Liability for robots I: legal challenges

Alice Guerra1,2*, Francesco Parisi2,3 and Daniel Pi4

1Department of Economics, University of Bologna, via Angherà 22, 47921 Rimini, Italy, 2Department of Economics, University of Bologna, Bologna, Italy, 3School of Law, University of Minnesota, Minneapolis, Minnesota, USA and 4School of Law, University of Maine, Portland, Maine, USA

*Corresponding author. Email: alice.guerra3@unibo.it

(Received 6 May 2021; revised 1 November 2021; accepted 2 November 2021; first published online 25 November 2021)

Abstract

In robot torts, robots carry out activities that are partially controlled by a human operator. Several legal and economic scholars across the world have argued for the need to rethink legal remedies as we apply them to robot torts. Yet, to date, there exists no general formulation of liability in case of robot accidents, and the proposed solutions differ across jurisdictions. We proceed in our research with a set of two companion papers. In this paper, we present the novel problems posed by robot accidents, and assess the legal challenges and institutional prospects that policymakers face in the regulation of robot torts. In the companion paper, we build on the present analysis and use an economic model to propose a new liability regime which blends negligence-based rules and strict manufacturer liability rules to create optimal incentives for robot torts.

Key words: Automated technology; electronic legal personhood; liability; robots

JEL codes: K13; K32

1. Introduction

The economic analysis of tort law assumes the existence of at least two human actors: an injurer and a victim (e.g. Miceli, 2017; Shavell, 1980, 1987). Nonetheless, this assumption becomes increasingly tenuous with the advancement of automated technologies (e.g. De Chiara et al., 2021; Shavell, 2020). Rather than still being the mere instruments of human decision-makers, machines are the decision-makers.

Since robots are insensitive to threats of legal liability, the question arises: how are we to regulate this new class of potential tortfeasors? The need for a theory to better understand robot torts is urgent, given that robots are already capable of driving automobiles and trains, delivering packages, piloting aircraft, trading stocks, and performing surgery with minimal human input or supervision. Engineers and futurists predict more revolutionary changes are still to come. How the law grapples with these emerging technologies will affect their rates of adoption and future investments in research and development. In the extremum case, the choice of liability regime could even extinguish technological advancement altogether. How the law responds to robot torts is thus an issue of crucial importance.

At the level of utmost generality, it is important to bear in mind that human negligence and machine error do not represent equivalent risks. Contrary to ordinary tools used by a human operator, robots serve as a replacement to the decision-making by a reasonable person.1 The social cost of...
machine error promises to be drastically lower than that of human negligence. We should therefore welcome the development of robot technology. Even if there was nothing that the law could do to reduce the risk of robot accidents, merely encouraging the transition to robot technology alone would likely effect a dramatic reduction in accident costs.

This paper comprises four sections. Section 2 discusses the novel problems posed by robot accidents, and the reasons why robots rather than other machines need special legal treatment. Section 3 reports the current legal approaches to dealing with robot accidents. Section 4 presents an overview of the companion paper (Guerra et al., 2021), where we build on the current legal analysis, to consider the possibility of blending negligence-based rules and strict liability rules to generate optimal incentives for robot torts. There, a formal economic model is used to study the incentives created by our proposed rules.

2. Rethinking legal remedies for robot torts

In an early article in Science, Duda and Shortliffe (1983) argued that the difference between a computerized instrument and a robot is intent. A computerized instrument – such as a computer program – is intended to aid human choice, while a robot becomes an autonomous knowledge-based, learning system, whose operation rivals, replaces, and outperforms that of human experts (Duda and Shortliffe, 1983: 261–268). Similar arguments on the dichotomy between mechanization and automation have been advanced in systems theory research. Among others, Rahmatian (1990) argued that automation ‘involves the use of machines as substitutes for human labor’, whereas ‘mechanization […] can take place without true automation’ (Rahmatian, 1990: 69). While computerized instruments are mere labor-saving devices (i.e. an extension of the human body in performing a work, mostly purely physical activities), robots are also mind-saving devices (i.e. an extension not only of the human body but also of the mind – hence performing both physical and mental activities). Robots are designed to have own cognitive capabilities, including ‘deciding (choosing, selecting, etc.)’ (Rahmatian, 1990: 69). Other scholars in systems theory research have put forth essentially the same arguments. For example, Ackoff (1974) defined automated technologies as machines that perform an activity for humans much as this latter would have done it themselves, or perhaps even more efficiently (Ackoff, 1974: 17). Thanks to the dynamic nature of the decision algorithm that drives their behavior, robots take into account the new information gathered in the course of their operation and dynamically adjust their way of operating, learning from their own past actions and mistakes (Bertolini et al., 2016; Giuffrida, 2019; Giuffrida et al., 2017).

