EconomiX

Which Liability Laws for Artificial Intelligence?

Eric Langlais
Nanxi Li

2024-22 Document de Travail/ Working Paper

EconomiX - UMR 7235 Bâtiment Maurice Allais
Université Paris Nanterre 200, Avenue de la République
92001 Nanterre Cedex

Site Web : economix.fr
Contact : secreteriat@economix.fr
Twitter : @EconomixU

Which Liability Laws for Robots?

Eric Langlais* and Nanxi Li†‡

July 8, 2024

Abstract

This paper studies how the combination of product liability and tort law shapes a monopoly's incentives to invest in R&D for developing risky AI-based technologies ("robots") that may induce accidental harm to third-party victims. We assume that robots are designed to have two alternative modes of motion (fully autonomous *versus* human-driven). In the autonomous mode, the monopoly (i.e., robot manufacturer) faces product liability and undertakes maintenance expenditures to mitigate victims' expected harm. In the human-driven mode, robot users face tort law and exert a level of care to reduce victims' expected harm. In this set-up, efficient maintenance by the robot designer and efficient care by robot users result regardless of the liability rule enforced in each area of law (strict liability, or negligence). However, the equilibrium output and R&D invesments levels are generally suboptimal, and no combination of tort law and product liability law provides higher output, higher R&D investment, and finally higher social welfare than the others. Combining regulation of the output with the general use of strict liability in both areas of law yields the efficient investment level only when the monopoly uses (perfect) price discrimination.

Keywords: Artificial Intelligence, Algorithms, Tort Law, Product Liability, Strict Liability, Negligence.
JEL Code: K13, K2, L1.

*Corresponding author: EconomiX, UPL, Univ. Paris Nanterre, CNRS F92000; elanglais@parisnanterre.fr.

†EconomiX, UPL, Univ. Paris Nanterre, CNRS F92000; nanxi_li@parisnanterre.fr

‡We have benefited of remarks by Eric Darmon and Axel Gautier on preliminary drafts of this work, as well as by participants to AFED 2023-Nice, the Workshop « Competing with Algorithms and Data: market outcomes and regulatory issues » 2023-Nanterre, and the 11th doctoral workshop of AFREN 2023-Paris Saclay. As usual, all remaining errors are ours.

1 Introduction

The rapid progress in the field of Artificial Intelligence has recently led to extensive applications of automated technologies in various sectors such as transportation, medicine, etc. The most advanced robots have the potential to learn from their past actions/experiences, gathering and analyzing new information, with the purpose of adapting their behavior. But algorithms operate as black boxes for a human brain, since their complexity often makes their operation/decision-making processes difficult to be foreseen and understood by robot users. In cases of robots-related accidents, this raises legal issues currently in debate (Abraham and Rabin [2019], Buiten [2024], Davola [2018], Guerra et al. [2022a], Lemley and Casey [2019]).

Since the first reported fatal accident between an autonomous Uber-Volvo XC90 and a pedestrian in USA (March 2018), hundreds of similar cases have occurred around the world (among which the sudden acceleration of a Uber-Tesla Model 3 in Paris, Dec. 2021); subsequent investigations in all cases have revealed that both robot design limitations and human failures in understanding robot operations are usually at stake. For human-driven cars, the legal provisions hold that the human driver faces a negligence rule (i.e., he is liable only when failing to exercise a due care standard), whereas the car manufacturer is liable only if the cause of the accident stems from a design flaw or a manufacturing defect. But autonomous vehicles point out some blind spots in the law. Known AI design limitations do exist but are not enough for robot manufacturers to be found at-fault from a legal perspective; in turn, human operators of robots may be unable to understand the robot's rapid decisions.

Such legal blind spots translate into the risk for victims of being unable to recover damages from their injurers in case of robots-related accidents, such that de facto, a no-liability regime prevails. Acknowledging this problem, the European Commission [2022a,b] has recently promulgated two directives on liability issues (for defective products and for non-contractual liability), urging Member States, first, to adapt their law to the increasing presence of AI in extended domains of economic activity, and in way warranting incentives a high R&D activity in the AI sector to develop safe AI-products; and second, to use more largely the strict liability rule, in order that victims recover damages in case of accidents more easily, i.e., without the need to prove the injurer is at-fault.

From a Law & Economics point of view, the debate about the design of liability regimes and the impact on AI designers' incentives to invest in safer robots must rely on the social benefits of AI technologies, and the social cost of AI-related accidents. So far, literature[1] has often focused on the case of Autonomous Vehicles as a laboratory of ideas, considering the impact of liability regimes on victims' incentives, AV designers' incentives, and AV users' incentives to undertake precautionary measures. On the one hand, Kim [2024], Shavell [2022], and Talley [2019][2] focus on fully autonomous vehicles, and analyze the resulting allocation of liabilities between AV users and third-party victims, while Guerra et al. [2022b] discuss different negligence-based rules when the manufacturer is the residual liability bearer. On the other hand, de Chiara et al. [2021] consider the transition period where human-operated vehicles still exist together with fully autonomous vehicles, and analyze the impact of liability regimes on investment in the sector of autonomous vehicles.

However, considering the long history of automatic pilots in aviation, one may expect a third scenario to be very likely, not only for the transportation sector but also for the health sector as well as the numerous automated processes employed in industry: at the maturity stage, robot users will still be involved to various degrees in the operation of the robot; hence AI-related technology will be only partially autonomous.[3] The main argument is that advanced AI-related technologies also face engineering defaults and design limits, as well as material flaws. At the engineering stage, it is thus acknowledged that the AI technology may perform poorly under some circumstances known and perfectly anticipated – thus undertaking erratic decisions with consequences that are worse than those resulting from a human decision-maker (losses, harms to some victims).

