Punishing Robots
Issues in the Economics of Tort Liability and Innovation in Artificial Intelligence

Alberto Galasso and Hong Luo

20.1 Introduction

A tort is an action that causes harm or loss, resulting in legal liability for the person who commits the act. The role of the tort system is to deter people from injuring others and to compensate those who are injured. Two important classes of tort law are product liability law that protects customers from defective or dangerous products, and medical malpractice law that governs professional negligence by physicians. Tort suits often make the headlines because of their large damages awards. For example, General Motors recently paid about $2.5 billion in penalties and settlements in a case involving faulty ignition switches linked to 124 deaths.¹

Rapid advancements in the field of artificial intelligence and robotics have led to lively debates over the application of tort law to these technologies. For example, the diffusion of autonomous vehicles is expected to shift the focus of motor vehicle accident litigation from driver liability to product (i.e., manufacturer) liability. Similar shifts are expected in health care because of advances in robot-assisted surgery and robot assistance for the elderly and disabled. These changes in the technological and economic landscape are also seen as an opportunity to redesign regulatory and liability rules. For

Alberto Galasso is associate professor of strategic management at the University of Toronto and a research associate of the National Bureau of Economic Research. Hong Luo is the James Dinan and Elizabeth Miller Associate Professor of Business Administration at Harvard Business School.

For acknowledgments, sources of research support, and disclosure of the authors’ material financial relationships, if any, please see http://www.nber.org/chapters/c14035.ack.

example, in February 2017 the European Parliament adopted—a resolution containing recommendations for EU-wide legislation to regulate “sophisticated robots, bots, androids and other manifestations of artificial intelligence” and to establish legislative instruments related to the liability for their actions (European Parliament 2017). An effective design and implementation of these policy changes require an understanding of how liability risk affects firms’ strategies and shapes future technological progress.

In an influential book, Porter (1990) concludes that “product liability is so extreme and uncertain as to retard innovation,” and he recommends a systematic overhaul of the US product liability system. A number of legal scholars share this view and warn about a potential “chilling effect” on innovation; that is, high damages awards may reduce firms’ willingness to develop new and riskier technologies, even if they are potentially superior to customary products (e.g., Huber 1989; Parchomovsky and Stein 2008). This idea that excessive liability may retard innovation also shaped high-profile legal cases such as the 2008 *Riegel v. Medtronic* Supreme Court decision and is a key argument for tort reforms currently discussed in the US Congress.

Despite the fundamental relevance of this issue, empirical work on the relationship between liability and innovation is scarce. Huber and Litan (1991) brought together a broad set of experts on five sectors of the economy where the liability system would have had the largest impacts. Based mostly on surveys and historical case studies, the authors were far from reaching a consensus. What were commonly agreed upon, however, were the dearth of data and systematic evidence, and a call for future research.

This chapter reviews the handful of empirical studies on the links between liability and innovation using a large sample of data. It aims to provide some insights into the potential impacts that liability laws and likely changes in the system may have on the rate and direction of innovation in robots and artificial intelligence, and to identify areas and questions for future research.2

### 20.2 Liability and Innovation: An Illustrative Theoretical Model

This section presents a simple, stylized model that explores the effects of liability risk on innovation incentives. Technologies are characterized by multidimensional heterogeneity. Specifically, a technology, \( i \), is characterized by two parameters: \( b_i \in [0,1] \) and \( r_i \in [0,1] \); \( b_i \) is the expected profit from incorporating technology \( i \) into the firm’s product, and \( r_i \) is the probability

2. It is important to note that this chapter focuses solely on the likely impacts on innovation and the direction of technological change. We refer interested readers to Hay and Spier (2005) and Polinsky and Shavell (2010) for an overall welfare discussion of the liability system and its features, and to Marchant and Lindor (2012) and Hubbard (2015) and the references within for details of tort law and an exploration of their applications to autonomous vehicles and sophisticated robots.
that the use of the product will result in personal injuries. The expected liability cost given that injury happens is $H$, which captures the (conditional) probability that a liability suit will be filed and the expected cost that the firm will face if involved in such a suit. We expect $H$ to be positive even if the firm is fully insured against claims for monetary damages because liability suits also invariably result in opportunity costs of employee time and firm resources, as well as in reputational damage.