In the face of the superior decision-making skills of a robot, the relationship between a robot and its operator is different from the relationship between an ordinary tool and its user. As the skills of a robot increase, the need and desirability of human intervention decreases. Even though there may
be special circumstances in which human judgment may outperform robots, robots outperform humans in most situations. Humans defer to the superior skills of a robot and delegate important decisions to them (Casey, 2019). However, as robots increase their skills, their ‘thinking’ becomes more ‘inscrutable’, falling beyond the human computational capacity (Michalski, 2018). Given the opacity of the robot’s decisions, it is very difficult – and often unwise – for operators to second-guess and override the decisions of a robot (Lemley and Casey, 2019).

The high complexity of the decision algorithm and the dynamic adjustment of the programing in unforeseen circumstances are what make robots different from other machines and what – according to many scholars – call for special legal treatment and a new approach to modeling accidents (Bertolini, 2014). Several legal and economic scholars across the world have argued for the need to rethink legal remedies as we apply them to robot torts (e.g. De Chiara et al., 2021; Lemley and Casey, 2019; Matsuzaki and Lindemann, 2016; Shavell, 2020; Talley, 2019). The proposed legal solutions to robot torts differ across jurisdictions (e.g. Europe versus Japan; Matsuzaki and Lindemann, 2016), yet the common awareness is that as the level of robot autonomy grows, under conventional torts or products liability law it will become increasingly difficult to attribute responsibility for robot accidents to a specific party (e.g. Bertolini et al., 2016). This problem is what Matthias (2004) called the ‘responsibility gap’. Matsuzaki and Lindemann (2016) noted that in both Europe and Japan, the belief is that product liability’s focus on safety would impair the autonomous functioning of the robot and slow down the necessary experimentation with new programing techniques. In a similar vein, in their US-focused article titled ‘Remedies for Robots’, Lemley and Casey wrote: ‘Robots will require us to rethink many of our current doctrines. They also offer important insights into the law of remedies we already apply to people and corporations’ (Lemley and Casey, 2019: 1311). Robots amount to a paradigmatic shift in the concept of instrumental products, which – according to Talley (2019) and Shavell (2020) – renders products liability law unable to create optimal incentives for the use, production, and adoption of safer robots as it is currently designed.

One of the challenges in the regulation of robots concerns accidents caused by ‘design limitations’, i.e. accidents that occur when the robot encounters a new unforeseen circumstance that causes it to behave in an undesired manner. For example, the algorithm of a self-driving car could not ‘know’ that a particular stretch of road is unusually slippery, or that a certain street is used by teenagers for drag racing. Under conventional products liability law, we could not hold a manufacturer liable for not having included that specific information in the software. Failing to account for every special circumstance cannot be regarded as a design flaw. However, we could design rules that might keep incentives in place for manufacturers to narrow the range of design limitations through greater investments in R&D and/or safety updates. In our example, we may be able to incentivize manufacturers to design self-driving cars that can ‘learn’ information and share their dynamic knowledge with other cars to reduce the risk of accidents in those locations.

4Mulligan (2017) refers to these as ‘black box algorithms’, ones that not even the original designers and programmers can decipher. ‘Machine learning’ and ‘genetic and evolutionary programing’ (see supra note 3) have further increased the complexity and opacity of the robot’s decision-making process.

5We thank Geoffrey Hodgson for encouraging us to elaborate on the discussion that follows.

6For a survey of the difficulties that legal scholars face when attempting to apply existing legal rules to robot torts, see Chopra and White (2011: 119–152).

7In their comparative study, Matsuzaki and Lindemann (2016) showed that both the legal framing and the concrete solutions to robot torts differ between Europe and Japan, especially in the legal construct of the robot as an ‘agent’ of the operator. While the European regulation debate explicitly addresses the degree of machine autonomy and its impact on legal institutions, this is not the case in Japan. See also, e.g. Leis (2006), MacDorman et al. (2009), and Sabanovic (2014).