[1] The literature reviewed here is focused on *performative algorithms*, i.e. that are able to collect and analyze information, before accomplishing autonomous actions. There is a different literature dealing with *advisory algorithms* (Chopard and Muzy 2023, 2024; Obidzinski and Oytana 2022, 2024), i.e. that provide decision-making support to humans decision makers.

[2] Di, Chen, and Talley [2020] afford a more technical extension of this work.

[3] Remind that in the tradition of the L&E literature, the kinds of activity at stake here are those dedicated to the prevention of accidental harm.

The purpose of this paper is to investigate in this third scenario the effects of alternative combinations of product liability and tort laws on R&D investments and the market for robots. We consider the following general problem: A robot manufacturer (i.e., Algorithms' developer) makes a costly initial R&D investment for developing robots. Post-investment, the manufacturer informs potential users that a robot has two alternative modes of operating: i) either an autonomous mode, recommended in some pre-determined circumstances or "states of nature",[4] or ii) the human-driven mode, when the latter circumstances are not satisfied. This reflects that "algorithm limits" do exist and are known at the engineering stage. In both cases, an operating robot may entail accidental harm for some victims. Then the robot manufacturer sets a monopoly price for its robots, and robot users consider whether they buy or not one unit of good (one robot). Finally, after the market for robot clears, the "state of nature" is observed (verifiable by courts and users owning a robot). This implies that when the autonomous mode is activated, victims' expected harm can be reduced by the maintenance efforts realized by the robot manufacturer (i.e., updating soft-wares, up-grading infrastructures: i.e., Wi-Fi, satellites, demarcation) who is subject to court; instead, when the human-driven mode is activated, victims' expected harm can be reduced through precautionary efforts by robot users (alertness, sobriety, compliance with the driving code) who are subject to tort law.

In this set-up where information is verifiable, we show that any combination of liability rules (strict liability or negligence) implemented in court and tort law provides efficient incentives to the robot manufacturer for his maintenance efforts, as well as to the robot users for their care levels. In contrast, the output distortion is regime-specific: if the standard

[4] In the case of applications to Autonomous Vehicles for example, a specific state of Nature summarizes the weather conditions, quality of the road and surrounding infrastructures, complexity of the network, traffic intensity and so on. The AI-based technology is based on a decision-making process controlled by algorithms, thus programmed to routinely/automatically adapt to all pre-specified situations (fully anticipated, and described), resulting in all precautionary measures required to avoid an accident in these instances (domain with Robot Full Capacities or Full Control). At the level of the Injurer equipped with such a technology, the performances/quality of the Robot (in the sense where the Robot fully performs under a domain of states of Nature, with no failure) are given: the expenditures in terms of R&D, engineering and design required in order to develop and expand the Robot domain represents up-front investment born at the level of the industry by the Robot Designer.

result (under provision) with a monopoly still holds when strict liability is used in both areas of laws, it turns out that other combinations of the strict liability/negligence rules may entail either under or over provision of output by the monopoly. Finally, we show that the distortion in R&D investment is also regime-specific; generally speaking the monopoly may overinvest as well as underinvest in R&D at equilibrium regardless of the combination of liability regimes in product liability and tort law, and whether the monopoly price discriminates or not.

These results are in a sense no surprise since the objective of liability laws is to provide potential injurers with incentives to minimize the cost of accidents (Calabresi [1970]). Remedying to market distortions must be achieved by other tools. The coordination between the enforcement of liability laws and other public interventions is an issue whenever market power exists (taxation: see Hamilton [1998]; market regulation: see Hamilton et al. [2023]; competition policy: see Charreire and Langlais [2021]). Nevertheless, de Chiara et al. [2021] suggest that allowing a AV monopoly to use perfect price discrimination would be sufficient for reaching the first best R&D investment level as long as strict liability is used both in tort law and product liability. Indeed we show here that this also requires output regulation by a public authority. In words, reaching the first best requires the consistency between the strategies of courts, the regulator and the monopoly. Failure to coordinate one side with the two others yields an inefficient outcome.

The rest of the paper is organized as follows. Section 2 provides the description of the model and characterizes the social optimum. Section 3 analyzes the decentralized equilibrium when the robots market is a non-discriminatory monopoly. We compare the impact of different combinations of liability regimes in product liability and tort law on monopoly incentives to invest in R&D, before marketing robots. Section 4 discusses the implications of combining output regulation together with liability laws. Section 5 concludes.

