help meet diversity goals.

#### Congress exempted collective bargaining from antitrust liability, BUT courts (specifically, the third and fourth circuit) are narrowly interpreting that exemption to exclude gig workers.

Bedoya ’24 [Alvaro and Bryce Tuttle; 2024; Commissioner of the Federal Trade Commission; Served as an honors paralegal in the office of Commissioner Bedoya at the Federal Trade Commission; Antitrust Law Journal, “‘Aiming at Dollars, Not Men’: Recovering The Congressional Intent Behind the Labor Exemption to Antitrust Law,” vol. 85]

We are aiming at the gigantic trusts and combinations of capital and not at associations of men for the betterment of their condition. We are aiming at the dollars and not at men. . . . Let us put the man above the dollar and exempt all associations of men organized for the betterment of their condition.

- Representative Thomas F. Konop (D., Wisconsin), June 1, 1914[1]

Antitrust law has been used by courts to stop a range of worker-organizing efforts in the late 20th and early 21st centuries-despite the existence of a labor exemption. This article explains that these actions contradict Congress's intent in passing the Sherman, Clayton, and Norris-LaGuardia Acts. To under stand legislative intent and how the courts misapplied the antitrust laws to labor, it is necessary to look at the history of the antitrust laws and how the labor exemption became law in the first place.

I. JOHN D. ROCKEFELLER AND TONY FERNANDEZ

At the time of its incorporation in Ohio in 1870, Standard Oil was already the wealthiest company in America, with one million dollars in assets and 10% of the country's oil refining capacity. That rose to 25% just two years later. In 1882, the company's stock was combined with the assets of three dozen other companies to form the Standard Oil Trust.[ 2 ] By 1890, it controlled 90% of U.S. oil refineries.[ 3 ]

The fundamental idea behind the "trust," the instrument of John D. Rockefeller and his partners' success, was that businessmen could make more money by not competing.[ 4 ] That idea made them rich. Rockefeller would become the first billionaire this country has ever seen.[ 5 ]

Copycats followed: the American Cotton Seed Oil Trust, with 75% of the country's production capacity; the Sugar Trust, with 85% of the capacity in the eastern United States; the Whiskey Trust; and the Beef Trust, which combined the four great Chicago meatpackers.[ 6 ]

But the public caught on. Outrage started to build, first in the state houses of New York, Kansas, and Iowa[ 7 ]-and then in Congress. By March of 1890, when he stood on the floor of the Senate to argue for the bill we now know as the Sherman Act, John Sherman was clear on who he was targeting and who he wanted to help. He was focused on:

a new form of combination . . . called trusts, that seeks to avoid competition by combining . . . corporations, partnerships, and individuals . . . often under the control of a single man called a trustee, a chairman, or a president.

. . . [Such a combination] can control the market, raise or lower prices, as will best promote its selfish interests . . . . The law of selfishness, uncontrolled by competition, compels it to disregard the interest of the consumer. It dictates terms to the transportation companies, it commands the price of labor without fear of strikes, for in its field it allows no competitors.[ 8 ]

That last line-"commands the price of labor without fear of strikes"-is telling; Senator Sherman was driven, in part, by the need to protect labor. From the beginning, American antitrust law aimed to protect worker organizing-not limit it.

This approach to labor within antitrust seems at odds with modern law enforcers' actions. One example from 1999, 110 years after Senator Sherman's floor remarks, helps underscore that disparity. It involves a group of truckers, the United Container Movers Association (UCMA), who worked the docks on the eastern and western seaboards, as well as the Gulf Coast-along with the authors' own agency, the Federal Trade Commission.

The truckers in question did not get a salary or an hourly wage. Most of them were technically independent contractors who got paid for each trailer they moved from one place to another at a rate of $35 per trailer. Over the years, the shipping companies raised their rates to account for the rising price of oil. But the truckers' $35 per trailer stayed the same.

After expenses, the truckers often barely made enough to live on. According to Tony Fernandez, president of one UCMA local, many harbor truck drivers worked "18 hours a day for less than the minimum wage."[ 9 ] Drivers sat "in long lines[,] . . . haul[ed] broken-down chassis[,] . . . and experience[d] verbal abuse . . . just to make the payments on [their] trucks and put food on the table."[10] "Our people are starving," said Mr. Fernandez. "They're losing their trucks."[11]

Another organizer traced the problem to an absence of collective action. "[T]he problem is that workers have no voice. They are not speaking collectively to any employers. They have no standing to articulate their concerns."[12]

Yet every time the truckers in one port got organized to ask for an increase on the $35 rate, the shipping companies would just divert their ships one port over. So, led by Robert Bates, the president of the UCMA in Charleston, the truckers in Baltimore, Charleston, Galveston, Jacksonville, Los Angeles, Long Beach, and Seattle decided to meet for the first time and try to get the port truckers a union contract.[13]

Then, ten days before the meeting, and two days before Thanksgiving, the FTC issued subpoenas to Mr. Fernandez, Mr. Bates, and their colleagues, directing them to testify before the Commission. According to press reports, the subpoenas explained that the Commission was investigating whether the truckers were "engaging in unfair methods of competitive pricing" because they were independent contractors rather than employees.[14]

Truckers used to earn a solid, middle class living. Now, their earnings are a shadow of what they used to be.[15] What happened to trucking is part of a much broader trend: companies are demanding more and more control over their workers while taking less and less responsibility for them. They do this, typically, by hiring or misclassifying them as independent contractors. This trend has disproportionately affected lower paid and more dangerous jobs.[16]

Unionizing might give these workers a way out-a way to fight for better wages, better benefits, and better working conditions. Unionized truckers, for example, typically earn 20% more than their non-union counterparts.[17] But as Mr. Bates and Mr. Fernandez discovered, when independent contractors organize, they are often accused of breaking our nation's antitrust laws.[18]

In other words, because of antitrust, the people most vulnerable to mis-treatment are the ones least capable of organizing to stop it. We need to ask ourselves: Is this really what Congress intended?

"Antitrust"-the body of law born to rein in John D. Rockefeller and the Oil Trust, the Beef Trust, and the Sugar Trust;[19] did Congress really mean for that law to target Tony Fernandez? Did it really mean to target uninsured truck drivers barely making minimum wage? And did it really aim to block their union contract?

These are the questions driving this article.[20] Part II will explore the back and forth of congressional action and judicial interpretations of how the antitrust laws apply to labor; the former seeking to exempt labor from antitrust enforcement and the latter undermining congressional intent by narrowly reading those exceptions. Part III then explores the approach to the labor exemption in a recent First Circuit decision, Confederación Hípica de Puerto Rico, Inc. v. Confederación de Jinetes Puertorriqueños, Inc.,[21] which, in the authors' view, reflects a reading of the labor exemption that is more faithful to Congress's intent in enacting it.

II. THREE WINS IN CONGRESS, THREE LOSSES IN THE COURTS

The need to protect worker organizing was at the center of congressional antitrust debates for forty years. In 1890, 1914, and 1932, Congress amended the law to make sure it was not used to stop worker organizing. But courts turned each effort on its head.

As early as February 1889, during a debate around the predecessor bill to the Sherman Act, Senator James George of Mississippi warned that, as written, the bill could be turned against "farmers and laborers," because language in the original bill banned any combination that raised prices to consumers.[22] Senator George thought that this language could be turned against "working men" organizing for better wages since increased wages may increase the price of goods.[23]

It was not only Senator George who issued this warning. The same warning also came from Senators Frank Hiscock, Henry Teller, and William Stewart. Each senator was specifically concerned that the language on increasing price to consumers would be used to stop worker organizing.[24]

During the ensuing floor debates, Senator Sherman shot back that organizing workers "are not affected in the slightest degree, nor can they be included in the words or intent of the bill."[25] Nevertheless, in response, the bill was amended twice to clarify that it would not apply to combinations of workers trying to reduce hours or increase pay.[26]

The bill was then referred to the Judiciary Committee. There, the language around consumer price-again, the key language that had driven Senators George, Hiscock, Teller, and Stewart to warn that the bill might be read to cover labor-was dropped. As a result, the labor exemptions intended to protect workers against that language were likewise dropped.[27]

The law passed the Senate on April 8, 1890, by a vote of 52-1. Who voted for it? Senators Teller and Stewart; even Senator George said he would vote for it (he was marked as "absent" for the vote).[28]

Unfortunately, courts effectively ignored that legislative history. Instead, they looked at the statute, saw no express labor exemption, and proceeded to turn the Sherman Act into a "savage weapon" against working people trying to organize.[29]

Shortly after its passage, the Act was used against longshoremen in New Orleans. In 1892, they organized black workers, white workers, "printers, hearse drivers . . . musicians[,] and carpenters" into a 20,000-person strike.[30] Federal prosecutors indicted the organizers under the Sherman Act.[31] A federal judge saw that the law was focused not on labor organizing but on "the evils of massed capital"-but then upheld the injunction against the workers anyway.[32]

The Act was used against people working 16 hours a day for the Pullman Palace Car Company in Illinois. In 1894, their wages were cut by 25%. They started running out of food. A delegation of workers asked to meet with the company's president, George Pullman. He fired them instead. So, they organized a strike.[33] Prosecutors used the Sherman Act against the union organizers. On a petition for a writ of habeas corpus, the Supreme Court voted 9-0 to uphold the contempt-of-court conviction based on that injunction and sent the lead organizer, Eugene Debs, to prison.[34]

The Act was also used against 250 hatters in Danbury, Connecticut. Hat making may seem quaint and harmless; in reality, industrial manufacture of felted fur hats required extensive use of mercury. A study would later find that, of 100 union hatters in Danbury, 43 had mercury poisoning. The report said that "[b]oys 20 and 21 years old are already so badly poisoned that their hands shake continually, while many of the men who have served longer at the trade cannot even feed themselves." People often accused the hatters of being drunk.[35]

In 1902, these men sought to unionize hat maker D.E. Loewe & Company. Mr. Loewe sued the men, along with national union leaders, under the Sherman Act.[36] The Supreme Court sided with Loewe, 9-0. The Court focused on the fact that the men had worked with non-employee union leaders to call for boycotts of anyone who did business with Mr. Loewe.[37]

Mr. Loewe won a settlement equaling $6.8 million in 2023 dollars. To satisfy the settlement, he seized the hatters' family homes and prepared them for auction. They were only saved after union members across the country donated an hour of their wages to support the men and allow them to buy back their homes.[38]

These narrow readings of the Sherman Act infuriated Congress; that fury crested in the passage of the Clayton Act in 1914. One senator said that the Sherman Act had been "tortured into a meaning" that transformed a law "intended for the relief of the plain people . . . into an instrument for their oppression."[39] Another member of Congress, Thomas Konop of Wisconsin, declared: "We are aiming at the gigantic trusts and combinations of capital. We are aiming at the dollars and not at men."[40]

Congress had seen in the Danbury Hatters case how courts were wrongly applying the Sherman Act to "the energies and activities of associated laborers."[41] So they drafted language to protect labor organizing. Section 6 of the Clayton Act insists that "[t]he labor of a human being is not a commodity or article of commerce."[42] Section 20 establishes protections for employees and then sets out an additional set of protections for "any person"; no antitrust injunction "shall prohibit any person or persons" from stopping work, from telling others to stop work, from telling others to boycott a business-the list goes on.[43]

This time, federal courts read the Clayton Act's labor exemptions so narrowly that they effectively deleted them from the law.44 Courts in the 1920s enjoined over 2,[100] strikes.[45] One scholar even concluded that the Clayton Act did not stop any injunctions from issuing under the Sherman Act.[46]

An emblematic case occurred in in Mingo County, West Virginia, in the early 1920s, where coal miners tried to unionize a historically non-union coal field. Their wages had been increasing, but they had not kept pace with infla tion, nor had they matched the 600% earnings increase for the mine's owners, the Red Jacket Coal Company.[47]

One of the union miners, Frank Ingham, would later testify before the Senate about the tricks the coal company used to keep their wages low. The men were paid by the carload of coal, but they were not paid in cash. Workers were paid in scrip redeemable only at the company store.[48] When the men asked for a 10-cent increase over their previous rate of 66 cents per car, the company gave them a 9-cent increase. But then, the next time the men came out of the mines, every item in the company store had been marked up by 5 to 25 cents.[49]

The men called a strike. The company sued under the Sherman Act. The Fourth Circuit said the union leaders involved in the dispute "are neither ex-employees nor seeking employment," and that therefore the labor exemption in the Clayton Act did not apply.[50]

Yet again, Congress was outraged. Congressman Fiorello LaGuardia denounced the "few . . . Federal judges" who had "willfully disobeyed the law."[51] He continued: "[T]hey emasculated it; they took out its meaning as intended by Congress; they made the law absolutely destructive of the very intent of Congress."[52 ]In response, Congress passed the Norris-LaGuardia Act of 1932.[53]

The Red Jacket opinion had relied on a Supreme Court case called Duplex Printing, where the Court said that the Clayton Act labor exemption only protected people who were "standing in [the] proximate relation" of employer and employee.[54] So Congress took those exact words and said the opposite. Thus, Norris-LaGuardia expressly directs that the antitrust exemption for labor disputes applies "regardless of whether or not the disputants stand in the proximate relation of employer and employee."[55]

As it had done in the Clayton Act, Congress also broadly declared as policy that the Norris-LaGuardia Act aimed to restore "actual liberty of contract" to the "individual unorganized worker."[56] However, in contrast to what happened after the Clayton Act, this time the courts were more restrained; the Lochner era[57] was coming to an end. But narrow readings of the labor exemption still followed.

In Columbia River Packers, the Court excluded from the protections of the labor exemption a group of fishermen on the grounds that they were "independent businessmen" selling commodities, not their labor.[58] That decision would be used to exclude from the exemption a range of other workers classified as independent contractors, some of whom did not sell commodities. In at least two circuit courts, this allowed for antitrust suits against people like the port truckers and Mr. Fernandez.[59]

Today, a lot of people like Mr. Fernandez may be technically classified as contractors, but they are not independent. Fifty years ago, "owner-operators" were much more likely to haul for multiple clients, with multiple trucks, and multiple employees. By the late '90s, the vast majority had one client, one truck, and one employee-themselves. In fact, 90% of them had permanent leases that required them to haul exclusively for one company.[60] The control exerted through those exclusive leases is exacting: "They can tell you when to breathe," said one trucker.[61]

"Control" matters in interpreting the labor exemption. When the Supreme Court declined to protect fishermen organizing in Columbia River Packers, it did so not just because they sold commodities, and not just because they were "independent businessmen." The Court declined to protect them because they were "independent businessmen, free from such control as an employer might exercise."[62]

III. THIRTY-SEVEN JOCKEYS IN CANÓVANAS

A recent First Circuit decision offers a different, and in our view more faithful, way of approaching the labor exemption. The facts underlying that case center on a horse track about an hour outside of San Juan, Puerto Rico.

Like hat making, horse racing may seem harmless. Consider, however, that racehorses can weigh almost 1,500 pounds, sprint at 55 miles an hour, and hit the ground with as much as 3,000 pounds of force. The average jockey is sidelined by injuries multiple times a year.[63]

If you want to work as a jockey in Puerto Rico, there is one place you can do it: the Camarero racetrack in Canóvanas. And when you race, unless you finish in the top five, you only get paid what's called a "mount fee." In Puerto Rico, it's $20, a fifth of what jockeys are paid in the United States. That rate has not changed since 1987.[64] These rates keep most jockeys in poverty.[65] For years, the jockeys' association had demanded pay and benefits that do justice to their dangerous profession.[66]

In June 2016, the jockeys threatened a strike and demanded higher pay.[67] The horse owners wrote the jockeys a letter stating that they are independent contractors, and as such, they are not a union and therefore they cannot go on strike, as that would violate the antitrust laws, in particular the Sherman Antitrust Act.[68]

Thirty-seven jockeys went on strike for three days. The horse owners and the racetrack sued under the Sherman Act. The jockeys lost in district court. The judge awarded the horse and racetrack owners treble damages of well over one million dollars.[69] Like with the Danbury hatters 100 years earlier, however, the owners did not just hold the jockeys liable. They also sued each jockey's spouse or domestic partner-making that award recoverable against not just the jockeys, but also their families.[70]

On appeal before the First Circuit, Judge Sandra Lynch did not dwell on whether the jockeys were correctly classified as independent contractors, pointing to language in Norris-LaGuardia saying that this question did not matter "by the express text of the Norris-LaGuardia Act," because "a labor dispute may exist 'regardless of whether or not the disputants stand in the proximate relation of employer and employee.'"[71] She focused instead on what she saw as the core question in Columbia River Packers: whether what was at issue is "compensation for [the jockey's] labor."[72] She found that the jockeys were clearly selling their labor and thus held that the Norris-LaGuardia Labor Exemption applied.[73] She and her colleagues nullified the judgment and dis missed the case. The jockeys and their families won.

IV. AIMING AT DOLLARS, NOT AT MEN

"Antitrust"- a body of law written to rein in the Oil Trust, the Sugar Trust, the Beef Trust, and other trusts. A body of law aimed at "the gigantic trusts and combinations of capital," at "dollars, and not at men."[74] Did Congress really mean for that law to target 20-year-old hatters with mercury poisoning? Coal miners paid in worthless scrip? Three dozen jockeys risking their lives for $20 a ride? Is that really what Congress intended?

The answer to that question is a very obvious "no." What's more, a review of the legislative history reveals that Congress answered that question not once, not twice, but three times-each time in a louder and clearer voice.

While many courts may share Justice Antonin Scalia's famous aversion to legislative history,[75] we rely on a clearer authority. On the question of how broadly to read the labor exemptions in Clayton and Norris-LaGuardia, the Supreme Court has stated a preference for a separate, more specific standard: one espoused by then-Judge Oliver Wendell Holmes Jr., who the Supreme Court quoted in 1941 on this very topic:

The Legislature has the power to decide what the policy of the law shall be, and if it has intimated its will, however indirectly, that will should be recognized and obeyed. The major premise of the conclusion expressed in a statute, the change of policy that induces the enactment, may not be set out in terms, but it is not an adequate discharge of duty for courts to say: We see what you are driving at, but you have not said it, and therefore we shall go on as before.[76]

When it comes to antitrust and the labor exemption, we know the history. We know what Congress was "driving at." Congress meant to strengthen labor's hand when it fought the trusts, not weaken it. Therefore, we cannot "go on as before." Congress has made clear that worker organizing and collective bargaining are not violations of the antitrust laws. When antitrust enforcers consider investigations and policy matters, that history should guide them.

#### That freezes gig workers bargaining power.

Keefe ’23 [Dylan; Editor of the Kentucky Journal of Equine, Agriculture, & Natural resources Law, “Jockeying for the Labor-Dispute Exemption: How the First Circuit Wrongly Exempted Independent Contractors from Antitrust Law, But Why Congress Should Expand the Exemption,” vol. 16]

From jockeys and trainers to horseshoers and stable workers, many equine workers are classified as independent contractors. 169 In all but the First Circuit, these workers are not exempt from antitrust law. 170 Thus, they cannot organize for higher pay or better working conditions without facing injunctions or liability for damages. Essentially, they are unable to collectively bargain. Without this tool, the workers have little to no leverage in negotiations. For example, as discussed earlier in Confederacion Hipica, the district court awarded the racehorse owners and racetrack owner over 1 million dollars in damages.171 The jockeys in that case, who are paid a mere $20 per mount, would not have been able to pay such high damages had the First Circuit affirmed the district court's holding.172

In the United States, the average mounting fee- the amount that a jockey is paid to ride a horse during a race-is between $25 and $100.173 On average, jockeys earn approximately 40 thousand dollars per year, which does not account for expenses like equipment and agent fees.1 74 Because the Sherman Act allows injured parties to sue for triple the damages they suffered, jockeys are likely to avoid organizing and remain a vulnerable class of workers. 175

B. Independent Contractors in the Emerging Gig Economy Need Protection

The need to protect independent contractors extends beyond the equine industry. The modern economy has seen immense growth in gig work.1 76 This surge can be attributed to the relative ease of finding gig work on the internet with apps such as Uber, Lyft, Airbnb, and Postmates.1 77 It can also be attributed to the lasting impact of COVID-19, as the pandemic "upended the traditional 9-5 working days" and caused many workers to turn to gig work.178 Under the Third and Fourth Circuits' rulings, these workers do not have the right to organize for higher pay or better working conditions without being subject to injunctions and liable for damages. 179 Consequently, these workers have little leverage to negotiate and remain vulnerable to exploitation.

Historically, there was not and did not need to be much concern over the rights of independent contractors. 180 Until recently, many independent contractors were "highly trained professionals-electricians, plumbers, lawyers, doctors-who were not as vulnerable to mistreatment by buyers of their labor as ordinary employees were.'" 181 But today, many independent contractors are gig workers for major corporations.1 8 2 The most effective tool for these workers in negotiations would be the ability to boycott or picket. Thus, Congress should consider expanding the labor-dispute exemption to protect gig workers.

The First Circuit misinterpreted the text and legislative history of the labor-dispute exemption and misread Supreme Court precedent and the precedent of the other Circuits. In doing so, the First Circuit wrongly exempted independent contractors from antitrust law. On the other hand, the Third and Fourth Circuits correctly interpreted the labor-dispute exemption and Supreme Court precedent to conclude that the exemption applies only to labor disputes in which an employer-employee relation is the matrix of the controversy.183 An association or union may be exempt if it represents employees. 184 In Confederacidn Hipica, Jinetes should not have been exempted because it did not represent employees-instead Jinetes represented independent contractors. 185 An employer-employee relationship was not the matrix of the controversy between the jockeys, racehorse owners, and racetrack owner.186 In sum, the First Circuit should have found that the labor-dispute exemption did not protect the jockeys or Jinetes and should have affirmed the ruling in favor of the racehorse owners and racetrack owner.

Finally, without the option to boycott, workers in the equine industry and the gig economy lack the bargaining power necessary to negotiate for higher pay and better working conditions. Accordingly, Congress should consider expanding the labor-dispute exemption to provide much needed protection for independent contractors.

#### Poor working conditions in the gig economy are the largest driver of poverty.

Zipperer ’22 [Ben, Celine McNicholas, Margaret Poydock, Daniel Schneider, and Kristen Harknett; June 1; Senior Economist at the Economic Policy Institute, Ph.D., Economics, University of Massachusetts, Amherst; Director of Policy/General Counsel at EPI, J.D., Villanova University School of Law; Senior Policy Analyst at EPI; Professor of public policy and professor of sociology at Harvard University, Ph.D. from Princeton University in sociology and social policy; professor of sociology at the University of California, San Francisco; Economic Policy Institute, “National survey of gig workers paints a picture of poor working conditions, low pay,” https://www.epi.org/publication/gig-worker-survey/]

A survey of gig workers in the spring of 2020 revealed that their jobs provided poor working conditions, even relative to other service-sector workers, who themselves typically receive low pay.

About 1 in 7 gig workers (14%) earned less than the federal minimum wage on an hourly basis. More than a quarter (29%) earned less than the state minimum wage that would be applicable if they were a W-2 service-sector worker.

Three out of every 5 gig workers (62%) lost earnings because of “technical difficulties clocking in or out,” compared with 19% of W-2 service-sector workers.

One in 5 gig workers (19%) went hungry because they could not afford enough to eat. Thirty percent used the Supplementary Nutrition Assistance Program (SNAP) within a month of the survey, twice the rate of W-2 service-sector workers (15%).

Nearly one-third (31%) of gig workers did not pay the full amount of their utility bills in the month prior to the survey.

In the most basic terms, gig work can be defined as work done by individuals who are classified as self-employed, freelancers, or independent contractors. However, in recent years the term “gig work” has become synonymous with working for digital platform companies, including driving for ride-share apps, making deliveries for restaurants, shopping or delivering groceries, and performing errands or household tasks. In this use, “gig work” is a misnomer that helps companies propagate the myth that these workers have more independence and control over their work than they actually do.

Digital platform companies have constructed a business model on the premise that they do not employ their workforce. These companies treat workers who perform the services they offer not as employees but as independent contractors. By classifying their workforce in this way, they deprive workers of fundamental rights under federal and state labor and employment laws, including wage and hour protections, anti-discrimination protection, workers’ compensation, unemployment benefits, and the right to organize and collectively bargain.

Digital platform companies claim that their workforce benefits from this classification, enjoying the benefits of entrepreneurship with good pay and more flexibility than workers classified as traditional W-2 employees. However, a survey of gig workers reveals that these workers often are paid low wages, in some instances less than the minimum wage; they face economic insecurity at high rates; and they routinely report losing earnings because of technical difficulties with digital platforms.

The impact of gig worker misclassification

The determination of whether an individual providing services to an employer is an employee or an independent contractor carries significant consequences for both the individual and the employer in terms of job protections, tax obligations, and eligibility for employment-based benefits and protections.

As Table 1 shows, individuals who are classified as independent contractors are not covered by federal or state wage and hour, anti-discrimination, health and safety, collective bargaining, or other worker protection laws. They do not receive employment-based health insurance or retirement benefits, and they do not qualify for paid sick or family leave in places where those benefits are statutorily prescribed. Nor are independent contractors eligible for unemployment insurance when temporarily unemployed, or workers’ compensation when injured on the job. This leaves independent contractors in a far more vulnerable status, as compared with employees, when it comes to basic rights and protections on the job.

#### Inequality is the driver of existential climate change, nuclear war, and emerging tech---it drives the globe to existential collapse---risks are empirically verified.

Carrington ’25 [Damian, citing Luke Kemp; August 2; Environmental editor at the Guardian, citing Luke Kemp, Postdoctoral Researcher at the Center for the Study of Existential Risk, Research Associate at Darwin College, Ph.D. in Political Science and International Relations from the Australian National University; The Guardian, “‘Self-termination is most likely’: the history and future of societal collapse,” https://www.theguardian.com/environment/2025/aug/02/self-termination-history-and-future-of-societal-collapse]

“We can’t put a date on Doomsday, but by looking at the 5,000 years of [civilisation], we can understand the trajectories we face today – and self-termination is most likely,” says Dr Luke Kemp at the Centre for the Study of Existential Risk at the University of Cambridge.

“I’m pessimistic about the future,” he says. “But I’m optimistic about people.” Kemp’s new book covers the rise and collapse of more than 400 societies over 5,000 years and took seven years to write. The lessons he has drawn are often striking: people are fundamentally egalitarian but are led to collapses by enriched, status-obsessed elites, while past collapses often improved the lives of ordinary citizens.

Today’s global civilisation, however, is deeply interconnected and unequal and could lead to the worst societal collapse yet, he says. The threat is from leaders who are “walking versions of the dark triad” – narcissism, psychopathy and Machiavellianism – in a world menaced by the climate crisis, nuclear weapons, artificial intelligence and killer robots.

The work is scholarly, but the straight-talking Australian can also be direct, such as when setting out how a global collapse could be avoided. “Don’t be a dick” is one of the solutions proposed, along with a move towards genuinely democratic societies and an end to inequality.

His first step was to ditch the word civilisation, a term he argues is really propaganda by rulers. “When you look at the near east, China, Mesoamerica or the Andes, where the first kingdoms and empires arose, you don’t see civilised conduct, you see war, patriarchy and human sacrifice,” he says. This was a form of evolutionary backsliding from the egalitarian and mobile hunter-gatherer societies which shared tools and culture widely and survived for hundreds of thousands of years. “Instead, we started to resemble the hierarchies of chimpanzees and the harems of gorillas.”

Instead Kemp uses the term Goliaths to describe kingdoms and empires, meaning a society built on domination, such as the Roman empire: state over citizen, rich over poor, master over slave and men over women. He says that, like the biblical warrior slain by David’s slingshot, Goliaths began in the bronze age, were steeped in violence and often surprisingly fragile.

Goliath states do not simply emerge as dominant cliques that loot surplus food and resources, he argues, but need three specific types of “Goliath fuel”. The first is a particular type of surplus food: grain. That can be “seen, stolen and stored”, Kemp says, unlike perishable foods.

In Cahokia, for example, a society in North America that peaked around the 11th century, the advent of maize and bean farming led to a society dominated by an elite of priests and human sacrifice, he says.

The second Goliath fuel is weaponry monopolised by one group. Bronze swords and axes were far superior to stone and wooden axes, and the first Goliaths in Mesopotamia followed their development, he says. Kemp calls the final Goliath fuel “caged land”, meaning places where oceans, rivers, deserts and mountains meant people could not simply migrate away from rising tyrants. Early Egyptians, trapped between the Red Sea and the Nile, fell prey to the pharaohs, for example.

“History is best told as a story of organised crime,” Kemp says. “It is one group creating a monopoly on resources through the use of violence over a certain territory and population.”

All Goliaths, however, contain the seeds of their own demise, he says: “They are cursed and this is because of inequality.” Inequality does not arise because all people are greedy. They are not, he says. The Khoisan peoples in southern Africa, for example, shared and preserved common lands for thousands of years despite the temptation to grab more.

Instead, it is the few people high in the dark triad who fall into races for resources, arms and status, he says. “Then as elites extract more wealth from the people and the land, they make societies more fragile, leading to infighting, corruption, immiseration of the masses, less healthy people, overexpansion, environmental degradation and poor decision making by a small oligarchy. The hollowed-out shell of a society is eventually cracked asunder by shocks such as disease, war or climate change.”

History shows that increasing wealth inequality consistently precedes collapse, says Kemp, from the Classical Lowland Maya to the Han dynasty in China and the Western Roman empire. He also points out that for the citizens of early rapacious regimes, collapse often improved their lives because they were freed from domination and taxation and returned to farming. “After the fall of Rome, people actually got taller and healthier,” he says.

Collapses in the past were at a regional level and often beneficial for most people, but collapse today would be global and disastrous for all. “Today, we don’t have regional empires so much as we have one single, interconnected global Goliath. All our societies act within one single global economic system – capitalism,” Kemp says.

He cites three reasons why the collapse of the global Goliath would be far worse than previous events. First is that collapses are accompanied by surges in violence as elites try to reassert their dominance. “In the past, those battles were waged with swords or muskets. Today we have nuclear weapons,” he says.

Second, people in the past were not heavily reliant on empires or states for services and, unlike today, could easily go back to farming or hunting and gathering. “Today, most of us are specialised, and we’re dependent upon global infrastructure. If that falls away, we too will fall,” he says.

“Last but not least is that, unfortunately, all the threats we face today are far worse than in the past,” he says. Past climatic changes that precipitated collapses, for example, usually involved a temperature change of 1C at a regional level. Today, we face 3C globally. There are also about 10,000 nuclear weapons, technologies such as artificial intelligence and killer robots and engineered pandemics, all sources of catastrophic global risk.

Kemp says his argument that Goliaths require rulers who are strong in the triad of dark traits is borne out today. “The three most powerful men in the world are a walking version of the dark triad: Trump is a textbook narcissist, Putin is a cold psychopath, and Xi Jinping came to rule [China] by being a master Machiavellian manipulator.”

“Our corporations and, increasingly, our algorithms, also resemble these kinds of people,” he says. “They’re basically amplifying the worst of us.”

Kemp points to these “agents of doom” as the source of the current trajectory towards societal collapse. “These are the large, psychopathic corporations and groups which produce global catastrophic risk,” he says. “Nuclear weapons, climate change, AI, are only produced by a very small number of secretive, highly wealthy, powerful groups, like the military-industrial complex, big tech and the fossil fuel industry.

“The key thing is this is not about all of humanity creating these threats. It is not about human nature. It is about small groups who bring out the worst in us, competing for profit and power and covering all [the risks] up.”

The global Goliath is the endgame for humanity, Kemp says, like the final moves in a chess match that determine the result. He sees two outcomes: self-destruction or a fundamental transformation of society.

He believes the first outcome is the most likely, but says escaping global collapse could be achieved. “First and foremost, you need to create genuine democratic societies to level all the forms of power that lead to Goliaths,” he says. That means running societies through citizen assemblies and juries, aided by digital technologies to enable direct democracy at large scales. History shows that more democratic societies tend to be more resilient, he says.

#### Economic precarity in the gig economy is an existential threat.

Parfitt ’20 [Claire and Tom Barnes; May 13; Ph.D. candidate in the Department of Political Economy at the University of Sydney; economic sociologist and Senior Research Fellow at the Institute for Humanities and Social Sciences, Australian Catholic University; Marxist Sociology Blog, adapted from an article in Critical Sociology, “Precarity and the Politics of Existential Crisis,” https://marxistsociology.org/2020/05/precarity-and-the-politics-of-existential-crisis/]

What new meanings does the concept of precarity adopt when society is suddenly plunged into a deep, prolonged, even existential crisis? While it has been used for decades, the publication of Guy Standing’s The Precariat in 2011 was critical to popularizing this concept. Today, precarity seems to be everywhere. The erosion of the 9-to-5 working day, the emerging gig economy, zero-hours contracts, rising self-employment and agency work are all signs that contingent work is the new normal. In Australia, where we write from, at least half of the workforce can now be regarded as contingent. This figure is much higher in many other places around the world.