8Specifically, Matthias (2004) wrote: ‘The rules by which [robots] act are not fixed during the production process, but can be changed during the operation of the machine, by the machine itself. This is what we call machine learning. […] [T]he traditional ways of responsibility ascription are not compatible with our sense of justice and the moral framework of society because nobody has enough control over the machine’s actions to be able to assume the responsibility for them. These cases constitute what we will call the responsibility gap’ (Matthias, 2004: 177).
Another challenge in the regulation of robots concern the double-edged capacity of robots to accomplish both useful and harmful tasks (Calo, 2015). As such, robots are increasingly perceived in society as social actors (Rachum-Twaig, 2020). Although legal scholars recognize that robots are mere physical instruments and not social actors, some have argued that from a pragmatic and theoretical perspective, granting them a legal personhood status – similar to corporations – might address some of the responsibility problems mentioned above. Eidenmüller (2017a, 2017b) observed that robots appear capable of intentional acts, and they seem to understand the consequences of their behavior, with a choice of actions. Furthermore, as Eidenmüller (2019) and Carroll (2021) pointed out, there is a ‘black box’ problem, and nobody, including manufacturers, can fully foresee robots’ future behavior because of machine learning and the dynamic programing of robots. This creates a difficult accountability gap between manufacturers, operators, and victims. The attribution of legal personhood to a robot is thus proposed by these scholars as a possible way to fill the accountability gap. The idea of attributing legal personhood to robots has been entertained in both Europe and the USA. The European Parliament has proposed the creation of specific status for autonomous robots, a third type of personhood between natural personhood and legal personhood, called ‘electronic personhood’ (European Parliament, 2017). The mechanics of how the electronic personhood of robots would operate is broadly presented by Bertolini and Episcopo (2021): ‘Attributing legal personhood to a given technology, demanding its registration and compliance with public disclosure duties, minimal capital and eventually insurance coverage would turn it into the entry point for all litigation, easing the claimants’ position’ (Bertolini and Episcopo, 2021: 14). The idea of giving some form of legal personhood to robots has also been voiced in the USA (Armour and Eidenmüller, 2020; Carroll, 2021; Eidenmüller, 2017a, 2017b, 2019; Jones, 2018; Kop, 2019), although it has never advanced to the legislative level.

Many challenges would arise in the application of existing tort instruments to robots with electronic personhood. Traditional legal rules refer to human-focused concepts such as willfulness, foreseeability, and the duty to act honestly and in good faith – concepts that no longer fit the new realities involving robots. Unlike humans, robots are insulated from self-interested incentives, which is intrinsically a good thing. However, the robots’ insulation from self-interested incentives can at times be a double-edged sword. Robots are not deterred by threats of legal or financial liability, since their personal freedoms and wealth are not at stake. To cope with this shortcoming, scholars and policymakers have investigated the possibility to make robots bearers of rights and duties, and holders of assets like corporations (Bertolini, 2020; Bertolini and Riccaboni, 2020; Giuffrida, 2019). In this respect, Eidenmüller (2017a) explicitly suggests that ‘smart robots should, in the not too distant future, be treated like humans. That means that they should […] have the power to acquire and hold property and to conclude contracts’. Future research should explore the extent to which these rights and financial entitlements could be leveraged by lawmakers to create incentives in robot tort situations.

3. Current legal status of robots

Robots are presently used in a variety of different settings. In some areas, they are already commonplace, while in others the technologies remain in their early stages (Książak and Wojtczak, 2020). There exists no general formulation of liability in case of accidents caused by robots, although some legislatures have attempted to anticipate some of the issues that could arise from robot torts. In this section, we survey some representative implementations to observe how legal rules have responded to the presence of robot actors to date.

3.1 Corporate robots

In 2014, a Hong Kong-based venture capital fund appointed a robot to its board of directors. The robot – named ‘Vital’ – was chosen for its ability to identify market trends that were not immediately
detectable by humans. The robot was given a vote on the board ‘as a member of the board with observer status’, allowing it to operate autonomously when making investment decisions. Although to our knowledge Vital in Hong Kong is the only robot benefiting from a board seat, and the recognition of personhood to a robot does not extend to other jurisdictions, the World Economic Forum released a 2015 report where nearly half of the 800 IT executives surveyed expected additional robots to be on corporate boards by 2025 (World Economic Forum, 2015). At present, Hong Kong and the UK already allow the delegation of directors’ duties to ‘supervised’ robots (Möslein, 2018).