The firm’s expected profit, net of liability risk from selling a product incorporating technology $i$, is

$$\Pi_i = b_i - r_i H.$$  

We denote the technology that the firm currently uses as $O$ and consider the firm’s decision to develop a new technology, which we denote as $N$. We assume a simple R&D process such that successful development takes place with probability $p(x) = x$ if the innovator incurs a research cost $C(x) = x^2/2$. As in Aghion et al. (2016), we refer to $x$ as the “innovation intensity,” which captures the likelihood of successfully developing a new technology. In this setting, the problem for the innovating firm is

$$\max_x x \Pi_N + (1 - x) \Pi_O - \frac{x^2}{2}$$

which yields the following:

$$x^* = \Pi_N - \Pi_O = b_N - b_O + (r_O - r_N)H.$$  

Formula (1) provides some basic insights into the relationship between liability and innovation. First, at the intensive margin, the sign of the derivative of $x^*$ with respect to $H$ captures the directional effect of an increase in liability risk on innovation intensity. Thus, an increase in liability risk suppresses innovation incentives for new technologies that are riskier than the current technology ($r_N > r_O$) but encourages new technologies that are safer ($r_N < r_O$). In other words, changes in liability risk affect the type of technologies in which a firm invests and influence the direction of innovation. Second, investment in innovation takes place if the profit potential of the new technology is greater than its liability risk, relative to the old technology—that is, $x^* > 0$ only if $b_N - b_O > (r_O - r_N)H$. Thus, at the extensive margin, marginal changes in liability risk will not affect whether the firm develops the new technology if it is expected to be highly profitable relative to the existing one (i.e., $b_N - b_O$ is very large) unless it is extremely risky. In contrast, liability concerns will matter more for technologies “at the margin” (i.e., the improvement in expected profitability is modest).

Galasso and Luo (2017) extend this stylized model to the medical setting, in which physicians (i.e., the direct users of technologies) face malpractice liability risk. Changes in their liability exposure affect innovation incentives in medical technologies through the demand channel. Assuming that ideas
for new technologies \((b_N r_N)\) are random draws from a bivariate distribution (as in Scotchmer 1999), they show that the overall effect of tort reforms that reduce physicians’ liability risk on innovation incentives is ambiguous and depends on the characteristics of the existing technology \((b_N - r_o)\).

The main message of this illustrative model is that the link between liability and innovation is more complex and nuanced than the simple view of “liability chills innovation,” which ignores the potential encouraging effect of liability risk on a potentially broad set of innovations that help firms and their customers manage risk.

20.3 Empirical Evidence on Liability and Innovation

In a pioneering study, Viscusi and Moore (1993) examine the relationship between product-liability insurance costs and firms’ research and development (R&D) investments, using a data set covering large US manufacturing firms in multiple industries between 1980 and 1984. They document a significant positive correlation between the expected liability insurance costs and firms’ R&D intensity when such costs are low or moderate. Only when liability costs are very high the correlation is negative. Furthermore, the liability-innovation link is driven mainly by product rather than by process R&D. They interpret these results as evidence that, on average, product liability, rather than discouraging innovation, promotes firm investment in product safety (likely through product design).

Galasso and Luo (2017) examine whether tort reforms that reduce physicians’ liability exposure to medical malpractice litigation affect incentives to develop new medical technologies. Different from the focus on product liability in Viscusi and Moore (2017), they examine how liability costs that users (physicians) face affect upstream research investment. It is worth noting that such a perspective broadens the scope of innovation from product safety design to include a wide variety of complementary technologies that help physicians manage risk, such as monitoring and diagnostic devices and devices used in complex procedures to reduce the likelihood of adverse events. Because these technologies are not themselves subject to product liability claims, they are more likely to be influenced by changes in user liability through the demand channel than by product liability.

Using a panel data set for the period of 1985–2005, Galasso and Luo (2017) find that, on average, the introduction of noneconomic damage caps in a state is associated with a 15 percent reduction in medical device patenting. The effect is, however, highly heterogeneous: tort reforms have the largest negative impact in medical fields in which the probability of a malpractice claim is the highest, and they do not seem to affect patenting of the highest or the lowest quality. These results are consistent with the idea that the decline in innovation is driven primarily by the reduced demand from
physicians for safer technologies or complementary technologies that help them manage risk. The welfare loss from such a large decline in quantity, however, appears not as worrying because patents with the highest impacts are not negatively affected.

Galasso and Luo (2018) study the medical implant industry in the early 1990s, during which the liability risk faced by raw material suppliers significantly increased relative to the risk faced by downstream producers. Vitek was a leading producer of jaw (temporomandibular joint) implants in the 1980s. Its Food and Drug Administration (FDA)-approved products were considered state of the art and safe for use by oral surgeons across the United States (Schmucki 1999). In the late 1980s, unexpected and widespread problems arose with Vitek’s products. Vitek filed for bankruptcy in 1990 under a deluge of lawsuits. Following Vitek’s bankruptcy, implant patients started to file a large number of lawsuits against DuPont, a raw material supplier for Vitek’s implants and a large firm with “deep pockets.”