But for many, there is nothing ‘new’ about this normal. Feminists in particular have pointed out that ‘standard’ employment relationships were always an exception enjoyed primarily by white men in wealthy economies. Arguments that hinge on the novelty of precarious work have unsurprisingly drawn criticism for their failure to acknowledge the diversity of economic lives across time and space.

In contrast to those who say that this concept has been stretched too far, our recent special issue in Critical Sociology emphasizes the value of a broad, multidimensional understanding of precarity. Following Nancy Ettlinger, we see precarity as a ‘condition of vulnerability relative to contingency and the inability to predict’. Although our work was compiled prior to the coronavirus pandemic, this broad orientation is useful in a context where lives and livelihoods are exposed to so many manifestations of risk. It invites us to think about what makes for vulnerability and resilience, the different types of risks we are exposed to, and how these can be negotiated individually and collectively.

What emerges through this kind of analysis is that exposure to risk and its social and economic impacts are widespread, even in wealthy economies and populations. Many of us find ourselves living in a ‘speculative life-world’ in which we are ‘condemned to decision making under uncertain levels of uncertainty, and to thus precarity and insecurity’. But just as we find exposure to risk in unlikely places, we also observe unexpected instances of resilience and collectivization of risk.

For us, a keynote example which preceded the coronavirus pandemic were the bushfires which ravaged Australia from September 2019 to January 2020. The destruction of forests released hundreds of millions of tonnes of carbon from the ground and increased emissions. In Australia, the coronavirus pandemic emerged on the back of the devastating experience of these bushfires, deepening a sense of widespread anxiety about the future.

Australia’s bushfires fires rendered material the existential threat of climate change in unprecedented ways. A textbook example of Ettlinger’s ‘condition of vulnerability relative to contingency and the inability to predict’, the fires revealed the limits of humanity’s control over nature, our inability to predict it, and the extent of our vulnerability to it.

The bushfires had a profound impact on the continent’s environment and population. Fires killed dozens of people, over one billion animals, and razed over 12.6 million hectares of forest. Around 3500 homes and countless livelihoods were destroyed while millions of people choked on smoke. The bushfires thrust the precarity of life to the forefront and generated a new national mood in which summer. Once a time to look forward to, it became a time to dread. This broader sense of fear intersected with the fields of work and education to generate new types of precarity. The fear of allowing children to play outdoors with hazardous air quality increased pressure on parents and educators. Outdoor-based workers in construction or horticulture were expose to the risk of respiratory illness from smoke haze. This reinforces the findings of one paper in our collection regarding the workplace as a site of ecological struggle.

Having endured the anxieties of the worst bushfire season on record, Australia was immediately drawn into the coronavirus pandemic. The full impact of this crisis is still being understood. But, like in many other countries, it has been met with an economic shutdown which has induced potentially the worst social and economic conditions since the Great Depression. Unemployment is predicted to triple to 15 percent, with some predicting a figure as high as 20 percent.

The depth of the crisis has brought about policy shifts which were unthinkable only weeks ago: the politically-conservative Australian Government has doubled the rate of payment for unemployment insurance, fully subsidised childcare for most households who need it, proposed a moratorium on evictions due to financial distress, and issued income support for businesses to continue paying their workers.

The policy response primarily defends capital from precarity by supporting on-going accumulation, while reinforcing established divisions within labor. The government has promised that the above measures are temporary features. A return to the old ‘normal’ is to be expected, including Australia’s punitive and workfarist model of unemployment insurance. There is no income support for more than a million migrant workers and short-term casual workers. Many migrant workers as well as international students have been denied access to healthcare. Refugees are confined to hazardous detention camps, as they are around the world.

For low-paid workers in Australia, the government’s asset-based approach to welfare is an empty gesture. Australia has a compulsory private pension system—known as superannuation – through which roughly ten per cent of workers’ wages is surrendered to financial institutions and held until retirement. This system enables those in secure, high-paid jobs to amass large savings with generous tax concessions, while denying low-paid workers cash when they need it. During the pandemic response, the Australian Government has encouraged workers to access their retirement savings. But those who are most in need of emergency funds, such as workers in tourism, hospitality and retail, tend to have some of the lowest savings balances. Furthermore, given the recent collapse in financial markets, superannuation accounts have been decimated. Workers who are required to draw on those funds will be forced to realize their losses rather than wait for a resurgence in the market.

The crisis is also generating precarity among seemingly stable sections of the population—so-called ‘Middle Australia’. Australia’s high levels of private home ownership are accompanied by unprecedented levels of household debt. Australia’s central bank has repeatedly warned of the risk these debt levels pose, not only to particular households, but to financial stability in the economy. Mass unemployment, and the prevalence of contingent work throughout the economy, has pushed millions of households to the brink of default on mortgages, rent and other debts.

While the financial dimensions of the crisis reflect familiar conflicts of interest between the wealthy and the poor, these interests manifest differently in an economy built on debt and financial assets. Several papers in our special issue consider the role of finance in a world of constantly shifting risks, exploring the ways in which finance can be both a tool for managing risk and a vehicle for its accentuation.

Perhaps unsurprisingly, those most able to manage the economic impacts will be those with access to household wealth, a conclusion that is brought into sharp relief by a paper in our special issue on the experiences of retrenched workers. At the same time, the social crisis of the pandemic, following the socio-ecological crisis of the bushfires, highlights different aspects of precarity. New iterations of vulnerability emerge and are filtered through familiar distinctions of class, gender and race.

#### The plan solves---Applying the anti-trust exemption to gig workers opens up avenues for collective bargaining. Gig workers must be able to collectively bargain. Anything else guarantees precarity.

Jacob ’23 [Josh; November 20; J.D.,Columbia Law School; Columbia Law Review, “Avenues for Gig Worker Collective Action After Jintes,” vol. 123]

Gig workers constitute an ever-increasing share of the American workforce, yet they are not afforded the rights to strike and bargain collectively under the National Labor Relations Act (NLRA) due to their independent contractor status. Independent contractors who attempt to act collectively face antitrust liability, whereas employees who are covered by the NLRA enjoy an antitrust exemption for the same collective action, known as the “labor exemption.” Observers have speculated that the First Circuit, in the recent case Confederación Hípica de Puerto Rico, Inc. v. Confederación de Jinetes Puertorriqueños, Inc. (Jinetes), 30 F.4th 306 (1st Cir. 2022), has begun to remedy the exclusion of gig workers from the labor exemption by holding that workers engaged in a labor dispute may benefit from the exemption regardless of their employment status.

This Comment argues that courts following the First Circuit’s lead may afford the Jinetes reasoning either a narrow or a broad interpretation and that the latter should be adopted because it would promote gig worker collective action. Under the narrow interpretation, most gig workers are still excluded from the labor exemption and face many of the same challenges as before. Under the broad interpretation, gig workers may enjoy new organizing avenues through striking, which has been successful for gig workers internationally, and through state and local regulatory frameworks, which have succumbed to antitrust scrutiny in the past. The contrasting interpretations reflect competing approaches to the antitrust laws. As the modalities of work change, so too should the understanding of the antitrust labor exemption.

INTRODUCTION

As work shifts away from the traditional employer–employee relationship,1 alternative forms of organizing are more important than ever. The COVID-19 pandemic has shown that workers rely on their employers to provide workplace safety measures,2 job and income stability,3 and health insurance benefits,4 especially in times of crisis. Yet with more workers taking part in the gig economy,5 these assurances are becoming harder to secure. Indeed, roughly half of gig workers feel that their gig platforms do not adequately provide unemployment, health care, and paid leave benefits.6 More than a third of gig workers say they have been harassed or have felt unsafe at work.7 With one in six Americans reporting that they have earned money from an online gig platform, these inadequacies affect a broad swath of the population but disproportionately impact young, Hispanic, and low-income workers.8

One way that workers have historically remedied precarity in the workplace is through collective action.9 Workers who coordinate their efforts and negotiate collectively with their shared employer are better positioned to determine the terms and conditions of their employment.10 This option is not available to gig workers, however, who are typically classified as independent contractors rather than employees and are thus excluded from the striking and collective bargaining protections of the National Labor Relations Act (NLRA), which apply only to employees.11 The law treats gig workers as independent businesspeople, so raising wages through collective bargaining is considered price-fixing, and striking a gig platform is considered a group boycott—both per se violations of section 1 of the Sherman Act.12 Gig workers who do attempt to act collectively therefore are likely to face antitrust liability.13 Workers have long been exempted from antitrust liability through statutory carveouts from the antitrust laws, collectively known as the “labor exemption,”14 but courts have traditionally excluded independent contractors from this exemption.15

The First Circuit in the recent case Confederación Hípica de Puerto Rico, Inc. v. Confederación de Jinetes Puertorriqueños, Inc. (Jinetes) made a step toward including gig workers in the exemption by holding that workers engaged in a labor dispute may benefit from the exemption regardless of their employment status.16 Courts in the First Circuit may construe this holding either narrowly or broadly, and other courts may similarly choose to adopt narrow or broad interpretations of the reasoning—or ignore it altogether. The scope and breadth of this interpretation will largely determine the policy effects of Jinetes. If the case is interpreted narrowly, most gig workers would still be excluded from the antitrust labor exemption, and their organizing options would be accordingly limited.17 If the case is interpreted broadly, not only would gig workers be able to strike and collectively bargain without inviting antitrust lawsuits, but states and municipalities would be able to enact affirmative protections granting these rights to gig workers.18 These state and local laws could protect workers from being fired or disciplined for engaging in collective action and even establish sectoral bargaining frameworks to set industry-wide standards.19

#### The plan is doctrinally constrained.

Estreicher ’25 [Sam and Jack Samuel; March 25; Dwight D. Opperman Professor of Public Law at NYU School of Law, where he directs the Center for Labor and Employment Law; associate at a global law firm; On Labor, “Labor’s Antitrust Immunity for Independent-Contractor Workers,” https://onlabor.org/labors-antitrust-immunity-for-independent-contractor-workers/]

The emergence of gig work is putting pressure on traditional notions of who is an employee and who is the employer. Workers classified as independent contractors rather than employees can lose state and federal protections for wages, overtime, whistleblowing, discriminatory firing, and more. They also lose federal labor law protections for group protest activity, union organizing, and collective bargaining. It is generally assumed that workers not classified as employees under federal labor law also face liability under the antitrust laws if they form unions, go on strike, or try to bargain collectively with those who hire them.

In a recent article, we challenge the assumption that only workers covered as employees by federal labor law are antitrust-exempt: employer classification of workers as “independent contractors,” whether well-founded or not, is irrelevant to the antitrust inquiry. As long as workers provide only their personal services without significant, non-fungible capital investment, they remain laborers for purposes of the exemption from the antitrust laws that permits collective action by workers.

In the late nineteenth and early twentieth century employers used various legal tools, including federal antitrust law, to blunt the momentum of labor organizing. Notwithstanding that the 1890 Sherman Act was passed to target the monopolistic excesses of large business trusts, courts relied on the vague language of the Act’s prohibition of agreements “in restraint of trade” to issue injunctions against strikers, picketers, and any form of labor action involving violence, social pressure, or “moral intimidation.”

In 1914 Congress attempted to clarify that antitrust law was not meant to police collective action by workers. Section 6 of the Clayton Act—dubbed “labor’s magna carta” by Samuel Gompers, head of the American Federation of Labor—declared that the labor of a human being was not an article of commerce. Section 20 provided that collective refusal to work, picketing, and boycotts could not be federally enjoined under the antitrust laws. At the behest of business groups federal courts pushed back, reading the Clayton Act’s prohibition of the labor injunction narrowly, until Congress enacted the 1932 Norris-LaGuardia Act and finally ended the abuses of the “labor injunction.” In the late 1930s, influenced in part by the 1935 passage of the National Labor Relations Act (NLRA), the Court established the “statutory labor exemption” from all antitrust scrutiny for actions by “labor acting alone,” not in combination with businesses, in any “dispute concerning terms and conditions of employment.”

The statutory labor exemption is often thought, without careful analysis, to apply only to the workers classified as employees under the NLRA, effectively carving out from antitrust liability only collective activity protected by labor law. Thanks to the 1947 Taft-Harley amendments to the NLRA, that would exclude independent contractors as determined by the common law’s “right to control” test. But the labor-antitrust exemption predates Taft-Hartley, and its statutory basis predates the NLRA. Taft-Hartley left the Clayton and Norris-LaGuardia Acts untouched. And, notwithstanding a line of cases in which the Court declined to apply the exemption to independent contractors selling goods or otherwise in business for themselves, the Court has never held that the common-law “right to control” test delimits the statutory labor exemption. To understand the exemption’s true scope we must turn to statutory language and context. As our article shows, the original meaning of the language used in the Clayton and Norris-LaGuardia Acts clearly establishes an exemption for all workers, not just employees within the meaning of the common law of agency.

An exempt “labor dispute” is defined in Norris-LaGuardia as “any controversy concerning terms or conditions of employment.” Based on this language, the Chamber of Commerce (among others) has argued that “[a] dispute between a business and independent contractors it has retained . . . does not concern ‘employment’ and thus is not a ‘labor dispute’ within the meaning of the Act.” But the Chamber’s gloss on “employment” is anachronistic. “[T]he dictionaries of the era,” the Supreme Court has observed, “consistently afforded the word ‘employment’ a broad construction, broader than may be often found in dictionaries today . . . . treat[ing] ‘employment’ more or less as a synonym for ‘work.’” “Work,” in turn, was treated as a category no less broad than it is now—the same dictionaries did not “distinguish between different kinds of work or workers: All work was treated as employment, whether or not the common law criteria for a master-servant relationship happened to be satisfied.”

This broad sense of “employment” finds support in cases decided at the time the Norris-LaGuardia and Clayton Acts were passed. In cases arising from labor disputes courts at that time routinely used “employee” interchangeably with “worker,” “workingman,” “laborer,” and “labor.” At a time of great labor unrest, courts routinely referred to controversies involving strikes and pickets as “labor disputes” without first inquiring into whether the workers involved were common-law employees—in some cases applying the term to industrial conflict between companies and workers like organ installers, plumbers, and sign-painters who would likely have been independent contractors at common law. We were unable to find a single case in which the distinction between an employee and an independent contractor appeared in a discussion of a labor dispute prior to or contemporaneous with the passage of the Norris-LaGuardia Act. It wasn’t until nearly a decade after the Noris-LaGuardia Act was passed that state courts began to consider whether a labor dispute could exist between independent contractors and the users of their services. There is no reason to believe that as a matter of original statutory meaning the Clayton or Norris-LaGuardia Acts limit their protection to common-law servants.

There are understandable reasons why the definition of covered employee in a given labor or employment statute does not alter the applicability of the antitrust exemption. Labor laws balance competing goals, such as the employees’ right to engage in concerted activities and the employer’s ability to manage the workforce, which may call for restrictions on statutory coverage that are generally irrelevant to the policies of the labor-antitrust exemption, which deals with what labor can do “acting alone,” before any agreements with employers or other users of their services. Agricultural workers are expressly excluded from the NLRA and its protections, yet those workers remain free to form unions and pursue collective bargaining without the intervention of antitrust laws. Supervisors are also expressly excluded from the protection of the NLRA because they generally function as management’s agents, but they can join unions and seek collective bargaining, outside of NLRA protection but free of antitrust liability.

The same is true of gig workers who, although they may be excluded from the protections of the NLRA, are within the shelter of labor’s statutory antitrust exemption. And if such workers are not statutory employees under the NLRA, federal labor law preemption does not apply, which means that states can legislate organizing protections on their behalf without running afoul of either labor law or antitrust restrictions (as illustrated by the 2024 Massachusetts initiative establishing a framework for union recognition and collective bargaining between “ride-hail” drivers and the platform companies). Recognition of the antitrust immunity for independent-contractor workers would mark an important step forward in the economic freedom of these workers, not requiring any statutory change or overturning of Supreme Court precedent, to engage in collective action for their betterment.

#### It's reverse causal---collective bargaining counterbalances monopsonist labor control.

Steinbaum ’19 [Marshall; September 5; Assistant Professor of Economics at the University of Utah and a Senior Fellow in Higher Education Finance at Jain Family Institute, PhD in Economics from the University of Chicago, trained as an empirical labor economist; Law and Contemporary Problems, “Antitrust, The Gig Economy, and Labor Market Power,” vol. 82]

The erosion of antitrust in the direction of permitting vertical price- and nonprice restraints has effectively legalized labor outsourcing, misclassification, and the gig economy. This has resulted in dominant firms having access to a wider range of profitable business models that exert greater power and control over workers than they once did. Fundamentally, this trend within antitrust is in the direction of increasing the power of the economy’s most powerful actors.

The flip side of this is that antitrust law has also made it more difficult for less powerful actors to collectively mitigate such power inequities. Sandeep Vaheesan refers to Albert Hirschman to make this point: not only has antitrust made it harder for workers and small businesses to exit in order to exercise countervailing power; it has also prevented them from using voice to do so.69 All of the mechanisms of concentrated power described in Part II could be categorized as curtailing workers’ use of exit strategies to evade the control of their employers. This Part focuses on antitrust’s dual opposition to worker voice.

Between the passage of the Sherman Act in 1890 and the Norris-LaGuardia Act in 1932, the federal antitrust enforcers used the former to curtail the collective bargaining activities of militant (and effective) labor organizing. In 1892, the Supreme Court ruled that the Workingmen’s Amalgamated Council of New Orleans was illegal coordination by labor groups in violation of the Sherman Act’s ban on restraints of trade.70 In 1894, the Cleveland Administration accused Eugene V. Debs, the head of the American Railway Union, of entering into criminal restraints of trade for organizing the Pullman Strike, including a nationwide boycott of trains carrying Pullman Cars. In fact, as the Supreme Court ultimately ruled on the case, the Sherman Act proved to be unnecessary: the Court held that the government could obtain an injunction against the strike and imprisonment of Debs without any statutory authorization, as it amounted to an exercise of its legitimate law enforcement powers to crush civil unrest.71

Likewise, in the 1908 case Loewe v. Lawlor, a company that had been targeted by a nationwide boycott on the part of the American Federation of Labor successfully sued the union trying to obtain recognition as its workers’ bargaining agent under the Sherman Act.72 The Supreme Court agreed that such a boycott was an illegal restraint of trade and forced the union and its members to pay treble damages to their employer.73 These three cases show that the antitrust laws were a potent weapon in the hands of employers seeking to prevent unionization. During the same period, the government struggled half-heartedly to find a way to use the Sherman Act to limit corporate power, but it moved decisively, with the full cooperation of the judiciary, to use it to curtail labor power.74

The Clayton Act of 1914 included an exemption for labor from the antitrust laws,75 but courts interpreted it narrowly such that secondary boycotts were still illegal.76 It was not until the Norris-LaGuardia Act that unions were entirely immunized from antitrust liability. But in the jurisprudence of the so-called labor exemption that developed in the decade or so thereafter, antitrust immunity came to be connected to statutory employment status, like the right to collectively bargain itself.77 Therefore, in the current era of the erosion of statutory employment, we also have the erosion of the antitrust labor exemption.

The Federal Trade Commission has undertaken a long-running campaign against collective action by associations of professionals who seek to constrain entry, and in some cases, to forbid their members from soliciting business away from fellow members and to set minimum prices for their services. The commission has brought such cases against doctors, church organists, ice-skating instructors, music teachers, and public defenders.78 This enforcement line has accompanied a push by the agency to reduce state action, meaning the regulatory authority of states or municipalities to displace competition in favor of some other legitimate policy goal, notwithstanding prohibitions in federal law. For example, municipal taxi regulatory regimes, which limit the total number of taxis on the road, impede entry into the taxi business. However, this impediment has the legitimate purpose of preventing market saturation, thereby ensuring that driving a taxi is a viable full-time job. It also hopefully ensures that coverage is universal in both time and space and that a customer unfamiliar with the city can obtain a licensed and qualified professional rather than an unsafe or just unqualified service provider. Some of the FTC’s campaign against restrictions on competition in the market for service professionals consisted of attacks on licensing regimes that effectively protect incumbents and limit competition. The FTC’s “economic liberty task force” is devoted to this, as was the case the FTC litigated to the Supreme Court in 2015 and won: North Carolina Board of Dental Examiners. 79 The ruling held that a state board consisting primarily of members of the profession being regulated could not benefit from the state action exemption.80

Even before the recent campaign against the state action exemption, the FTC involved itself in efforts by independent contractors to organize themselves in response to the trucking deregulation that de-unionized the sector in the late 1970s. Port truckers aspired to bargain collectively against logistics companies that were coordinating trucking services on behalf of powerful wholesalers and retailers and subjected them to low pay, long hours, and thus high turnover. The FTC as well as quasi-public entities like port authorities intervened on behalf of those companies and accused the truckers of violating the Sherman Act during organizing drives in the late 1990s and 2000s.81 This stance is consonant with the rationale behind trucking and transportation deregulation in the first place: that inefficient suppliers, middlemen, and stakeholders were preventing efficiencies attending to unitary control from being realized in regulated industries. Therefore, competition in the deregulation era would drive down the rents being earned by those insiders. Permitting truckers to bargain collectively as contractors once they had been de-unionized would have sacrificed all those supposed gains. The FTC also dissuaded Ohio from passing a state law that would have allowed independent contractor home health aides to bargain collectively with staffing companies and their clients in 2008.82

The logic for these types of enforcement decisions can be found in the consumer welfare standard, just as was the case for vertical restraints imposed by dominant firms: protecting consumers is all that matters, and consumers are protected best when the most efficient firms have sufficient power and discretion to control the market, including at multiple levels of the supply chain, without having to reckon with any other stakeholder. On this reasoning, collective action by port truckers, home health aides, church organists, or gig economy workers is inefficient rent-seeking, and antitrust action against it “protects competition, not competitors.” The superior efficiency to be found in, for example, Uber having the power to surveil, direct, and fix prices for its drivers, despite their independent contractor status, would be threatened if drivers had the power to mediate that surveillance or price-setting through any kind of co-determination.

It is against this background of hostility to state and local regulation and collective bargaining that the FTC and the Department of Justice intervened in another antitrust case involving Uber. After the lawsuits alleging employment misclassification against Uber had been sent to arbitration, the Seattle City Council passed an ordinance granting collective bargaining rights to ridesharing drivers who are not employees.83 The Chamber of Commerce, acting on behalf of Uber, filed an antitrust claim against the city for facilitating a violation of the Sherman Act: collective bargaining over wages and working conditions by nonemployee drivers.84 After the federal district court sided with the city that its ordinance was covered by the state action exemption,85 the federal antitrust agencies filed an amicus brief in circuit court alleging that the state action exemption was limited to the customer-facing side of the taxi market and thus did not cover anti-competitive regulation of ride-sharing drivers.86 The Ninth Circuit Court of Appeals overturned the district court,87 setting up an antitrust trial about whether the ridesharing collective bargaining ordinance was, in fact, anti-competitive. The federal agencies further suggested in their brief that if not covered by the state action exemption, driver collective bargaining is a per se violation of the Sherman Act, a naked restraint with no possible pro-competitive justification.88

Rather than fight the case on the merits, Seattle modified the ordinance to remove collective bargaining over wages, in the hope of at least salvaging some version of collective bargaining without running afoul of antitrust laws.89 But the Chamber has apparently not been satisfied by that significant concession; in renewed filings, it demanded that the ordinance be wholly abandoned, because naming an exclusive bargaining agent for ridesharing drivers amounts to an illegal group boycott against any driver who does not wish to be represented collectively.90

At this point, it is clear that the federal agencies are fully behind the use of antitrust laws to undermine worker bargaining power, just as much as they are behind the non-use of the antitrust laws against employer power and control in the fissured workplace. As Sanjukta Paul has pointed out, if the church organist professional organization had, instead of publishing guidelines preventing its members from underbidding one another for gigs, programmed an app to match organists to churches seeking their services, and prevented their members from performing at a church matched to another member via the app, the antitrust authorities would have been just as solicitous of the organists’ app as they have been of Uber’s price-fixing and market-division business model—provided the organists’ app had been operated in the interest of a “for-profit hiring hall” like Uber, as opposed to in the interest of the organists themselves.91 That jurisprudence and disposition of enforcement resources effectively means that Uber drivers or organists as workers are paying a significant price to the unitary platforms coordinating the labor market in which those workers sell their services for the privilege of restraining trade and avoiding a free-for-all.92 Should any antitrust case against Uber for price-fixing, exclusion, or market division ever see the light of day after Meyer v. Kalanick, it’s likely that the agencies would take the view that its restraints are vertical and hence subject to the Rule of Reason (as the DOJ has argued in the litigation over no-poaching clauses in franchising contracts93), rather than per se illegal like non-employee driver collective bargaining. As Sandeep Vaheesan has phrased it, antitrust is about “accommodating capital and policing labor.”94

#### Lack of bargaining in the gig economy greenlights unchecked corporate surveillance that coalesces into digital authoritarianism.

Belli ’25 [Luca; June 4; Founder of Sator Labs, a responsible AI consultancy, Tech Policy Fellowship at UC Berkeley, Visiting AI Fellow at the National Institute ofe Standards and Technology (NIST), PhD from Vergata University; Tech Policy, “AI, Gig Workers, and the Erosion of Democracies,” https://www.techpolicy.press/ai-gig-workers-and-the-erosion-of-democracies/]

In this new environment, the tech broligarchy and ideologues like Marc Andreessen can finally assert their authority with the masks off. Their demand: unquestioned, outright rule by venture capital and CEOs. At the same time, they seek to be praised as visionary problem-solvers, even as their efforts contribute to the erosion of democracy, the expansion of surveillance, and the destruction of any semblance of a social safety net. Additionally, their goals for AI seem to involve transforming the entire working class into the precariat by making all jobs gig-based. Meanwhile, there is a vast, precarious, marginalized global underclass of exploited workers who label, annotate, content moderate, and in general enable all of our Western AI wonders.

This remaking of government, the economy, and society is justified under the guise that everything must be made more efficient, an aspiration that Silicon Valley shares with the fascists of the past, from Mussolini’s on-time trains to Hitler’s elimination of all unemployment. The fantasy of AI boosters is that AI will "solve" all economic and social problems. It is the efficiency of and for the algorithm, not for the people. The (un)spoken goal is to have as few workers as possible who are as obedient as possible.

The recent ubiquity of large language models is making it possible for companies' executives to shift the labor market back to their advantage after it was more favorable to engineers. Until two years ago it was not uncommon for a skilled engineer to demand high salaries, including stock options that could potentially be worth millions of dollars.

Now, CEOs can convince themselves that they can run companies with fewer spoiled engineers, because of the “AI productivity boost.” Coupled with a lower union membership rate among tech workers than the nation's average, the resulting uncertainty makes internal organizing in companies less likely, as everyone wants to be quiet to keep their job. Shareholders can rejoice as less worker organizing means fewer problems for companies: no one complaining about selling technology to the military, no one asking questions if POTUS was given too much latitude on Twitter, or demanding more investment in green energy. And if the people who work the closest with this technology don't get a say in it, who will? The tech CEOs are not going to save us: they already have plans to live in luxury bunkers, or maybe just escape to Mars.

The fact that a small group of extremely wealthy individuals is fighting against workers' rights, trying to influence policy, and colluding among themselves is nothing new: Robber barons such as Carnegie, Rockefeller, Vanderbilt, and Morgan have charted this course before. This time seems different—not because of the amount of wealth, influence, and power accumulated, but because their rise has been actively encouraged by policymakers and those in charge of protecting the public’s interests.

Modern tech companies were able to expand to their current size against the backdrop of two trends. On the policy side, neoliberal ideology, favoring free and unregulated markets, minimum to no government regulation, and privatized services, went from a set of niche academic ideas to mainstream discourse. These ideas were translated into policies that either deregulated existing industries or didn’t regulate new ones. That helped the small internet start-ups of the early 2000s avoid oversight and expand unchecked into the super-national entities they are today—companies capable of influencing elections, predicting influenza outbreaks, and facilitating genocide.

Secondly, following 9/11, the global public sector, especially in the US, sought to expand surveillance and data collection, coupled with predictive analytics as a means to prevent future attacks. New laws gave the state broad authority to collect data on citizens. However, the collection of data solely by the public sector was insufficient and less useful without adequate tools for interpretation, so the government increasingly turned to private tech companies. The government was more than happy to learn and apply the techniques of surveillance capitalism to mass surveillance. Online platforms, in addition to selling ads, had the opportunity to help surveil their citizens 24/7 in a low-cost, low-personnel, and efficient way.

#### Digital authoritarianism causes global miscalc, information wars, and cascading tech vulnerabilities.

Manstead ’20 [Katherine; May 28; Nonresident Fellow at the Alliance for Securing Democracy and Senior Adviser for Public Policy at the Australian National University’s Security College; Alliance for Securing Democracy, “Strong Yet Brittle: The Risks of Digital Authoritarianism,” https://securingdemocracy.gmfus.org/wp-content/uploads/2020/05/Strong-Yet-Brittle-The-Risks-of-Digital-Authoritarianism.pdf]

The Vulnerabilities of Digital Authoritarians

While digital authoritarianism can enhance regime durability and national power, it also introduces deep-seated vulnerabilities, eight of which are considered below. Significantly, digital authoritarians may find themselves in a state of constant contest with other regime types, trapped in cycles of overreach and backlash, and prone to strategic miscalculations that pull them into interstate conflict. The current turn to digital authoritarianism therefore also has broader implications for international peace and stability.

Brittle Legitimacy

Reliance on information control makes authoritarians brittle. Small ~~chinks~~ [cuts] in their information control armor could have existential consequences, particularly during political or economic crises (i.e. when the regime needs to rely on control for legitimacy because it is not delivering for citizens). The information and ideas most dangerous to authoritarians include:

* the identity of opposition groups and leaders and their levels of support; 17
* technical means for subverting control of communications and surveillance technologies; 18
* ideas about values that transcend state sovereignty, such as liberalism and human rights; 19
* evidence that the central government is not delivering efficient outcomes; 20 and
* ideas that undermine the myths and narratives used to legitimize authoritarian rule or the power of the ruling elite. 21

Constant Contest

Since technologies and ideas are dynamic, the battle for information control is a constant struggle. It can never be ‘won.’ Authoritarians are therefore in a perpetual state of information warfare, inside and outside their regime, and feel perpetually insecure. This dynamic may lead authoritarian governments to assess that it is worth engaging in information or cyberattacks to discredit liberal ideas at their foreign source or to shape or disable systems that jeopardize their information control—despite real risks of conflict escalation and global pushback.

Overreach and Backlash

The fundamental importance of information control to authoritarians increases the likelihood of overreach, leading to cycles of backlash and reprisal. Many perceive China’s heavy-handed narrative warfare in Hong Kong and confrontational efforts to control narratives about coronavirus to be strategic missteps. For example, CCP efforts to stifle dissent by punishing online gaming company Blizzard and the National Basketball Association (NBA) arguably aided Hong Kong protester narratives;22 while CCP obfuscation about coronavirus has prompted unprecedented diplomatic rebukes from world leaders.23 Despite rising international awareness and condemnation of China’s sharp power tactics,24 China is accelerating, not muting, these behaviors.25 One explanation for this is that the CCP calculates that the risks of international backlash (and occasional overreach by its officials) are acceptable, compared with the risk of letting domestic information control falter.

Impaired Feedback Mechanisms

Authoritarians embrace technology to increase the legibility of their societies. But legibility requires cooperation from society. It is facilitated by an open information ecosystem, robust civil society, mechanisms of transparency, and protections for political speech.26 Conversely, information control and technology-enabled systems of surveillance and enforcement discourage accurate reporting and punish whistleblowing, while incentivizing officials to conceal failures and exaggerate successes.27 In 2007, Le Keqiang (before he became China’s premier) described China’s national income figures as “man-made” and unreliable, and noted that more objectively verifiable proxies should be preferred to official statistics collected by provinces.28 Without elections, authoritarians can also struggle to understand public sentiment, a problem highlighted by the Chinese government’s mismanagement of massive ongoing protests in Hong Kong. Party leaders wrongly assessed that the protestors’ grievances were primarily economic rather than political and that they did not enjoy broader public support.29 As Zeynep Tufekci has observed, the costs of China’s “authoritarian blindness” have been immense: a solvable issue (demands to withdraw a relatively unimportant extradition treaty) became “a bigger, durable crisis” with ongoing political consequences.30

China’s delayed reaction to coronavirus is a stark example of the authoritarian legibility and feedback problem. Local officials and hospital administrators in Wuhan suppressed information about the outbreak and punished doctor whistleblowers—depriving other provinces and the central government (not to mention international authorities) of vital signals that would have allowed swifter action to control the pandemic.31 Once authorities acknowledged the pandemic, China deployed the full weight of its digital surveillance capabilities. It was able to implement top-down lockdowns quickly; marshal its tech sector to build health apps; force citizens to download these apps; and access vast commercial holdings of personal data to cross-check compliance. However, it lacked critical bottom-up feedback systems that may have obviated the need for such draconian measures in the first place.32 Indeed, controlling for income and population size, authoritarian regimes appear to be more lethal than democracies during epidemics, arguably because of their closed information ecosystems.33

Overreliance on Technological Systems which ‘Fail Hard’

Many authoritarian governments are embracing AI-driven surveillance and control methods—from ‘smart cities’ to digital currencies, e-payment platforms and social apps. However, when AI systems fail, they tend to fail in unpredictable, often catastrophic ways. While citizens in democracies lament slow adoption of digital governance, authoritarians’ speed comes with the risk that authorities roll out unsafe or vulnerable systems.34 Imagine a critical failure of China’s social credit system—whether by accident or sabotage—which affected the integrity of records. The implications for regime stability could be significant.