The adoption of robots in corporate boardrooms will unavoidably raise legal questions on the liability arising from directors’ use of robots and for losses to corporate investors and creditors caused by robots’ errors (Burridge, 2017; Fox et al., 2019; Zolfagharifard, 2014). As Armour and Eidenmüller (2020) point out in their article titled ‘Self-Driving Corporations’, when robot directors become a reality, corporate law will need to deploy ‘other regulatory devices to protect investors and third parties from what we refer to as “algorithmic failure”: unlawful acts triggered by an algorithm, which cause physical or financial harm’. However, as of today, these questions remain without proper answers and the regulation of corporate robots has been left within the discretionary shield of corporate charters.

### 3.2 Aircraft autopilot

Aircraft autopilot systems are among the oldest class of robot technologies. The earliest robot flight system – a gyroscopic wing leveler – was implemented as far back as 1909 (Cooling and Herbers, 1983: 693). After a century of development, autopilot technology has progressed to nearly full automation. Aircraft autopilot systems are presently capable of taking off, navigating to a destination, and landing with minimal human input.

The longevity of the autopilot technology in aviation affords us a clear exemplar of how the law can respond to the emergence of robot technology. Early treatment of autopilot cases was mixed. The standard for liability was not negligence, but rather strict liability. However, the cases were not litigated as a species of products liability. Aircraft and autopilot systems’ manufacturers were therefore rarely found liable (see Goldsmith v. Martin (221 F. Supp. 91 [1962]); see also Cooling and Herbers, 1983; Eish and Hwang, 2015). Relatively early onward, it was established that operators (i.e. the airlines) would be held liable when an accident was caused by an autopilot system (see Nelson v. American Airlines (263 Cal. App. 2d 742 [1968])). There were two main reasons why aircraft and autopilot manufacturers were generally successful in avoiding liability: first, they were punctilious in crafting enforceable disclaimers and safety warnings, which effectively shielded them from products liability claims; and second, manufacturers aggressively litigated any claims against them, rarely settled, and thereby established favorable precedents (Leveen, 1983).

The legal outcome is largely unchanged today. It remains the airlines – not the manufacturers – that are liable for harms caused by autopilot systems. However, although the result has not changed, the legal justifications have evolved. Products’ liability law has undergone a radical transformation since the early autopilot accident cases, yet manufacturers continue to successfully avoid liability, for two reasons. First, in order for a products’ liability claim to succeed, the risk of harm must be reasonably foreseeable. Present-day aircraft manufacturing is heavily regulated, and an autopilot system that satisfactorily meets Federal Aviation Administration requirements is unlikely to be susceptible to any ‘reasonably foreseeable’ risk of harm. Direct regulation thus pre-empts tort liability. Second, even when an autopilot system is engaged, pilots have a duty to monitor and override it if operation becomes unsafe. The logic is

---

11The regulation of autopilots and other aviation equipment in Europe and Japan is equally nuanced. See, e.g. ‘Easy Access Rules for Airworthiness and Environmental Certification (Regulation (E.U.) No 748/2012) for Europe, and the ‘General Policy for Approval of Types and Specifications of Appliances’ for Japan (available at https://www.mlit.go.jp/common/001111795.pdf; last accessed October 2021).

1214 Code of Federal Regulations §91.3 (‘The pilot in command of an aircraft is directly responsible for, and is the final authority as to, the operation of that aircraft’).
that the human operator is legally responsible for anything that a robot does, because the human ultimately chooses to engage (and not override) the machine.

### 3.3 Self-driving cars

Self-driving cars are the most salient future use of robot technology. For quite some time, prototypes have demonstrated the feasibility of the technology, and fully autonomous vehicles are now part of the daily reality, from private cars to commercial taxi transportation, delivery robots, and self-driving trucks.\(^{13}\) In September 2016, the Department of Transportation published the *Federal Automated Vehicles Policy*, providing legislative guidance for states contemplating the regulation of self-driving cars (National Highway Traffic Safety Administration, 2016). A growing number of jurisdictions have enacted laws regulating the use of self-driving cars. At present, in the USA, 50 states and the District of Columbia have introduced autonomous vehicle bills.\(^{14}\) However, legislative efforts thus far have principally focused on determining whether an autonomous vehicle may be operated on public roads.\(^{15}\) Few jurisdictions have attempted to address the tort issues relating to self-driving cars. The *Federal Automated Vehicles Policy* suggests various factors that lawmakers should consider when formulating a liability rule (National Highway Traffic Safety Administration, 2016: 45–46):