2 The model

2.1 Assumptions

We consider a set-up with three kinds of agents: a robot manufacturer/developer (i.e.,, algorithms' developer, car or aircraft producer, medical robot manufacturer, etc), a population of potential robot users (i.e., firms, drivers or pilotes, surgeons, etc), and victims suffering from an exogenous harm (i.e., consumers, pedestrians, passengers, patients, etc).[5]

The manufacturer makes an initial R&D investment a at a cost $\gamma(a)$, where $\gamma'(a) > 0$ and $\gamma''(a) > 0$ (this latter being large enough for the SOC to hold in the rest of the paper). Once the R&D investment is realized, the manufacturer warns potential robot users that the optimal mode of motion/use of a robot is conditional to the realization of the state of Nature ω (verifiable). For the sake of simplicity, let us assume that $\omega \in \{Aut, NAut\}$ is obtained with probability $p(\omega)$: if $\omega = Aut$ is observed, then it is common knowledge that optimal use of robots is in autonomous mode; instead, if $\omega = NAut$ is observed, then the human-driven mode is the optimal mode.[6] Since it is reasonable to consider that the purpose of R&D investment is to expand the set of circumstances where the autonomous mode applies, we assume wlog that $p(Aut) = a$ and $p(NAut) = 1 - a$. More generally, the technology/robot developed by the monopoly defines a "general technology of accident" for society with the following characteristics:

i) In state $\omega = NAut$, a robot user with type $v \in [0, 1]$, receives a benefit v from using the robot in the non-autonomous mode, and has to undertake its precautionary measure/care x at a cost $c(x)$, where $c'(x) > 0$, $c''(x) > 0$; an accident occurs with probability $q(x)$, where $q'(x) < 0 < q''(x)$, and some third-party victims are injured, bearing the harm h. We

[5]Our framework has a straightforward interpretation in terms of market for Autonomous Vehicles; however, it has broader application since it captures any situation where victims are "passive", i.e., victims do not undertake precautionary measures. Typically, the L&E literature distinguishes between, on the one hand, the accident model with unilateral precautions (where injurers only undertake care), such as accidents in urban areas between a car and a pedestrian, and on the other hand the accident model with bilateral precautions (where injurers and victims do undertake precautions), such as a collision between two vehicles on a highway.

[6]The erratic behaviors/responses of robots in state $\omega = NAut$ could be captured by the assumption that the marginal cost of maintenance measures tends to infinity.

assume that v is distributed according to a density $f(v) > 0$ and a cumulative function $F(v)$, satisfying the usual property that $\frac{f}{1-F}$ is increasing on $[0, 1]$.

ii) In state $\omega = Aut$, the autonomous mode implies that each robot user receives the same benefit (regardless of his own type) from using the robot (in the autonomous mode) normalized to 1; the robot manufacturer makes efforts for maintenance y at a cost $\mathbb{C}(y)$, where $\mathbb{C}'(y) > 0$, $\mathbb{C}''(y) > 0$; an accident occurs with probability $q(y)$, where $q'(y) < 0 < q''(y)$, implying that some third-party victims are injured, $h > 0$ being the size of the harm.

For any user-type $v \in [0, 1]$, the expected benefit of being endowed with a robot, denoted $B(a; v) \equiv a + (1 - a)v$, is increasing with a and v. Also, we consider that the following assumption is satisfied:

Assumption 1: For any $e > 0$, then: $\mathbb{C}(e) < c(e)$, and $\mathbb{C}'(e) < c'(e)$.

Assumption 1 requires that given the same intensity of effort, the autonomous mode for the robot proceeds at a lower cost and a lower marginal cost than the human-driven mode. This implies that $\min_e(\mathbb{C}(e) + q(e)h) < \min_e(c(e) + q(e)h)$, meaning that the autonomous mode for the robot proceeds at a lower cost of accidents than the human-driven mode.

2.2 Social optimum

The problem of the social planner is to select the set of human users who will hold a robot, $[v_{sw}, 1]$ (or equivalently, a level of output $Q_{sw} = 1 - F(v_{sw})$), to choose a level of care for each human user served, x, and a level of maintenance for the manufacturer per robot, y, and finally to choose a level of R&D investment, a, to maximize the expected social welfare function defined as follows (saving on notations for the decision variables)

$$W = a.\int_{v_{sw}}^{1} (1 - \mathbb{C}(y) - q(y)h)dF(v) + (1 - a).\int_{v_{sw}}^{1} (v - c(x) - q(x)h)dF(v) - \gamma(a). \quad (1)$$

The first (second) integral is the probability of state $\omega = Aut$ ($\omega = NAut$) times the expected benefit of using a robot net of the cost of accidentss in the autonomous (resp. human-driven) mode. It is easy to see that the decision of care and maintenance can be solved independently of the others. For each robot delivered to a human user, the social planner chooses

– the level of maintenance expenditures, required in state $\omega = Aut$, that minimizes the cost of accidents in this state: $y_{sw} \equiv \arg\min_y (\mathbb{C}(y) + q(y)h)$. We denote the associated expected cost of accidents is $C_M(y_{sw}; h) = \mathbb{C}(y_{sw}) + q(y_{sw})h$.

– the level of care, required in state $\omega = Aut$, that minimizes the cost of accidents in this state: $x_{sw} \equiv \arg\min_x (c(x) + q(x)h)$. We denote the associated expected cost of accidents is $C_H(x_{sw}; h) = c(x_{sw}) + q(x_{sw})h$.

A consequence of Assumption 1 is that the inequality: $C_M(y_{sw}; h) < C_H(x_{sw}; h)$ always holds.

Let us denote the expected social cost of accidents per robot (at efficient care and maintenance levels) as: $SC(a; h) = aC_M(y_{sw}; h) + (1 - a)C_H(x_{sw}; h)$. Social welfare can thus be written as: $\int_{v_{sw}}^{1} (B(a; v) - SC(a; h))\, dF(v) - \gamma(a)$. Hence, the social planner provides a human of type-v with a robot only if the expected benefit for this latter is larger than the expected cost associated: $B(a; v) \geq SC(a; h)$. This inequality may be written equivalently as $a(1 - C_M(y_{sw}; h)) + (1 - a)(v - C_H(x_{sw}; h)) \geq 0$; hence, a condition necessary for that inequality to hold, at least for a subset of users' types, is the following:

Assumption 2: $1 > C_H(x_{sw}; h)$.