AI systems do not need to fail to produce problematic results. They draw insights and make predictions based on correlations in vast datasets but are not good at identifying causal mechanisms. This means that AI systems often produce outcomes which humans cannot reverse engineer or routinely evaluate. Like using asbestos to build a city, AI governance systems might produce good results in the short-term, but inconsistencies or oversights in their approaches could lead to cascading failures that humans struggle to identify, let alone rectify.35

Unintended Consequences from High-Tech Modernism

Fixation by central governments on achieving targets or deploying certain technologies creates incentives for local officials to deploy “technology placebos” that do little to address underlying economic and social concerns. For example, many so-called smart city projects in authoritarian societies have failed to meet development and economic goals. They are fraught with issues such as “unclear strategic goals” (e.g. they often optimize for surveillance, not development) and “inadequate implementation.”36 This problem may be particularly pronounced for less-developed authoritarian governments which have been persuaded, for strategic reasons, to buy Chinese-exported digital surveillance tools that are not customized to local circumstances. These cities may also become locked into unstable or insecure technical architectures 37 and economic dependence on China.38

Commitments to targets, and ideological fervor about technology, can also distort commercial decisions and raise unrealistic public expectations. Analysis of China’s AI industry, for example, suggests that companies are eschewing investment in basic research and focusing on quick wins in applied research.39 Additionally, China is already behind on meeting a number of its technology targets40—a lag that will likely be exacerbated by the global economic downturn following the coronavirus pandemic, and rising security fears in foreign markets about the security of Chinese technology and IP theft by its companies.

From a strategic perspective, there are risks that authoritarian governments’ fixation on technology-centric strategies will lead them to overestimate what technology can in fact achieve. For example, Chinese military strategists have posited that AI could lift the ‘fog’ of war and eliminate uncertainty and confusion on the battlefield. This is an ahistorical and unlikely prediction that could inspire miscalculation.41 Russian strategists theorize about how psychological operations might subdue adversaries without a shot being fired—an approach that may overestimate what cognitive warfare can achieve, at least without being combined with other elements of national power.42

Challenges to Social Cohesion

The medium- and long-term social consequences of digital authoritarianism are yet untested. Overreliance on surveillance and enforcement systems could attenuate relationships within a society, exacerbating authoritarians’ underlying low trust problems. Since they tend to reduce citizens to data inputs, these systems may deny citizens’ intrinsic desire for dignity and identity—with unexpected results.43 Information control tactics—such as flooding—can repress opposition, but long-term may exacerbate public uncertainty and decrease business confidence and trust in official information, with implications for social cohesion and economic progress.44

Dysfunctional Innovation Ecosystems

Information control and state-led pushes for technology dominance risk hampering innovation. For example, to achieve Xi Jinping’s ‘Made in China 2025’ goals, the CCP is supporting high-tech monopolies, restricting international collaboration, and yoking the state and market together.45 However, monopolies are notoriously inefficient and cross-border collaboration is an important driver of innovation. Further, innovation works best under free market conditions and in open societies.46 Some analysts argue that China’s success in deploying AI applications is an exception to this rule. However, there is a risk that Chinese companies are prioritizing short-term breakthroughs (e.g. analyzing existing datasets to find new insights) at the expense of long-term investment in basic research.47 While authoritarians may excel at developing and deploying AI applications, conceptual research is arguably the real engine of AI advancement—and something that will continue to thrive in open societies.

Summary and Further Research

All states face risks in the information age, but the extent to which regime type affects the relative likelihood of these risks materializing, and their magnitude, is understudied. For example, much has been written about liberal democracies’ vulnerabilities to propaganda and foreign interference via social media.48 But while information warfare against open societies is more likely, arguably it is a higher magnitude threat for authoritarians, where control of information is core to regime survival. Similarly, analysts often lament that democratic governments have been slow to digitize governance systems and craft forward-looking technology policy.49 But while digital authoritarians might outcompete democracies in the roll-out of advanced technologies, this creates new vulnerabilities and risks. Inappropriate safeguards and accidents may result in cascading failures, while heavily digitized governance systems may be susceptible to foreign attack. Regime type may also affect the relative ability of authoritarians and democracies to mitigate their information age risks. For example, a democracy can build resilience to cyber and information threats through a variety of civil society and market-based interventions. Digital authoritarians must rely on a more limited set of top-down policy tools. Ultimately, a more systematic effort to map the comparative strengths and vulnerabilities of authoritarians and democracies in the information age could help both to better understand the other’s threat perceptions and manage escalation risks. It might also highlight ways in which democracies can hold digital authoritarians’ core interests at risk, in order to deter authoritarian interference in their own digital environments.

#### Privacy checks extinction.

Levesque ’16 [Jordan; August 2016; Head of corporate governance and policy at Westpac New Zealand Limited, LLM in Internet Privacy from the Peter A. Allard School of Law at UBC, Bachelors in Commerce and Laws from the University of Auckland; UBC, “The Right to be Forgotten: No Solution to the Challenges of the Digital Environment,” https://dx.doi.org/10.14288/1.0308713]

2.1.4.1 Privacy as a Requirement for Human Survival

The above discussion of the value of privacy can easily be criticised for being too subjective. An individual might argue, for example, that personal autonomy has little value in their life and this may be true. This criticism becomes stronger if we set this premise in the remote Java culture. However, subjectivity alone cannot be seen as sufficient evidence to discredit the value of privacy altogether. In response to such a claim, and without taking away from the aforementioned importance of the individual and social values, privacy can be better valued as an essential part of human flourishing.

To understand the need individuals have for privacy, we must first accept an individual’s need for seclusion and separation. In doing so, the path becomes lit showing the evolution from seclusion to control over oneself with a focus on personal information. Alan Westin suggests that a human’s need for privacy may be found in the animal world.62 He illustrates this by explaining that many animals display their desire for seclusion through the act of establishing their territory.63 Animals require their own territory to ensure the survival of their species, as without a private space to breed and nest, many animals will not survive beyond one generation. If an animal cannot lay claim to its own territory due to overpopulation, their survival is jeopardized as the flow-on effects from limited space impact on breeding, social interaction and sense of smell.64

Worryingly, overpopulation may lead to animals killing each other or “biochemical die-off”.65 A study of a population of 150 rats roaming freely in a cage saw in-fighting increase as the population increased.66 Eventually the fighting became so widespread that rearing of the young rats deteriorated to a level where many did not survive. When the same sized cage was altered to include privacy enhancements a population of 5,000 was able to be supported.67 Assuming that human evolution took place from the animal world, it is not too implausible to see the link with the animal kingdom’s need for privacy to the same need being essential in human society.

#### Only certain protection of gig workers from antitrust empowers worker voice.

Hafiz ’23 [Hiba and Iona Marinescu; Assistant Professor of Law, Boston College Law School; Thurman Arnold Project Fellow, Yale University; Expert Advisor, Federal Trade Commission; Associate Professor, School of Social Policy & Practice, University of Pennsylvania; Research Associate at the National Bureau of Economic Research; Principal Economist, U.S. Department of Justice, Antitrust Division; The University of Chicago Law Review, “Labor Market Regulation and Worker Power,” vol. 90]

The antitrust agencies impact worker voice when they target worker coordination as cartel activity unprotected by the labor exemption to the antitrust laws.45 To the extent the agencies (and reviewing courts) subject worker coordination to criminal sanctions, injunctions, or treble damages liability, they can reduce worker voice and chill worker organizing due to litigation risk.46

Antitrust agencies have charged independent contractors with unlawful collusion, and courts have generally held their coordination ineligible for the labor exemption.47 But while the agencies have expressed interest in expanding the exemption,48 they have yet to establish a policy on enforcement in misclassification cases. Agency and court analyses of the labor exemption’s scope lack clear metrics for determining when independent contractors may coordinate or withhold services free of liability, reinforcing the uncertainty of the exemption’s application.

### Plan---1AC

#### The United States Federal Government should substantially strengthen collective bargaining rights for gig economy workers.

### ADV---Entrepreneurship---1AC

#### Advantage 2 is ENTREPRENEURSHIP.

#### The gig economy stimulates entrepreneurship, but it must balance labor standards.

Cheng ’25 [Zhi; January 7; Assistant Professor of Information Systems and Innovation in the Department of Management at the London School of Economics; Forbes, “It’s Time We Recognised How The Gig Economy Is Stimulating Entrepreneurialism,” https://www.forbes.com/sites/londonschoolofeconomics/2025/01/07/its-time-we-recognised-how-the-gig-economy-is-stimulating-entrepreneurialism/]

Gig economy platforms, which match freelance workers to a local demand for services, such as Uber and TaskRabbit, are often seen as either a “disruptive threat” or an “innovative opportunity”.

Those in the first camp argue that the platforms intensify competition among traditional businesses by offering more cost-effective services and do away with many workers’ rights. Those in the second camp believe the platforms create new job opportunities by connecting clients and workers.

However, what’s much less known about these platforms is how they redistribute local labor.

Gig economy platforms encourage a broader rethink of our economic participation

Research I carried out with colleagues at Georgia State University and the University of Miami found that these platforms are allowing middle-skilled workers, such as managers and supervisors, to move into self-employment and start small-scale ventures, which are not always tied to gig economy platforms.

We discovered, for example, that the emergence of TaskRabbit may well be encouraging many middle-skilled workers to launch their own home repairs and removals businesses. The gig economy, while disrupting traditional employment patterns, is also catalyzing a broader rethinking of career paths and economic participation.

We used employment data from the US Census Bureau, which revealed that when TaskRabbit entered a metropolitan area — a region that includes a central city and its surrounding suburbs — in the United States, the overall wage-based housekeeping workforce decreased by an average of 7.1%, compared to areas without TaskRabbit and before its entry. However, this impact was not uniform across skill types. Middle managers or first-line supervisors in housekeeping experienced the largest reduction, with a 24.7% drop in employment. In contrast, the number of cleaners and janitors remained stable, indicating that frontline service occupations were not significantly affected.

Looking at how individual workers transitioned to new roles, we analyzed data from the US Current Population Survey and found no evidence that TaskRabbit entry forced workers out of housekeeping or into other skill-related jobs, as there were no significant changes in job shifts to other occupations. Interestingly, unemployment in the local housekeeping sector actually decreased by 2.9% after TaskRabbit entered, and 6.9% more workers transitioned to self-employment. Many of these individuals became small-scale business owners, suggesting that TaskRabbit might encourage local entrepreneurial activity among workers.

Lastly, our analysis of data from the County Business Patterns (CBP) database shows an increase in small housekeeping businesses (those with fewer than 50 employees) by at least 3.9% after TaskRabbit’s entry. Larger businesses, however, showed no significant change. This further supports the idea that TaskRabbit’s presence is associated with a shift toward local entrepreneurship and self-employment within the housekeeping industry

Overly cautious policy-making

Why then have so many policy-makers approached the gig economy with such caution?

While some countries, such as the UK and Australia, have embraced the potential of gig economy platforms, many others have imposed stringent regulations and limited the integration of these platforms into local - and national - economies.

Overly cautious policy-makers have focused on risks like worker exploitation, unstable incomes and the erosion of traditional employment protections. All are valid concerns that deserve our attention but ignore the nuances of what is happening in the labor market. By focusing primarily on regulating platforms and addressing immediate labor concerns, these policy-makers risk neglecting the entrepreneurial ripple effects the gig economy has triggered.

An alternative to over-regulation

Rather than over-regulating or forcing gig work into traditional frameworks, policy-makers could focus on supporting this broader wave of small business creation.

This might include helping workers using the platforms to access resources by offering training, funding, or mentorship programs to support former traditional employees transitioning into self-employment. They could also encourage flexibility in benefits through portable benefits systems that cater to both gig workers and self-employed business owners, ensuring security without compromising independence. Policy-makers could also promote innovation-friendly policies by avoiding overly restrictive measures that could stifle platforms’ ability to drive economic transitions and entrepreneurial opportunities.

Overly restrictive measures risk stifling the gig economy’s broader transformative potential, while overly lax policies fail to protect vulnerable workers.

As our research shows, the gig economy does more than provide jobs on platforms — it inspires workers to reimagine their economic roles, often transitioning into self-employment and small-scale entrepreneurship. Policy-makers should focus not just on worker protections within gig platforms, but also on fostering an environment that supports entrepreneurial growth and innovation. A balanced approach can harness the gig economy’s dual potential to protect workers and catalyze economic evolution.

#### Two internal links:

#### 1. INCOME-SMOOTHING---The gig economy enables risk-taking by backstopping entrepreneurs’ income. Bargaining power secures income security necessary to facilitate entrepreneurship.

Wolla ’24 [Scott A; April 1; economic education officer at the St. Louis Fed, PhD from Saint Louis University; Federal Reserve Bank of St. Louis, “How Does the Gig Economy Support Entrepreneurship?” https://www.stlouisfed.org/publications/page-one-economics/2024/04/01/how-does-the-gig-economy-support-entrepreneurship]

Entrepreneurs develop new products and start new businesses. They recognize opportunities and the prospect of financial rewards. They enjoy working for themselves and will accept challenges. There are certainly risks for entrepreneurial activity: Entrepreneurs often risk their own money or borrow against their future earnings; they forgo other opportunities (say, taking a job as an engineer, nurse, or accountant) to pursue their dreams.

Of course, chasing your dreams doesn't always pay the bills. So how do they make ends meet?

Enter the Gig Economy

For people who want to work for themselves, the gig economy provides some exciting opportunities. It is made up of companies whose business model usually involves a smartphone app that serves a key role in two-sided markets, matching people who want and will pay for a service with those who will provide a service for a fee. In fact, an app plays a key role in how these businesses function; it can make launching a new business a little easier because it takes care of many of the details that a small business would often have to do on their own, such as attracting clients, collecting fees, and compensating producers for the jobs they perform.

App-based ride-hailing businesses such as Uber and Lyft are some of the most widely known in the gig economy. Ride-hailing apps offer low barriers to entry because drivers use their own or rented cars to offer rides whenever they choose. These businesses offer workers true flexibility; there are no minimum hour requirements, so drivers can work however many hours they want to.

Making Ends Meet with Gigs

Entrepreneurs can both work in the gig economy and establish a new business at the same time. This enables them to smooth their income while they develop their product and work to get their business up and running; that is, they can supplement their income when their new business is slow. In fact, the vast majority of Uber drivers work only 8, 12, or maybe 15 hours a week driving,2 using these gigs to supplement their income from their startup or another job. So how do entrepreneurs balance working on their business venture and working their gig job?

Balancing Gig Work and Entrepreneurship

Let's again use Uber as an example. Research suggests that an Uber driver's willingness to work is related to the driver's reservation wage,3 which is the earnings level below which the person will not work. For example, if a worker's reservation wage is $25 per hour, then the worker will not accept work for less than that: Wages above $25 per hour result in surplus, so the worker will gladly take the work.

Consider an entrepreneur who just started a small business but also wants to supplement income while the product catches on, so she drives Uber when her schedule allows. Her reservation wage might be very high during hours in which her new business is open—say, $75 per hour. But her reservation wage might be lower in the evenings when she would otherwise be socializing or relaxing—say, $10 per hour. In this case, an entrepreneur would be less likely to drive Uber during regular business hours (9 a.m.-5 p.m.) and more likely to drive in the evenings and on weekends.

Data suggest the same pattern; that is, Uber drivers provide more rides outside of conventional business hours. The figure below shows the working habits of employed males over 20 years of age surveyed in the American Time Use Survey (darker areas) and the working habits of Uber drivers (lighter areas). As suggested, Uber drivers put in more hours in the early morning, in the evenings, and on weekends.

Economist Judith Chevalier tells the story of being picked up in a very nice, very large car on a recent business trip. In addition to driving for Uber, the driver also ran a funeral service. When Chevalier asked her about her reasons for driving, the driver said, "If nobody dies, I drive. If somebody dies, I don't."4 Her answer shows that she knew her reservation rate and that she valued the extra work and wages to help smooth her income over time.

Conclusion

The gig economy supports entrepreneurship by providing entrepreneurs with opportunities to earn extra income. In fact, researchers suggest that the gig economy supports and encourages more entrepreneurial activity, finding an increase of 4-6% in new business registrations following the arrival of the gig economy in a city.5 In short, the gig economy provides services that people value, and it also has a spillover: It encourages entrepreneurial activity by supplementing and smoothing the income of entrepreneurs.

#### 2. EXPERIMENTATION---Gig economy participation enables skill-building and human capital acquisition---participation is directly correlated with entrepreneurship.

Lagaras ’25 [Article Summarizing study by Spyros Lagaras, Matthew Denes, and Margarita Tsoutsoura; April 28; Assistant Professor of Finance at the Gies College of Business at the University of Illinois at Urbana-Champaign, PhD in Finance from the Gies College of Business, MS and MBA in Finance; Assistance Professor of Finance at Carnegie Mellon University Tepper School of Business, PhD in Finance from the University of Washington, and an MA in Finance from the University of Pennsylvania; Associate Professor of Finance at the Olin Business School, Washington University in St. Louis, CEPR Research Fellow, PhD in Finance and Economics from Columbia University; Gies College of Business, “New study shows gig economy as viable pathway to entrepreneurship,” https://giesbusiness.illinois.edu/news/2025/04/28/new-study-shows-gig-economy-as-viable-pathway-to-entrepreneurship]

It’s not new, but it is a phenomenon that picked up steam in the early 2010s and has carried into the mid-2020s – Americans choosing to enter the gig economy, a free-market system in which individuals are hired for short-term “gigs.” Examples include drivers for a ride-sharing company, short-term rentals, freelancers, and other one-off jobs. More recently, researchers were curious to see whether those involved in the gig economy were using their participation to become entrepreneurs.

Spyros Lagaras, an assistant professor of finance in Gies College of Business at the University of Illinois Urbana-Champaign, joined researchers Matthew Denes from Carnegie Mellon University and Margarita Tsoutsoura from Washington University in St. Louis to take a closer look at whether the gig economy was, in fact, a common pathway to entrepreneurship. Their findings, which confirmed their suspicion, were published in the National Bureau of Economic Research’s (NBER) Working Paper Series and are scheduled to be published in the Journal of Financial Economics.

“I have a strong interest in how finance correlates with labor markets and how technological advances or innovation in the labor market affect entrepreneurship in general,” Lagaras said on his interest in this subject. “Because the gig economy was a huge innovation in the labor market, we set out to understand if somehow participation in the gig economy could act as a pathway to entrepreneurship. Beyond that, we tried to see who benefits and what the characteristics of the people are who respond to this specific innovation.”

More than 10 million individuals received income from the gig economy from 2012-2021, with the gig income accounting for $120 billion in one year during that period.

According to this paper, the researchers speculated that “the gig economy might mirror the experiences of an entrepreneur, allowing individuals to learn about entrepreneurship and accumulate industry-specific experience.”

A summary of their findings includes:

Individuals who participated in the gig economy were significantly more likely to start new firms.

First-time entrepreneurs make up 75 percent of those starting new firms.

They tend to be relatively younger and have dependents, a demographic that is seeking more flexibility.

They typically start firms in a similar industry to those in which they were participating in the gig economy.

While gig-founded firms were less likely to survive, those that did grew at a higher rate and had better performance than those entrepreneurs who had not participated in the gig economy, suggesting that the gig economy allows individuals to experiment with entrepreneurship.

Those who participated in the gig economy were more than twice as likely to start a business as those who hadn’t.

Of those entrepreneurs who transitioned from the gig economy, there is an 8-16 percent higher likelihood of those businesses employing at least five people.

“The number one barrier in entrepreneurship is always access to capital,” Lagaras said. “You can see the gig economy as a way to secure the capital and the tools you need to start the company.”

Lagaras made a few observations that are likely to explain the trend, although he is careful to point out that these aren’t specifically outlined in the paper. First, he says that the gig economy allows people to experiment and learn a new relevant skill. Secondly, he notes that even if those individuals’ firms fail, there is always the fallback of going back to the gig activities, making becoming an entrepreneurial venture much less risky. Next, he says younger people are more likely to make this move because they tend to be early adopters, and even if they don’t succeed, they have more of their career remaining. Finally, he notes that younger people with families appreciate flexibility in their work schedule.

“The gig economy is a substantial part of the economy in the US now and has considerably grown in the last 15 or 20 years,” Lagaras said. “While we are careful not to assign a cause-and-effect relationship with what we discovered, we can say there is a positive relationship between the gig economy and entrepreneurial entry.”

Confidential data from the IRS was at the heart of the research project. The researchers wanted to see if individuals involved in the gig economy were creating sole proprietorships with a separate business identification number. They had to weed out those who filed business returns for the sole purpose of reporting gig income. If the gig income reported was the same as the income of the sole proprietorship, they were not considered as creating a business and were dropped from the study.

“We focused only on sole proprietorships for this paper for two reasons,” Lagaras said. “First, unlike a corporation, sole proprietorship involves one owner, so we know the owner spends time and money as the actual entrepreneur. We also expect gig workers to be more likely to form a sole proprietorship when starting a new firm relative to other firm types like corporations.”

Lagaras said policymakers have shown interest in these findings.

“Some policymakers express the idea that using the gig economy as a stepping stone for entrepreneurship can be an engine of growth, used as a place-based policy to support entrepreneurship or as a less risky way to experiment with entrepreneurship,” Lagaras said.

#### Accelerating innovation bolsters societal resilience---that caps existential black swans.

Aschenbrenner ’24 [Leopold and Phillip Trammell; June 9; B.A. in Mathematics, Statistics, and Economics from Columbia University, Research Affiliate at the Global Priorities Institute; Postdoctoral Researcher at Stanford Digital Economy Lab, B.A. in Economics and Mathematics from Brown University; Phillip Trammell, “Existential Risk and Growth,” https://philiptrammell.com/static/Existential\_Risk\_and\_Growth.pdf]

Technology brings prosperity. Its impact on existential risk—the risk of human extinction, or, equivalently for decision purposes, of an equally complete and permanent loss of human welfare—strikes many as ambiguous at best.1 Advances in vaccine technology render us less vulnerable to devastating plagues; advances in gain-of-function virology arguably make them more likely (Millett and Snyder-Beattie, 2017)

This raises the possibility of a tradeoff: concern for the survival of civilization may motivate slowing development, at least outside some narrow domains. Environmentalist sentiments along these lines go back at least to the Club of Rome’s 1972 report on the “Limits to Growth”, and have recently reemerged with calls to pause AI development (Future of Life Institute, 2023). Jones (2024) explores how to make the tradeoff between AI development and AI risk, assuming the tradeoff exists.

To shed light on this question, we begin in Section 2 with a simple model in which the hazard rate—the probability of catastrophe per period—is a positive function of the technology level. Here, an existential catastrophe must occur eventually unless in the long run higher technology levels carry hazard rates that fall toward zero.

This leaves two possibilities. If advanced technology does not eventually drive the hazard rate toward zero, then a catastrophe is inevitable, so accelerating technological development cannot increase its probability. Otherwise, however, a catastrophe is avoidable, and acceleration can lower its probability by hastening the arrival of safety.

This simple model formalizes two observations. The first is that if we believe that the hazard rate is currently high, our only hope for a long and valuable future is the hope that we are living through a temporary “time of perils”. This view was famously expressed by Sagan (1997), who coined the phrase, and its implications for those especially concerned about the long-term future are emphasized by Parfit (1984), Ord (2020), and others. The second is less widely appreciated: that if we are in a time of perils, with the hazard rate a positive function of the technology level, then, though technological development has increased the hazard rate so far, deceleration for the sake of long-term survival is misguided. Speeding technological development may be temporarily risky, but it is safer in the long run.

The model of Section 2 is not “economic”. It studies the impact on risk of quickly escaping risky states, not optimal policy under constraints. It thus leaves open the possibility that, when consumption–risk tradeoffs are navigated by a policymaker with little concern for long-term survival, technological acceleration can increase risk after all. Section 2 also offers no reason to believe that future states will be safe. If one believes that technology has historically increased the hazard rate, the hope that this relationship will reverse in the future may seem naive.

In Section 3, we therefore introduce an environment in which technology grows exogenously and its risks can be mitigated by policy. As new potentially dangerous technologies are introduced, a planner, discounting the future at an arbitrary rate, decides how much consumption to sacrifice to lower the hazard rate.

We illustrate that, even if technological advances in isolation always raise the hazard rate, optimal policy can generate an “existential risk Kuznets curve”, with the hazard rate rising and then falling as technology advances. Early, when the expected discounted value of the future of civilization is low and the marginal utility of consumption is high, it is worthwhile to adopt risky technologies as they arrive. Later, when the discounted future is more valuable and the marginal utility of consumption has fallen, substantial risk mitigation becomes worthwhile.

The possibility of a policy response thus offers an economic justification for the view that we may indeed be living through a once-in-history time of perils. Safety is a luxury good, and technological development generates a wealth effect. If the wealth effect is strong enough, then optimal policy eventually lowers the hazard quickly enough that the probability of escaping the time of perils is positive.

This insight mirrors the logic of Stokey (1998) and Brock and Taylor (2005), on which environmental damages rise and then fall with economic development, and of Jones (2016, 2024), on which growth yields increases in the value of life relative to marginal consumption. Like the analysis presented here, these papers find that, given sufficiently concave utility functions, wealth increases motivate large reallocations from consumption to safety. None of these sources solve for the optimal path of a hazard rate over time, however, or characterize conditions under which the probability of a binary event (here, existential catastrophe) under optimal policy is less than 1.

Our model of catastrophic risk differs more significantly from those of Martin and Pindyck (2015, 2021) and Aurland-Bredesen (2019). That literature studies a society’s willingness to pay to reduce the risk of catastrophes that are, or are equivalent to, proportional consumption cuts. In such a context there are no wealth effects: the fraction of consumption one is willing to sacrifice to avoid a proportional consumption cut is, by definition, independent of one’s consumption.

For the reasons that policy facilitates survival on a given (increasing) technology path, optimal policy tends to magnify the extent to which technological acceleration decreases long-term risk. As in the policy-free model, acceleration decreases the time spent in any given risky state. Under optimal policy, however, the wealthier future states pulled forward by an acceleration are systematically inclined to be safer, due to the wealth effect. Furthermore, given an increase to the future growth rate, even before actual productive capacity has yet increased, the anticipation of a more valuable future motivates more stringent safety policy in the present.

This analysis might be compared with that of Baranzini and Bourguinion (1995). Baranzini and Bourguinion find conditions under which the growth path that maximizes expected discounted utility also minimizes the probability of existential catastrophe. In our model these objectives never perfectly align, but we explore how technological advances, when regulated with a view to maximizing expected discounted utility, can lower the probability of an existential catastrophe.

Sections 2 and 3 explore models in which the state of technology at a given time contributes to the hazard rate. Section 4 considers the possibility that risk is “transitional”, increasing in the rate of technological development.

Absent policy, the effect of acceleration on long-term transition risk is ambiguous. In particular, acceleration has no effect on long-term risk under the assumption that the “experiment” associated with developing a given technology poses a risk that is independent of how many experiments happen concurrently. This is the assumption of e.g. Jones’s (2016) “Russian roulette” model of risky technological development. If the future contains a sequence of experiments, each of which will induce some probability of existential catastrophe, then stagnation can lower risk by avoiding advanced experiments altogether, but a technological acceleration that only pulls forward their date leaves the probability of catastrophe unchanged.

As in Section 3, introducing an optimal policy response facilitates survival due to wealth effects, potentially replacing an ever-increasing hazard rate with a Kuznets curve. Also, though the effect of acceleration on long-term transition risk remains ambiguous given policy, policy can shift the conditions under which acceleration has a given effect on risk. At least in the particular model of transition risk studied in Section 4, the existence of a policy response significantly widens the conditions under which acceleration lowers transition risk.