States are responsible for determining liability rules for HAVs ['highly automated vehicles']. States should consider how to allocate liability among HAV owners, operators, passengers, manufacturers, and others when a crash occurs. For example, if an HAV is determined to be at fault in a crash then who should be held liable? For insurance, States need to determine who (owner, operator, passenger, manufacturer, etc.) must carry motor vehicle insurance. Determination of who or what is the ‘driver’ of an HAV in a given circumstance does not necessarily determine liability for crashes involving that HAV. For example, States may determine that in some circumstances liability for a crash involving a human driver of an HAV should be assigned to the manufacturer of the HAV.

Rules and laws allocating tort liability could have a significant effect on both consumer acceptance of HAVs and their rate of deployment. Such rules also could have a substantial effect on the level and incidence of automobile liability insurance costs in jurisdictions in which HAVs operate.

The few jurisdictions addressing the problem of tort liability merely push the problem back. For example, Tenn. Code Ann. §55-30-106(a) (2019) states that ‘[l]iability for accidents involving an [Automated Driving System]-operated vehicle shall be determined in accordance with product liability law, common law, or other applicable federal or state law’. Other states have enacted similarly opaque boilerplate that fails to delineate the applicable liability rule and define the legal ‘driver’ in the context of self-driving cars as being the robot itself.

In Europe, driverless vehicle policy proposals have evolved into a multinational policy initiative, under the United Nations. EU member states – and other countries, including Japan and South

---

\(^{13}\)The Society of Automotive Engineers (SAE) defines six levels of autonomy, ranging from 0 (fully manual) to 5 (fully autonomous). Most automakers currently developing self-driving vehicles are seeking level 4 autonomy, which does not require human interaction in most circumstances. Usually these vehicles are limited to routes or areas that have previously been mapped. See SAE International. 2021. Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles (https://www.sae.org/standards/content/j3016_202104/; last accessed October 2021).

\(^{14}\)The National Conference of State Legislatures maintains a database of state autonomous vehicle legislation, which is regularly updated. The searchable database is available at https://www.ncsl.org/research/transportation/autonomous-vehicles-legislative-database.aspx (last accessed October 2021).

\(^{15}\)For example, several states require that a self-driving car (1) satisfies the vehicular requirements of human-operated cars, (2) is capable of complying with traffic and safety laws, and (3) in case of a system failure, is capable of entering a ‘minimal risk condition’, which allows the car to achieve a reasonably safe state (e.g. by pulling over to the shoulder of the road when feasible). See, e.g. Iowa Code §321.515 (2020); La. Stat. Ann. §32:400.3 (2019); N.C. Gen. Stat. §20–401(h) (2020); Tenn. Code Ann. §§55-30-103 (2019).
Korea – have agreed to common regulations for vehicles that can take over some driving functions (e.g. mandatory use of a black box; automated lane keeping systems). Nonetheless, unlike in the USA, those countries currently do not have specific regulations for fully-automated cars. In the UK, policies on driverless vehicles are still evolving, and press releases from the UK Department of Transport refer to a regulatory process that has been underway since the summer of 2020 and will reach a point of greater completion in 2021 and the following years.

In Japan, the Road Transport Vehicle Act and the Road Traffic Act were revised to account for the possibility of autonomous vehicles driving on public roads (Imai, 2019). Those revisions have significantly reduced the legal obstacles to the operation of quasi-autonomous driving vehicles (SAE level-3), but not for self-driving vehicles (SAE level-4). The legalization of fully autonomous vehicles is still being debated, mainly due to issues related to the determination of the rules for criminal and civil liability in the event of traffic accidents.

The existing regulations of automated vehicles specify safety standards and mark the boundaries of the legalization of the levels of SAE automation, but leave questions open on how existing liability rules should be tailored to allocate accident losses. For example, the interaction of negligence torts and products liability is indeterminate when the driver of a vehicle is a robot. In an ordinary car accident, the human driver is liable under negligence torts if he/she failed to exercise due care, and the manufacturer is liable if the accident was caused by a manufacturing defect or design defect. If there is neither negligence nor a product defect, then the victim is left uncompensated for the accident loss. On the one hand, it could be argued that robot torts fall within the domain of products liability because the self-driving software is simply part of the car. It is well established that automobile manufacturers have a duty to ensure that the design of an automobile mitigates danger in case of a collision (Larsen v. General Motors Corp. (391 F.2d 495 [1968])). This rule would naturally extend to self-driving cars, where manufacturers are afforded greater opportunity to avert or mitigate accidents, thereby expanding their duty of care. The standard for demonstrating a defect in a self-driving car can be inferred from existing case law. For example, in In re Toyota Motor Corp. Unintended Acceleration Mkts., Sales Practices & Prods. Liab. Litig. (978 F. Supp. 2d 1053 [2013]), vehicles produced by Toyota automatically accelerated without any driver action, and the plaintiffs were granted recovery.