The consensus among industry observers is that these events generated a substantial increase in the perceived liability risk faced by firms that supplied materials to producers of permanent implants, many of which had withdrawn from this market. This view is well summarized in a 1994 report on the status of the biomaterial market (Aronoff 1995), which links this fear of product liability suits to the jaw implant litigation. Eventually, DuPont won all the lawsuits, but the process took ten years and cost over $40 million (House of Representatives 1997). In contrast, DuPont’s revenue from these implants totaled only a few thousand dollars.

Galasso and Luo (2018) compare the rate of patenting in implant devices—excluding technologies involved in these litigations—to patenting in a control group of nonimplant medical technologies whose suppliers were not affected by the heightened litigation risk. The difference-in-differences (DID) results show, overall, a substantial decrease in the number of new patents for implants in the five years after Vitek’s bankruptcy in 1990. Time-specific effects show that implant and nonimplant technologies exhibited parallel increasing trends before 1990, and that the negative effect on implant technologies was immediate after 1990 and increased in magnitude over time. The significant drop in innovation in medical implants appears to have been largely driven by device producers’ expectation of a supply shortage of material inputs.

To address this problem, in 1998 the US Congress passed the Biomaterial Access Assurance Act (BAAA), which exempts biomaterial suppliers for medical implants from liabilities as long as they do not engage in the design,  

---

3. In parallel, problems also surfaced with silicone breast implants. Also in this case, a leading implant manufacturer filed for bankruptcy, and silicone suppliers were named as defendants in numerous lawsuits (Feder 1994).
production, testing, and distribution of the implants. The BAAA is one of the few federal liability reforms, an area of legislation typically reserved for the states (Kerouac 2001). Together, the empirical evidence in Viscusi and Moore (1993) and Galasso and Luo (2017, 2018) challenges the simple view that “liability chills innovation.” All three papers suggest that the link between liability and innovation depends on the context, including the nature of the innovation, the level of the liability risk, and the value of the technology. Furthermore, liability risk affecting one area may impact innovation incentives in other, vertically related segments. More research is needed to understand the complex and nuanced links between liability and innovation, and whether targeted policies can address these issues.

20.4 Tort Liability and the Development of AI Technologies

The liability system may affect innovation incentives of AI technologies and sophisticated robots in multiple ways, and the development of these technologies may, in turn, demand adjustments to the law. Below, we focus on a number of areas and highlight some of the economic trade-offs that deserve further examination, both theoretically and empirically.

20.4.1 Allocation of Liability Risk between Producers and Consumers

A central question in designing a liability system for AI technologies is how liability risk should be allocated between producers and consumers, and how this allocation might affect innovation. Effective policies would require a basic understanding of the relationship between humans and AI technologies—for example, whether they are substitutes or complements (Agrawal, Gans, and Goldfarb, chapter 3, this volume).

A key promise of AI technologies is to achieve autonomy. With less room for consumers to take precautions, the relative liability burden is likely to shift toward producers, especially in situations in which producers are in a better position than individual users to control risk. For example, the operator of a fleet of self-driving cars would have the data and predictive capability to provide instantaneous warnings of an adverse event. The cost

4. Examples of such federal policies include the General Aviation Revitalization Act of 1994, which exempts makers of small aircraft from liability for planes after eighteen years, and the National Childhood Vaccine Injury Act of 1986, which limits liability for drug companies and creates a no-fault compensation system for those injured by vaccines.

5. It is important to note that the likely impacts on innovation incentives would also depend on firms’ ability to write contracts and the development of the insurance markets (Schwartz 1988). We leave the discussion of these important topics and the interplay between liability law and contract law for future work. In situations in which externality (harm to third parties) is high and in the early stage of AI technologies, during which the insurance market may not be well developed or even exist, the roles of these systems are likely to be more limited than for mature technologies.
of observing systematic, hazardous user behaviors may also become sufficiently low such that it would be more efficient for producers to take precautions through product redesign. How such a shift might affect innovation incentives would depend on how producer liability is specified, especially whether the long-term social benefits are included in the analysis of the producer’s liability.