<<TEXT CONDENSED, NONE OMITTED>>

Section 5 summarizes these analyses and their limitations. 5 2 State risk without mitigation 2.1 Model The hazard rate — The “hazard rate” δt is the flow probability at t of anthropogenic existential catastrophe. Assume that it is a continuous function of the technology level At: δt = δ(At). Assume that At is exogenous and strictly increasing without bound in t. Fur ther assumptions on the technology path can be made without loss of generality, as they simply amount to re-indexing technology levels without changing their ordering. Assume therefore without loss of generality that A(·) is differentiable and that its derivative is everywhere positive. Finally, assume that δ(A) > 0 for all A. Survival — The probability that civilization survives to date t is given by St ≡ e− t 0δsds ⇐⇒ ˙ St = −δtSt, S0 = 1. The probability that human civilization avoids an [anthropogenic] existential catas trophe and, at least in expectation, enjoys a long and flourishing future3 is S∞ ≡ lim t→∞ St = e− ∞ 0 δsds. (1) We will refer to {δt}∞ t=0 as the hazard curve, to the area under the hazard curve ( ∞ 0 δtdt) as cumulative risk, and to S∞ as the probability of survival. Note that 3In the face of natural existential risk, this will entail succumbing to a natural existential catas trophe instead. From very-long-run historical data on large-scale natural catastrophes, and the typical survival rate of other mammalian species, Snyder-Beattie et al. (2019) estimate that human ity’s natural existential hazard rate is below one in 870,000 per year. Throughout this paper we ignore the possibility that technological advances may mitigate natural existential risks. Accounting for this possibility would only strengthen the headline results. 6 the probability of survival decreases in cumulative risk, and that survival is possible (S∞ > 0) iff cumulative risk is finite. 2.2 How does acceleration affect risk? Absent a negative shock severe enough to induce stagnation or recession, technology crosses every value from A0 to ∞ exactly once. So the area under the hazard curve can be defined by integrating with respect to technology instead of time: ∞ 0 ∞ δ(At)dt = A0 −1 δ(A) dA dt dA = ∞ A0 δ(A) ˙ A−1 A dA, (2) where, somewhat abusing notation, ˙ AA denotes the value of ˙ A when the technology level equals the subscripted A. This change of variables makes it easier to see how various shocks to the growth path affect cumulative risk. Instantaneous level effects — Consider a shock to the technology level for a short period beginning at t, so that the technology level over this period is approximately ˜ Arather than At (and the subsequent technology path is unchanged). The sign of the impact of this shock on cumulative risk depends on whether δ( ˜ A) is greater or less than δ(At). From the leftmost integral of (2), we see that the impact on cumulative A, equals risk per unit time of an instantaneous shock to the technology level at t, from At to ˜ δ( ˜ A) −δ(At). Instantaneous accelerations — Consider the impact on cumulative risk per unit time of an instantaneous shock to technology growth at t, so that the technology growth rate at t is ˙ ˜ A rather than ˙ At, and the subsequent technology growth rate at each level of technology is unchanged. From the rightmost integral of (2), we see that the impact of this shock on cumulative risk per unit of increase to the technology level during the acceleration is δ(At)( ˙ ˜ A−1 − ˙ A−1 t ). Multiplying this by the new rate of 7 technology growth per unit time, the impact on cumulative risk per unit time is −δ(At) ˙ ˜ A/ ˙ At − 1 . Accelerations — Choose technology level A > A. Since the baseline technology path increases continuously and without bound, A = AT for some T > t. Consider the effect of increasing the technology level at t from A to A, and sub sequently maintaining the technology path As = As+(T−t) (s ≥ t). This shock to the technology path amounts to a “leap forward in time”. The impact of this shock on cumulative risk is therefore to cut a slice cut out of the hazard curve. Cumulative risk falls from (2) to At A0 That is, it falls by ∞ δ(A) ˙ A−1 A dA+ A AT δ(A) ˙ A−1 δ(A) ˙ A−1 A dA. A dA. At More generally, define a temporary acceleration as an increase to ˙ A at some range of technology levels: say, from At to AT. Because the exponent on ˙ A in the integral is negative, the acceleration lowers the risk endured at the given range of technology levels. A discontinuous jump in the technology level amounts to raising ˙ AA to ∞, and thus lowering ˙ A−1 A to 0, from A = At to AT. A jump in the technology level from At to AT temporarily increases the hazard rate if δ(AT) > δ(At). Likewise, an acceleration to technology growth accelerates an increase to the hazard rate if δ(·) is increasing around At. It may therefore appear to contemporaries that a given permanent level effect decreases the probability of survival. Here, that would be incorrect. If (2) is infinite, the probability of survival is zero with or without the permanent level effect.4 If (2) is finite, the permanent level effect increases decreases cumulative risk and increases the probability of survival. 4 AT At δ(At)dt is finite by the continuity of δ in A and of A in t. 8 Define a permanent acceleration to be a permanent increase to ˙ A from some time t— or, equivalently, some technology level At—onward. As is plain from (2), a permanent acceleration, like a temporary acceleration, must lower cumulative risk if cumulative risk is finite on the baseline technology path. Unlike temporary accelerations, however, permanent accelerations can render sur vival possible when it would otherwise be impossible. Shrinking a heavy-tailed curve with an infinite integral can yield a thin-tailed curve with a finite integral. To state this lesson in reverse, consider stagnation: a permanent “negative accel eration” setting As = At for all s ≥ t. The hazard rate is then permanently positive, and survival impossible, even if it might have been possible at any positive technology growth rate. More concretely, consider the implications of a large negative technology shock today which returned the world to a state of ignorance about every technol ogy developed since 1924. Perhaps the hazard rate was much lower in 1924 than today, but even if so, this reset would largely doom us to relive the nuclear standoffs, emissions-intensive industrializations, and biotechnological hazards of the past. With enough replays of the past century, a catastrophe would presumably be inevitable. 3 State risk with mitigation 3.1 Motivation If technological progress has historically increased the hazard rate, the message of the previous section is that those who wish to reduce existential risk should accelerate technological progress in the hope that the relationship between risk and technology eventually reverses. This may seem naive. Perhaps the more natural assumption is that, all else equal, technological progress will only increase the hazard rate, bringing the inevitable catastrophe sooner. But the hazard rate presumably depends also on policy. If the hazard rate has increased historically, this represents a failure of policy to keep up with new risks as 9 they have arisen. In light of the interaction between technology and policy, could existential risk be lowered by developing technology more slowly? If the path of policy is not optimal, yes. E.g. if policy is exogenous, cumulative risk is lower when periods of especially risky technology are timed to coincide with periods of especially stringent policy. For illustration, suppose that δt = Atxt, xt = (1+t)−2, where x denotes a policy variable. Then consider an acceleration from the technology path At = (1 + t)k to the technology path At = (1 + t)˜ k, where k < 1 < ˜ k. This acceleration increases cumulative risk from ∞ 0 (1 +t)k−2dt to ∞ 0 (1 +t)˜ k−2dt. The former is finite, because k − 2 < −1. The latter is infinite, because ˜ k − 2 > −1. In this case, acceleration lowers the probability of survival to zero. Less obvious is whether acceleration can increase cumulative risk when the policy re sponse is optimal, within a plausible model of the feasible policy set. One might worry that, during an interval in which more advanced technology carries higher hazard, a planner will adapt policy to the degree of risk, but too weakly for acceleration to lower cumulative risk on balance—perhaps in part because she cares too little about the future to sacrifice much present consumption for safety. To evaluate this possibility, this section introduces a policy channel through which a planner, discounting the future at an arbitrary rate, can sacrifice consumption to lower the hazard rate. As we will see, when policy is set optimally—with respect to any discount rate—the conclusion that acceleration lowers cumulative risk is generally not only maintained but strengthened. As in the tech-only model of Section 2, survival can only be achieved by pulling forward a future that asymptotically approaches perfect safety. Whereas the earlier 10 model is agnostic about whether more advanced technology will in fact carry a lower hazard rate, however, a policy response introduces a tendency for faster technological development to carry lower risk in the long run. This is because technology increases consumption, which both decreases the utility cost of a marginal consumption sacrifice and increases the value of life. Furthermore, the prospect of a future acceleration now lowers the present hazard rate, because when the value of the future is greater, it is worth sacrificing more today to prevent its ruin. These dynamics are illustrated in a simple model of technology and optimal policy in the rest of this section. Generalized results are given in Appendix A.3. 3.2 The economic environment 3.2.1 Technology Themaximumfeasible level of consumption at t equals the technology level At. Actual consumption is At multiplied by a policy choice xt ∈ [0,1]: Ct = Atxt. (3) The tradeoff at the heart of this section is that a technologically advanced civilization can risk self-destruction, but that this risk can be lowered at some cost to consump tion, as represented here by a choice of x below 1. (We denote the choice variable x to remind the reader that higher choices of x come with higher existential risk.) Choices of x below 1 may constitute bans on the adoption of consumption-increasing but risky production processes and/or allocations of resources to the production of safety-increasing goods and services like pandemic monitoring. The technology frontier A grows at a constant rate g: ˙ At = Atg, g >0,A0 >1. (4) 11 3.2.2 Hazard rate The hazard rate δt is now a function of the technology level At and the policy choice xt ∈ [0,1], and is increasing in xt. In this simple model, the elasticities of the hazard rate in A and in x are constant, so that δ(At,xt) = ¯ δAα txβ t, ¯ δ > 0, β > α > 0, β > 1. We impose the three inequalities of β > α > 0, β > 1 to satisfy three desiderata.5 (5) The first is that, fixing xt > 0, δt increase in At. This imposes α > 0. The assumption that δt increase in At is necessary if we are to concede that technological development has rendered existential catastrophe more likely now than it was long ago, and that this trend would continue absent a change in policy. The proportion 1 −x of potential consumption sacrificed for the sake of existential safety has only increased alongside technological development: having once been zero, it is a small but positive share today.6 If it had remained fixed, the hazard rate would presumably have followed a weakly higher path. Second, the elasticity of δt with respect to xt is assumed to exceed the elasticity of δt with respect to At; i.e., β > α. This is equivalent to the condition that, when technology advances, it is always feasible to lower the risk level by retaining the former consumption level, allocating all marginal productive capacity to existential 5Hazard function (5) is closely analogous to the environmental damage function of Stokey (1998). While Stokey focuses on the implications of the damage function for the chosen path of x (or “z” in her notation), we will study how accelerations to the path of A affect the probability of a binary event: the occurrence of an anthropogenic existential catastrophe at any time. 6Ord (2020, p. 313) estimates that, as of 2020, approximately $100M/year was spent specifically on reducing existential risk. This is likely a great underestimate of existential safety expenditures in the sense relevant here, for two reasons. First, explicit expenditures do not include foregone con sumption due to regulatory barriers. Second, many catastrophic risk reduction efforts are motivated both by the desire to reduce existential risks and by the desire to reduce smaller-scale damages. By contrast, Moynihan (2020) argues that the very concept of an anthropogenic existential catastrophe essentially did not exist 300 years ago; it appears there were then no efforts taken to prevent one. 12 safety measures. This may be seen by substituting xt = Ct/At (from (3)) into the hazard function (5), yielding δt = ¯ δAα−β t Cβ t . Fixing C, the hazard rate falls over time iff β > α. If it is (indefinitely) infeasible to lower the hazard rate while fixing consumption, as it is in this model if β ≤ α, then an existential catastrophe is unavoidable unless consumption falls to zero. This degrowth would amount to the destruction of advanced civilization by other means. If β ≤ α, therefore, speeding or slowing growth can have no impact on the probability of an existential catastrophe broadly construed. Third, fixing At > 0, δt is assumed to be strictly convex in xt. This imposes β > 1. The convexity implies diminishing returns to existential risk mitigation efforts. We take this to be a reasonable assumption both from first principles and from Shulman and Thornley’s (2024) recent estimates of the cost-effectiveness of existential risk mitigation efforts (Appendix A.1). The relationship between a hazard curve and the corresponding probability of survival S∞ is described in Section 2.1. 3.2.3 Preferences A planner seeks to maximize ∞ 0 e−ρtSt u(Ct)dt; u(Ct) = C1−γ t −1 1 −γ , γ>1. (6) That is, flow utility u(·) is CRRA in consumption for some coefficient of relative risk aversion γ > 1. Flow utility is discounted at exponential rate ρ > 0, representing the sum of some rate of pure time preference, if any, and some rate of natural and unavoidable existential risk.7 7One valid interpretation of these preferences would be that the population is fixed and (6) is the expected utility of a representative household. Another would be that population grows 13 The utility of death is implicitly normalized to 0 and the death-equivalent con sumption level to 1. Equivalently, we are normalizing to 1 the technology level at which, when consumption is maximized, flow utility equals 0. The planner chooses the path of x to maximize (6) subject to (3)–(5). Like Martin and Pindyck (2015, 2021), we assume that γ > 1 throughout the rest of the paper, except in Section 3.3.4. We assume this in part because it appears to be true, as documented by Hall (1988), Lucas (1994), Chetty (2006), and others. Also, however, the results are otherwise relatively uninteresting. This is for two reasons. First, observe that when γ > 1, flow utility is upper-bounded by 1 γ−1 > 0. Acceler ating consumption growth, from a baseline of positive consumption growth, therefore yields a stream of utility benefits that eventually shrinks over time. This dynamic produces the key tradeoff: concern for the future may cast doubt on the value of speeding technological development, because the consumption benefits of doing so primarily accrue in the short run, whereas the costs of an existential catastrophe are everlasting. By contrast, when γ ≤ 1, flow utility can grow without bound, so acceler ations to consumption growth and reductions in existential risk can have comparable long-term benefits. Second and relatedly, when γ ≤ 1, the marginal utility of consumption does not decline quickly enough (relative to the rising value of civilization) to motivate rapid increases in consumption sacrifices for the sake of safety. As a result, the probability of long-term survival is always zero on the planner’s chosen path, and accelerations or decelerations to technological development have no impact on the probability. This is detailed in Section 3.3.4. exponentially at rate n < ρ, that the rate of pure time preference and exogenous risk is in fact ρ+n, and that the planner uses the total utilitarian social welfare function. 14 3.3 The existential risk Kuznets curve 3.3.1 Optimality Summarizing the environment of Section 3.2, the planner chooses {xt}∞ ∞ e−ρtSt u(Ct)dt, 0 u(Ct) ≡ C1−γ t −1 1 −γ , γ>1 subject to A0 >1, ˙ At = gAt (g > 0), Ct = Atxt, S0 = 1, ˙ St = −δtSt, δt = ¯ δAα txβ t (¯ δ > 0, β > α > 0, β > 1). t=0 to maximize (7) (8) (9) This section finds the path of the hazard rate in the planner’s solution, observing that it rises and then falls with time. In the next section we will explore what this implies for the impact of acceleration on cumulative risk. The planner faces one choice variable, xt, and one state variable, St. Her (expected) f low payoff at t is Stu(Ct). Her problem can be represented by the following current value Lagrangian: Lt = Stu(Ct) +vt ˙ St + µt(1 −xt) =St (Atxt)1−γ − 1 1 −γ −vt ¯ δAα txβ tSt +µt(1−xt). µt is the Lagrange multiplier on x, positive iff the xt ≤ 1 constraint binds. ∞ vt = e−ρ(s−t)Ss St u(Cs)ds t is the costate variable on survival: the expected value of civilization as of t.8 (10) (11) 8The fact that the costate variable on survival must equal (11) can be seen immediately by reflecting on the fact that, in effect, the value of saving the world must equal the value of the world. It is also derived formally in Appendix B.1. 15 On an optimal path, the first-order condition on (10) with respect to the choice variable xt is satisfied. Differentiating (10) with respect to xt, we have StA1−γ t x−γ t −¯ δAα tβxβ−1 t vtSt ≥ 0, (12) with inequality iff the left-hand side is positive at xt = 1, in which case xt = 1 is optimal.9 Thus, • As long as (12) is nonnegative at xt = 1, the optimal xt ∈ [0,1] equals 1. Any consumption sacrifices would carry greater flow costs than expected benefits. • When (12) is negative at xt = 1, the optimal choice of xt is interior. It sets (12) equal to zero, maintaining the condition that the marginal loss of flow utility from lowering consumption equals the expected benefit via risk reduction.10 In fact there is a unique11 optimal path, characterized by first-order condition (12), a first-order condition corresponding to the state variable St, and identity (11). This is shown in Appendix B.1. For now, our discussion will rely only on the observations that (12) is satisfied on any optimal path, and that vt is upper-bounded by ¯ v ≡ 1 ρ(γ −1) . 3.3.2 Initial risk increases (13) The condition that (12) is nonnegative at xt = 1 is equivalent to the condition that A−(α+γ−1) t ≥ ¯ δβvt. (14) The continuation value of civilization at t given survival to t, vt, always strictly rises over time. This follows from the fact that, given the optimal paths {Cs}s≥t and 9The second derivative with respect to xt is negative by the assumption that β > 1. 10We can ignore the possibility that optimal xt equals 0 because this yields infinite flow disutility. 11Given piecewise continuity. If path x is optimal, measure-zero deviations from x are of course also optimal. 16 {δs}s≥t achievable at a given initial technology level At, a higher initial technology level allows for a path with an equal hazard rate but more consumption at each future period, by the assumption that β > α. A higher initial technology level always enables the planner to implement a preferred future. Suppose inequality (14) is satisfied strictly at t = 0. Then early in time, when At is low, the optimal policy choice is x = 1, and the hazard rate rises at rate gδt = αg. 3.3.3 Eventual risk declines and survival As the left-hand side of (14) falls exponentially with At and the right-hand side rises, there is a unique time t∗ at which (14) holds with equality. After t∗, the optimal choice of xt is interior and sets (12) equal to zero. Setting (12) equal to zero and rearranging, we have the optimal choice of xt after t∗, and thus the optimal choice of xt in general:    xt =   1, ¯ δβAα+γ−1 β+γ−1 t vt − 1 t ≤t∗; , t >t∗. (15) Taking the growth rate of each side, we can find the growth rate of the policy choice variable after t∗: gxt = −α+γ−1 β +γ−1g− 1 β +γ−1gvt, (16) where, given a time-dependent variable y, gyt ≡ ˙ yt/yt denotes its proportional growth rate at t. The hazard rate in turn grows as gδt = αg +βgxt = −(β −α)(γ −1) β +γ−1 g− β β +γ−1gvt. Because β > α and γ > 1, (17) is negative. (17) Furthermore, though gvt is always positive, gvt → 0. This roughly follows from 17 the fact that the expected value of the future vt is bounded above by ¯ v.12 This gives us the asymptotic long-run negative growth rates gx and gδ. Finally, since Ct = Atxt, we have gCt = g +gxt = β−α β +γ−1g− 1 β +γ−1gvt. Because β > α, long-run consumption growth is positive: x declines to 0, but A grows more quickly than x declines. Indeed, the growth of consumption is key to the growth in sacrifices for safety. With decreasing marginal utility to consumption and decreasing marginal returns to sacrifices for safety, potential consumption is split between the former and latter so that the marginal value of each stays equal. To summarize: Proposition 1. The existential risk Kuznets curve On the path defined by (7)–(9), there is a time t∗ ≥ 0 such that for t < t∗, xt = 1, gCt =g >0, gδt =αg >0 and for t ≥ t∗, lim t→∞ gxt = −α+γ −1 β +γ−1g 0, lim t→∞ gδt = −(β −α)(γ −1) β +γ−1 g 0. (18) (19) (20) The corollary follows from (20) and the definition of S∞. Because δt ultimately falls exponentially, ∞ 0 δtdt < ∞, so S∞ ≡e− ∞ 0 δtdt > 0. Note that δt → 0 is insufficient for survival. If δt fell to 0 too slowly, the integral would diverge, and we would have S∞ = 0. 12The gvt → 0 limit is shown formally in Appendix B.2. 18 3.3.4 No survival with γ ≤ 1 As noted in Section 3.2.3, one reason for assuming γ > 1 is that, when the marginal utility of consumption declines too slowly, a rapid shift from consumption to safety effort is not implemented, and the probability of long-term survival is always zero. This result recalls the “Russian roulette” model of Jones (2016). There, it is found that a planner will choose to sacrifice enough consumption for safety that a technologically induced catastrophe is not inevitable iff γ ≥ 1. In that model, however, risk is posed by the development, rather than the existence, of advanced technologies. It is thus more closely analogous to (indeed, essentially a special case of) our “transition risk” model of Section 4, and is discussed further there. The result also recalls, and sharpens, Jones’s (2024) observation about the impor tance of the coefficient of relative risk aversion for the willingness to avoid existential risk. Jones finds in a single-period setting that when γ is low, the planner is willing to tolerate a high risk of existential catastrophe in exchange for a spurt to consumption growth. In a single-period setting, the tolerated risk is continuous in γ; no disconti nuity is observed at γ = 1. In the dynamic setting studied here, however, a planner effectively chooses how much risk to tolerate period after period. When γ ≤ 1, enough risk is tolerated each period that an eventual catastrophe is guaranteed. Proposition 2. Policy choice and risk with γ ≤ 1 Suppose a planner faces problem (7)–(9), but with utility function (8) replaced by    u(Ct) =   log(Ct), γ = 1; C1−γ t −1 1−γ , γ ρ≡ (β−α)(1−γ) β Then there is a time t∗ ≥ 0 such that for t < t∗, g. xt = 1, gCt =g >0, gδt =αg >0 (21) (22) 19 and for t ≥ t∗, lim t→∞ gxt = −α βg < 0, lim t→∞ gCt = β −α β g>0, lim t→∞ δtt = δ∗ ≡ lim ρ (β −α)g > 0, γ =1; t→∞ δt = (ρ−ρ)(1−γ) β +γ−1 >0, γ< 1. This is essentially because, when γ < 1, consumption and thus flow utility grow at a higher exponential rate in the long run when g is higher, so the effect of raising g is similar to the effect of decreasing the discount rate ρ. Understanding the path of policy choice and risk is somewhat more complex when γ ≤1 than when γ > 1, because we do not have the result that vt is asymptotically constant, but a sketch is as follows. As in the γ > 1 setting, early in time inequality (14) holds and it is optimal to 20 set xt = 1. Likewise, later in time, optimality requires setting xt < 1 to maintain Atu′(Ct) = ∂δ =⇒ AtxtC−γ t ∂x ·vt = ¯ δAα tβxβ tvt =⇒ δt = C1−γ t βvt . (26) Observe from (11) that vt grows roughly with flow utility u(Ct). Flow utility, for large Ct, then grows approximately like C1−γ when γ < 1. So, though consumption grows t exponentially in the long run for any γ, δ is asymptotically constant when γ < 1. Intuitively, for the policy path to be optimal, it must maintain a) the flow utility to proportionally increasing consumption, Ct · C−γ t = b) the damage done via proportionally raising the hazard rate, which equals the hazard rate × the value of civilization. When the value of civilization also grows like C1−γ , as it does when γ < 1, the hazard t rate must be constant for (a) and (b) to grow at the same rate. When γ > 1, the value of civilization is asymptotically constant, so the hazard rate falls like C1−γ . t When γ = 1, given that consumption grows exponentially, log(Ct) and thus vt grow linearly. The hazard rate then falls proportionally to 1/t. 3.3.5 Simulation The paths of policy choice and the hazard rate are simulated below, for the parameter values listed in Table 1. The values of ρ, γ, and g have been chosen as central estimates ρ 0.02 γ 1.5 g 0.02 A0 2 α 1 β 2 δ 0.00012 Table 1: Simulation parameters for Figure 1 from the macroeconomics literature. A0 = 2 is chosen so that the value of a statistical life-year at t = 75 is four times consumption per capita, roughly matching estimates 21 from Klenow et al. (2023).13 That is, the first year of the simulation might be taken to denote 1949, when a nuclear war between superpowers first became possible, in which case the 75th year denotes the present. ¯ δ, α, and β are chosen so that the hazard rate today is approximately 0.1%, matching Stern’s (2007) oft-cited figure; so that the hazard rate begins to fall at approximately t = 100; and so that the growth rate and then the decay rate of the hazard rate are non-negligible, for clarity in illustration. The probability of survival S∞ under these parameters, from t = 75 onward, is approximately 65%. 1 0.8 Policy choice x 0.6 0.4 0.2 0 0 50 100 150 200 250 300 350 0.2 0.15 0.1 0.05 0 400 Time (%) Hazard rate Figure 1: Evolution of the policy choice and the hazard rate along the optimal path Calculations and code for replicating the simulation and corresponding probability of survival may be found in Appendix C. As Figure 1 illustrates, one potentially unappealing feature of this simple model 13They estimate that this ratio was roughly 5 in the United States in 2019. The figure must be adjusted upward for economic growth since 2019, but downward insofar as we are considering optimal policy across all countries advanced enough to be deploying existentially hazardous technology. 22 is that it implies that, on the optimal path, the hazard rate only rises while no sacrifices whatsoever are made for existential safety. In this it resembles Stokey’s (1998) “environmental Kuznets curve”, whose damages also rise exponentially with growth and then fall sharply once it becomes optimal to take action. As in Stokey (1998), this dynamic is driven by the lack of a lower Inada condition on 1−x. If marginal “safety expenditures” lower the hazard rate infinitely per unit spent at x = 1, then as long as vt > 0 it is optimal to set xt < 1, even if at first the hazard rate is allowed to rise. Rising δ can thus be found alongside falling x by tweaking the hazard function around x = 1. Such tweaks do not affect the long-run behavior of policy or risk as given by (18)–(20), which are set by the shape of the hazard function around x = 0. This is discussed further in Appendix A.4.1. 3.4 Acceleration and state risk As in the tech-only model of Section 2, the impact on cumulative risk of a temporary shock is ambiguous, but the impact of an acceleration—e.g. a permanent level or growth effect—is always to lower cumulative risk. 3.4.1 Preliminaries Let A(·) denote the baseline technology path, given by (4). Let A∗ ≡ At∗, where t∗ is defined as in Proposition 1. Absent a negative shock severe enough to induce stagnation or recession, A crosses every value from A0 to ∞ exactly once, so the area under the hazard curve can be defined by integrating with respect to A instead of t. We will let X denote cumulative risk given that the technology path is A(·) and the policy path x is optimal given A(·): ∞ X ≡ 0 ∞ ¯ δAα txβ tdt = = A0 ∞ A0 −1 ¯ δ Aαxβ A dA ¯ δ Aαxβ A ˙ A−1 dA dt A dA, (27) 23 where we will again abuse notation somewhat by letting xA and ˙ AA denote, respec tively, the optimal value of x (given technology path A(·)) and the value of ˙ A when the technology level equals the subscripted A. We will define vA and δA likewise. Note that δA ≡ ¯ δAαxβ A, without dividing this expression by ˙ AA. That is, it is still a hazard rate: it represents the probability of catastrophe per unit time at technology level A, not the probability of catastrophe per unit of technological development. ˜ A(·) is an acceleration from A ∈ [A0,∞) to A ∈ (At,∞] if ˜ A0 = A0 and ˙ ˜ AA = ˙ AA, A ˙ AA, A∈(A,A); = ˙ AA, A≥A. The acceleration is permanent if A = ∞ and temporary otherwise. Let ˜ A(·) be an acceleration from A. Define ˜ vA such that at A < A, ˜ vA = vA, and at A ≥ A, ˜ vA is the costate variable on survival at A given that the subsequent technology path is ˜ A(·). Then ˜ xA is defined to equal (15) with A,˜ vA in place of At,vt; ˜ δA ≡ δ(A,˜ xA); and ˜ X ≡ ∞ A0 ˜ δA ˙ ˜ A−1 A dA. Given a baseline technology level A and a technology growth rate ˙ ˜ by ˜ A(·)[ϵ] the acceleration from A to A + ϵ with ˙ ˜ AA = ˙ ˜ A, A ∈[A,A+ϵ). A> ˙ AA, denote Then the effect on cumulative risk, per unit of technological development, of instan taneously accelerating to ˙ ˜ A at A is defined to be ∆A, ˙ ˜ A ≡ lim ϵ→0 where ˜ X[ϵ] is cumulative risk ˜ ˜ X[ϵ] −X/ϵ, X, as defined above, given acceleration ˜ A(·)[ϵ].14 14The effect on an instantaneous acceleration on cumulative risk per unit time is ∆A, ˙ ˜ A ˙ ˜ A, since 24 3.4.2 Three shocks Instantaneous level effects — The effect per unit time of a positive shock to the tech nology level At, letting policy adjust instantaneously, depends on whether the shock occurs before or after the regime-change time t∗. At t < t∗, temporarily multiplying the technology level by m > 1 has no impact on the optimal choice of x.15 The hazard rate thus rises. The future hazard rate is unaffected, so cumulative risk increases by δt(mα −1) > 0 per unit of time that the technology level is raised. At t ≥ t∗, temporarily multiplying the technology level by m > 1 multiplies the policy variable by m−α+γ−1 β+γ−1 , by (15). In combination, the positive shock to tech nology and the negative impact on the policy variable multiply the hazard rate by mα−βα+γ−1 β+γ−1 =m−(β−α)(γ−1) β+γ−1 1 lowers cumulative risk (per unit of time that the shock lasts) regardless of t. It does so only because the shock decreases the time spent at technology levels around At. The shock has no impact on the policy associated with any technology level. As in the tech-only model, therefore, we see that the impact of this shock on cumulative risk per unit of increase to the technology level during the acceleration is δt((m ˙ At)−1 − ˙ A−1 t ) < 0. Accelerations — Consider a sharp temporary acceleration, in which the technology level jumps at t from At to A > At and exponential technology growth is subsequently maintained. Since in this model optimal policy is independent of history, this technol ogy shock amounts to a “leap forward in time”. The resulting impact on cumulative risk is A − At δA ˙ A−1 A dA. More generally, an acceleration from A to A can lower the risk endured at the given range of technology levels for two reasons. 1. As in the tech-only model, increasing the technology growth rate at A always lowers cumulative risk directly because the exponent on ˙ AA in integral (27) is negative: ˙ ˜ A−1 A < ˙ A−1 A . 2. Going beyond the tech-only model, given A ∈ [At,A), the value of the future at A is higher given faster future technology growth: ˜ vA > vA. By (15), this motivates weakly more stringent policy ˜ xA ≤ xA and thus a weakly lower hazard rate ˜ δA ≤ δA. If ˙ ˜ AA ˙ AA decreases cumulative risk per unit of technological development during which it endures: a) ∆A, ˙ ˜ A = δA( ˙ ˜ A−1 − ˙ A−1 A ) < X, with equality strict only if A ≤ A∗. The impact of a shock to growth on the probability of survival is explored in the strictly more general model of Appendix A.3. The generalized results are given and proved there in Proposition 8. 3.4.3 Simulation The effects of a sharp temporary acceleration are illustrated in Figure 2. The pa rameter values used to illustrate the baseline path are the same as those used to simulate Figure 1. The acceleration takes place “today”, at t = 75, and multiplies A by e0.2 ≈ 1.22, so that at g = 0.02, it amounts to a 10-year leap forward. Recall from Section 3.3.5 that the probability of survival (from t = 75 onward) on the baseline path is approximately 65%. The proportional increase in the probability of survival can be found analytically. Cumulative risk X declines by precisely the area under the baseline hazard curve from t = 75 to 85; and since δ75 = 0.1%, g = 0.02, and α =1, this difference equals ∆X =−0.001 10 0 e0.02tdt = −0.05(e0.2 − 1). S∞ =e−X is then multiplied by e−∆X ≈ 1.011, so that in absolute terms S∞ rises by approximately 0.65 · 0.011 ≈ 0.7%. 27 1 0.8 Policy choice x 0.6 0.4 0.2 0 0 50 100 150 200 250 300 350 0.2 0.15 0.1 0.05 0 400 Time Figure 2: Acceleration shrinks cumulative risk (%) Hazard rate Calculations and code for replicating the simulation may be found in Appendix C. 3.4.4 Discussion

<<PARAGRAPH BREAKS RESUME>>

Slow growth makes catastrophe inevitable — As noted in Section 2.2, a permanent negative acceleration, or “deceleration”, can render survival impossible: e.g. if it induces stagnation.

In this simple setting, the technology conditions necessary for survival can be stated more precisely. Consider a permanent deceleration after which technology grows power-functionally, so that ˜ At = tk for some k > 0. The exponential growth rate of ˜ A, denoted ˜ g, is then not constant at g but time-varying, with ˜ gt = k/t. By (17) and since ˜ gv → 0, δt then falls to 0 like t−(α−β)(γ−1) β+γ−1 k. Since cumulative risk is finite for δt ∝ t−κ iff κ > 1, the probability of survival is positive iff k > β+γ−1 (α −β)(γ −1).

Growth vs. patience — Faster growth increases the willingness to pay for safety. By contrast, those concerned about the safety of the long-term future often attempt to increase others’ willingness to pay for safety via ethical arguments for a low rate of time preference. Consider e.g. the Stern–Nordhaus debate (and the long debate since) over the discount rate to use in climate policy, or the arguments for concern for the future made by philosophers such as Parfit (1984), Cowen and Parfit (1992), Ord (2020), and MacAskill (2022).