Assume there exists a cut-off type $v_{sw} \in]0, 1[$ defined by the condition $B(a; v_{sw}) = SC(a; h)$. Rearranging; yields:

$$v_{sw} = C_H(x_{sw}; h) + \frac{a}{1 - a}(C_M(y_{sw}; h) - 1) \tag{PO}$$

For $0 < v_{sw} < 1$ to hold, it must be that: $C_H(x_{sw}; h) \in]c, 1+c[$ with $c \equiv \frac{a}{1-a}(1 - C_M(y_{sw}; h))$, meaning that it must be that the social cost of accidents in the human-driven mode is large enough, but not too large, compared with the net benefit of the autonomous mode. Assuming this is verified, only users-types $v \in [v_{sw}, 1]$ obtain a robot, such that the socially efficient number of robots is equal to $Q_{sw} = 1 - F(v_{sw})$.

Finally, the social planner chooses the level of R&D expenditures that maximizes

$$W(a) = \int_{v_{sw}}^{1} (B(a; v) - SC(a; h))dF(v) - \gamma(a) \tag{2}$$

where v_{sw} satisfies condition (PO). The derivative of social welfare w.r.t. a is given by:

$W'(a) = \int_{v_{sw}}^{1} (1 - v - C_M(y_{sw}; h) + C_H(x_{sw}; h))\, dF(v) - \gamma'(a)$.

Thus, the FOC for the optimal R&D investment, a_{sw}, is written as[7]

$$Q_{sw}.\left((1 - E(v|v \geq v_{sw})) + (C_H(x_{sw}, h) - C_M(y_{sw}, h))\right) = \gamma'(a_{sw}) \tag{3}$$

where v_{sw} satisfies condition (PO), and $E(v|v \geq v_{sw}) \equiv \int_{v_{sw}}^{1} v \frac{f(v)}{1-F(v_{sw})} dF(v)$ denotes the (conditional) average human-type endowed with a robot. The LHS term of (3) is the socially efficient output level times the unit social marginal benefit of R&D expenditures, which has two components: the first one is the social marginal benefit of allocating robots for the average human' type served (i.e., for a human having the conditional average type of the population served), $\frac{dB}{da}(a; E(v|v \geq v_{sw}))$ – this is the "composition effect", i.e., it is related to the selection of the population of robot users, and its size depends on the properties of the (marginal) distribution of users' types who are endowed with a robot; the second component is the marginal effect of R&D expenditures on the unitary expected total cost of accidents $\frac{dSC}{da}(a; h)$, and is equal to the difference between the cost of accidents related to the autonomous mode minus the cost of accidents related to the human-driven mode: increasing R&D investments (the probability that state $\omega = Aut$ is observed) is socially worthwhile because this allows a switch from a high-cost human-driven mode for robot use to a lower-cost autonomous mode. Finally, the RHS of (3) corresponds to the marginal cost of R&D expenditures.

To sum up, the social marginal benefit of R&D is positively related to three components (up to the marginal cost γ'): the first best output level; a composition effect, captured by one minus the conditional average robot user (which depends on specific properties of the probability distribution of human types); and finally, the saving in the social cost of

[7] It can be verified that; $W''(a) = -\gamma''(a) - \left(\frac{dv_{sw}}{da}\right)_{(OP)} f(v_{sw}) \left(1 - v_{sw} + C_H(x_{sw}; h) - C_M(y_{sw}; h)\right)$. Hence, given that $\left(\frac{dv_{sw}}{da}\right)_{(PO)} < 0$, the Second Order Condition is satisfied whenever γ'' is large enough.

accidentss associated with a switch from human-driven mode to autonomous mode of robots motion. The higher any one of these components, the higher the efficient level of R&D investment.

We then turn to the equilibrium analysis under alternative liability regimes.

3 The market for robots under different liability laws

The timing of decisions and events is as follows. At stage 0, courts set a liability regime for robot production (product liability law) and human use of robots (tort law). At stage 1, the robot manufacturer chooses a at a cost $\gamma(a)$; a is verifiable. At stage 2, the robot manufacturer sets a (monopoly) price P, and each human v considers whether to buy a robot (and buys whenever it is individually beneficial to be equipped with one). At stage 3, the robot manufacturer and robot users observe $\omega \in \{Aut, NAut\}$; either $\omega = Aut$ is realized, and thus the robot manufacturer chooses a level of maintenance y for each robot delivered; or $\omega = NAut$ is realized, and each robot user chooses his level of care x. At stage 4, the victim's harm is realized in case of accident, and the liability regime is enforced at trial.

tort law \ product liability	strict liability	negligence
strict liability	$D_H = h = D_M$	$D_H = h$, $D_M = \begin{cases} 0 & if\ y \geq y_{sw} \\ h & otherwise \end{cases}$
negligence	$D_H = \begin{cases} 0 & if\ x \geq x_{sw} \\ h & otherwise \end{cases}$, $D_M = h$	$D_H = \begin{cases} 0 & if\ x \geq x_{sw} \\ h & otherwise \end{cases}$, $D_M = \begin{cases} 0 & if\ y \geq y_{sw} \\ h & otherwise \end{cases}$

Table 1 – Compensation schemes under product liability law and tort law

We will assume that courts i) award full compensation for victims' harm, and ii) when the negligence rule is implemented, it is associated with due standards of precaution set at their socially optimal level (i.e., at y_{sw} for maintenance, and at x_{sw} for care activities). The relevant combinations of compensation schemes are described in Table 1, where $D_H(D_M)$

denotes the compensation paid to victims by the robot users (resp. the robot manufacturer) in the different liability regimes.