On the other hand, during the transitional period of an AI technology, substantial human supervision may still be required. Such interaction between AI and humans may not be obvious and difficult to predict. For example, it may actually be more difficult for drivers to sustain a safe degree of concentration levels and reaction speed when they are not actively engaged in driving. Human-machine interactions may also become more extensive and span increasingly complex domains as technologies are developed to enhance human skills. In the case of robot-assisted surgeries, for example, physicians may not have enough incentive to obtain sufficient training or to be sufficiently prepared for back-up options if the machine were to malfunction.

In many of these situations, it may be impractical or too costly for producers to monitor individual users and to intervene. Therefore, it would be important to maintain consumer liability to the extent that users of AI technologies have sufficient incentives to take precautions and invest in training, thus internalizing potential harm to others. When negative externalities are sufficiently high, regulators may find it necessary to mandate such investment. For example, a special driver’s license may be required to operate a self-driving car. Similarly, doctors may be required to take a minimum number of training sessions with the robotic system before being allowed to perform certain types of procedures on patients.

Consumer liability may incentivize users themselves to innovate in ways that help them take more effective precautions (Von Hippel 2005). For example, hospitals may redesign the operating room process or reorganize physicians’ training and work schedules. Furthermore, consumer liability may also incentivize producer innovation because users would demand safer and easier-to-use design features (Hay and Spier 2005), and mandatory training would favor “easier to teach” designs in order to reduce adoption costs.

20.4.2 Federal Regulation

Another key issue is whether Congress should pass federal regulations on the safety of AI and robotic technologies that preempts state laws and how such regulation would affect innovation. This would involve the creation of

---

6. The National Transportation Safety Board determined that the 2016 fatal Tesla crash was partly due to the driver’s inattention and overreliance on vehicle automation despite manufacturer safety warnings.

7. Some expert robotic surgeons and many surgical societies have voiced the need for basic, standardized training and certification in robotic surgery skills (O’Reilly 2014).
a centralized regulatory system similar to the FDA for drugs and high-risk medical devices: federal regulatory bodies would specify the safety standards, and approved products would be exempted from state liability claims under certain conditions.\(^8\) For autonomous vehicles, the House passed a version of such a regulation with bipartisan support in September 2017 (the SELF DRIVE Act, H.R. 3388).

From the perspective of innovation, a centralized AI regulatory system presents a number of trade-offs. On the one hand, relative to tort laws that examine liability cases ex post through judges and juries, ex ante regulations and safety preemption would significantly reduce the degree of uncertainty regarding liability risk.\(^9\) Reduction in uncertainty, in general, increases R&D and other complementary investment. Furthermore, harmonizing different, slow-moving state-wide regulations could also speed up experimentation and adoption. In the case of autonomous vehicles, as of September 2017, some testing was explicitly allowed in less than half of the states with different degrees of restriction and safety standards.

On the other hand, federal regulation could trade off certainty with flexibility. With the fast-changing landscape of AI technologies, federal agencies may not have sufficient information in the early development stage to set effective standards.\(^10\) If such regulations were hard to change, they could influence the rate and direction of innovation in undesirable ways.

20.4.3 Allocation of Liability Risk across the Vertical Chain

Artificial intelligence and sophisticated robotics are often complex technologies that involve multiple suppliers of software and hardware and that may require high degrees of integration between different components. Furthermore, AI technologies, like other general purpose technologies (such as polymers), once developed for the first few areas, may later be developed for a wide variety of applications at a lower cost.

Current laws, such as component parts and sophisticated purchaser doctrines, stipulate that component suppliers are not liable unless the component per se is defective or the process of integrating the component has caused the adverse effect (Hubbard 2015). In practice, however, these laws may be inadequate in certain circumstances and may expose component

---

8. Federal preemption of state laws may be explicit or implicit, with the former providing significantly greater clarity. In the case of FDA preemption in *Riegel v. Medtronic, Inc.* (2008), the US Supreme Court ruled that manufacturers of FDA-approved devices that went through the pre-market approval process are protected from liability claims under state laws. In *Wyeth v. Levine* (2009), however, the US Supreme Court ruled that Vermont tort law was not preempted.

9. Kaplow (1992) provides a general economic analysis of rules versus standards; that is, whether laws should be given content ex ante or ex post. The basic trade-offs depend on factors including the frequency and the degree of heterogeneity of adverse events, as well as the relative costs of individuals in learning and applying the law.