<<TEXT CONDENSED, NONE OMITTED>>

At any value of γ, stagnation at a low technology level A yields a permanent hazard rate of ¯ δAα. This may be arbitrarily low, so the expected duration until catastrophe 1/(¯ δAα) may be arbitrarily high. When γ < 1, an acceleration can quickly yield a hazard rate that permanently approximates δ∗ (25). The acceleration can thus lower civilizational life expectancy to approximately 1/δ∗. 29 4 Transition risk 4.1 Motivation A hazard function of the form δ(At,xt) captures what we have called “state risk”: δ depends on the level of technology. On this framing, it is perhaps unsurprising that escaping risky states more quickly lowers cumulative risk. But risk may instead be “transitional”: posed by technological development. This is the intuition captured by Jones’s (2016) “Russian roulette” model of technological development and (2024) model of AI risk, and by Bostrom’s (2019) analogy to drawing potentially destructive balls from an urn. Perhaps stagnation at a given level of tech nology is essentially safe, and risk arises in the process of discovering and deploying new technologies with unknown consequences. If so, given a positive-growth baseline, does accelerating technological development further increase cumulative risk? 4.2 A transition-risk-based hazard function To explore this possibility, suppose δ increases in ˙ At instead of, or as well as, in At. We will again restrict our consideration to a constant elasticity hazard function: δt = ¯ δAα t ˙ Aζ txβ t, ¯ δ > 0, ζ ≥ 0, β > 1. (29) Our original hazard function (5) is the special case of (29) with ζ = 0 (and β > α > 0). This model is thus a generalization of hazard function (5), complementary to that of Appendix A.3. If ζ > 0, however, the model is most naturally interpreted as one in which risk is posed by the introduction of new technologies—new “draws from Bostrom’s urn”— which consist of absolute increases to A. Fixing policy, introducing multiple technolo gies can pose more, less, or equal risk if done concurrently than if done in sequence, depending on the sign of ζ − 1. Introducing more advanced technologies can pose more, less, or equal risk than less advanced, depending on the sign of α. 30 Alternatively, to interpret one “new technology” as a proportional increase to A, simply rewrite the hazard function as δt = ¯ δAα+ζ t ˙ At At ζ xβ t. On this interpretation, α + ζ > 0 is the condition under which developing more advanced technologies poses more risk than developing less advanced technologies. Because ˙ A/A has been roughly constant through the last century, the view that the hazard rate has risen must be attributed to the increasing danger of each “technolog ical development” in this sense. Finally, consider the case of α = −1, ζ = 1, so that δt = ¯ δ ˙ At At xβ t. Here, fixing x, each proportional increase to A induces a constant hazard, indepen dently of how quickly the increase occurs. In the absence of policy—with x = 1 (or any other constant) permanently—this model is essentially equivalent to the “Russian roulette” model of Jones (2016)16 and the AI risk model of Jones (2024). 4.3 Acceleration and transition risk 4.3.1 Without mitigation Suppose that the baseline technology path A(·) is continuously differentiable, with a positive derivative. Let ˆ A≡limt→∞At be finite or infinite. As implied above, fixing policy, whether acceleration increases or decreases cumu lative risk depends on whether ζ is greater or less than 1. This can, as usual, be seen most clearly by integrating the hazard curve with respect to A: ∞ X = 0 16Our ¯ δ is the variable there denoted π. ˆ A ¯ δAα t ˙ Aζ tdt = A0 ¯ δAα ˙ Aζ−1 A dA. 31 Given acceleration ˜ A(·) from A ∈ [A0, ˆ A) to A ∈ (A, ˆ A], cumulative risk equals A ˜ X =X+ A ¯ δAα ˙ ˜ Aζ−1 − ˙ Aζ−1 A dA. The integral is negative if ζ < 1, zero if ζ = 1, and positive if ζ > 1. In the Russian roulette model, for instance, though there is a technology level ˆ A < ∞at which it is optimal to halt technological development (Appendix A.5), accelerating technological development before ˆ A does not affect cumulative risk. 4.3.2 With mitigation In Section 2, we saw that acceleration weakly lowered cumulative state risk absent policy. In Section 3, we saw that the tendency of acceleration to lower cumulative state risk was amplified by the presence of optimal policy. Here, we have seen that the impact of acceleration on cumulative transition risk is ambiguous absent policy. We will now see that it remains ambiguous given optimal policy, but that policy can reintroduce a tendency for acceleration to lower cumulative risk. For simplicity, we will now again impose the assumption that A grows at a constant exponential rate g. Also, since given exponential growth g ˙ A = g, we will impose β >α+ζ, (30) which, rather than β > α, is now the condition necessary for survival without Ct = Atxt → 0. Under these conditions, since ˙ A is proportional to A, the planner’s problem is precisely as described in Section 3.3, with α + ζ taking the place of α (up to a coefficient gζ that can be rolled into ¯ δ). Baseline x and δ paths, and S∞, are unchanged. The existential risk Kuznets curve remains. Let A∗ denote the uppermost technology level at which it is optimal to set x = 1 on the baseline technology path. Since the first-order condition ∂u ∂xt (At,xt) ≥ ∂δ ∂xt (At,xt)vt =⇒ A1−γ t x−γ t ≥ ¯ δAα tβxβ−1 t vt 32 must be satisfied everywhere and hold with equality for x < 1, we have    xA =   1 A≤A∗, ¯ δβAα+γ−1 ˙ Aζ AvA − 1 β+γ−1 A >A∗. Substituting (31) into the expression for cumulative risk ∞ X = we have X = A∗ ∞ ¯ δAα ˙ Aζ−1 A dA+ A0 ¯ ¯ δ Aα ˙ Aζ−1 A xβ AdA, δ1−γββA(β−α)(γ−1)vβ A − 1 β+γ−1 ˙ A ζ γ−1 β+γ−1−1 (31) (32) A dA. (33) A0 A∗ Recall that a technology path ˜ A(·) is an acceleration if ˙ ˜ AA > ˙ A for technology levels A ∈ [A0,∞) to A ∈ (A ≤ ∞]. With or without policy, an acceleration affects cumulative risk directly, by changing the technology growth rate from A to A. With policy, an acceleration also affects cumulative risk indirectly by affecting vA for A ∈ [A,A), which affects policy at this range of technology levels. Under the hazard functions of the previous sections, as we have seen, faster tech nology growth is always weakly preferred. This follows from the fact that it is feasible to offset higher values of At with lower choices of xt, such that the original consump tion path is maintained, and from the assumption from β > α that given this policy response, the hazard curve is weakly lowered. Since a future with faster growth is more valuable, an acceleration from A to A raises vA for A ∈ [A,A). Under hazard function (29), this argument is no longer valid. This is because, un like an increase to At, an increase to ˙ At brings no contemporaneous benefit, though it imposes risks that can still be mitigated only with less contemporaneous consump tion. And indeed, under hazard function (29), faster technology growth is no longer always preferred. We can see this most straightforwardly in the case of α = −1, ζ =1: again, this is the Russian roulette model of Jones (2016), and as Jones finds, with γ > 1, it is optimal for technology to grow only to a finite level. In the more 33 general model here, the result that stagnation is optimal is knife-edge, as discussed in Appendix A.5. Nevertheless, the result that an acceleration from A to A does not necessarily yield ˜ vA ≥ vA for A ∈ [A,A) holds more generally. These complexities are avoided when we focus on instantaneous accelerations. The impact of an acceleration from A to A on vA, for A ∈ [A,A), falls to zero as A−A → 0. The impact of a brief acceleration on cumulative risk is therefore approximately the impact found when we ignore impacts on vA. Proposition 4. Instantaneous acceleration and transition risk Given hazard function (29) and technology path (4), choose a technology level A > 1 and growth rate ˙ ˜ A> ˙ AA. If a. A ≥A∗ and ζ )1+ β γ−1 , or b. A )1, and ˙ ˜ A maintains (31) = 1 at A = A, then ∆A, ˙ ˜ A < (=,>)0. Proof. See Appendix B.3. The result follows essentially immediately from the exponent on ˙ AA in (32). In particular, instantaneous acceleration after A∗ lowers cumulative risk as long as ζ γ−1 β +γ−1−1< 1. The α ≥ −1, γ ≤ 2 case follows from the fact that if α ≥ −1, then, by (30), ζ < β + 1, so ζ β+1 < 1. Since macroeconomic estimates of γ ≤ 2 are standard, this result suggests that accelerations lower cumulative risk on the optimal path in the context of transition risk, at least if they occur late enough in time that mitigation is already underway. 34 Furthermore, this is, again, even without considering the fact that an increase to future growth can change the value of the future. Though the direction of this change is in principle ambiguous, most observers today take it for granted that, at least on a conventional discount rate, faster technology growth would be a benefit on the current margin. This would then be another channel through which a (positive duration) acceleration would motivate greater concerns today for safety. It may be counterintuitive that instantaneous acceleration reduces risk only when γ lies below a bound, because when γ is higher, the marginal utility of consumption is lower and it is optimal to shift resources from consumption to safety more rapidly. The result stems from the fact that, when γ is high, the marginal utility of consump tion rises rapidly as x is cut, so following an acceleration, a small cut to x suffices to equalize the marginal utility of consumption with the marginal value of safety spend ing. The higher γ is, the more quickly x falls as A rises, but the less sensitive x is to a change in ∂δ/∂x—e.g. an increase due to higher ˙ A—at a given value of A. 4.3.3 Discussion The nonrivalry of safety effort — Hazard function (29) is explored here mainly for its simplicity and similarity to (5). One valid criticism of this functional form is that it overemphasizes a channel through which the risks posed by a series of technological developments can be cheaper to mitigate if they occur at once than if they occur in sequence. Suppose that β ≈ 1, that ζ = 1, and that two small increases to ˙ A—let us call them two “experiments”—can occur in sequence or simultaneously. If they occur in sequence, halving the risk posed by each requires halving x and thus consumption for two periods in a row. If the experiments occur simultaneously, the same reduction in cumulative risk only requires halving consumption once. For some kinds of experiments and some kinds of safety infrastructure, the as sumption that safety efforts are “nonrival” in this sense is reasonable. Wastewater monitoring for the sake of early pandemic detection reduces the risk posed by poten 35 tially pandemic-inducing biological experiments by a proportion independent of how many experiments are underway. In other cases, however, it is not reasonable. It does not apply, for instance, to the costs of the safety equipment that must be used at each lab. Safety efforts of this kind might be better modeled by a modified version of the “safety in redundancy” model of Appendix A.4.2. A thorough attempt to shed light on the relationship between growth and transition risk would require further study. Nevertheless, the basic model explored here offers two lessons. First, in the absence of policy, the effect of acceleration on transition risk is ambiguous, and there is no effect in the arguably central “ζ = 1” case assumed by Jones (2016, 2024). Second, the presence of an optimal policy response can shift the relevant “ζ” threshold, in particular significantly shifting it upward to the extent that safety efforts are nonrival across contemporaneous risks. Stagnation vs. deceleration — When ζ > 0, complete stagnation ( ˙ A = 0) is always the safest path of all. Nevertheless, we have seen with and without policy that given a positive growth rate, faster growth can decrease risk. The key to this puzzle is that, given stagnation at ˆ A, levels of A > ˆ A are never attained. Cumulative risk is therefore not (32) but (32) with the ∞ replaced with ˆ A. Absent stagnation, however slow the growth rate, all levels of A are attained. The growth rate only determines the risk endured at each one. The direct cost of faster progress during a given range of A-values (higher risk per unit time, to the extent that ζ > 0) is partially, and may be more than fully, outweighed by the fact that faster progress motivates more mitigation at each point in time, in combination with the now familiar fact that when progress is faster we do not linger in a given range of A-values as long.

<<PARAGRAPH BREAKS RESUME>>

Human activity can create or mitigate existential risks. The framework presented here illustrates that, under various conditions, existential risk should be expected to exhibit a Kuznets curve. This observation offers a potential economic explanation for the claim by some prominent thinkers that humanity is in a critical “time of perils”. We may be advanced enough to be able to destroy ourselves, but not yet enough that we are willing to make large sacrifices for the sake of safety. If we are indeed living through the time of perils, reductions to existential risk today have massive expected long-term consequences.

At the same time, this framework highlights a channel through which some ef forts intended to reduce existential risk may backfire. In the absence of policy, when risk is posed by the existence of advanced technologies, broad-based decelerations to technological development typically worsen or do not affect the odds of long-term survival. Given an optimal policy response, even by a policymaker with little care for the long-term future, this impact is magnified. The impact can be significant, with proportional consumption decreases having comparable impacts to proportional in creases in the planner’s rate of time preference. In the extreme, permanent stagnation can make a catastrophe inevitable that might otherwise have been avoided.

#### Technological upsides buffer societal resilience to inevitable Great Filter events---extinction.

Jiang ’23 [Jonathan, Phillip Rosen, Kelly Lu, Kristen A. Fahy, and Piotr Obacz; February 2; Senior Research Scientist at the Jet Propulsion Laboratory, California Institute of Technology; academic researcher; Researcher at NASA’s Jet Propulsion Laboratory; Professor in the Faculty of International and Political Studies at Jagiellonian University; Journal of Humanities and the Social Sciences, “Avoiding the ‘Great Filter’: Extraterrestrial Life and Humanity’s Future in the Universe,” vol. 6]

But, if “life as we know it is merely an afterthought in the global scheme of the cosmos” [1], we may well be led to conclude the Earth’s bounty is of truly extraordinary – perhaps even unique – nature. Are humans, just one among the millions of species sharing this remarkably hospitable but fragile oasis in the cosmos, a kind of multiply improbable instance of fused ash first assembling to primitive life then, much later, tumbling into self-awareness followed by spacefaring technological prowess? One school of thought posits this ‘Rare Earth’ hypothesis [2]: given a Universe stretching approximately 92 billion lightyears and existing for nearly 14 billion years, intelligent life can be both inevitable but still exceedingly rare. Hence, present era Earth is merely the particular time and place such extraordinarily long odds, in effect, paid off and we are the lucky beneficiaries. While such a notion may come as comfort to some as they (philosophically speaking) claim universal ownership, this scenario would also leave us profoundly isolated and stunted. The great scientists, mathematicians and artists our civilization has produced achieved their historic feats through collaboration and competition. Extending this notion beyond our home world, how could humanity as a species ever truly realize our full potential if there are no other technological civilizations with whom to interact?

Technological developments in the years following Enrico Fermi’s famous question, posed to colleagues in 1950 and forming the Paradox [3] which came to bear the name of the great 20th century physicist, enabled the search for extraterrestrial intelligence (SETI) to commence. Among these innovations are radio astronomy and dramatic developments in rocketry and computational power. Astronomers, long confined to making observations within just the very narrow band of the electromagnetic spectrum afforded by human vision, could now view and measure a cosmos whose radiation signatures extended from long wave radio through high energy x-rays. For centuries the light reaching Earthly observers from the planets and other bodies of our Solar System was no more than tiny atmosphere-distorted points and weakly resolved discs flickering in glass lenses. Many of these worlds and other celestial objects have now been physically visited by humanity’s robotic emissaries while Earth’s Moon has been touched by humans themselves. Additionally, our vantage point in observing the Uni verse has taken up positions off our home world, using ever more sophisticated and sensitive instruments to peer across billions of lightyears and even in a few cases imaging worlds orbiting other suns. Since 1992 over 5000 exoplanets have been confirmed with several thousand candidates additionally pending, attesting to the ubiquitous nature of planetary systems. Moreover, modeling from such early works as the Drake Equation [4] to more recent investigations suggest extraterrestrial intelligence may well have arisen in the Milky Way [5], [6], [7]. Pursuing the question still further, in recent years serious exploration of the complex implications for human society upon coming into contact with life off the Earth has moved into the realm of mainstream scientific inquiry [8]. If resolution of the Fermi Paradox does not to condemn humanity to a lonely Universe, it is only logical that life must inevitably strive to seek other life wherever it may exist.

Just as technology enables humanity to push back the boundaries of our knowledge of the cosmos, it tempts as well with the means of self-destruction. Abruptly realized, we find ourselves the sole stewards of a resource rich world – and socially ill-prepared for the job. A worrying sense of humanity’s technological cleverness outpacing our better judgment pervades, for even as we reach far beyond Earth’s gravity well, we are being pulled down by internal strife. This presents a potentially universal question, as well as a most unsettling solution to Fermi’s Paradox. As depicted in Figure 1, is there a “Great Filter” [9] of sorts awaiting every civilization which sets out on the path of technological development and if so, has humanity yet to confront this ultimate rite of passage? An optimistic perspective would point out our continued existence de spite first developing the capability for self-annihilation in 1945. Caution, however, is well warranted as these past 77 years (in the developed world, merely an average human lifespan) have been fraught with near misses such as the Cuban Missile Crisis in 1962 and persistent flaring of armed conflicts around the globe. Additionally, human activity has unsettled the Earth’s otherwise highly accommodative environment for supporting life, casting a dark shadow over the prospect of endlessly advancing technological innovation opening up an unlimited future for humanity to spread across space and time. Returning to the Drake Equation, recent modeling suggests that it is the lifespan of civilizations capable of interstellar communication, “L”, which is the most influential among its seven variables [5], [10]. Taking this claim as stipulation, it follows immediately that the sub-factors comprising L must be identified and studied in detail if we are to maximize humanity’s lifespan.

In pragmatically simplified terms, the notion of a filter suggests a structure specifically configured to allow one or more constituents of a multiple component flow to pass through while the remaining constituents are prevented from fully crossing to the far side. In this analogy, the overall development (i.e., forward-moving flow) of species’ evolution to sentience, as well as the follow-on development of civilization itself, is the bulk fluid and the critical characteristics particular to each component species determines if it passes this semi-permeable barrier or is captured, irreversibly halting progress. The key to humanity successfully traversing such a universal filter is found in understanding what characteristics the barrier will constrain, identifying those attributes in ourselves and neutralizing them in advance. Human civilization over the past 5000+ years, and in particular since 1945, has revealed much of what would surlily impede, if not outright arrest, our aspirations to colonize other worlds in the Solar System and beyond. It seems as though nearly every great discovery or invention, while pushing back the borders of our technological ignorance, is all too quickly and easily turned to destructive ends. Examples such as splitting the atom, biomedical innovations and resource extraction and consumption come to mind with disconcerting swiftness. Still, some have suggested artificial intelligence (AI) as yet another factor which, pending substantial technical hurdles, may yet have its chance to prove friend or foe. The implications of AI pose complexities perhaps only an AI itself could understand, and has even been suggested as an alternative explanation [11] for Fermi’s Paradox. Although alien AI is, inescapably, a doubly strange notion, some perspective may be gained when compared to still more exotic hypotheses such as involving dark matter [12]. Finally, nature itself has the potential to extinguish human civilization, often envisioned as an asteroid or comet impact similar to that which Earth suffered 66 million years ago, triggering the Cretaceous-Paleogene extinction and wiping out roughly 75% of all species worldwide.

Rationale

If life arisen on Earth is ever to know of life elsewhere, assuming such exists, we as the Earth’s sole technological species must first come fully to terms with ourselves and our environment. The struggle for survival, security and dominance - all rooted in human passions - drives creativity and with it, civilization and invention [13]. As history has shown time and again, however, this cleverness comes at great cost. The human brain, still orders of magnitude more complex in terms of synaptic connections than the transistor-based structures underlying the most advanced super computers, holds the key. Using our demonstrated inventiveness to proactively recognize, diagnose and formulate countermeasures to the most serious threats to our existence, humanity may yet avert the Great Filter. In so doing we would likely emerge downstream of the Great Filter as a near Type I civilization on the Kardashev scale [14], ready to seek our place in a future greater than what we could realize if confined to just our home world. Indeed, recent modeling suggests human-crewed exploration of our Solar System beyond Mars may well be possible within this century [15], [16]. Analysis of these leading threats has found them to include large-scale nuclear warfare, pathogens (both naturally occurring and engineered), artificial intelligence, impacts from asteroids and climate change [17].

Unchecked population growth is a factor in the aforementioned scenarios, excepting that of asteroid impact, threatening human civilization and life on Earth in general. At present world population is at 8 billion, an exponential rise from about 1.6 billion at the start of the 20th century, and has doubled over the last 49 years. While Malthusian-inspired worst-case predictions of such a rapid increase [Ehrlich, 1968] have so far been averted, thanks in large part to technological advancements in farming, energy production and distribution, invention cannot be expected to indefinitely offset the multifaceted stresses imposed by ever escalating population. However, further improvements in modeling and better-informed controls, with education in developing nations as a critical factor for success, suggest a pathway to towards reducing population at a modest pace after a projected peak of slightly less than 10 billion is passed in the 2060s [18]. If such a prediction is at least directionally correct, it is not unreasonable to cautiously expect a moderating effect will ensue in many of the major challenges humanity now faces as we move towards the latter decades of the 21st century.

Nuclear War

Warfare has beset humanity long before civilization began taking hold approximately 5200 years ago. An outgrowth of tribalism and our innate sense of competition for resources, invention has been harnessed to make successively more deadly weapons, as well as defenses. With the passage of time, the horizon has effectively drawn closer to home and no more geographical frontiers remain to be to conquered here on our native world. War cannot be thought of as “something bad” happening to unfortunate people a safe distance away from where one has chosen to live. That said, there are encouraging signs of rationality emerging: peace agreements in the historically troubled Middle East, a vast reduction in nuclear warheads since the height of the Cold War and a wide coalition of nations rallying their support for the besieged in Eastern Europe. Amid continued strife it should also be noted that at the start of the 20th century the U.S. was among a very small handful of constitutional democracies in the world. The present finds a clear majority of the world’s nations are, albeit imperfectly and often only partially, democracies with at least nominally representative governments [19]. Though strained, the Democratic Peace Theory - which holds that democracies are hesitant to go to war with each other [20] - has historically been borne out. In parallel to this hopeful trend, the end of WWII saw the age of colonial rule finally begin its long ebb into tragic history. Hence, to estimate the threat of large-scale warfare one may take as proxy the quantified extent of constitutional democracies across the world with time. Weighting relatively more heavily the trends towards or away from functionally representative democracy in those nations con rolling significant stockpiles of nuclear weapons would logically follow. The past, it has often been said, is prologue - but it need not be prediction. A future free of catastrophic warfare remains, for now, within humanity’s grasp and with it, avoidance of perhaps the most obvious of Great Filters.

Pathogens & Pandemics

As very recent events have painfully reminded us, biologically-based threats remain at the forefront of humanity’s many concerns. Pathogens, microscopic but with potential for causing death on a planetary scale, have continually emerged throughout history. Although the vast majority of viruses and bacteria are either harmless or nearly so, combinatorial biochemistry incubated in large populations and integrated across a great many iterations have nonetheless repeatedly given rise to the rare deadly strain capable of rapid transmission. Catalyzing this threat in modern times is human civilization’s ever-increasing interconnectedness, shrinking the vast distances between continents to effectively that which might have existed between neighboring medieval villages. It is not unreasonable to suggest humanity has actually been fortunate to have only encountered two particularly serious pandemics since WWI. Here again, our ingenuity holds the means of survival. Whereas past generations were at the mercy of deadly pathogens, modern diagnostic and pharmaceutical techniques are powerful allies in containing and, ultimately, defeating this recurring foe. The current struggle with SARS-CoV-2 offers a silver lining: an opportunity to model quantitatively the factors which comprise the threat of pandemic in the setting of a technologically well equipped society. Data from past pandemics such as Influenza, which struck in the years immediately after WWI, may well add some useful (albeit limited) context to current modeling of pandemics. That said, it must be emphasized there were no vaccines quickly developed to battle that pandemic and it ran its horrific course unencumbered just as pandemics had done long before the turn of the 20th century. In summary, application of reliable data in the present [21] must take centerstage in predicting how future pandemics will spread, how deadly they will be and how quickly and effectively we will be able to leverage our knowledge of the life sciences to counter this manifestation of the Great Filter.

Artificial Intelligence

Once confined to the realm of speculative fiction in popular works such as Arthur C. Clarke’s 2001: A Space Odyssey (1968), William Gibson’s Neuromancer (1984) and James Cameron’s The Terminator (1984), the practicality of achieving artificial intelligence has moved methodically towards realization with advances in microcircuit technology. While dramatic screen and print depictions range from that of friendly androids to malevolent supervillains bent on world destruction, a sober assessment of the actual risks posed by AI remains as elusive to full comprehension as the minds of AI’s presumptive human inventors. Taking the guarded view, if and when AI does come to fruition it may well be too late to rely on empirical evidence gathered from its actual attitudes and behaviors towards the species which brought about its existence. Prudence then strongly suggests we perform sooner rather than later what modeling can be done, evaluate the necessarily preliminary conclusions drawn and proactively plan for a peaceful approach to the possibility of sharing the Earth with a new technological entity. To start, an assumption is required which theorizes arrival of AI is conditionalized, though not guaranteed, on achieving with hardware the same level of structural complexity as that of the human brain, which itself encompasses ~1014 synaptic connections among its ~1011 neurons [22]. Presently, microprocessors can contain up to ~1010 transistors - electronic gateways which serve a roughly comparable function to the bioelectrically driven synaptic connections between neurons in the brain [23], [24]. As the density of transistors per microprocessor has increased exponentially since the 1960s, this corresponding to what is widely characterized as Moore’s Law [25], one can project when computer sophistication may rival that of the human mind. Of course, impediments posed by material limitations and quantum effects would need to be over come if this rapid, decades-long trend towards brain-like complexity is to be maintained beyond the mid-2020s. As for whether AI would be benign or otherwise, self-imposing a Great Filter of our own invention, that will depend on the evolving nature and disposition of Earth’s first high-tech species.

Asteroid & Comet Impacts

For centuries astronomers considered the movements of the Solar System’s visible planets and the stars to mimic that of clockwork. Indeed, primitive timekeeping depended on the comings and goings of the Sun, Moon, “wandering stars” (i.e., planets) and those flickering specs of light - many of which we now know warm worlds many lightyears distant. The lesser remains of our Solar System’s formation still orbit the Sun as asteroids in their un counted billions, carbonaceous or stony objects while still others contain significant percentages of metals. Gravitational perturbations sometimes send these remnants sunward, typically originating from outer regions such as the Kuiper Belt and the Oort Clouds, where they can occasionally tumble across planetary orbits. Analogously, the paths of comets – icy bodies sheathed in frozen gases – may very occasionally intersect planets and moons. Most objects are relatively small and upon encountering Earth’s atmosphere at high velocity, disintegrate into harmless bits or cinders. There are, however, a non-zero percentage which are large enough to survive passage through the atmosphere and, impacting the surface, cause catastrophic destruction to our sensitive biosphere. Risk modeling for such ultra-low frequency, high severity events are, by their very nature, challenging. At the fundamental level such calculations involve the product of especially low annualized likelihoods multiplying difficult-to-fathom severity factors to produce time dependent cumulative risk trends. A useful information source which may be leveraged for this purpose is found in one of the major initiatives addressing this challenge to planetary defense: NASA’s Near-Earth Object (NEO) Observations Program [26]. A simplified procedure would envision mining public-facing data from this resource and making use of readily available risk management modeling techniques to generate quantifiable results. The end product would be weighted risk curves versus time for mass extinction event (MEE) level impacts – i.e., a vanishingly tiny risk of a MEE in the next year ranging asymptotically towards 100% as likelihood is integrated over time into the very distant future. It should be noted that as with any statistically derived result there is no guaranteed interval of complete safety, only a likely window to prepare in the near term. Fortunately, NEO and other projects such as the Double Asteroid Redirection Test (DART) mission are examples of humanity using our technological capabilities to proactively address this possible version of the Great Filter.

Climate Change

In recent decades there is perhaps no large-scale threat to life on Earth which has been studied more intensely than climate change. While public opinion on the bottom-line implications of a warming biosphere continues to vary, general acceptance of the basic contention that surface temperatures are rising and human activity is a significant driver has largely moved beyond doubt among national governments. Central to this focus is the United Nations’ Intergovernmental Panel on Climate Change (IPCC), whose ongoing investigations include updated predictions for rising temperatures resulting from emissions of greenhouse gases (GHGs) using a multitude of climate models [27]. Many independent studies have also been performed such as those seeking an empirically-based relationship between surface temperature rise and escalating con centration of the chief GHG, CO2, in the lower atmosphere [28]. Given this plethora of models, logic strongly suggests the most re liable method for making analytically sound predictions for climate change is not in piecing together new models but rather determining where the most widely accepted established models converge – i.e., a preponderance of the evidence approach. The relatively slow nature of climate change, along with disagreements between some of the models’ predictions, continues to present headwinds to engaging wider efforts to blunt the effects of GHGs. The major impediment to taking more decisive actions, however, are the challenges imposed by transitioning to non-carbon-based energy sources such as solar, wind, and nuclear power. Here again, rapidly advancing technologies in areas such as modularized nuclear power plants [29] and carbon capture and sequestration (CCS) are among the best hopes for avoiding slow-motion ensnarement by this lulling but lethal Great Filter.

#### Specifically, gig work arrangements facilitate open innovation.

Redlich ‘25 [Alexander, Kathleen Diener, Frank Piller; Research Associate at RWTH Aachen University; Professor for Business Informatics - Digital Innovation; professor of technology management at RWTH Aachen University; Four Waves, “Leveraging the Gig Economy: How Firms utilize Open Talent Platforms to Innovate,” https://event.fourwaves.com/ouiconference/abstracts/5578d56b-8654-4551-a58d-54becf77a984]

Today, the gig economy is disrupting traditional employment relationships and how firms get access to talent. The gig economy is characterized by temporary work, project-based compensation, flexibility and non-organizational membership of freelancers. Also referred to as on-demand workforce or external talent crowd, it enables firms to hire independent individuals through digital platforms like Upwork, fiverr or Freelancer. This trend is driven by both firms needing more flexible access to specialized talent, while achieving operational agility and reduced costs, as well as gig workers choosing the best paid and most interesting tasks while having benefits such as the flexibility of working from home. Despite not being traditional employees, gig workers can be viewed as strategic talent, with a growing trend towards longer-term relationships between skilled gig workers and client firms. In the U.S. alone, already 64 million people are working part or full-time in the gig economy, with a majority of the U.S. workforce estimated to become self-employed by 2027. As a result, the gig economy has the potential to completely change how firms employ talent to innovate.

The gig economy's impact on innovation in firms is multifaceted. It can potentially enhance knowledge sharing and transfer between organizations through freelancers' project experiences. Yet, relying on gig workers can also undermine knowledge continuity and cultural coherence within firms, potentially hindering sustained innovation. Therefore, it is crucial for firms to understand for which types of innovation activities (radical or incremental) firms can seek and employ gig workers to innovate successfully, while benefitting from the advantages utilizing an on-demand workforce. Additionally, gig workers have to balance whether they further specialize in more exploitation related tasks like in the past or build new skills related to exploration and innovation tasks. Drawing from the theories of open innovation, innovation ecosystems and ambidexterity, with this study we therefore aim to develop a framework to answer the following overarching research questions:

How do gig workers balance the promotion of exploration vs. exploitation skills?

For what different types of innovation do firms use gig-platforms?

In order to answer this question, we analyze the tasks and profiles posted on gig platforms, how these tasks are formulated and structured to attract talent capable of driving innovation and what strategic innovation objectives of firms are reflected in job postings. We also aim to analyze how they match with the offered skills and expertise of gig-worker profiles. For this purpose, we will use a quantitative approach, by collecting primary data from a set of leading gig work platforms. On these platforms firms can post job offerings including a detailed task description, including required skills, level of experience, timeline and offered compensation. Gig workers in turn offer their services through a profile including information like skills, experience and availability. We aim to collect primary data using web scraping algorithms to extract the information from both job offerings and profiles, enhanced by dictionary-based text analysis to extract data from texts. We hypothesize that task phrasing and design on gig-platforms provide insights into firms’ priorities, such as fostering radical innovation versus optimising existing processes.

This study contributes to the literature on gig work and innovation management by uncovering how task structuring on gig platforms reflects strategic innovation goals. It bridges a gap in understanding the operationalization of gig work for innovation, highlighting opportunities and limitations in leveraging external talent. Findings offer new perspectives on designing tasks that align with broader organizational innovation strategies. This is relevant for managers to understand for which type of innovation activities to utilize gig work effectively and give insights into the design of gig-jobs for different types of innovation objectives.

#### Cumulative existential risk from a failure to implement safeguards makes extinction inevitable and outweighs any single scenario.

Piper ’22 [Kelsey citing Toby Ord; September 22; senior writer at Future Perfect; PhD, senior research fellow in philosophy at Oxford; Vox, “How to stop rolling the dice on the destruction of human civilization,” https://www.vox.com/future-perfect/23362175/un-human-development-report-ord-existential-security]

“The war in Ukraine reverberates throughout the world,” the report opens, “causing immense human suffering, including a cost-of-living crisis. Climate and ecological disasters threaten the world daily. It is seductively easy to discount crises as one-offs, natural to hope for a return to normal. But dousing the latest fire or booting the latest demagogue will be an unwinnable game of whack-a-mole unless we come to terms with the fact that the world is fundamentally changing. There is no going back.”

Those words ring true. Only a few years ago, we lived in a world where experts had long warned that a pandemic was coming and it could be devastating — now, we live in a world that a pandemic has clearly devastated. Only a year ago, there hadn’t been a large land war in Europe since World War II, and some experts optimistically assumed that two countries with McDonald’s in them would never go to war.

Now, not only is Russia occupying stretches of Ukraine, but the destruction of Russia’s army in the fighting there has kicked off other regional instability, most notably with Azerbaijan attacking Armenia earlier this month. Fears of the use of nuclear weapons in wartime, quiet since the Cold War, are back as people worry about whether Putin could turn to tactical nukes if faced with a total defeat in Ukraine.

Of course, all of those situations are possible — even likely — to resolve without catastrophe. The worst rarely happens. But it’s hard to avoid a feeling that we’re just rolling the dice, hoping that we somehow won’t eventually hit on an unlucky number. Every pandemic, every minor war between nuclear-armed powers, every new and uncontrolled technology, may pose only a small chance of escalating to a catastrophic-scale event. But if we take that risk every year without taking precautions, humanity’s lifespan may be limited.

Why “existential security” is the opposite of “existential risk”

Toby Ord, senior research fellow at Oxford’s Future of Humanity Institute and the author of the existential risk book The Precipice: Existential Risk and the Future of Humanity, explores this question in an essay in the latest UNDP report. He calls it the problem of “existential security”: the challenge not just of preventing each individual prospective catastrophe, but of building a world that stops rolling the dice on possible extinction.

“To survive,” he writes in the report, “we need to achieve two things. We must first bring the current level of existential risk down — putting out the fires we already face from the threats of nuclear war and climate change. But we cannot always be fighting fires. A defining feature of existential risk is that there are no second chances — a single existential catastrophe would be our permanent undoing. So we must also create the equivalent of fire brigades and fire safety codes — making institutional changes to ensure that existential risk (including that from new technologies and developments) stays low forever.”

He illustrates the point with this fairly terrifying graph:

<<GRAPH OMITTED>>

The idea is this: Say we go through a situation where a dictator threatens to use nuclear war, or where tensions between two nuclear powers seem to be hitting the breaking point. Maybe most of the time the situation is defused, as indeed was the case during the many, many Cold War close calls. But if this situation recurs every few decades, then the probability we’ll defuse every single prospective nuclear war will get steadily lower. The odds that humanity will still be around in 200 years eventually become quite low, just as the odds that you can keep winning at craps drop with every roll.

“Existential security” is the state where we are mostly not facing risks in any given year, or decade, or ideally even century, that have a substantial chance of annihilating civilization. For existential security from nuclear risk, for instance, perhaps we reduce nuclear arsenals to the point where even a full nuclear exchange would not pose a risk of collapsing civilization, something the world made significant progress on as countries slashed nuclear arsenal levels after the Cold War. For existential security from pandemics, we could develop PPE that is comfortable to wear and provides approximately total protection against disease, plus a worldwide system to detect diseases early — ensuring that any catastrophic pandemic would be possible to nip in the bud and protect people from.

The ideal, though, would be existential security from everything — not just from the knowns, but the unknowns. For example, one big worry among experts including Ord is that once we build highly capable artificial intelligences, AI will dramatically hasten the development of new technologies that imperil the world while — because of how modern AI systems are designed — it’ll be incredibly difficult to tell what it’s doing or why.

So an ideal approach to managing existential risk doesn’t just fight today’s threats but makes policies that will prevent threats from arising in the future too.

#### Open innovation safeguards humanity from existential emerging tech.

Callaghan ’18 [Chis; December; Professor of Management in School of Business Sciences at the University of the Witwatersrand; Futures, “Surviving a Technological Future: Technological Proliferation and Modes of Discovery,” vol. 104 p. 100—16]

“Almost all the problems we face nowadays are complex, interconnected, contradictory, located in an uncertain environment and embedded in landscapes that are rapidly changing” according to Sardar (2010:183). Over and above the use of nuclear or biological weapons, potential risks are associated with emergent technologies such as artificial intelligence (AI), biotechnology, geoengineering, and nanotechnology (Baum, 2015). Unlike the threats of nuclear, biological or chemical weapons of mass destruction, however, novel technologies such as genetics, nanotechnology and robotics (GNR) do not require large-scale activities to pose threats to humankind, but only require knowledge, posing the threat of knowledge-enabled mass destruction, “amplified by the power of self-replication” (Joy, 2000:1).