3.1 Equilibrium for a given liability law

We solve backward. At stage 3, it is straightforward to see that the decisions made by robot users and/or the robot manufacturer are consistent with those of the standard unilateral care model, whether they face strict liability or negligence:

Proposition 1. If $\omega = NAut$ is observed, robot users undertake x_{sw} whether they are subject to the strict liability rule or to the negligence rule. If $\omega = Aut$ is observed, the robot manufacturer undertakes y_{sw} whether it is subject to the strict liability rule or to the negligence rule.

The formal proof is a straightforward application of the standard analysis of the unilateral care model of accidents (left to the reader). The argument is that whenever the strict liability rule is enforced (in tort law and/or product liability), the injurer (either robot user or robot manufacturer) faces the expected social cost of the accident entailed by his decisions, like the social planner. Instead, whenever negligence is used (associated with the socially efficient due standard), then the injurer abides by the standard to avoid the payment of damages to victims.

Below we denote robot users' expected liability burden as $C_H(x_{sw}; \rho h) \equiv c(x_{sw}) + \rho q(x_{sw})h$ where $\rho = 1$ (resp. $\rho = 0$) if robot users are subject to the strict liability rule (resp. the negligence rule); and similarly the robot manufacturer' expected liability burden (per robot) will be denoted as $C_M(y_{sw}; \sigma h) \equiv \mathbb{C}(y_{sw}) + \sigma q(y_{sw})h$ where $\sigma = 1$ (resp. $\sigma = 0$) if the robot manufacturer is subject to the strict liability rule (resp. the negligence rule).

At stage 2, the market for robots clears as follows. Given the price P chosen by the manufacturer, a potential robot user with type v will buy a robot given the liability regime prevailing under tort law as long as its individual expected benefit is larger than its expected cost (total expected liability cost); i.e., if the following inequality holds: $B(a; v) - (1 -$

$a)C_H(x_{sw};\rho h) - P \geq 0$. Let us define the marginal buyer, denoted $v_e \in [0,1]$, as the human-type that satisfies the condition: $B(a;v_e) - (1-a)C_H(x_{sw};\rho h) = P$. Rearranging yields:

$$v_e = C_H(x_{sw};\rho h) + \frac{P-a}{1-a}. \tag{4}$$

Thus, all human types $v \in [v_e, 1]$ will buy a robot, such that the number of humans owning a robot at a given price P (i.e., the market demand) is equal to $1 - F(v_e)$.

The robot manufacturer sets a price P that maximizes the stage 2-expected profit, given the liability regime under product liability law, defined as

$$\pi(P) = \int_{v_e}^{1} (P - aC_M(y_{sw};\sigma h))\, dF(v) \tag{5}$$

where v_e is defined by condition (4). The derivative w.r.t. P is $\pi'(P) = \int_{v_e}^{1} 1dF(v) - (P - aC_M(y_{sw};\sigma h))\, f(v_e) \left(\frac{dv_e}{dP}\right)_{(4)}$, which interprets as the difference between the marginal market proceeds, minus the marginal cost corresponding here to the monopoly's expected liability burden per robot (which depends on the product liability regime that prevails). For an interior solution, the stage 2-monopoly price $P_e = P(a)$ is thus given by the following FOC[8]

$$P_e = aC_M(y_{sw};\sigma h) + (1-a)\left(\frac{1-F}{f}\right)_{|v_e}. \tag{6}$$

Condition (6) shows, as usual, that the monopoly price covers the robot manufacturer marginal cost (i.e., the expected liability burden, per robot), plus a mark-up. Thus, using (4) and (6), the stage 2-equilibrium marginal-type conditional to any given level of R&D investment, denoted as $v_e(a)$, is defined by:

$$v_e(a) = C_H(x_{sw};\rho h) + \frac{a}{1-a}\left(C_M(y_{sw};\sigma h) - 1\right) + \left(\frac{1-F}{f}\right)_{|v_e(a)}. \tag{ME}$$

[8]It can be verified that: $\pi''(P) = f(v_e)\frac{1}{1-p(a)}\left(\left(\frac{1-F}{f}\right)'_{v_e} - 1\right) < 0$. Hence, given that the condition $\left(\frac{1-F}{f}\right)' < 0$ holds at any v by assumption, the Second Order Condition is satisfied.

Once more, for $0 < v_e(a) < 1$ to hold, it must be that: $C_H(x_{sw}; \rho h) \in]\bar{c}, 1 + \bar{c}[$ where $\bar{c} \equiv \frac{a}{1-a}(1 - C_M(y_{sw}; \sigma h)) - \left(\frac{1-F}{f}\right)_{|v_e(a)}$, meaning now that it must be that users' liability burden in the human-driven mode is large enough, but not too large, compared with the benefit net of the manufacturer's liability burden of the autonomous mode. This implies that the stage-2 equilibrium output is $Q_e(a) = 1 - F(v_e(a))$. Using Assumptions 1, 2, and $\left(\frac{1-F}{f}\right)'_{v_e} < 0$, it can be verified that: $\left(\frac{dv_e}{da}\right)_{(ME)} = -\frac{1-C_M(y_{sw};\sigma h)}{(1-a)^2} \times \left(1 - \left(\frac{1-F}{f}\right)'_{v_e}\right)^{-1} < 0$ (meaning: the higher the R&D investment, the higher the stage 2-output).