Similar reasoning could be transposed, mutatis mutandis, to self-driving vehicles.

On the other hand, it could be argued that robot torts fall within the domain of negligence torts, because autonomous driving is not qualitatively different from earlier innovations in automobile technology. Automation is not a discrete state, but rather a continuum. The electric starter, automatic transmission, power steering, cruise control, and anti-lock brakes have all increased the control gap between the operator and vehicle. Nonetheless, none of these technological innovations have excused the operator of tort liability. The move to autonomous driving will not be instantaneous, and it is


unlikely to be total.\textsuperscript{21} It is likely that for the foreseeable future operators will have the option to disengage autonomous operation. Indeed, it is plausible that there will be conditions where it would constitute negligence to engage autonomous operation.\textsuperscript{22} As long as the operator is ultimately in control – even if that control only extends to whether autonomous operation is engaged or not – traditional tort doctrine identifies the operator rather than the manufacturer as the party that should be the primary bearer of liability.

Thus, reasonable arguments can be advanced for assigning liability to the manufacturer as well as the operator. However, claiming that robot torts should be adjudicated ‘in accordance with product liability law, common law, or other applicable federal or state law’ merely begs the question. Tort law is a blank slate with respect to self-driving cars. The Federal Automated Vehicles Policy merely suggests factors to consider when formulating a rule, whereas it does not recommend any particular liability rule. Indeed, the few states that have acknowledged the issue have merely booted the problem to be resolved by existing law, despite the existing law’s indeterminacy on the novel question.

\subsection*{3.4 Medical robots}

Another recent and promising use of robot technology is in the field of medicine. Robots have been utilized in surgical operations since at least the 1980s, and their usage is now widespread (e.g. Lanfranco \textit{et al.}, 2004; Mingtsung and Wei, 2020).\textsuperscript{23} Due to their better precision and smaller size, robots can reduce the invasiveness of surgery. Previously inoperable cases are now feasible, and recovery times have been shortened.

Some surgical robots require constant input from surgeons. For example, the da Vinci and Zues robotic surgical systems use robotic arms linked to a control system manipulated by the surgeon.\textsuperscript{24} While da Vinci and Zues systems still require input from a human operator, in other areas of medicine there is a general trend toward even greater robot autonomy. Many healthcare providers are beginning to use artificial intelligence to diagnose patients and propose treatment plans. These artificial intelligence systems analyze data, make decisions, and output results, although the results may be overridden by a human operator or supervisor (Kamensky, 2020). As the technology further develops, it is plausible that surgical robots will require even less input from operators.

The applicable tort regime for medical robots is still evolving (see, e.g. Bertolini, 2015 on liability regimes for robotic prostheses). Allain (2012) provides an overview of the tort theories that victims have used in cases involving surgical robots, including medical malpractice, vicarious liability, products liability, and the learned intermediary doctrine. In instances where medical professionals actively control surgical robots, victims often assert medical malpractice claims that focus on the negligence of the medical professional, with reasonableness standards evolving over time based on advances in technology and knowledge. If the surgical robot or artificial intelligence is deemed a medical product – and therefore subject to Food and Drug Administration regulations – victims also often assert a products liability claim against manufacturers (Marchant and Tournas, 2019). However, this area of law remains relatively undefined, especially in cases involving software only.\textsuperscript{25}

\textsuperscript{21}SAE Level 5 autonomous vehicles are defined as being able to drive under all conditions, but there may still be limits to certain components, e.g. heavy rain making it difficult to distinguish objects, under which an autonomous vehicle may not operate, but a human could.

\textsuperscript{22}For example, see Plumer, Brad (2016). ’5 Big Challenges That Self-Driving Cars Still Have to Overcome’. \textit{Vox}. Available at https://www.vox.com/2016/4/21/11447838/self-driving-cars-challenges-obstacles (last updated April 21, 2016).