10. For example, for autonomous vehicles state regulators currently differ in their opinions about whether cars without steering wheels or brake pedals should be allowed on public roads for testing and operation purposes.
suppliers to disproportionately high liability risk relative to their expected revenue. Evidence from Galasso and Luo (2018) suggests that in such situations, downstream innovation may suffer. It would be interesting for future research to examine how liability costs should be allocated across component producers and its impacts on innovation, and under what conditions policymakers should consider applying exemption regulations, such as the BAAA enacted in the medical implants industry.

It would also be interesting to examine how liability rules, apart from their direct impacts on innovation, influence firm boundaries, which, in turn, could affect innovation. For example, rules such as the BAAA may discourage vertical integration because its exemption applies only to component material suppliers sufficiently removed from downstream activities. Similarly, liability rules may also influence how products and services are designed. For example, they may encourage more modular designs to better insulate liability risk across different components.

20.4.4 Liability Risk and Market Structure

Relatedly, it would be interesting to better understand the interplay between liability risk and industry market structure and whether changes in market structure driven by liability risk have long-term consequences for innovation (Agrawal et al. 2014).

How liability risk affects firms of different sizes is likely to depend on the empirical context. Plaintiffs may be more likely to target larger and cash-rich firms (Cohen, Gurun, and Kominers [2014] find this pattern in patent litigation cases). At the same time, larger firms are better at withstanding high liability risk because they have greater resources both to self-insure and to provide more generous indemnification contracts to suppliers.

One may argue that liability insurance could insulate producers from potential liability concerns. However, in the early stages of AI technologies the market for liability insurance may not be fully developed, or even exist, due to insufficient data on adverse events of a particular nature and their damages. Even with well-developed insurance markets, high liability risk may result in high premiums, which can be prohibitively expensive for smaller firms and, thus, deter entry.

20.4.5 Liability Litigation

An important feature of an effective liability system is that disputes are resolved quickly. Longer settlement delays are typically associated with higher transaction costs for the negotiating parties. More importantly, delays and uncertainty in the process mean slower diffusion of the AI technology at the center of the dispute.

It is not obvious whether liability suits related to AI technologies will be easier to settle than those involving other technologies. In particular, the complexity of these new technologies and certain types of human-machine
interactions may reduce the litigants’ ability to find a compromise. That said, classic models of pretrial negotiations predict a higher likelihood of settlement when information asymmetries between litigating parties are reduced (Spier 2007). Manufacturers of AI technologies may find it in their own interest to design the machine’s data-recording capability in ways that facilitate the discovery process and speed up settlement. In cases where manufacturers lack such incentives, mandates of certain designs may be necessary if they are clearly efficiency enhancing. Once again, how effective these data capabilities of AI technologies are in facilitating dispute resolution would also depend on the ability of the court system to understand and interpret data, on private incentives for data sharing, and on whether policies are in place to discourage misrepresentation and manipulation of data.

20.4.6 Liability Risk and Intellectual Property Protection

The likely impacts of liability risk on innovation would also depend on the strength of intellectual property (IP) rights. Intuitively, when IP rights are strong, firms can invest in safer products and recover their investments by charging a price premium. However, if competitors can easily copy and sell these products or features, the incentive to innovate in the first place would decrease. The above considerations may be different, however, if consumers cannot easily distinguish between safer and less safe products, and their fears about dangerous products suppress their demand for the entire product category. For example, Jarrell and Pelzman (1985) show that one firm’s product recalls may have negative reputational impacts on competitors and producers of related products. When such negative spillover is strong, firms with other means of extracting rents (e.g., larger firms) may have the incentive to invest in safety features and share them with firms in the industry so as to maintain consumer demand for the whole industry.

Finally, in a cumulative innovation environment, as Green and Scotchmer (1995) have shown, the allocation of IP rights among sequential innovators may have important effects on their respective innovation incentives. Related trade-offs are likely to also emerge for the allocation of liability damages among sequential innovators.

20.5 Conclusion

This chapter has examined some of the basic economic trade-offs linking liability risk with innovation incentives and the direction of technological progress in the context of artificial intelligence and sophisticated robots. Features of the liability system, such as the allocation of risk between producers and consumers and the level of centralization in regulation, may have a significant impact on the development and diffusion of these new technologies, as well as on the products and services that apply them. The
extent of these effects is likely to also depend on the market structure and the organization of the vertical chain of innovation.

More broadly, our analysis supports the idea that the liability system and its reforms can affect the rate and the direction of technological change, indicating that these policies have dynamic effects on innovation incentives that go beyond their short-term impact on the safety of the users and others. As Finkelstein (2004) stresses, recognizing and estimating these dynamic effects is crucial to evaluating the costs and benefits of policy reforms.

References