At stage 1, the robot manufacturer chooses a level of R&D investment that maximizes $\Pi(a) = \pi(P(a)) - \gamma(a)$. Its derivative w.r.t. a is

$$\Pi'(a) = \int_{v_e}^{1} \left(\frac{\partial P(a)}{\partial a} - C_M(y_{sw}; \sigma h)\right) dF(v) - (P(a) - aC_M(y_{sw}; \sigma h)) f(v_e) \left(\frac{dv_e}{da}\right)_{(ME)} - \gamma'(a)$$

which means that increasing R&D expenditures yields an effect on the marginal profit per robot (integral term), plus a change in the composition of the robot users population (second term), and finally, an increase in the physical marginal cost of R&D. Rearranging and assuming an interior solution exists, the equilibrium level of R&D expenditures is defined by the following condition (assuming $\Pi(a) \geq 0$)[9]

$$Q_e.\left((1 - v_e) + (C_H(x_{sw}; \rho h) - C_M(y_{sw}; \sigma h))\right) = \gamma'(a). \quad (7)$$

where v_e satisfies condition (ME). The LHS term is the equilibrium output level times the private marginal benefit of R&D expenditures (per robot). Again, this later has two components: the first one is also a selection effect, corresponding here to the monopoly's marginal benefit of selling robots to the last human' types served by the monopoly (i.e., for the weakest type buying a robot), $\frac{dB}{da}(a_e; v_e)$; this selection effect goes in the same direction as the output effect. The second one captures that the level of R&D investment depends on the

[9] It can be verified that: $\Pi''(a) = -\gamma''(a_e) - \left(\frac{dv_e}{da}\right)_{(ME)} f(v_e) \left(1 - v_e + C_H(x_{sw}; \rho h) - C_M(y_{sw}; \sigma h) + \left(\frac{1-F}{f}\right)_{|v_e}\right)$. Hence, given that $\left(\frac{dv_e}{da}\right)_{(ME)} < 0$, the Second Order Condition is satisfied whenever γ'' is large enough once again.

difference in liability burdens between robots users and the manufacturer – the higher this difference (for a given output level), the higher the marginal benefit of R&D investment, and thus, the higher the level of R&D investment. Finally, the RHS is the marginal cost of R&D expenditures.

Hence, the monopoly's incentives to invest in R&D are also driven by three main components: the equilibrium output level; the composition effect which is now channeled by the marginal consumers, $1 - v_e$; and finally, the difference in the liability burden between the human-driven mode and the autonomous mode for robots.

The choice of the legal regime in tort law and product liability law affects each one of these components. Hence, the legal design has an influence on R&D investment at equilibrium through the allocation of the cost of accidents between the victims, the robot users (the demand for robot), and the manufacturer (the supply for robots), as we discuss now.

3.2 Comparison of alternative legal designs

First, we focus on the stage-2 equilibrium. This allows to illustrate how the allocation of the cost of accidentss between injurers and victims (through the combination of liability regimes) influences the supply and demand for robots. Let us denote $v_e^{i,j}(a)$ the solution to (ME) when tort law is associated with the liability rule $i \in \{sl, neg\}$, while product liability is associated with the liability rule $j \in \{sl, neg\}$. We obtain the following result:

Proposition 2: Conditional on any given level of R&D investment: i) The stage 2-output level when strict liability is used both in tort law and in product liability is smaller than the levels associated with each "mixed" legal regime combining tort law in one domain of law and negligence in the other; these latter are also smaller than the stage 2-output when negligence is used both in tort law and in product liability (i.e., $Q_e^{sl,sl}(a) < \inf\left\{Q_e^{sl,neg}(a), Q_e^{neg,sl}(a)\right\} < \sup\left\{Q_e^{sl,neg}(a), Q_e^{neg,sl}(a)\right\} < Q_e^{neg,neg}(a)$). ii) The stage 2-output level when tort law uses strict liability while negligence is applied in product liability is smaller than under the reverse legal design (tort law is based on negligence, while product liability uses strict liability) if the probability of using the robot in the autonomous mode is smaller than $\frac{1}{2}$ (i.e., $a \leq \frac{1}{2} \Rightarrow Q_e^{sl,neg}(a) < Q_e^{neg,sl}(a)$). In contrast, the stage 2-output level when tort law uses strict liability while negligence is applied in product liability is larger than under the reverse legal design only if the probability of using the robot in the autonomous mode is large enough (i.e., $a > \frac{q(x_{sw})}{q(x_{sw})+q(y_{sw})} \Rightarrow Q_e^{sl,neg}(a) > Q_e^{neg,sl}(a)$).