How does one address such threats? To do so, a choice must be made in terms of what scientific methodology to use, and a definition is required, as to what it means to ‘address’ such threats. Given the almost unimaginable uncertainties associated with technological advancement and its consequences (Vinge, 1993; Szerszynski, Kearnes, Macnaghten, Owen, & Stilgoe, 2013; Bostrom, 2017; Tegmark, 2017), and therefore the need for responsible innovation (Grunwald, 2011; Stilgoe, Owen, & Macnaghten, 2013), this paper takes recourse to the approach of future studies to develop a theoretical framework relating to how to prepare for such scenarios. In doing so, this paper also seeks to advance the argument that only through improvements in the capacity of humans to manage technology, can human agency survive into a technological future. This argument therefore draws from Tegmark’s (2017) logic, that given uncertainty about technological outcomes, an important goal is to immediately undertake technology-safety research, and make this research mainstream. As an organising framework, literature is used to derive six primary technological threats. These threats are used to guide and anchor discussions and arguments throughout different sections of the paper.

This paper therefore seeks to make a contribution to the future studies literature, in the following ways. First, it seeks to relate the threat of technological development to Sardar’s (2010) four laws of future studies. Whereas most approaches to such problems suffer from the constraints associated with disciplinary lenses, this work seeks to understand these problems through a systematic approach that does not privilege any specific disciplinary approach over another. Future studies “is not just multi- and trans-disciplinary, it is unashamedly un-disciplinary: that is, it consciously rejects the status and state of a discipline while being a fully fledged systematic mode of critical enquiry” (Sardar, 2010:183). This is Sardar’s first law, and it is particularly important in light of certain of the arguments made here, which draw on Nielsen’s (2012) theory of networked science, and its predictions that human collaboration enabled by novel technologies can result in radical innovations in the scientific discovery process itself. Nielsen argues that it is changes in the scientific discovery process itself that account for the great scientific leaps, or scientific advances through history. It is argued here that Nielsen’s theory offers useful insights into how radical innovations in the research, or R&D process itself can improve the capacity of humans to manage technological advancement, or other ‘wicked problems’ associated with technological proliferation, considered here in terms of six primary technological threats facing humankind.

Second, although seemingly alarmist, the potential threats of technological proliferation require consideration for those seeking to develop useful future scenarios for how technology may impact society going forward. Technological advancement has seemingly considerable potential for technological development and proliferation to contribute to the betterment of society through elevating the health and wellbeing of populations (and through offering outright solutions to catastrophic problems). However, it also has the potential for catastrophic consequences of technological development itself and the proliferation of harmful outcomes. Effective theory needs to be developed that can help stakeholders capture benefits while mitigating risk. This paper therefore poses an important question, namely what are the key theoretical dimensions of this problem? This problem is defined as the need to manage the seemingly unlimited opportunities of technological change while at the same time also managing the proliferation of dangerous technological applications.

In light of this problem, and the lack of sufficient knowledge of how to address the tensions between potential benefits and risk (and the failure of humankind to solve many of its present contemporary problems), this paper seeks to present an argument that two aspects of technological change account for most of the variance associated with this problem. It is also argued that a focus on the management of these two aspects, or dimensions, can provide important guidelines for the development of a discovery system that is relatively more robust to threats posed by both innovation failure as well as from dangerous technology proliferation.

These two dimensions relate to the urgent need (i) to manage the openness of the discovery system itself, and (ii) to manage the power relationships associated with a rapidly developing system of technological innovation. Sardar’s (2010:183) second law states that future studies are characterized by Mutually Assured Diversity (MAD), whereby MAD “is the proposition that full preservation of our humanity requires that [human diversity] is assured, that it not only survives but thrives in any desired future, and that future generations mutually recognise and appreciate each others’ diversity.” It is argued here that key to the survival of humans in the face of technological proliferation, following Nielsen’s (2012) theory, is the need for openness in the discovery process itself, whereby openness is a necessary condition for the knowledge creation necessary for the co-evolution of our capabilities to manage the accelerating rate of emergent technologies. Indeed, according to Tegmark (2017:89), the challenges of technological advancement “transcend all traditional boundaries- both between specialties and between nations.”

Sardar’s (2010:183) third law relates to how futures studies are sceptical, and “the reality of things is inaccessible to the human mind and certitude is impossible to attain.” Given the uncertainties associated with technological proliferation, in developing theoretical insights it is necessary to take recourse to theory that predicts how diversity in human action has seemingly been checked by power throughout human history (see Foucault, 1982). Thus, to develop a theoretical frame that is sceptical of truth claims, the initial structure developed here focuses first on a key defining issue of the knowledge age, namely the radical changes in the discovery process on account of the democratisation of knowledge (Callaghan, 2016). These changes seem to be fundamentally enabled by the increasing openness of the knowledge creation process itself. Second, the analysis focuses on the way that power relationships have seemed to pose a fundamental threat to the diversity of views in human history (Foucault, 1982). In this way, analysis is related to primary technological threats.

Third, in light of Sardar’s (2010:184) fourth law, that future studies are futureless, and the influence of futures studies is on thought and scientific behaviour in the present, this paper seeks to “change peoples’ perceptions, make them aware of dangers and opportunities ahead” and to “galvanise them into collective social action” to provoke discussion and consideration of important and uncomfortable scenarios that will unfold on account of our actions in the present.

Having introduced the paper, and having provided a brief overview of its key arguments, literature is now reviewed, and certain key technology threats are identified. After this, interrelationships between these threats are discussed. A theoretical framework is then developed, and four modes of discovery are derived from the use of two propositions as heuristics. Certain issues drawn from the literature are then discussed in relation to each of these modes of discovery. The paper concludes with a discussion of these modes.

Go to:

2. Theory and technological threats

The industrial and digital (information) revolution has had far-reaching impacts on “practically all aspects of our society, life, firms and employment” (Makridakis, 2017:46). As stressed previously, these effects are expected to intensify over time. For Makridakis (2017:50), those predicting different influences of technology on society can be differentiated on the basis of their optimism or pessimism:

[O]ptimists predict a “science fiction,” utopian future with Genetics, Nanotechnology and Robotics (GNR) revolutionizing everything, allowing humans to harness the speed, memory capacities and knowledge sharing ability of computers and our brain being directly connected to the cloud. Genetics would enable changing our genes to avoid disease and slow down, or even reverse aging, thus extending our life span considerably and perhaps eventually achieving immortality. Nanotechnology, using 3D printers, would enable us to create virtually any physical product from information and inexpensive materials bringing an unlimited creation of wealth. Finally robots would be doing all the actual work, leaving humans with the choice of spending their time performing activities of their choice or working, when they want, at jobs that interest them.