\textsuperscript{23}It is also worth mentioning the spate of peer-reviewed scholarly journals dedicated to the topic that arose during this period. For example, the \textit{International Journal of Medical Robotics and Computer Assisted Surgery}, established in 2004, the \textit{Journal of Robotic Surgery}, established in 2007, the \textit{American Journal of Robotic Surgery}, established in 2014, and the \textit{Journal of Medical Robotics Research}, established in 2016.

\textsuperscript{24}In a recent case involving the da Vinci robotic surgical system, a Florida man died as a result of a botched kidney surgery. The family claimed negligence based on the surgeon’s lack of training and experience with the system, but the case was later settled out of court (Allain, 2012).

\textsuperscript{25}In one recent decision involving a surgical robot, \textit{Thomas v. Intuitive Surgical, Inc.} (389 P.3d 517 [2017]), the court sought to decide whether the surgeon – i.e. the operator – or the manufacturer of the robotic surgical system would be liable.
As with self-driving cars, victims currently have no clear liability regime under which to seek compensation from operators or manufacturers for autonomous medical robots. At present, fully autonomous medical robots are still relatively uncommon; however, machines are taking on an ever-increasing share of decision-making tasks (see Kassahun et al., 2016). The tort issues that have been litigated thus far have tended to revolve around operator error (see, e.g. Taylor v. Intuitive Surgical, Inc. (389 P.3d 517 [2017])). Thus, for our purposes—much like self-driving car accidents—the law of medical robot torts is a tabula rasa.

3.5 Military robots

Military drones and robotic weapons are another area where robot torts are implicated. These machines are already being used to identify and track military targets. Additionally, weaponized drones have been used extensively in lethal combat. The UN Security Council Report of March 8, 2021 (UN S/2021/229) regarding a Turkish military drone that autonomously hunted humans in Libya without any human input or supervision in March 2020 is just the first of possibly many instances of autonomous attacks by military robots. During recent years, media speculation about this topic has been rampant and the recent Libya incident has revived the debate.26 It is easy to imagine other circumstances in the near future where constant communication with a human operator may not be possible and the identification and killing of an enemy target will be conducted autonomously. Should military technology continue to develop along this trajectory, it seems inevitable that other innocent targets will be attacked and eventually killed.

At present, no legal framework exists in the USA to address a mistaken killing by a military robot. Regarding the civilian use of non-military drones, the Federal Aviation Administration has begun to address ways to regulate drone usage within the USA in recent years, although it has not yet systematically addressed liability for physical harm (Hubbard, 2014).27 In an August 2021 Report released by the Human Rights Watch and the Harvard Law School International Human Rights Clinic, a proposal has been presented for a normative and operational framework on robotic weapons. States that favored an international treaty regulation of autonomous weapon systems agreed that humans must be required to play a role in the use of force, with a prohibition of robotic weapons that make life-and-death decisions without meaningful human control.28

3.6 Other uses

Robots are also used in factories and other industrial settings due to their ability to quickly and efficiently execute repetitive tasks (Bertolini et al., 2016). When an industrial robot injures a victim, it often occurs in the context of employment. In such instances, workers are typically limited to claiming workers’ compensation and barred from asserting tort claims against their employer. Many states include exceptions to this rule for situations where the employer acted with an intent to injure or

and whether the manufacturer had a duty to warn. The court held that medical device manufacturers have a duty to warn hospitals and operators of the systems of the risks of the system.


27For a review of unmanned aerial vehicles regulations on the global scale, see Stöcker et al. (2017) and Jones (2017); for European cases, see Esteve and Domènech (2017); for military drone laws and policies in Japan, see Sheets (2018).

with a ‘deliberate intention’ of exposing the worker to risk. However, thus far most of the cases brought by victims have proven unsuccessful (Hubbard, 2014). Due to the relatively controlled environment of factories and other industrial settings, operators can typically ensure a relatively high probability of safe operation and prevent injuries to potential victims.