Proof. Under tort law (regardless of the product liability regime $j \in \{sl, neg\}$), a switch from negligence $\rho = 0$ to strict liability $\rho = 1$ corresponds to an increase in $C_H(x_{sw}; \rho h)$; this implies that the RHS in (ME) becomes larger: hence $v_e^{sl,j}(a) > v_e^{neg,j}(a)$. Similarly, under product liability (regardless of the liability regime under tort law $i \in \{sl, neg\}$), a switch from negligence $\sigma = 0$ to strict liability $\sigma = 1$ yields an increase in $C_M(y_{sw}; \sigma h)$; again this means that the RHS in (ME) becomes larger; hence $v_e^{i,sl}(a) > v_e^{i,neg}(a)$. As a consequence: i) both $v_e^{sl,sl}(a) > v_e^{sl,neg}(a) > v_e^{neg,neg}(a)$ and $v_e^{sl,sl}(a) > v_e^{neg,sl} > v_e^{neg,neg}(a)$ hold, however $v_e^{sl,neg}(a)$ and $v_e^{neg,sl}(a)$ cannot be compared generally. ii) Taking the difference between the expression of the RHS of (ME) under the combination (sl, neg) and under the combination (neg, sl), we obtain that (for any given $a > 0$) it is equal to: $\epsilon \equiv \left(q(x_{sw}) - \frac{a}{1-a}q(y_{sw})\right) h$, where by construction $q(x_{sw}) > q(y_{sw})$. Hence: $a \leq \frac{1}{2} \Rightarrow \frac{a}{1-a} < 1 \Rightarrow \epsilon > 0$, whereas $a > \frac{1}{2} \Rightarrow \epsilon \lesseqgtr 0$; more specifically, one may verify that for $\epsilon > 0$ to hold, it must be that $a > \frac{q(x_{sw})}{q(x_{sw})+q(y_{sw})}$. Hence the results. ■

Figure 1 represents conditions (4) and (6) in the plan (v_e, P_e) for the different combinations of liability regimes (see the Appendix for more details). It illustrates how the market for robots clears when liability laws change.

The expected benefit for robot users if strict liability is used in tort law is below the expected benefit when negligence is used in tort law; switching from negligence to strict liability in tort law pushes downwards the demand-price curve (condition (4)); simultaneously, the expected cost for the robot manufacturer when strict liability is used in product liability law is above its value when negligence is used in product liability; switching from negligence to strict liability in product liability pushes upwards the supply-price curve (condition (6)).

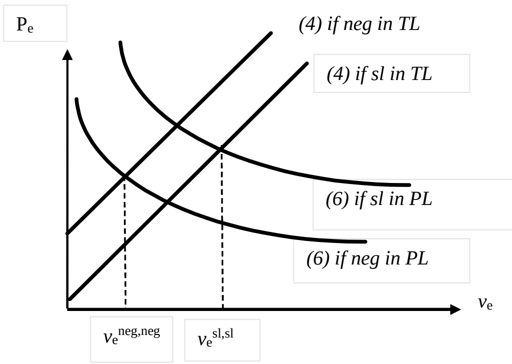


Figure 1 – The market for robots under alternative liability regimes
(sl: strict liability – neg: negligence – TL: tort law – PL: product liability)

Hence, the implementation of strict liability in both laws leads to the lowest number of robots, since both robot users and the robot manufacturer must consider the full expected cost of accident associated with each robot, which yields the lowest expected benefit for users, and highest expected cost for the monopoly. In contrast, the adoption of negligence, at least in one area of liability law, allows either robot users or the robot manufacturer to avoid compensating victims in case of accident. As a consequence, applying negligence in both areas of law provides the highest expected benefit for users and the lowest expected cost for the monopoly, which results in the highest output level.

Finally the comparison between both legal designs combining strict liability in one domain of the law and negligence in the other, is generally indeterminate; we only know that for each one, the equilibrium output is between the legal regimes where the same rule (either strict liability or negligence) is applied in both areas. A possible interpretation of the restrictions appearing in Part ii) of Proposition 2 relies to different stages of development/maturity of robot technology, as follows. When the technology is in its preliminary stage of development, the probability of autonomous mode activation is low, in the sense that $a < \frac{1}{2}$ (the human-driven mode is used more often); then the market for robots will reach a higher output level if negligence is applied in tort law while strict liability is applied in product liability, than if the reverse combination is used. The opposite occurs only when the technology is sufficiently advanced (i.e., the probability of activating the autonomous mode is much higher than $\frac{1}{2}$).

We now analyze how the combinations of liability regimes affects both the equilibrium levels of output and R&D investments (i.e., the solution to the system (EM)-(7)), denoted as $(Q_e^{i,j}, a_e^{i,j})$, where $i, j \in \{sl, neg\}$.

Proposition 3. i) Regardless of the liability regime in tort law, the equilibrium levels of output and R&D investment are higher when negligence rather than strict liability is implemented in product liability (i.e., $Q_e^{i,neg} > Q_e^{i,sl}$ and $a_e^{i,neg} > a_e^{i,sl}$ where $i \in \{sl, neg\}$). ii) Regardless of the liability regime in product liability: a) When the market for robots is weakly expanded, the equilibrium levels of output and R&D investment are higher when negligence rather than strict liability is implemented in tort law (i.e., $Q_e^{neg,j}$ is low enough $\Rightarrow Q_e^{neg,j} > Q_e^{sl,j}$ and $a_e^{neg,j} > a_e^{sl,j}$ where $j \in \{sl, neg\}$). b) When the market for robots is highly expanded, the equilibrium levels of output and R&D investment are lower when negligence rather than strict liability is implemented in tort law (i.e., $Q_e^{neg,j}$ is large enough $\Rightarrow Q_e^{neg,j} < Q_e^{sl,j}$ and $a_e^{neg,j} < a_e^{sl,j}$ where $j \in \{sl, neg\}$). iii) The equilibrium levels of output and R&D investments are higher when tort law uses strict liability while product liability uses negligence, than when the opposite combination is implemented (negligence in tort law, strict liability in product liability), only if the probability of using the robot in autonomous mode is large enough (i.e., $a > \frac{q(x_{sw})}{q(x_{sw})+q(y_{sw})} \Rightarrow Q_e^{sl,neg} > Q_e^{neg,sl}$ and $a_e^{sl,neg} > a_e^{neg,sl}$).