<<<TEXT CONDENSED NONE OMITTED>>>

Pessimists, however, offer a more dystopian view while others are more nuanced in the way they frame opportunities or threats associated with technological advancement. With regard to AI, Tegmark (2017) differentiates between different schools of thought on the basis of when they consider AI to ultimately be able to surpass human levels of intelligence, and whether they consider this ‘superhuman’ AI to be a threat or not. This event has been termed the ‘singularity.’ 2.1. The control threat According to Vinge (1993:1), if the technological singularity “can happen, it will,” and we should therefore primarily be concerned about issues of control. The challenge of losing control of technological advancement is termed the ‘control threat.’ Longstanding debates in the literature seem to suggest that this control threat warrants consideration as a primary technological threat. Vinge (1993:1) defines the singularity as the “imminent creation by technology of entities with greater than human intelligence”. Bostrom (2017:300) refers to the singlularity as an “intelligence explosion,” whereby “we humans are like small children playing with a bomb;” such “is the mismatch between the power of our plaything and the immaturity of our conduct.” Here, issues of human agency are central to the dangers inherent in this loss of control. Loss of control of dangerous technologies, and the need for responsible innovation (Grunwald, 2011; Stilgoe et al., 2013) are therefore key to discussions of the control threat. For Bostrom (2017:300), in light of the threat of an intelligence explosion the “most appropriate attitude may be a bitter determination to be as competent as we can, much as if we were preparing for a difficult exam that will either realize our dreams or obliterate them.” In light of increasing uncertainty related to technological advancement, it is necessary to consider all voices, however abhorrent, or diversity of opinion, in “all of its mindboggling forms” (Sardar, 2010:183). According to Kaczynski (1996:1), as “society and the problems that face it become more and more complex and machines become more and more intelligent, people will let machines make more of their decisions for them, simply because machine-made decisions will bring better results than man-made ones.” Kaczynski (1996) argues that the increasing complexity associated with decision making will then ultimately outstrip the capacity of humans to make them, and human control of decisions would be lost. The control threat therefore relates to the problem that humankind will simply not be able to manage these changes, primarily due to increasing complexity. Tegmark (2017) argues that many in the scientific community subscribe to the beneficial AI movement, whereby survival of a technological future depends on threats being identified now, and safety research so as to ensure that technological development will be beneficial to humankind. Central to this argument is the notion that the competence of the management of technological advancement will determine how beneficial is this advancement. Accordingly, binary conceptions of technological advancement as either utopian or dystopian might yield unhelpful and erroneous heuristics. Tegmark (2017) stresses that discussions of malevolent machines are a red herring, and that the control issue reduces to one of competence, whereby the true danger would exist if the goals of a competent AI, for example, diverged from ours. It is argued here that technological advances need to be harnessed to enable human collaborative problem solving, or the capabilities associated with human collective intelligence so as to address this control threat, and to ensure that human management of technology is not eclipsed by machine intelligence or outpaced simply by the increasing complexity of the management challenge itself. 2.2. The power inequality threat There are also challenges that might arise due to this loss of control. Certain of these challenges relate to how power can influence technological change, and in turn societies, and particularly how elites might behave if their power is unchecked. The loss of control of dangerous technologies to powerful elites, be they business, political or otherwise, can also have important consequences, just as it would if control is lost to researchers, absent principles of responsible innovation. Having the management of dangerous complex technologies subject to market profitability logics can be problematic, as shown in research on nuclear technology (Osborne & Jackson, 1988). Losing control of dangerous technologies to market power, or exposing them to the vagaries of executive risk seeking are therefore yet other dimension of the control threat. The question naturally arises as to who should then be in control of technological advancement? According to Olsen, Kruke, and Hovden, (2007:69), societal safety is “a sensitive political issue containing dilemmas and value choices that are hardly possible to perceive or solve as pure scientific problems.” Central to such a perspective is the role of power in how such issues come to be managed. Given multiple perspectives and stakeholder interests, the full reality of things as they relate to dangerous technologies may indeed be impossible to perceive accurately (Sardar, 2010:183). The dispersion of power, including that associated with knowledge, may ultimately be key to, at the very least, ensuring societal scrutiny of the issues associated with dangerous technologies. Power is of central importance in understanding the consequences of human behaviour (Foucault, 1982). There are therefore also perhaps increased dangers posed by elites if human control is maintained, and if technological change increases this power. According to one dystopian perspective, human work may ultimately no longer necessary, and “the masses will be superfluous, a useless burden on the system” (Kaczynski, 1996:1). The following passage is from Kaczynski (1996:1): If the elite is ruthless they may simply decide to exterminate the mass of humanity. If they are humane they may use propaganda or other psychological or biological techniques to reduce the birth rate until the mass of humanity becomes extinct, leaving the world to the elite. Or, if the elite consists of soft-hearted liberals, they may decide to play the role of good shepherds to the rest of the human race. They will see to it that everyone’s physical needs are satisfied…Of course, life will be so purposeless that people will have to be biologically or psychologically engineered either to remove their need for power processes or make them “sublimate” their drive for power into some harmless hobby. These engineered human beings may be happy in such a society, but they will most certainly not be free. The will have been reduced to the status of domestic animals. The notion of the need to biologically or psychologically engineer away the need for power differs radically from Foucault’s (1982) approach. To consider future technological scenarios, however, one cannot shy away from unpleasant narratives, no matter how ‘wicked’ or seemingly impossible to solve (Sardar, 2010), and this is a case in point. Joy (2000) also cites this paragraph, acknowledging Kaczynski’s role as a contemporary Luddite, and stressing the need to confront the arguments made therein. The problem of power inequality that may result from technological change is termed here the power inequality threat. Technologies such as human reproductive cloning and inheritable genome modification (Bostrom, 2017; Isasi & Knoppers, 2015) also raise important issues about inequality that may result of genetic engineering, including concerns about eugenics (Gilding, 2002). There is clearly a need to address these issues, but in a different way than Kaczynski, who as the Unabomber sought to physically attack (bomb) those involved in the advancement of technology, including those in universities (Joy, 2000). Indeed, the power inequality threat might not extend simply to elites, and the power to unleash destructive forces would perhaps give those (ironically, such as Kaczynski himself) who seek destruction (for any reason) the means to do so, and ultimately, perhaps, the ability to threaten humankind itself. 2.3. The destructive empowerment threat Advances in technology can alter the balance of power between nations and between individuals and security institutions. Tegmark (2017:107) stresses that “those who stand to gain most from an arms race aren’t superpowers but small rogue states and non-state actors such as terrorists, who gain access to the black market once they’ve been developed.” Cyberwar, and its potential to disable of critical infrastructure, will become increasingly likely between belligerent states as technology advances. Joy (2000:1) stresses that genetic engineering “gives the power- whether militarily, accidentally, or in a deliberate terrorist act- to create a White Plague.” According to Joy (2000:1), nanotechnology has “clear military and terrorist uses, and you need not be suicidal to release a massively destructive nanotechnological device- such devices can be built to be selectively destructive, affecting, for example, only a certain geographical area or a group of people who are genetically distinct.” Further, replicating nanotechnology can be dangerous on its own, for example in the way they can crowd out other life, and this could result from a single laboratory accident (Joy, 2000). Given a world population of billions of individuals, under the assumption that there are very few at the far end of a normal distribution of harmful intentions, there would perhaps always be some that would meet Joy’s definition of ‘extreme individuals.’ With the proliferation of increasingly dangerous technologies, and increasing access to them, these scenarios are perhaps increasingly likely. This threat is considered here the destructive empowerment threat. Another example offered by Kurzweil (1999) relates to the dangers of nanotechnology, which might be more dangerous than nuclear, as its consequences are not localised, but can spread. According to Kurzweil (1999:160), once “the basic technology is available, it would not be difficult to adapt it as an instrument of war or terrorism.” 2.4. Intrinsic challenge displacement threat Issues related to the loss of purpose, such as those suggested by Kaczynski (1996), that might be experienced by those who’s need to work has been displaced by more effective and efficient technologies is considered the intrinsic challenge displacement threat. Indeed, the psychological impact of such radical technological change is uncertain, and consideration of this issue seems warranted. For example, Tegmark (2017:89) stresses the need to “grow our prosperity without leaving people lacking income or purpose.” This threat is also taken to relate to how technology might replace humans in the workplace, for example in ways that result in ‘technological unemployment amidst digital plenty.’ Brynjolfsson refers to this future scenario as a ‘digital Athens,’ whereby in the same way that Athenian citizens lived lives of leisure with labour performed by captives, a highly productive and automated economy might free up human labour without reducing living standards (in Regalado, 2012). The reality, however, may be very different. Some have argued that we are entering an era of machine intelligence that ultimately heralds the end of human employability (Brynjolfsson and McAfee, 2011:9). Rifkin (2011) argues that the world is experiencing a third industrial revolution related to computer power (following the first, associated with steam power, and the second, related to the rise of oil and electricity as sources of power). According to Rifkin, machines are increasingly displacing human jobs and making blue collar work obsolete, giving rise to a ‘silicon-collar’ workforce, of machines that have replaced humans in the work place. An ‘end of work’ era of worker economic irrelevance, and extensive joblessness might give rise to problems like rising levels of crime and feelings of irrelevance and alienation (Rifkin, 1995). Closely related to this threat is that of resource competition, in the form of competition from machines for human jobs. 2.5. The resource competition threat The resource competition threat can become an unintended societal consequence of technological advancement. Joy (2000) points to how the design and use of technology has resulted in unintended consequences, such as the overuse of antibiotics which has given rise to antibiotic resistance, and drug resistant genes in Malaria parasites. Drawing from Moravec’s (1999) work, he also points to an argument that biological species “almost never survive encounters with superior competitors,” suggesting further that in a “completely free marketplace, superior robots” would outcompete humans, as robotic industries “would compete vigorously among themselves for matter, energy, and space, incidentally driving their price beyond human reach,” whereby humans, unable to afford the necessities of life, could be “squeezed out of existence” (Joy, 2000:1). This type of problem relates to the problem of crowding out, primarily related to resources, or the resource competition threat. Brynjolfsson and McAfee (2014) point to how technological change may ultimately displace jobs. Advances in technologies are creating an unprecedented reallocation of wealth and income. However, whereas wages have increased alongside productivity for the previous two centuries, median wages have recently stopped tracking productivity, with important societal implications. Brynjolfsson and McAfee (2014) stress that on account of these changes, technological change has altered certain structural economic relationships and that this is in turn driving rapidly-increasing inequality. A small elite are therefore benefitting from growth in GDP and productivity but the median income is diverging from the mean. The main driver of this increasing inequality is therefore exponential, digital and combinatory technological change driven by the new economics associated with near zero marginal cost (Rifkin, 2014) which is creating ‘winner takes all’ markets where leading providers (with a fraction of traditional employment costs) can capture most of a market through digitization. Brynjolfsson and McAfee (2014) argue that it is no longer true that a rising tide of technical progress will ‘lift all boats’ because technology acts as a multiplier, in that while it produces more with limited inputs, it also substitutes for workers in lower-skilled work, increasing returns to high skilled work. This then results in skill-based technical change associated with increasing inequality. Technology is also shifting the returns to physical capital versus labour. The corporate profit share of GDP has now surpassed that of the wage share, bolstered by winner-take-all markets enabled by the low marginal costs of digital goods and their low capacity constraints which allow substantial economies of scale (Brynjolfsson & McAfee, 2014). Others however have argued that previous dystopian predictions of devastating technological unemployment have in every historical case failed to materialise. According to Tegmark (2017) it may be different this time, as those arguing this might not have considered what will occur when machine intelligence begins to perform the creative work at which humans currently outperform machines. 2.6. The reproduction threat According to Joy (2000:1), robots, engineered organisms, and nanobots differ from all previous technologies, as they share the ability to self-replicate, and with this will necessarily come the risk of substantial damage to the physical world; with each of these technologies, a “sequence of small, individually sensible advances leads to an accumulation of great power and, concomitantly, great danger.” There is therefore an issue related to the power of humans versus machines, and the dimensions along which these differences in power can result in different scenarios for society. Although this also derives from the control threat, this threat is considered more specifically to relate to the reproduction of technology, and the potential for exponential increases in the harmful effects of technology. This threat is defined here as the reproduction threat. The unchecked reproduction of nanomachines is a particular concern, according to Kurzweil (1999:158), due to the fact that to “be effective, nanometer-sized machines need to come in the trillions” and the “only way to achieve this economically is through combinatory explosion: let the machines build themselves.” He points to the risk of an exponentially exploding nanomachine population, and the risk of even minor software problems that fail to halt self-replication. The same is true for other technologies such as biotechnology. According to Kurzweil (1999:176), we are “very close to the point where the knowledge and equipment in a typical graduate-school biotechnology program will be sufficient to create self-replicating pathogens.” A theoretical model that seeks to contribute useful insights regarding the management of these threats therefore also needs to take into account the amount of time available to do this. 2.7. Scenario timelines How long do we have until we can no longer manage technological advancement and its proliferation? According to Tegmark (2017), two conferences of AI researchers have collectively estimated that human-level artificial general intelligence will be created by the year 2055 (the first conference) and 2047 (the second, two years later). With regard to when dangerous technologies may no longer be manageable, Joy’s perspectives are included here as being representative of many dystopian commentaries. Unlike the 20th century weapons of mass destruction, GNR technologies are being rapidly developed commercially by corporate enterprises, and their promises are being aggressively pursued; according to Joy (2000:1), this “is the first moment in the history of our planet when any species, by its own voluntary actions, has become a danger to itself- as well as to vast numbers of others.” Although it is simply not known when the machine intelligence ‘explosion’ will occur, other dangerous technologies are currently also developing; therefore, in terms of timelines, technology safety research is an immediate imperative according to Bostrom (2017) and Tegmark (2017). This is also the argument of Vinge (1993:1), whereby if the technological singularity “can happen, it will;” we should therefore primarily be concerned about issues of researching how this can be prepared for. Our scientific research capacity is therefore key to our ability to undertake effective technology safety research, and to our ability to manage dangerous technologies so as to not lose control of them. Recent literature suggests that timelines in scientific research production may be about to shorten. In terms of the research capacity needed to support technology safety research, there is another perspective that argues that the human capability to research, and therefore to manage, a threatening context is about to be radically enhanced. Whereas the literature to date seems to have given much attention to scenarios similar to Joy’s (2000) commentary, lacking in these debates seems to be the argument that science itself is on the cusp of a reorganization, which some have termed the ‘reinvention of discovery.’ According to Nielsen (2012:19) the “reinvention of discovery is one of the great changes of our time” whereby to “historians looking back a hundred years from now, there will be two eras of science: pre-networked science, and network science” as we are currently “experiencing a time of transition to the second era of science.” The theoretical model to be introduced in the following sections will draw on this body of literature [networked science] that suggests the emergence of a new dawn of knowledge creation, whereby novel technological developments make it possible to take advantage of hitherto unimagined economies of scale in both (big) data collection and, importantly, analysis. If the threats posed by technology are real, and if relinquishing science (as advocated by Joy) will not solve the problem of technological proliferation, and only exacerbate the current inequalities in innovation outcomes, then there is perhaps only one other alternative, namely to enhance our ability to manage it. According to networked science theory, this might be possible. For the purposes of this work, the opposite of relinquishment is taken to be uptake of open modes of science, or open innovation. Other alternatives arguably fall into these two categories, or along a continuum between the two. At this nexus, interrelationships between the technological threats discussed above are discussed, and then ‘real life’ examples are considered that specifically relate to the complexities associated with the management of dangerous technologies. Go to: 3. Interrelationships between the technological threats To develop a theoretical model of technological threats and their potential impact on society it is first necessary to relate these threats and to derive underlying regularities between them as the basis for a problem solving response that can to some extent address aspects of them all. Before discussing this model it is then necessary to consider existing proposed solutions to the threats discussed above. The only realistic alternative [to the dangers of technological advancement], according to Joy (2000:1) is “relinquishment: to limit the development of the technologies that are too dangerous, by limiting our pursuit of certain kinds of knowledge.” Tegmark (2017:169) acknowledges that although the term Luddite is typically used as a derogatory epithet for those who are perceived as technophobes “on the wrong side of history,” the notion of relinquishment has nonetheless found “new support” in the environmental and anti-globalisation movements. This is an important argument, because to revalue knowledge as either ‘bad’ or ‘good’ would seemingly have important implications for almost every aspect of human life. Whereas six broad threats were derived in the previous sections, the notion that knowledge can be harmful is an important proposition, and this idea (relinquishment) as a proposed solution requires interrogation according to Joy’s criteria of potential harm, using these same six categories of threat as heuristics to sharpen the discussion. The nature of innovation itself, and how it works through different channels, is not independent of these issues, and its channel of open innovation presents exponential advantages for knowledge creation, whether from its ability to harness very large volumes of data, or to harness large scale data analysis, or problem solving opportunities (Callaghan, 2016). This potential has been described in terms of the contributions of collective intelligence (Bernstein, Klein, & Malone, 2012) and networked science (Nielsen, 2012), which share a focus on the opportunities offered by open science, and open systems of innovation. In contrast, Joy (2000:1) argues further that “despite the strong historical precedents, if open access to and unlimited development of knowledge henceforth puts us all in danger of extinction, then common sense demands that we reexamine even these basic, long-held beliefs.” One therefore has to question how realistic Joy’s solution (relinquishment) really is, to what is clearly a wicked problem according to Sardar’s (2010) futures studies approach. There are arguably three potential problems with the relinquishment approach. Firstly, if open access gives way to closed access, and some (not all) relinquish knowledge, but not others, then who would access to this knowledge be limited to? The power inequality threat is not addressed by shutting down open access. Control of technology might then shift to elites, and as we surely know by now, state powers to limit or suppress activity can be captured by powerful interests, essentially creating the conditions for a monopoly in knowledge. Thus, secondly, the control threat brings with it similar problems as those associated with losing control of technology to intelligent machines. Thirdly, there is the problem posed by slow response to threats under closed systems of innovation (discovery). Under closed systems, destructive empowerment threats would perhaps be more difficult to defend against, without the social and institutional infrastructure that open systems of innovation are rapidly developing. Open systems of knowledge creation have demonstrated their increasing effectiveness in enabling timely disaster response (Callaghan, 2016). To contextualise certain of these ideas prior to discussions of the theoretical model it is necessary to first identify certain real life examples of the tension between openness and closure in decisions about dangerous technologies. 3.1. Contextualising the reality of technological threats Certain decision making issues related to dangerous technologies have already been considered in real world contexts. Dual-use research of concern (DURC) is perhaps a useful example of how issues related to the control, power inequality, and destructive empowerment threats have been considered to date. DURC research is “research that, based on current understanding, can be reasonably anticipated to provide knowledge, products, or technologies that could be directly misapplied by others to pose a threat to public health and safety, agricultural crops and other plants, animals, the environment, material, or national security” (NIH, 2017:1). An example of DURC research can be found in the debates concerning the publication of two papers revealing how to genetically engineer strains of the H5N1 avian influenza virus (Resnik, 2013). Those arguing against publication have cited concerns, particularly since 2001, about the use of this knowledge by terrorists, or others with destructive motives (Resnik, 2013), to create a bioweapon and set loose a global pandemic (Cohen & Malakoff, 2012). According to Specter (2012:1), the decision to allow publication of this knowledge “fundamentally altered the scope of the biological sciences.” The U.S. National Institutes of Health in 2011 initially recommended redaction of these papers, but after careful consideration the National Science Advisory Board for Biosecurity (NSABB) recommended (notwithstanding a lack of unanimity) they “should be made public, in full,” as the potential public health benefits were considered to outweigh the potential harm (Cohen and Malakoff, 2012:19). This decision (creating an important precedent) was therefore taken in support of openness rather than closure, notwithstanding the destructive empowerment threat. Over and above the issue of freely available dangerous information, accidental release of pathogens is another threat that is known to occur, if not regularly, but then often enough to warrant consideration here. Evidence-based examples of fatalities exist in the form of research-related accidental release of smallpox, severe acute respiratory syndrome (SARS), and Ebola pathogens (Specter, 2012:1). Prior to the NSABB decision, in 2002, someone had already “stitched together hundreds of DNA fragments, mostly acquired via the Internet, then used them to create a fully functional polio virus,” and in 2005 academic papers published the genomic sequence of the 1918 Spanish flu, but these have both (notwithstanding much initial condemnation) ultimately been considered valuable contributions to knowledge (Specter, 2012:1). The threat of biological terror seems real, as even Al Qaeda have called for its supporters with degrees in microbiology or chemistry to develop a weapon of mass destruction (Specter, 2012). This threat is of great concern, given proof of concept of how relatively easy it has been to reconstitute an extinct poxvirus, costing approximately $100 000 using only mail-order DNA (see Kupferschmidt, 2017). Some also argue that such knowledge can help those developing vaccines or drugs to know if these are effective. Additionally, the scientific method and “the entire edifice of institutional research depends on such openness; without it, progress would slow dramatically” (Specter, 2012:1). However, unlike the all-or-nothing decisions to research or produce pandemic strains of pathogens, the threats of GNR technologies are unclear, even as they currently proliferate, mostly behind the closed doors of commercial enterprises. At the same time, the artificial intelligence (AI) revolution will bring extensive changes to all aspects of society and life, and additionally to firms and employment, “resulting in richly interconnected organizations with decision making based on the analysis and exploitation of “big” data and intensified, global competition among firms” (Makridakis, 2017:46). Previous research has sought to make sense of the complexity of human engagement with technology through the use of metaphors to describe technological futures. This literature is now also briefly considered here to contextualise discussions in the above sections. 3.2. Metaphors as a heuristic for understanding technological threats The evolution of technology and its core threats can be taken to be reflected in the metaphors people use when considering technological futures. Metaphors used by stakeholders reflect the evolution of technologies, as for the past two centuries the ‘technology is good’ metaphor has persisted, related to improvements in productivity; this metaphor has also been associated with another, namely that ‘more is good’ (Carbonell, Sánchez-Esguevillas, & Carro, 2016). Joy’s (2000) perspective might be read as a metaphor, that ‘technology is dangerous,’ conflicting with the metaphor that ‘technology will solve our problems.’ Drawing directly from this is the binary conflict between the metaphors ‘technology should be relinquished,’ and therefore that ‘closed models of development are best’ versus ‘technology should be shared, and open models are best to keep us safe.’ However, the danger here is that subscribing to these metaphors simply puts us at risk of creating unhelpful binaries. It goes without saying that there are always graduations between these extremes, and Carbonell et al.’s (2016) use of technology metaphors are useful in order to simplify explanations of conflicts between different perspectives. These metaphors are then useful as heuristics, in that they can be related to the six technology threats, encouraging dialectical tensions that give rise to a more considered discussion of scenarios. On the basis of these conceptions, the answering metaphor is perhaps that ‘technological dangers can be successfully managed,’ juxtaposed against its counterpoint ‘technological dangers cannot be successfully managed.’ The relinquishment argument of Joy (2000) might needlessly echo historical Luddite arguments if there are no other options with which to frame our response to technological dangers. We might have no other choice but to embrace open systems of discovery in order to improve our ability to manage technology, with the hope that improved systems of discovery will ultimately be key to the successful management of technological change itself. Historical Luddite protests associated with the metaphor ‘the job is up’ rather than ‘technology is up’ offer an early example of debates about the trade-offs some argue are to be made when technology advances (Carbonell et al., 2016). This is perhaps an example of the resource competition problem and the threat posed by technology to resources in the form of jobs. Other examples include those of religious groups that have also advanced metaphors conflicting with technology, as reflected in longstanding historical tensions between science and religion. Similarly, there are now tensions between societal values like equality, respect or privacy (reflected in concerns about the digital divide, harassment and other outcomes) and the capacities the Internet now offers (Carbonell et al., 2016). These tensions can perhaps be related to the control threat, as individuals face losing control over privacy, over the continuity of their lifestyles, or as societies lose control over widening inequality on account of the digital divide. The latter issue also relates to the power inequality threat. If the cat is out of the bag already (as the example of the publication of the H5N1 papers shows), and relinquishment may no longer be a useful strategy (as countries and individuals differ in their moral propensity to develop and use dangerous technologies (Bostrom, 2017)), the only way out may be to radically increase our ability, as humans, to collectively manage these threats. As it stands, it is unlikely however that we have this capacity at present or will be able to develop it quickly. How then, could this capacity be developed? And what future scenarios would result from failure to successfully manage these challenges? Alternatively, what future scenarios would result if such successful management of technological proliferation were possible? Successful management is defined here as effective research and knowledge creation that enables the threats of technological development to be mitigated in a sustainable way (Bostrom, 2017; Tegmark, 2017), and which results in a relatively equitable distribution of the outcomes of discovery (Rifkin, 1995). One would need to ask, however, what is the role of the state in such successful management? Other metaphors relevant to debates about the impact of technology on society are those related to the tensions between ‘big brother dystopia’ versus ‘state as protector,’ and ‘equality is up’ versus ‘market is up’ (Carbonell et al., 2016). The enhanced surveillance abilities of the state might also lend themselves to a change in power dynamics, and an increase in power inequality, as this power might be part of the trade-off for safety in an era in which public gatherings, for example, are increasingly vulnerable to attack. Indeed, the same technological advances can also enable destructive empowerment, as individuals can use technology to amplify the damage they can cause. The key then, might be to therefore consider such management according to the principles of openness, and the democratisation of science, and its attendant ethical framework. According to the principles of maximum transparency and accountability, power inequality is reduced, and power over knowledge is made to be more equitable. In this way, inequality in the outcomes of knowledge is also reduced, maximising benefits to all affected by science as well as the problems it is tasked to solve. A key feature of the theoretical model proposed here to offer certain insights into the societal impact of technology is therefore open knowledge creation, and an ethical framework that is fundamentally suited to open systems of innovation and discovery. Given scarce resources (including time), however, it is unclear as to which of the six threats require more urgent attention than others. What then are the relationships between these threats? 3.3. Ordering relationships between threats Fig. 1 shows a possible ordering of technological threats. As discussed above, these threats reflect primary technological concerns in the technology futures literature. <<<FIGURE OMITTED>>> Criteria for inclusion was based on the perceived relative seriousness of a threat. Threats were not considered for discussion on their own if they fell within another of these categories, other than the control threat category. This iterative and inductive process of review resulted in the six categories included here. A brief sketch is now provided, of how these threats might relate to each other. Technological futures are by definition uncertain, and the relationships discussed here are necessarily speculative, but such a discussion is necessary in order to draw out an ordering of these threats and to better understand which are more urgent. If the control threat is considered the ‘origin’ of the other threats considered here, then the management of this threat would require an immediate and proactive response. This threat is therefore ‘immediately urgent’ while those that derive from it are ‘urgent,’ in that the control threat would need to be considered together with the others. Such an ordering might have important implications for which societal stakeholders should be more involved in managing these threats. If the control threat is considered the dominant threat, then this places technology safety researchers at the source of the problem of managing dangerous technologies. Indeed, if technological development to date has typically followed the trial-and-error model, then we will “inevitably reach the point where even a single accident could be devastating enough to outweigh all benefits” (Tegmark, 2017:90). Having private or corporate stakeholders drive the technological development process without the engagement of independent research stakeholder groups may no longer be safe in an era that transcends trial and error approaches to dangerous technologies. According to the logics described in Fig. 1, the control threat, or losing control of the management of technology, can therefore lead to other threats. We are now perhaps in an era in which the consequences of practice-based trial and error make it necessary to elevate the status of technology safety researchers. This is an immediately urgent imperative. Certain research findings seem to support the necessity of a change in societal stakeholder relationships (as they relate to dangerous technologies) to include technology safety researchers. Insights from the use of nuclear power suggest that certain risks can arise from organisational structures of large corporate organizations themselves. Fewer coordinative mechanisms between functional departments, more levels of administration, centralization and higher numbers of employees may constrain an organization’s ability to respond to safety issues (Osborne & Jackson, 1988). Risk preferences are also not constant over different types of decision making (Osborne & Jackson, 1988). Indeed, under conditions of growing losses, decisions are typically more risky than they are under conditions of gains (Kahneman & Tversky, 1979). Osborne and Jackson (1988:930) therefore suggest that the proportion of a utility’s investment in a dangerous technology like nuclear power “partially reflects the technological risk preferences of its senior executives.” Developing and empowering technology safety researchers as an important stakeholder group may therefore be an urgent need, so as to ensure that the post-trial-and-error paradigm is safely managed through more inclusive engagement going forward. To manage the control threat it may be important to therefore shift the locus of power related to decision making about dangerous technologies from corporate or other interests to proactively include technology safety researchers. This dispersion of power might however be at odds with historical practice and the autarky of corporate R&D. The management of, and decisions about, dangerous technologies require openness and societal inclusion, according to the principles of responsible innovation (Grunwald, 2011; Stilgoe et al., 2013). According to Douglas (2000:559), “value-free science is inadequate science; the reasoning is flawed and incomplete.” The task of managing the control problem cannot therefore simply be left to corporate market incentives, or even to science on its own. Thus, the threat of losing control of technology, whether to human elites or to machine intelligence, is perhaps the most important of the threats, and is considered immediately urgent, requiring inclusive engagement across society. This implies openness and power dispersion to mitigate against loss of control of dangerous technology to any set of particular interest groups. Given the centrality of the control threat, what then of the relationships between the others described in Fig. 1? The control threat, if unsuccessfully managed, may contribute to the power inequality, resource competition and destructive empowerment threats. The resource competition threat may in turn contribute to increased power inequality through two channels. The first is arguably the way digitisation is creating a winner-takes-all economy (Brynjolfsson & McAfee, 2014) as the new economics of near zero marginal costs (Rifkin, 2014) allow a few producers to capture substantial market shares with very few human workers. Another channel might be through the erosion of jobs that fall below the ‘waterline’ of advancing machine intelligence, empowering a class of workers in areas that machines have not yet mastered (Tegmark, 2017). If technology creates or exacerbates such class differentials, and if these classes are able to prioritize their own interests at the expense of others, they might seek solutions associated with power inequality. Intrinsic displacement might be considered a derivate threat, arising from a lack of purpose in a world in which machines do most forms of work, or (alternatively) a state of powerlessness as people are excluded from meaningful opportunities by a technologically-enabled human elite. Thus, the four threats of control, resource competition, power inequality and intrinsic displacement might benefit from further research that considers their potential interdependencies. The destructive empowerment threat is an ever-present one, as advancing technology necessarily provides more options for both individuals and states to pursue destructive goals. Although the link is not shown in Fig. 1, power inequality can result in destructive empowerment if elites or elite countries use these technologies in war. A global power hierarchy held in place by technology might be such an outcome. Key to managing the relationships between these threats, however, seems to be the need for proactive engagement with technology safety researchers, and the use of technology to improve our research capabilities and, thereby, our ability to manage technological change. Thus an order of importance seems to exist amongst these threats. A focus of resources and attention on the control threat without neglecting relationships between these threats is important. This approach seems to also find support in the literature. Bostrom (2017) and Tegmark (2017) stress the urgency of technological safely research, to be able to control the trajectory of technological change, and ensure its ‘beneficial’ use and contribution to human society. Their arguments capture the essence of what the management of the control problem entails. If managing the control threat is key to the management of the others, what then are the principles most likely to empower this control, or management of technology? The theoretical model that follows draws from novel ideas and theory that suggest certain principles that might be useful in this task. Go to: 4. Theoretical model According to Tarko and Aligica (2011:987), Kahn’s conceptualisation of the ‘institutionalisation of interdisciplinarity’ is reflected in “a phenomenon that has as its core a process-based approach to knowledge and method aggregation” reflected in novel developments enabled by web-based techniques. According to Tarko and Aligica (2011)), terms associated with this phenomenon include Wikinomics (Tapscott & Williams, 2006), the wisdom of crowds (Surowiecki, 2004), the “army of Davids” (Reynolds, 2006), and “collective intelligence” (Malone, Laubacher, & Dellarocas, 2009). Malone et al. (2009:2) also acknowledge other additional descriptions of the phenomenon in the literature, such as radical decentralization, crowd-sourcing, and peer production, arguing that the term collective intelligence is the most useful, defined broadly as “groups of individuals doing things collectively that seem intelligent.” The phenomenon has also been defined as crowdsourced R&D, and considered in terms of its roots in the seminal knowledge aggregation problem with a view to formalising this both as a body of theory and as a new scientific methodology in its own right (Callaghan, 2016). Nielsen (2012) argues that dramatic breakthrough periods in science have typically followed changes, or improvements in the way discovery is conducted. For Nielsen (2012:19): This change is important. Improving the way science is done means speeding up the rate of all scientific discovery. It means speeding up things such as curing cancer, solving the climate change problem, launching humanity permanently into space. It means fundamental insights into the human condition, into how the universe works and what it is made of. It means discoveries we’ve not dreamt of. Over the next few years we have an astonishing opportunity to change and improve the way science is done. Rapid acceleration of the pace of scientific discovery, however, requires an ethical framework that is robust to the range of different issues that can be encountered. Over time there have been increasing calls for increased democratization of science, and for greater stakeholder involvement (Siune et al., 2009). Concerns about the role of science in society and its impacts have contributed to the rise of new research fields. These fields include risk studies, impact studies, technology assessment, [Science and Technology Studies (STS)], and applied ethics, which are increasingly integrated into research programmes (Siune et al., 2009). Governance of science and R&D processes is changing, opening up “new possibilities and opportunities for involving new actors and new types of reflection” (Grunwald, 2011:9). This literature highlights a growing movement advocating the democratization of science premised on open models of knowledge creation. The democratization of science movement stresses the increasing importance of disclosure and transparency issues not only in the contemporary bioethics field, but in broader areas of technology governance. The concept of ethical practice in this literature highlights the importance of increased transparency together with increased accountability to stakeholders. This perspective echoes the emergence of new movements, such as those associated with citizen science (CS) (Bonney et al., 2009), public participation in scientific research (PPSR) (Shirk et al., 2012), and participant-led biomedical research (PLR) (Vayena & Tasioulas, 2013), which all relate to increasing access of citizens, or populations to the research process itself. These movements are in turn related to post-normal science (Funtowicz & Ravetz, 1994), and its ethical framework premised on the need for maximized transparency and accountability. The need for the post-normal science ethical approach arose from the tensions between different perspectives of climate science, whereby only through maximized transparency could the necessary scrutiny of research findings result (Funtowicz & Ravetz, 1994). These bodies of theory extend stakeholder theory (Freeman, 1984) and may form the basis for a complementary model of ethics in science that is robust to technological change. Synthesis and integration of this literature suggests certain core ideas. The first is that the growing literature on technology governance seems to be able to provide ethical frameworks that might be sufficiently robust to support a rapid acceleration of the pace of scientific discovery. However, key to this is the need for maximised transparency and accountability, and for the full inclusion of societal stakeholders in technology governance. The second is that, as suggested by Nielsen (2012) and documented in his work on networked science, there seems to be a coming ‘revolution’ in science itself, whereby we are on the cusp of a ‘second era’ of science, in that the nature of the research process itself is changing. In fact, Nielsen’s (2012) theory is perhaps foreshadowed by prior examples of the same phenomenon in the futures literature (see Tarko & Aligica, 2011). These changes also seem to echo Sardar’s (2010) principles, in that networked science transcends disciplines (first law), incorporates maximised diversity and inclusivity (openness) across society (second law), thereby mitigating the uncertainties (third law) inherent in the interactions of human agency with technological change though the dispersion of power and the empowerment of the scientific citizen. Such changes in the processes of science, or scientific research itself may make it possible to develop the management capabilities to address the control threat. A synthesis of this literature suggests, however, that there are two necessary (but not sufficient) conditions to the successful management of the control threat, namely the need for openness as a primary mode of discovery, and the need for dispersion in the power relationships around the management of dangerous technologies. From this body of theory, the following proposition is derived. Proposition 1 The successful management of technology is fundamentally related to openness as the primary mode of discovery The need for maximized transparency and accountability is therefore taken to necessarily be related to openness, or open access to knowledge and information as a necessary condition. There is a need however to articulate the tensions between the six problems, or potential technological consequences, as knowledge of the complex interrelationships between these threats is important, given that solving one problem might exacerbate another. It is therefore necessary to construct solutions that address a substantial aspect of these problems at the same time. Given a framework that maximises accountability and transparency, ethical management of technological change may be possible. Under closed modes of discovery, relinquishment of technology might not occur, as those with more power would not have to disclose what they have not relinquished. The relinquishment option may therefore not be as effective as an open mode of discovery in addressing technological threats, as long as transparency and accountability is ensured in an open environment. However, because transparency and accountability, as well as the ethical framework related to this, is but a necessary condition, and not a sufficient condition for the effective management of technological change, it is argued that a further condition is necessary, namely the need for dispersed power relationships, whereby dominant elites do not gain control of the discovery process, resulting in inequitable access to it, and unequal access to its outcomes. It is with regard to the need for openness and for dispersed power relationships in discovery that we then need to weigh up Joy’s (2000) other alternative, namely to give up the goals of perpetual economic growth as they may be inseparable from the dangers of technological growth. Joy (2000:1) suggests that material progress and the pursuit of the power of knowledge are problematic goals, arguing that “we must find alternative outlets for our creative forces, beyond the culture of perpetual economic growth; this growth has largely been a blessing for several hundred years, and we must now choose between the pursuit of unrestricted and undirected growth through science and technology and the clear accompanying dangers.” Indeed, openness, even with its attendant ethical framework, might not on its own be enough to address this threat, but it is arguably only through control that human incentives for progress can be subdued. A better solution therefore might not be the curtailment of material progress but the mitigation of the power of knowledge associated unequal concentration. Whereby openness ensures access to information and knowledge for affected populations, what is additionally needed is a mechanism to ensure the dispersion of power, or a mechanism to address power inequality, and to address threats of control. In the seminal words of Foucault (1982:780): I would like to suggest another way to go further toward a new economy of power relations, a way which is more empirical, more directly related to our present situation, and which implies more relations between theory and practice. It consists of taking the forms of resistance against different forms of power as a starting point. To use another metaphor, it consists of using this resistance as a chemical catalyst so as to bring to light power relations, locate their position, and find out their point of application and the methods used. In order to optimise the ability of collaborative human networks to manage rapidly developing technologies, dominance of the network by any set of stakeholders needs to be kept in check, lest openness gives rise to this dominance. Whereas closed modes of discovery may favour incumbents, openness may also give rise to new, or emergent configurations of power. In light of this, Proposition 2 is offered: Proposition 2 The successful management of technology is fundamentally associated with the dispersion of power, whereby control over the research process itself (and its outcomes) is, and remains, accessible At this current time, with R&D models, and particularly healthcare discovery models, at the mercy of the need for high levels investments under conditions of uncertainty about returns to these investments, the discovery process cannot be considered to be entirely accessible, and the outcomes of such a process are therefore also unequally distributed. Pharmaceutical investments, for example, will be skewed towards wealthy markets, and poorer populations will typically be disadvantaged if there is no mechanism through which firms can obtain returns on investment without targeting only markets wealthy enough to recoup investment costs. Private firms, however, can take advantage of openness to lower their costs of R&D, but may have few incentives to do so if market power is concentrated. Arguably, at the extremes of low or high openness, and of low and high power relationships, the societal impacts of technology will be very different, and it might be in conditions of high openness and dispersed power relationships that collaborative networks of human stakeholders would have an improved ability to manage rapidly increasing technological change, and to more effectively mitigate its threats. To better ground the propositions derived here in relation to the scenarios they predict, and the scenarios associated with their opposite extremes, Fig. 2 relates the extremes associated with each proposition. The extreme states of Proposition 1, namely openness as a mode of discovery versus closed modes of discovery, are related to the extremes associated with Proposition 2, or the dispersion of power versus its opposite orientation, the intensification of power. Four modes of discovery are taken to result. These four modes are now discussed. <<<FIGURE OMITTED>>> 4.1. Innovation closure Conditions associated with closed modes of discovery and relatively high power dispersion are taken to be associated with a state of innovation closure, or a failure to made dramatic breakthroughs in important socially important areas. This is the state predicted by probabilistic innovation theory, whereby innovation failure, or gridlock persists on account of a failure to taking advantage of the exponentially increasing economies of scale in data analysis that are currently offered by technologies that already exist (Callaghan, 2016). Some have argued that in pharmaceutical innovation, for example, returns on investment have been stagnant for decades now. Although power is concentrated in markets, and innovation outcomes are inequitably distributed, the monopoly structure does not explicitly shut out new entrants, and the discovery system is considered to be open to disruption. This is broadly considered to reflect the current state of affairs. Because there is no explicit closure of the discovery process, the outcomes of discovery might be considered to be probabilistically related to investments in the discovery process. In others words, there is investment risk associated with innovation investments, but this risk can largely be quantified. Investment in innovation is not fundamentally uncertain in its outcomes. 4.2. Dystopian control Under conditions of power intensification, the resources that dictate relationships within modes of discovery, and the outcomes of the discovery process, are within the power of certain agents, typically industry incumbents, or elites, because closed modes of discovery are expected to allow for the control of knowledge, and also its outcomes. Dystopian control is taken to represent a mode of discovery associated with high power differentials and low levels of openness. Under these conditions, inequality in discovery outcomes and in access to the discovery process is maximized. The power of knowledge creation is in the hands of elites, and both human progress and the threat of technological advancement are held in check, but at great cost to disadvantaged populations who are denied the benefits of innovation. This is arguably a feasible outcome if Joy’s (2000) strategy of technological relinquishment, or abandonment were adopted, as those less committed to it would not relinquish, and in so doing might increase their power over others. 4.3. Captured future Under conditions of openness with high power relationships, it is still possible that industry incumbents, or new emergent groups might take control of the discovery process, in that openness might not be a sufficient condition for optimum effectiveness in the management of discovery. Given the efficiencies of shared knowledge, the consequences of concentrations of power in a context of openness, which facilitates the disruption of business models, are uncertain. Under conditions of such uncertainty, it may simply be not possible to calculate risk. Under the uncertainty associated with this quadrant, there might therefore be a shift in this quadrant toward any of the other three quadrants. Opportunities for Internet-based global trade in goods and services and the exploitation of “unlimited, additional benefits” may result from AI inventions, but these “vast opportunities” for trade and productivity improvements need to be considered in relation to “dangers and disadvantages in terms of increased employment and greater wealth inequalities” (Makridakis, 2017:46). These advances may conceivably result in what Kurzweil (1999) has termed singularity, where nonbiological intelligence matches that of humans, and distinctions between human, machine, real reality and virtual reality disappear. Given the high levels of uncertainty associated with this mode of discovery, these outcomes need to be carefully considered. Indeed, there might come a time where computers will choose those who serve in public office, given the poor choices humans often make in this area (Makridakis, 2017). As with the heralded advent of driverless vehicles, under conditions of openness and high power knowledge advantages that can be seized by the most powerful, technological change will be expected to accelerate, and attempts to manage it may be thwarted by powerful elites, perhaps in the form of a commercial arms race as technological advances fuel the pursuit of profitability. It is this mode that perhaps best captures the spirit of Joy’s (2000) criticism of material progress as a cause of the problem of dangerous technological advancement itself. Joy’s solution of relinquishment, however, might simply result in a shift toward dystopian control, as it is unlikely that elites will relinquish power. Under conditions of high openness coupled with dispersed power relationships, on the other hand, the mode of discovery might be uniquely suited to more effective management of societal problems, including that of dangerous technological change. 4.4. Age of effectiveness The mode of discovery associated with a high level of openness and a high level of power dispersion is termed the ‘age of effectiveness’ as it is taken to offer the conditions most likely to contribute to the effective management of technologies. Digital technologies have “rendered new opportunities for learning that transcend barriers of time and space,” and harnessing the potential for robots as social agents in synergistic human-robot learning exchanges is distinct from many descriptions of AI which relegate humans to a “secondary role in the learning community” (Bricout et al., 2017:92). What such conceptions suggest is that technological advances can be harnessed in support of human learning and human agency in a world of AI. Advances in AI learning capabilities themselves show a dramatic increase over time. Milestones in this process include the reading of handwriting digits by the neural net device (1990), vision-based navigation (1993), the development of speech (1998), and self-driving cars (2009) (Makridakis, 2017). How then could human connectedness leverage human problem solving ability to the point at which it would be up to the challenge of effectively managing the complexities and dangers inherent in technological advancement and proliferation? Monat (2017) argues that the current level of human interconnectedness is growing, but in terms of ‘connections’ or ‘synapses’ is well short of the number of these connections in the human brain. He suggests that collective intelligence is emergent, in much the same way as the connections in an individual’s brain exhibit ‘emergent’ intelligence. He offers the notion that the brute number of connections in a human brain account for an individual’s intelligence, and that if the brute number of human connections in the world matched this number of brain connections then collective human behaviour would relatively be as intelligence as an average human. Although only a useful analogy, this notion suggests that collective intelligence might offer useful opportunities to leverage emergent human intelligence in the quest to manage problems like technological change. If there are billions of people, however, why then has the world seemingly not developed more collective intelligence (currently about that of a chimpanzee, according to Monat, 2017)? He argues that this is because there are too few nodes (individuals that are connected), and there are too few connected to the Internet or news media globally, and because much information, if not biased or sensationalised, is filtered by the media. According to Monat (2017:27): A fundamental difference between humans and other animals is that humans are highly self-aware while other complex animals are less so; and simple creatures like mosquitos are not self-aware at all. Some researchers believe that self-awareness is an emergent property of a complex neural network. If this is so, then high self-awareness should appear when a neural network approaches the complexity of the human brain (∼90 billion neurons and 1014 synapses). If one takes a much broader view and considers all of humanity as a neural network, then today there are ∼7 billion individual elements, of whom ∼3 billion are interconnected via computers, smart phones, tables, and the Internet. By morphological analogy, as human interconnectivity continues to grow and strengthen, eventually humanity will approach ∼70 billion interconnected humans, at which point we will become highly self-aware as a single human super-organism. This organismal self-awareness may manifest as the elimination of wars, hunger, and strife, and as the collaboration of all individual elements working together for the greater good of humanity. The lesson that emerges from this concept is that it is human collaborations and working together that might be key to leveraging human problem solving abilities, in the form of collective intelligence. According to Nielsen (2012), innovations in the discovery process can amplify human collective intelligence. The notion that humans can only stay ahead of the threats of technology by improving their ability to learn and manage technological change is associated not with technology pessimists, but with technology pragmatists. The key argument of pragmatists is that by focusing on intelligence augmentation the dangers of AI can be managed, while “providing the means to stay ahead in the race against thinking machines and smart robots” (Makridakis, 2017:52). Some pragmatists have argued that AI technologies can be controlled using OpenAI together with regulation, as open systems that are not hidden behind proprietary doors will inherently offset risks (Peckham, 2016). High openness and high power dispersion might create the best conditions for humans to be able to manage technology, but this will necessitate taking advantage of technology itself to do this. Humans may indeed have creativity advantages over intelligent machines. According to Jankel (2015:1), artificial intelligence has “raced forward in the last few years, championed by a libertarian, tech-loving and science-driven elite,” or “transhumanists who pronounce the eventual victory of the machine over nature.” He argues, however, that the belief that human brains are computers is “rooted more in metaphor than reality,” because algorithms act according to rules, and creative human disruptive innovators typically break rules, as breakthroughs are, by their nature, unpredictable; breakthrough “creativity is fundamentally organic, not algorithmic” (Jankel 2015:1). Within the next twenty years, however, rapid developments in AI are expected to result in breakthroughs based on deep learning that reflects the way children learn. Creativity might therefore not ultimately be the exclusive domain of humankind. There is no limit to deep learning, on account of three factors, namely (i) open source software makes progress available to all and encourages the development of more powerful algorithms and cumulative learning, (ii) deep learning algorithms will use memory to apply problem solving to new contexts, and (iii) intelligence programmes will themselves write new programmes (Makridakis, 2017). According to Bricout et al. (2017:91) assistive technologies in the form of socially assistive robotics (SAR) can augment learning and action, and human-robot learning communities can develop, the success of which is contingent upon “how human users engage the networking capacity” of those communities. Thus, in the future, this level of machine intelligence might be unavoidable, and the key to successfully negotiating such an environment might be the ability we have to utilize technology to leverage human management capabilities. Some might find these ideas unpalatable, given that they draw from a literature that engages with problems that are not yet part of our everyday experience. However, the fact that certain problems are wicked (Sardar, 2010) makes it necessary to confront them, as a consideration of future scenarios can help better manage these issues in the present. The arguments considered here are considered far future arguments. Baum (2015) argues that the far future argument, that “people should confront catastrophic threats to humanity in order to improve the far future trajectory of human civilization,” is important, notwithstanding the lack of motivation many have to do so, given their overriding concern for the near future rather than the far future, and the fact that there is little likelihood that they will experience the far future. Can a technological future be a meaningful place for human life? Bricout et al. (2017:102) invoke Amartya Sen’s notion of capabilities relating to freedom, choice, and ability to act, to highlight the potential impact of vertical integration of technologies, or of a nexus future with universal accessibility in which the flow of information is unchecked. This future would give rise to “major ethical concerns of users around confidentiality, privacy and autonomy,” and therefore human capabilities (p. 102). Again, these potentialities might be a function of the extent to which technological advances can be successfully managed. Many of the changes, however, may be difficult to negotiate. An example is the effect of AI and computerisation on the nature of human work, which also requires the effective management of technological change. Using an analytic Markov chain model, Kim, Kim, and Lee (2017:6) analysed the effect of advances in big data, machine learning and robotics that have reduced human employment opportunities, concluding that “even if computerization proceeds at an uncontrollable pace and renders all previously non-susceptible jobs susceptible, a healthy portion of the future economy will consist of new jobs that permit a peaceful coexistence between humans and machines.” Kim et al. (2017) however caution that their results demonstrate that “legal and social limitations on computerization are key to ensuring an economically viable future for humanity.” Therefore controlling the crossover rate of occupations between susceptible and non-susceptible states “will help reduce the proportion of susceptible occupations in the economy (p.6).” Kim et al. (2017:8) also suggest that with regard to employment loss due to technology, the “most viable solution for long-term success, however, may be a large-scale revision of the education system, in order to better equip future employees with the skills that will be necessary in a human-machine hybrid economy.” It is argued here that an age of effectiveness is perhaps possible, as long as openness is used to increase connectivity and collaboration between humans, which might allow collective intelligence to be used to leverage human management capabilities. It is also argued that human agency is also key to this challenge, and that there are ways to meet these challenges, but these might require action in the present. A careful consideration is necessary now, to understand how the education system, for example, and other human systems, can be reconfigured to meet these future challenges. Table 1 summarises concepts derived from the discussions above, and relates certain key challenges to each of the technological scenarios, or modes of discovery identified in Fig. 2. Further discussion of these relationships is offered in this table. <<<TABLE OMITTED>>> In summary, it is argued that at high levels of openness and high power dispersion, the low concentration of power over the discovery process is expected to enable effective management of discovery. This is considered a probabilistic era as outcomes can be calculated as risk. Dispersed power relationships mitigate the control and the power inequality threats. Destructive empowerment in the form of harmful activities are more effectively managed using the enhanced response capabilities associated with openness and distributed networks of collaborators. Accelerated problem solving may result under systems of collaborative problem solving, with lower power asymmetries providing the ability to harness the economies of scale of collective problem solving. Under conditions of technological change, which might be unforeseen, and therefore difficult to predict, it is arguably this quadrant which provides the most effective response to these potential dangers. Threats related to resource competition are also perhaps more effectively managed by the approach described by this quadrant. Arguably, the self-reproduction of artificial intelligence is a state that is subject to the extent to which technological change can be managed, and it is also this quadrant that provides the most useful approach to this. Similarly, the societal changes that influence human work are also, to some extent, a function of the effectiveness of the management of these changes. Go to: 5. Limitations Certain limitations of this work need to be acknowledged. This article sought primarily to provoke further engagement with certain issues surrounding the development of dangerous technologies and their ultimate societal impact. The analysis undertaken here is however based on a critical review of literature, and therefore premised on subjective judgements of what aspects to prioritise in discussions. What insights were gained by the choices made here were necessarily at the cost of that lost by not considering other aspects. For example, anchoring the work on discussions of transparency and accountability as aspects of openness was based on the growing literature on responsible innovation, which was given priority. The choice to prioritise these perspectives was taken due to their accordance with the primary arguments of technology futures experts such as Bostrom and Tegmark. Given the need to subjectively provide an ordering, according to importance, of ideas and theory in this area, the analysis sought to draw on only what was taken to be the most salient work. In so doing, the analysis also provides insights that are at a certain level of abstraction. To cover the necessary conceptual ground it was necessary to sacrifice depth of discussion in certain areas. Future work might address these deficiencies. Consideration of the six primary technological threats also necessitated subjective decisions as to which threats to recognise as primary and which to relegate to within-threat discussions of others. Given the uncertainties associated with attempts to make sense of technological futures, further work is invited, to improve on the categorisations made here. Indeed, it is hoped that further conceptual and data-driven work will draw out more detailed causal relationships between these threats (and highlight others), and ultimately provide tests of the predictions of the theoretical framework. Go to:

<<<PARAGRAPH BREAKS RESUME>>>

6. Conclusions

Given the substantial promise of technological advancement for the improvement of human lives, and given the threat of the proliferation of dangerous technologies, the objective of this paper was to offer certain insights for how these threats could be better managed. Certain key threats associated with the future proliferation of technology were identified. A theoretical model was developed, on the basis of theoretical propositions derived from the literature. Using these propositions as a heuristic frame, four future scenarios were identified, predicting different societal outcomes for different permutations of openness and the power of elites. Under conditions of high power and low openness, it was predicted that powerful elites might control innovation at the expense of relatively less powerful populations. The current global state of discovery was considered to be categorised by a mode of low power and low openness, associated with innovation gridlock, whereby few have access to the discovery process and slow innovation, particularly healthcare innovation, results in high inequality in outcomes, as only wealthier markets attract substantial R&D investments from firms. Under conditions of high power and high openness, however, the consequences of technological advancement and proliferation were taken to be uncertain, as the discovery process might be dominated by powerful elites who have the power to either curtail innovation or enable the proliferation of dangerous technologies. It was finally argued that conditions of high openness and high power dispersion might be optimal for the development of the management capabilities required to successfully manage technological change, and that technology itself may hold the key to developing these capabilities. According to this pragmatic perspective, an important avenue for future research is how collective intelligence might be leveraged using technology, as this might offer a useful approach to keeping pace with machine intelligence and other threats associated with a technological future. Ironically, it is typically only in the face of a common threat that humans become united, and seek radically improved collaborations. Uniting now, in the present, to develop radically enhanced collaborative capabilities might be our saving grace, and it is the responsibility of future studies to lead the way.

#### A thriving gig economy is key to the sharing economy.

Lao ’18 [Marina; May 13; Professor of Law, Seton Hall University School, former Director, Office of Policy Planning, Federal Trade Commission; UC David Law Review, “Workers in the 'Gig' Economy: The Case for Extending the Antitrust Labor Exemption,” vol. 51]

The proliferation of the so-called “sharing economy” platforms in the past decade has given rise to a still small,15 but rapidly increasing, gig economy workforce.16 Though there is no universally accepted definition, the term “sharing economy” generally refers to marketplaces created by companies that take advantage of broadband internet, mobile phones, and other technologies to develop new online platform models that efficiently link potential suppliers and customers on a large scale.17 While some of these platforms facilitate the sale or leasing of assets — for example, AirBnb links homeowners interested in renting their homes or spare rooms and those seeking short-term accommodations — other platforms match those who provide personal services with customers.18 It is this latter group of sharing economy businesses — essentially a subset of the sharing economy dubbed the “gig economy” — and the work arrangements arising out of them that is the focus of this Article.

The economic benefits of the gig economy, and the sharing economy more generally, are widely acknowledged.19 They stem mostly from low transaction costs made possible by modern technological advances that enable online platforms to match participants on different sides of multisided platforms (suppliers and customers) efficiently and on a scale that was not possible before.20 The sharing economy business model also allows underutilized assets to be put to more productive use, thus reducing entry costs on the supplier side. For example, an individual who owns an underused car can relatively easily, and without incurring many fixed costs, enter the market to drive passengers at times of her own choosing to generate income.21 And the innovations introduced by platforms increase consumer welfare by improving the consumer experience and offering more options.22 In short, online platforms clearly provide significant benefits to consumers and society.

#### The sharing economy stops extinction.

Schor ’21 [Juliet B. and Steven P. Vallas; March 26; American economist and Sociology Professor at Boston College, PhD in Economics from the University of Massachusetts Amherst; Professor Emeritus of Sociology at Northeastern University in Boston, PhD in Sociology from Rutgers University; Annual Review of Sociology, “The Sharing Economy: Rhetoric and Reality,” vol. 47]

While there is ample evidence to support the foregoing pessimistic view on the sharing economy, recent developments may be creating new openings for change. As the world faces pandemic and economic catastrophe, postcapitalist discourses are increasingly compelling. In the USA, the epicenter of the tech sector, a rising tide of activism and protest on racial, climate, and economic justice is shifting the policy conversation to the left. These movements align in important ways with the aspirations of early sharing-economy advocates, who aimed to transcend conventional market principles and create an egalitarian, communal, and sustainable alternative to capitalism (Scholz & Schneider 2016, Schneider 2018). The increasingly predatory and antisocial actions of the large platforms have led to renewed efforts to chart a new direction in which sharing technologies are retained but the social and economic models of the corporate apps are transformed. These alternate structures include worker-owned platforms, cashless for-profits, and the many types of community sharing entities that have developed in the last 10 years.

The idea that has attracted the most attention is probably the platform cooperative, which retains the digital features of sharing platforms but is owned by earners. Early advocates such as Trebor Scholz, Nathan Schneider, and Janelle Orsi have been developing infrastructure and networks to support the formation of these entities (Scholz & Schneider 2016, Schneider 2018). One of the earliest, Stocksy United, a stock photography co-op, achieved financial success early and has developed an admirable record of high remuneration and satisfaction for its artists (Sulakshana et al. 2018). SMart, a European freelancers’ cooperative, has more than 35,000 members and continues to expand (Charles et al. 2020). An alternative to Airbnb called Fairbnb plans to donate revenue to the communities it operates in, but the pandemic has hampered its expansion (Foramitti et al. 2020). Smaller co-ops offering local services include a ride-hail co-op in Colorado as well as Up & Go, a New York City cleaning co-op for immigrant women. Platform cooperatives have unique challenges, such as the fact that work performed is typically individualized, which leads to unequal earnings distributions (Schor et al. 2020) or, in some cases, a globally dispersed workforce. Furthermore, successful offline cooperatives are often buoyed by preexisting forms of occupational community, such as those formed among bicycle couriers, artists, photographers, programmers, and other creative class workers, or in some cases by solidary ties among immigrant groups. It remains unclear whether cooperatives can be sustainable in the absence of such bonds. The Fairbnb study reveals that managing the loci of governance and control is complex, and the ecological impacts of its model are also a challenge. A study of Freegle, a UK breakaway from the donation platform Freecycle, which has instituted democratic governance, found that it has been successful despite tensions between funders and participants (Martin et al. 2017). However, the authors of this study note the relevance of their case for smaller groups of breakaways from large platforms. Van Doorn (2017) cautions that this movement can lapse into “technological solutionism,” with insufficient attention paid to issues of racism and sexism, as well as to the state, which is a necessary actor to achieve the aims of a true sharing economy. While cooperatives are a promising alternative to predatory platforms, they have also failed to scale in comparison to well-funded corporate entities.