4. Looking forward

In the companion paper (Guerra et al., 2021), we develop a model of liability for robots. We consider a fault-based liability regime where operators and victims bear accident losses attributable to their negligent behavior, and manufacturers are held liable for non-negligent robot accidents. We call that rule ‘manufacturer residual liability’, and show that it provides a second-best efficient set of incentives, nearly accomplishing all the four objectives of a liability regime, i.e. incentivizing (1) efficient care levels; (2) efficient investments in developing safer robots; (3) the adoption of safer robots; and (4) efficient activity levels. In our analysis, we bracket off the many interesting philosophical questions that commonly arise when considering autonomous robots’ decision-making. For example, a self-driving car may be faced with a situation where the vehicle ahead of it abruptly brakes, and the robot must choose whether to collide with that vehicle or swerve onto the sidewalk, where it risks hitting pedestrians. Alternatively, a robot surgeon may be forced to make split-second decisions requiring contentious value judgments. In such instances, should the robot choose a course of action that would result in a high chance of death and low chance of healthy recovery, or one that would result in a lower chance of death but a higher chance of survival with an abysmally low quality of life? While these moral questions are serious and difficult (Giuffrida, 2019; Sparrow and Howard, 2017), we exclude them from our inquiry because we do not consider them critical for the solution to the incentive problem that we are tackling. First, as a practical matter it cannot seriously be entertained that the design of rules governing such a critical area of technological progress should be put on hold until philosophers ‘solve’ the trolley problem or the infinitude of thought experiments like it. Second, even if ‘right answers’ exist to the ethical problems that a robot may face, its failure to choose the ‘morally correct’ course of action in some novel circumstance unanticipated by its designers can be construed by courts or lawmakers as a basis for legal liability. The objective of tort law is to minimize the social cost of accidents, and if the compliance with virtuous conduct in ethical boundary cases helps to accomplish that social objective, ethical standards should be incorporated into the legal standards of due care. Finally, if it is mandated as a matter of public policy that a certain approach to moral problems should be implemented, then this can be effected by direct regulation of robot manufacturing, outside of rules of tort liability.

Future research should consider that with some of the new programing techniques, the improvement of the robot can be carried out by the robot itself, and robots can evolve beyond the design and foresight of their original manufacturers. With these technologies, legal policymakers face what Matthias (2004) described as the ‘responsibility gap’, whereby it is increasingly difficult to attribute the harmful behavior of ‘evolved’ robots to the original manufacturer. In this context, models of liability in which robots could become their own legal entity with financial assets attached to them, like a corporation, could be considered. This could, but should not necessarily require the granting of (‘electronic’) legal personhood to robots, as discussed in Eidenmüller (2017b) and Bertolini (2020).

The issue has several implications which deserve future investigations. For example, a simple bond or escrow requirement for robots likely to cause harm to third parties could create a liability buffer to provide compensation. Robots could be assigned some assets to satisfy future claims, and perhaps a small fraction of the revenues earned from the robot’s operation could be automatically diverted to the robot’s asset base, improving its solvency. Claims exceeding the robot’s assets could then possibly fall on the manufacturer or the robot’s operator. An institutionally more ambitious alternative would be to conceive robots as profit-maximizing entities, just like corporations, owned by single or multiple investors. More efficient and safer robots would be yielding higher profits and would attract more capital on the market, driving less efficient and unsafe robots out of the markets. This natural selection
would mimic the natural selection of firms in the marketplace, and decentralize the decisions to acquire better robots and to invest in optimal safety to corporate investors. Liability would no longer risk penalizing manufacturers, but reward forward-looking investors, and possibly foster greater levels of innovative research.

As a final note, we should observe that the mere design of an applicable liability regime for robot technologies is not the only mechanism by which to incentivize further automation. There are also other means available, including regulation and mandatory adoption requirements, intellectual property rights, prizes, preferential tax treatments, or tax premiums. Insurance discounts for individuals adopting automated technologies can mitigate potentially high adoption costs. An optimal combination of these policy instruments may foster a widespread use of safer automated technologies.

Acknowledgements. The authors are indebted to Geoffrey Hodgson and the anonymous referees for their insightful and valuable comments. The authors are grateful to Carole Billiet, Emanuela Carbonara, Andrew Daughety, Herbert David, Luigi A. Franzoni, Anna Guerra, Fernando G. Pomar, Roland Kirstein, Peter Krebs, Jennifer F. Reinganum, Enrico Santarelli, Eric Talley, Gerhard Wagner, and the participants of the ZiF Research Group 2021 Opening Conference for discussions and helpful suggestions, and to Scott Dewey, Ryan Fitzgerald, Anna Clara Grace Parisi, and Rakin Hamad for their research contribution. An early draft of this idea by Alice Guerra and Daniel Pi was circulated under the title 'Tort Law for Robot Actors'.

References