CONCLUSION

When it launched, many believed that the sharing economy prefigured an alternative form of economic practice to neoliberal capitalism. But the power of that system has all but overwhelmed it. The growth of giant “sharing” firms has cast doubt on the status of its utopian rhetoric. Its claims of generating greater inclusivity and an ethos steeped in mutuality have been contradicted by evidence demonstrating its tendency to reinscribe social inequalities through digital means. Its ability to generate trust among strangers has been revealed to be complex. And expectations of environmental sustainability have been undermined by increasing evidence of its contributions to carbon emissions and other pollutants. These considerations challenge the aspirations that have driven the sharing economy from its very beginning. Yet arguably, the most pronounced challenge to a genuine sharing economy may be an exogenous one: the COVID-19 pandemic that has swept across the globe. This is not to say that the large, for-profit firms cannot adjust. They are nothing if not flexible, consisting of technology rather than physical capital. They can shift their focus nimbly, as Uber has redirected its efforts from ride hailing to food delivery, and will likely survive the pandemic. But the threats to smaller sites specializing in face-to-face community are different. These progenitors of the sharing economy envisioned a form of consumption that transcends capitalism's long-standing emphasis on property ownership and individual ownership of goods. That vision has been seriously challenged by the advent of a pandemic that has transformed the sharing of goods and services into a source of fear and dread rather than mutuality and reciprocity. As social contact has become perilous, and strangers have become sources of potential infection, people may well shun the kinds of physical connections that are the foundation of the sharing economy. But if societies are to survive the existential threats posed not only by the pandemic but also by the climate and financial crises, they will need to reclaim the fundamental values of true sharing economies—ensuring the safety and security of all in a spirit of reciprocity and generosity. For surely the other way lies barbarism.

#### States are pre-empted by antitrust and labor law.

Jacob ’23 [Josh; November 20; J.D.,Columbia Law School; Columbia Law Review, “Avenues for Gig Worker Collective Action After Jintes,” vol. 123]

If Jinetes were interpreted to stand for the propositions that workers must be able to engage in collective action regardless of their employment status and that therefore disputes with platform companies over wages, hours, and working conditions constitute “labor disputes,” then most gig workers would likely be protected by the decision. This protection could expand the organizing options available to gig workers, including actions like the strikes and collective bargaining seen internationally. This broad interpretation of Jinetes could also allow states and cities to pass laws affirmatively protecting this collective action, attempts at which have been ensnared by antitrust problems in the past.

In 2015, Seattle became the first city to adopt an ordinance allowing collective bargaining by rideshare drivers.87 The city’s collective bargaining framework permitted an elected driver representative to meet with platform representatives and negotiate certain standards related to the drivers’ work, including wages, hours, and working conditions, in the form of a written agreement.88 After approval by the City, the agreement would become binding on the parties.89 The ordinance went into effect in 2016, and the U.S. Chamber of Commerce filed suit against the City of Seattle in 2017 alleging NLRA preemption and Sherman Act preemption and violation.90

Seattle claimed state-action immunity on the Sherman Act preemption claim.91 Under state-action immunity (or Parker immunity), a state may enact an anticompetitive restraint if the restraint is a product of “an act of government.”92 But courts are skeptical of antitrust defendants claiming Parker immunity.93 Without a showing that a putatively anticompetitive restraint flows directly from sovereign state action, a party claiming Parker immunity must satisfy the two-pronged Midcal test, which requires that (1) the challenged restraint be clearly articulated and affirmatively expressed as state policy, and (2) the policy be actively supervised by the State.94 On appeal, the Ninth Circuit held that the city ordinance did not pass the Midcal test, finding that the State of Washington had not clearly articulated a policy authorizing price-fixing by private parties in the rideshare driver market and that the State played no role in supervising the bargaining process, approving the agreements, or enforcing the ordinance.95 The ordinance was thus struck down,96 and no collective bargaining framework exists for rideshare drivers in Seattle today.

Because city or state collective bargaining frameworks for gig workers generally involve private parties organizing themselves, usually through another private party like a union, restraints on trade resulting from such arrangements do not flow directly from sovereign state action and thus must be able to pass the Midcal test in order to survive antitrust suits. In California, labor-friendly state legislators introduced a bill in 2016 that was substantially similar to the Seattle ordinance.97 A state legislature (rather than a city council) passing the bill would satisfy the “clearly articulated state policy” Midcal prong,98 but such a bill may still run into antitrust issues regarding the “active supervision” requirement, depending on the State’s role in reviewing the resulting collective bargaining agreement.99

The California bill’s author later removed the proposed legislation from consideration due to antitrust concerns.100 When asked why they pulled the bill, a spokesperson for the bill’s author noted that the bill contained “a number of untested legal theories,” implying that the bill may not have been able to pass muster before the California Assembly Judiciary Committee.101 The spokesperson averred that their office needed to “really explore all the legal issues that could be involved with this bill” before putting it to a vote,102 but, so far, California legislators have not reintroduced the bill.

#### Alternative mechanisms are also preempted by both employment and labor law.

Yousif ’20 [Marilyn; Labor and employment attorney, former attorney at the National Labor Relations Board, Clerked at the Third Judicial Court of Michigan for Judge Noah P. Hood, J.D. from Wayne State University Law School; The Journal of Law in Society, “The (Mis)Classification Issue: Should Federal Laws Preempt State Statutes that Classify Gig Economy Workers as Independent Contractors?” vol. 19]

III. ANALYSIS

A. Federal Preemption: Applying the Supremacy Clause and Worker Classification Tests to Employees in the Gig Economy

The "Supremacy Clause," of the United States Constitution (Article VI, Clause 2) establishes that federal laws and United States treaties negotiated with the United States are superior to state laws. (63) Article VI establishes that the "federal constitution, and federal law generally, take precedence over state laws, and even state constitutions," (64) thus prohibiting "states from interfering with the federal government's exercise of its constitutional powers, and from assuming any functions that are exclusively entrusted to the federal government." (65) When a state and a federal law explicitly conflict, the state law cannot be enforced. (66) The Supremacy Clause gives Congress the authority to preempt state law (67) when local law stands "as an obstacle to [the] accomplishment and execution of the full purposes and objectives of Congress." (68) Therefore, when it is impossible to comply with both state and federal requirements, federal law preempts state law. (69)

While the intricacies of contracting are too numerous for a comprehensive solution regarding the classification of employment status, tests created through the Internal Revenue Service (IRS) and Department of Labor (DOL) offer useful guidelines as to who may or may not be classified as an "independent contractor." (70) These tests look to various factors; two of these factors are the employer's control as to the manner in which the work is performed, and the extent to which the service rendered is integral to the business. These factors help to determine the nature of an individual's relationship with his or her employer. (71) Courts use determinative tests to clarify the issue of misclassification. The tests are grouped into three general categories. (72) The first category is, "(1) the IRS test, often termed the common law "right-to-control" test, serving as the most dominant test;" the second category is, "(2) the "economic realities test," which is "used in circumstances where a potential employment relationship has been created by social legislation" and is widely used by various jurisdictions, including the DOL; the third category is, "(3) hybrid tests that combine the control and economic realities tests." (73)

The common law "right-to-control" test is dominant, and the NLRB and the courts apply it to determine employment relationships. The IRS uses this test to determine whether a worker should be classified as an "employee" or an "independent contractor." (74) The "right-to-control" test uses a "20-factor test" to assess the employer's degree of control over the way the work is performed in order to determine whether workers are "independent contractors" or "employees." (75) The more control a company exercises over how, when, where, and by whom the work is performed, the more likely the workers are classified as "employees" and not "independent contractors." (76)

The "20-factor test" has been compressed into three main categories: behavior control, financial control, and the relationship of the parties. (77) The first category, "behavioral control," gathers facts that clarify whether the business has a right to control or direct the worker. (78) If a worker is generally told "when, where, and how to work; what tools or equipment to use; what workers to hire or to assist with work; where to purchase supplies and services; what work must be performed by a specified individual; and what order or sequence to follow," then the worker is deemed to be an "employee." (79) The second category, "financial control," questions whether the business aspects of the worker's job is controlled by the payer, i.e., how the worker is paid, whether expenses are reimbursed, and the extent of the worker's investment. (80) The third category, "type of relationship," is used to identify whether there are written contracts describing the relationship the parties intended to create, whether the worker is provided with employee-type benefits, the permanency of the relationship, and how integral the services are to the principal activity. (81) This "right-to control" test concentrates on the degree of control the employer may exert over the worker, and in reaching a conclusion about control, the IRS balances these factors to each individual worker's case and determines whether the worker is an "independent contractor" or an "employee." (82)

The Department of Labor's "economic realities test" is different than the IRS control test, since it is specific to determining the employer-employee relationship under the FLSA. It has an interest in ensuring accurate classification because only individuals classified as "employees" receive FLSA benefits. (83) The DOL uses the test to "determine who is an employee and, thus, eligible for FLSA benefits, by trying to establish whether the worker is economically dependent on the supposed employer." (84) According to the DOL, an "employee," as distinguished from a person who is engaged in a business of his or her own, is one who, as a matter of economic reality, follows the usual path of an "employee" and is dependent on the business which he or she serves." (85) The test stresses seven factors the Court has considered significant:

1. The extent to which the services rendered are an integral part of the principal's business.

2. The permanency of the relationship.

3. The amount of the alleged contractor's investment in facilities and equipment.

4. The nature and degree of control by the principal.

5. The alleged contractor's opportunities for profit and loss.

6. The amount of initiative, judgment, or foresight in open market competition with others required for the success of the claimed independent contractor.

7. The degree of independent business organization and operations. (86)

The test does not establish any bright-line rules, but instead determines whether "the worker is economically dependent on the employer (and thus an employee) or is in business for herself (and thus an independent contractor)." (87)

Under the hybrid test, a combination of both the common law "right-to-control" test and the "economic realities test," the Court considers additional factors relevant to the type of work performed and the employer-employee relationship. The courts consider "the skills required, the degree to which the work could be considered a separate operation from the employer's business, and extent of the worker's expected individual liability." (88) Additional considerations include: whether the work that is performed is a regular part of the employer's business, the regularity of the work performed, and whether the work that is being performed could be considered continuing services or contracting from the completion of each specific job. (89)

All three of these control tests indicate that a critical factor in determining whether an employment relationship exists between the ride share workers and their employers, in addition to carrying out core functions of the business, is the issue of control. (90) Rideshare companies have repeatedly argued that they lack sufficient control over drivers in order to consider them to be "employees." Drivers, however, have argued that certain elements of control, such as disallowing termination without cause and deactivation for low acceptance rates, constitute control over the workers. (91) They further argue that companies such as Uber [and Lyft] "retain control over the manner and means by which drivers perform their jobs through continuous monitoring and the driver rating system, behavioral and performance rules and evaluations, scheduling management and unilateral financial control over rates." (92) In reality, Uber has, to a significant extent, meaningful control over its drivers. (93) The complexities in determining the employment relationship within the ridesharing industry is indicative of the fact that the issue has yet to be resolved, and until it is, control tests established under federal laws and regulations, such as the NRLA and FLSA, are to make distinctions rather than state laws and regulations, such as the Michigan and Florida TNC statutes.

B. Preemption under the National Relations Labor Act

Within the context of the NRLA, Congress and the courts have the power to expressly preempt states from passing laws that fall within a defined scope of a statute. The NRLA, a continually expanding preemption doctrine, prevents states and cities from passing laws that touch upon anything related to labor, involve the interpretation of collective bargaining agreements, or involve issues that the courts believe Congress intended to leave to the free play of market forces. (94) The Supreme Court's preemption doctrines began with a 1959 case, San Diego Trades v. Garmon. (95) In this case, the Court established the "Garmon preemption," which precludes "state interference with [the] National Labor Relations Board's [NLRB's] interpretation and active enforcement of the integrated scheme of regulation established by the NLRA." (96) Doubling down on preemptions in 1976, the Supreme Court case, Machinists v. Wisconsin Employment Relations Commission, established the "Machinists preemption," which "forbids both the [NLRB] and States to regulate conduct that Congress intended be unregulated [and] left [it] 'to be controlled by the free play of economic forces.'" (97) Both of these doctrines are principles of labor law with the intention to preempt potential state interference with national labor policy.

Zoning in on the "Garmon preemption," it is well established that there is a requirement forcing state regulations to be preempted if they involve conduct that is meant to be prohibited or protected by the NLRA. Section 7 of the NLRA guarantees "employees the right to self-organization, to form, join, or assist labor organizations, collectively through representatives of their own choosing, and to engage in other concerted activities for the purpose of collective bargaining or other mutual aid or protection." (98) A state regulation that interferes with the NLRA's overall purpose, such as those purposes listed in section 7, would be an unfair labor practice (99) in violation of section 8 of the NLRA and thus in violation of the Supremacy Clause.

Section 2 of the NLRA does not, however, afford section 7 and 8 rights to "any individual having the status of an independent contractor." (100) When a state statute expressly classifies individuals as "independent contractors," as is done in Michigan's 2017 TNC statute, it is essentially creating its own worker classification test. By doing so, the Michigan TNC statute not only ignores the statutory provisions of the NLRA but also federally implemented control tests in place to properly define an employer-employee relationship. NLRA, The Michigan TNC Statute took away NLRA protections by classifying TNC drivers as "independent contractors." This is impermissible given the NLRA's complex and "interrelated federal scheme of law, remedy, and administration," (101) where the Court held that "due regard for the federal enactment requires that state jurisdiction must yield," (102) when the activities sought to be regulated by a state are clearly or may fairly be assumed to be within the purview of section 7 or 8 of the Act. Therefore, section 37 of the Michigan Statute, classifying TNC drivers as "independent contractors," should be preempted by the "Garmon doctrine." (103)

C. Preemption under the Fair Labor Standards Act (FLSA)

Moreover, under the Supremacy Clause, the Fair Labor Standards Act (FLSA) preempts state regulations that take away an individual's right to minimum wage and time-and-a-half overtime pay when such individual may be classified as an "employee." (104) The FLSA contains a savings clause which permits states to set more stringent wage-and-hour laws. (105) Where states set their own wage laws, they generally provide greater substantive rights, such as higher minimum wage than federal laws guarantee, or provide the same substantive rights as the FLSA. (106) Both federal and state minimum wage laws also govern and mandate that employers classify workers in order to determine eligibility for minimum wage and/or overtime pay. (107) In terms of classification, "exempt" workers and "independent contractors" do not qualify for overtime and/or minimum wages, and non-exempt workers must receive minimum wage and overtime for all hours they work. (108)

The employment status of Uber drivers has been litigated continuously over the last several years, with plaintiffs bringing forth claims under FLSA and their respective state wage-and-hour laws. In Razak v. Uber Technologies, Inc., drivers employed as "independent contractors" for the ride-share service, Uber, brought a putative class action against the ride-sharing service, claiming that Uber violated federal minimum wage and overtime requirements under the FLSA and parallel Pennsylvania state wage and labor laws because they were misclassified as "independent contractors" rather than "employees." (109) When Uber moved for Judgment on the Pleadings, the court found that with respect to the degree of control exercised by Defendants, Plaintiffs had alleged sufficient facts that they qualify as "employees" rather than "independent contractors" under the "economic realities" test. (110) The court left it up to the parties to complete an appropriate discovery on the issue of whether Plaintiffs are "employees" or "independent contractors" under the FLSA, and therefore, have allowed the Plaintiffs' minimum wage claims to proceed as pled. (111) The issue remains unresolved.

#### Defense ignores compounding future AND self-reinforcing risks.

Salmon ’24 [Paul et al; October 22; Professor of Human Factors and creator of the Centre for Human Factors and Sociotechnical Systems at the University of the Sunshine Coast; Ergonomics, “Tomorrow’s Demons: A Scoping Review of the Risks Associated with Emerging Technologies,” vol. 1]

While multiple technologies were identified in the included articles, it is interesting to also consider technologies and applications that have not been considered. Examples of emerging technologies discussed elsewhere but not covered in the reviewed articles include AGI, brain-computer interfaces, passenger autonomous aerial vehicles, technologies to support space colonisation and tourism and xenotransplantation to name only a few. This suggests that current research may be overly focused on the risks associated with ‘near-term’ emerging technologies, and that there is a clear gap around ‘far-future’ technologies. Whilst this is likely because insufficient details are available for far-future technologies, it is our view that the development of methods and processes to support valid risk assessments for far-future emerging technologies is a critical area of research. This will enable a proactive approach to the development and testing of risk controls and could support discussions regarding the extent to which future technology concepts should be pursued. Arguably, the later prospective risk assessment occurs in the design lifecycle, the less opportunity there is to prevent a high-risk technology from being pursued. Furthermore, modifications required to prevent risks from being realised will be more costly.

4.2. The risks associated with emerging technologies

Ten risk themes were identified, showing that there are a diverse set of known risks associated with emerging technologies. Whilst distinct themes were identified, it is noted that two of the themes cover process risks and eight cover outcomes risks. For example, sub-standard technology risks and malicious use risks could in turn create risks to human health and safety when the technology does not perform as intended or is used in a malicious manner. Beyond these interdependencies it is also likely that risks within the outcome risks theme influence each other. For example, legal and ethical risks could create privacy and security risks, socioeconomic impacts and geopolitical risks. This points to a need for studies to consider knock-on or ‘emergent risks’ arising from identified risks. This form of analysis can be achieved through methods such as Net-HARMS (Dallat, Salmon, and Goode Citation2018) which include an emergent risk analysis phase (Salmon et al. Citation2022). Further work exploring the interactions between different risks would be useful and could support the development of a general model of emerging technology risks.

Before discussing specific risks, it is worth noting that many of the risks identified are different to those typically identified in formal risk assessments which tend to focus on ‘sharp-end’ risks within the safety critical domains (Dallat, Salmon, and Goode Citation2018). For example, most prospective risk assessment methods are used to identify risks that emerge when front line operators such as control room operators, pilots and train drivers interact sub-optimally with technology in some way (Dallat, Salmon, and Goode Citation2019). In the aviation context this might include a pilot making inappropriate control inputs, operating the wrong control, or not taking a required action. This is in contrast with many of the risks identified in the current review which instead relate to either the design of the technology (eg sub-standard technology risks), the behaviour of the technology itself (eg existential threats), malicious use of the technology (eg malicious use risks), management of the technology (eg legal and ethical risks), or risks which are beyond the immediate context in which the technology is being used (eg ecological and environmental risks, geopolitical risks). As such, it may be that the nature and capabilities of emerging technologies requires a shift in thinking about the kinds of risks that could emerge and how they can be identified. A simple example of this is AGI, where decisions and actions taken by the technology itself will create risks (McLean et al. Citation2023; Salmon et al. Citation2023). Furthermore, identification and management of most of the risk categories identified requires analysis beyond the sharp-end and consideration also of the broader sociotechnical system in which future technologies will be deployed. Concerningly, the review indicated a lack of sociotechnical systems methods and a use of older methods such as FMEA when assessing the risks associated with future technologies. FMEA has various limitations when used for prospective risk assessment including difficulties in identifying risks beyond the sharp-end of system operation (Simsekler et al. Citation2019).

The most frequently identified risks were risks to human health and safety, with almost three quarters of the included articles identifying risks such as physical harm (Farcas et al. Citation2019; Murino et al. Citation2023), human health impacts (Finkel et al. Citation2018; Hartley et al. Citation2021; Hienuki et al. Citation2020) and traffic safety risks (Yao et al. Citation2020). This is perhaps not surprising given that most formal risk assessments are focused on identifying and preventing hazards to human health and safety. Encouragingly there is a strong existing knowledge base on the management of safety risks in complex sociotechnical systems (eg Leveson Citation2004; Rasmussen Citation1997). Application of contemporary safety science models, methods and principles during the design of emerging technologies to minimise risks to human health and safety is therefore critical.

Sub-standard technology risks where technologies do not achieve their intended aims were the next most frequently identified; for example, poorly designed or performing hydrogen energy (Hienuki et al. Citation2020), supercritical water gasification systems (Z. Liu, Bi, and Liu Citation2022), automated vehicles (Yao et al. Citation2020) and algorithmic policing models (Oswald et al. Citation2018). These risks should be considered during the design, testing and certification of emerging technologies and managed through conventional design lifecycle processes. Given some of the risks identified, the use of simulation and computational modelling techniques to simulate the behaviour of emerging technologies will be critical to enable the safe evaluation of prototype designs.

The third most frequently identified category was legal and ethical risks, which are particularly relevant for AI-based emerging technologies. Ensuring that AI technologies are designed and operate in an ethical manner, an area known as ‘machine ethics’ (Brundage Citation2016), is currently receiving much attention in areas such as defence, healthcare, security and law enforcement. In defence, for example, various sets of principles for the responsible design and use of AI technologies have been proposed, including those recently outlined by NATO (Citation2022) and the UK (Ministry of Defence Citation2022), US (U.S. Department of Defence Citation2020) and Australian (Department of Defence Citation2021) Defence forces. Though these principles support a strong commitment to the ethical design and use of AI technologies, there is little guidance available on how these principles can be incorporated and assessed during AI design, implementation and operation. This represents a critical gap in the knowledge base around how to design and implement ethical AI technologies (Salmon et al. Citation2023; Shneiderman Citation2022). Further work developing more specific and detailed guidance around the ethical design and operation of emerging technologies is recommended.

Of most concern are existential threats whereby the introduction of a new technology could eventually lead to human extinction. This is cited as a potential risk for atomically precise manufacturing (Umbrello and Baum Citation2018), and outside of this review has also been discussed in relation to AI and AGI (Bostrom Citation2014; Campbell Citation2022; Critch and Krueger Citation2020; Hancock Citation2022; McLean et al. Citation2021, Citation2023; Salmon, Carden, and Hancock Citation2021; Salmon et al. Citation2023). It is worth noting that there is debate over the veracity of such projections (Krueger Citation2023; Shneiderman Citation2022) and hence robust analyses are required to ensure such threats are taken seriously. The application of accepted and proven prospective risk assessment methods to technologies that could potentially create existential threats is therefore recommended (see Dallat, Salmon, and Goode Citation2019 for a review). As noted above, the focus of such risk assessments should go beyond the traditional sharp-end front line operator focus to consider interactions across the broader organisational, sociotechnical and societal systems in which the technologies will be deployed. This requires a shift in thinking from deterministic and linear models of risk towards a systems view that incorporates economic, social, cultural and ethical dimensions (Calay et al. Citation2022). Resilience engineering and systems thinking-based prospective risk assessment methods such as the Functional Resonance Analysis Method (FRAM; Hollnagel Citation2012), Systems Theoretic Process Analysis (SPTA; Leveson Citation2011), the Event Analysis of Systemic Teamwork Broken Links approach (EAST-BL; Stanton and Harvey Citation2020) and the Networked Hazard Analysis and Risk Management System (Net-HARMS; Dallat, Salmon, and Goode Citation2018) would be suited to this form of assessment (eg Salmon, Carden, and Hancock Citation2021; Salmon et al. Citation2023). Such methods consider the technology as part of a broader work, sociotechnical and societal system and seek to identify risks arising from interactions between the technology and other system components. This enables analysts to go beyond traditional risk assessment methods and provide a more comprehensive assessment of potential risks (Dallat, Salmon, and Goode Citation2018; Hulme et al. Citation2021). Further applications of the methods above to assess emerging technologies is therefore recommended.

It is acknowledged, however, that the conduct of prospective risk assessments for technologies that do not yet exist is highly challenging. Most prospective risk assessment methods are applied to a description or model of an existing system, developed based on data regarding current system functioning (Hulme et al. Citation2021; Salmon et al. Citation2022). For example, the STPA (Leveson Citation2011) method requires the development of a system control structure detailing key stakeholders and the control and feedback mechanisms used to maintain safety. The Net-HARMS (Dallat, Salmon, and Goode Citation2018) and EAST-BL (Stanton and Harvey Citation2020) methods use a task network of system operations that is developed based on data regarding how work is currently done (though recent work has used EAST to examine work as imagined eg Banks et al. Citation2018). Given that most technologies with potential existential threats do not yet exist, important questions to pose include whether contemporary methods can provide valid risk assessments for emerging technologies or whether there needs to be a shift in thinking around how risks are identified and regulated.

Recent work in EHF has involved the use of ‘envisioned world’ modelling whereby in-depth analyses of future systems are undertaken on models developed based on a range of data sources including information on existing system functioning, SME knowledge and analyses of early iterations of the technology (eg McLean et al. Citation2023; King et al., Citation2025; Salmon et al. Citation2024). Though this work is useful, it may be pertinent for EHF researchers and practitioners to explore methods described in the futures literature (eg Miller Citation2018; Voros, Citation2006). For example, creative processes such as imagination (Ward Citation1994), divergent thinking (Acar and Runco Citation2019), analogical thinking (Bonnardel and Marmèche Citation2004) and counterfactual thinking (Lin et al. Citation2024) could be useful. Voros (2006) presents a review of methods that have been used to support prospection, or the creation of ‘forward views’ or ‘images of the future’. These include evolutionary methods that seek to develop future views that evolve from a distinct point in time (usually the present) and revolutionary methods which seek to project forward to distinctly different future state (Voros Citation2006), with the latter being similar to the concept of envisioned worlds described above. Both approaches are supported by various methods including wildcards, visioning, backcasting, microhistory, counterfactuals, alternative histories and scenarios (Voros Citation2006). It may be worthwhile exploring the integration of such methods with prospective risk assessment methods such as STPA, Net-HARMS and EAST-BL. For example, wildcards, visioning, counterfactuals and scenarios could be used to test and validate STPA, Net-HARMS or EAST-BL models which in turn would heighten the validity of risks identified. Furthermore, in the present review, the use of SME-based knowledge elicitation methods was identified in six studies. Building on this it may be appropriate to explore the development of an SME-based envisioned world modelling approach that can be used to develop useful models of future systems and technologies (eg King et al., Citation2025; Salmon et al. Citation2022). Given the recommended sociotechnical systems level focus (Voros Citation2006), a key requirement will be to involve SMEs from all levels of the system in which the emerging technology will be deployed (eg government, regulatory, tech design/manufacturer, tech distributor/operator, supervisory and work levels). It should be noted that this sociotechnical systems level focus will not discount or overlook risks at the sharp-end of technology use and operation, rather it will expand the lens to identify other risks as well as the factors that interact to create the sharp-end risks.

Given some of the limitations in existing EHF methods, an important area of future research is to identify the methodological requirements for a prospective emerging technology risk assessment method. Whilst some requirements are clear, such as the need for a systems perspective described above, others are less well defined. Voros (Citation2006) provides a useful starting point by outlining a five-layer framework for prospection based on the systems iceberg metaphor. Voros’s framework includes events (ie discrete events that occur in the world), trends (ie the level that patterns and trends are observed), the system (ie the systemic structure), worldviews (ie mental models, worldviews and types of thinking) and historical layers (ie the level of societal, historical and macrohistorcial change). Arguably any prospective emerging technology risk assessment method should be able to assess future envisioned worlds across each of these five layers.

The advanced capabilities of future technologies such as AGI necessitate a shift in the way technologies are regulated and risks are managed. Governance frameworks tend to lag technological progress and are unable to cope with step-change breakthroughs in capability (Bengio et al. Citation2024). More proactive and adaptive regulatory frameworks have therefore been recommended for technologies such as AI (Bengio et al. Citation2024). The findings from this review should serve as a reminder that proactive and flexible regulation is required for all advanced technologies, even those considered to be far future in nature. As technologies become more advanced and risks become more extreme, the capability to manage them in a reactive manner is diminished. The provision of frameworks that support the proactive development of governance approaches for advanced emerging technologies are therefore recommended.

<<TEXT CONDENSED, NONE OMITTED>>

The current risk threshold ‘As Low as Reasonably Practicable’ (ALARP)-based approach to risk management may also require modification. The ALARP approach requires operators of potentially hazardous facilities to demonstrate that a facility is fit for purpose, that the risks associated with its functioning are sufficiently low, and that sufficient safety and emergency measures are in place (Melchers Citation2001). Giving the dynamic nature of capabilities and risks, the extent to which ALARP fits advanced technologies such as AI has been questioned (eg Koessler, Schuett, and Anderljung Citation2024) and more responsive and inclusive approaches may be required. For example, the potential use of capability thresholds has been suggested whereby each threshold has a set of corresponding requisite risk controls (Bengio et al. Citation2024; Koessler, Schuett, and Anderljung Citation2024). These thresholds could be identified and set in advance based on modelling and then used to support proactive development of risk controls for each threshold. This would ensure that risks are considered and that controls are in place before advanced capabilities are reached. A final consideration may be to have a more comprehensive and inclusive approach that considers risks beyond a narrow set of application areas and includes the perspectives of multiple stakeholders, including end users. In the case of brain computer interfaces for example, rather than limit the identification of risks to narrow application areas, King et al., Citation2025 has demonstrated the utility of a more comprehensive consideration of risks and requisite controls across domains and stakeholders. 4.3. Towards a model of emerging technology risks There are few models of emerging technology risks in the peer reviewed literature, and most are based on specific technologies such as AI (eg Macrae Citation2022). Such a model will be useful throughout design lifecycles to support technology developers, distributors and operators in identifying potential risks and developing suitable controls. The ten themes identified in the present review could form the basis for such a model or classification scheme, whereby each theme could be expanded to include specific sub-themes and suggested controls. Further research is therefore recommended to develop and validate such a model, including the conduct of case study assessments focused on the risks introduced by contemporary unruly emerging technologies (eg automated vehicles, uncrewed combat aerial vehicles, hypersonic weapons, ChatGPT and Watson IBM), the development of a classification scheme and the use of suitable experts to review, refine and validate an emerging technology risks classification scheme (eg via a Delphi study). 4.4. Risk controls A wide range of risk controls were proposed in the included articles, with the most frequently identified including regulation, stakeholder education and cooperation, risk assessments, monitoring and evaluation, and the use of measures to ensure that the technology performs optimally. The repetition of certain controls across different emerging technologies suggests that a generic framework of risk controls for emerging technologies could be useful; however, it should be noted that some of the less frequently reported controls are applicable only to specific technologies. For example, reducing blockchain user anonymity (Akartuna, Johnson, and Thornton Citation2022; Zhao and Chan Citation2020) is a suggested control for blockchain-related technologies only, and the use of personal protective equipment is a suggested control for anaerobic digestion technologies only (Pöther et al. Citation2021). This suggests that effective risk management for emerging technologies will require the conduct of technology-focused prospective risk assessments in addition to the use of generic taxonomies of risks and risk controls. 4.5. Study limitations Four limitations are important to note. First, to prevent the pre-empting of technologies identified, a set of non-technology specific search terms (eg future technologies, advanced technologies) were used instead of terms relating to specific technologies. This may have resulted in some articles being overlooked. Second, whilst the review adopted a comprehensive search strategy in-line with PRISMA guidelines, the focus was on peer-reviewed journal articles only. Consequently, studies reported in conference articles (eg Salmon et al. Citation2022) and the grey literature (eg Bugos and Reif Citation2021) was not included. Whilst we feel the review provides a comprehensive overview of the risks identified to date, other omitted literature will also be useful for researchers and practitioners wishing to understand the risks associated with future technologies. Third, the focus only on articles involving the application of a specific method to identify risks meant that some useful commentary and review articles were not analysed (eg Kupferschmidt Citation2018, McLean et al., Citation2021; Salmon et al. Citation2023; Scimeca and Verron Citation2022). These articles will also be useful for researchers and practitioners to develop a comprehensive understanding of the potential risks associated with emerging technologies. Finally, studies not available in English language were excluded from the review. As a result, findings from studies undertaken in geographically and culturally diverse contexts may have been omitted. This limitation should be considered when generalising the findings of the review beyond English-speaking countries, where there may be significant developments in technologies. 5. Conclusion

<<PARAGRAPH BREAKS RESUME>>

Whilst typically offering significant benefits, emerging technologies often also pose a broad spectrum of individual, organisational and societal risks. As technologies become more advanced and more integral to societal functioning, this emerging technology paradox represents a critical threat to humanity. Understanding not only the risks associated with emerging technologies, but also how best they can be identified and controlled, is therefore critical. This review has demonstrated that a diverse set of emerging technologies and risks have been identified, but that further research is required to ensure a more comprehensive and rigorous assessment of future risks. This includes increasing the range of emerging technologies assessed and expanding the analytical lens beyond the immediate context of use to consider emergent risks at an organisational, sociotechnical and societal system level. Further, the integration of methods from the futures literature with prospective risk assessment methods offers an opportunity for EHF researchers and practitioners to enhance the validity of future technology risk assessments. It is recommended that further work be undertaken to develop a generic method to support the assessment of emerging technology risks across domains.