Proof. a) First, let us ignore the influence of the output and the composition effects on stage-1 R&D investments. On the one hand, according to the LHS in (7), for any given v_e the higher $C_H(x_{sw}; \rho h) - C_M(y_{sw}; \sigma h)$, the higher the level of investment $a(v_e)$. On the other hand, the following inequalities always hold (under tort law, negligence corresponds to $\rho = 0$, while strict liability is associated with $\rho = 1$; while under product liability, negligence is represented by $\sigma = 0$ and strict liability is the case where $\sigma = 1$):

$$\begin{aligned} C_H(x_{sw}; 0) - C_M(y_{sw}; h) &< C_H(x_{sw}; 0) - C_M(y_{sw}; 0) \\ &< C_H(x_{sw}; h) - C_M(y_{sw}; h) < C_H(x_{sw}; h) - C_M(y_{sw}; 0). \end{aligned} \quad (8)$$

Thus, this yields: $a^{neg,sl}(v_e) < a^{neg,neg}(v_e) < a^{sl,sl}(v_e) < a^{sl,neg}(v_e)$.

b) Let us now consider the influence of the output and the composition effects. Condition (7) shows that the lower v_e, the higher the marginal benefit of R&D, implying a higher level of R&D investment at equilibrium; in turn (given that $\left(\frac{dv_e}{da}\right)_{(ME)} < 0$), this later exerts a positive feedback effect on the (perfect Nash) equilibrium output. Hence, using Proposition 2, we may verify that:

i) $Q_e^{sl,sl}(a) < Q_e^{sl,neg}(a)$ and $a^{sl,sl}(v_e) < a^{sl,neg}(v_e)$ are both satisfied; hence we obtain $Q_e^{sl,sl} < Q_e^{sl,neg}$ and $a^{sl,sl} < a^{sl,neg}$. Similarly, $Q_e^{neg,sl}(a) < Q_e^{neg,neg}(a)$ and $a^{neg,sl}(v_e) < a^{neg,neg}(v_e)$ are both satisfied; thus, we also obtain $Q_e^{neg,sl} < Q_e^{neg,neg}$ and $a^{neg,sl} < a^{neg,neg}$. Hence the result.

ii) $Q_e^{sl,sl}(a) < Q_e^{neg,sl}(a)$ while $a^{sl,sl}(v_e) > a^{neg,sl}(v_e)$; similarly, $Q_e^{sl,neg}(a) < Q_e^{neg,neg}(a)$ while $a^{sl,neg}(v_e) > a^{neg,neg}(v_e)$. Hence, in this case the resulting influence of stage-2 output on R&D investment, together with the feedback effect on R&D investment on the output level is complex. We perform the formal analysis for comparative statics in the Appendix. Denoting

$$\begin{aligned}\alpha &= 1-\left(\frac{1-F}{f}\right)'_{|v_e^{neg,j}} > 0; \beta = \frac{1-C_M(y_{sw};\sigma)}{\left(1-(a_e^{neg,j})\right)^2} > 0;\\ \lambda &= f(v_e^{neg,j})\left(1-v_e^{neg,j}+C_H(x_{sw};\rho h)-C_M(y_{sw};\sigma h)+\left(\frac{1-F}{f}\right)_{|v_e^{neg,j}}\right) > 0,\end{aligned}$$

we obtain that, for any $j \in \{sl, neg\}$:

– if $Q_e^{neg,j} > \frac{\gamma''(a_e^{neg,j})}{\beta}$, then $Q_e^{neg,j} < Q_e^{sl,j}$ and $a^{neg,j} < a^{sl,j}$;

– if $\frac{\gamma''(a_e^{neg,j})}{\beta} > Q_e^{neg,j} > \frac{\lambda}{\alpha}$, then $Q_e^{neg,j} > Q_e^{sl,j}$ and $a^{neg,j} < a^{sl,j}$;

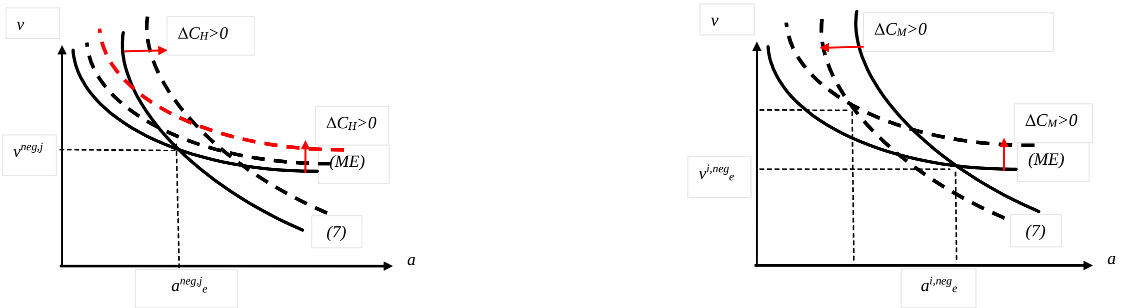
– if $\frac{\lambda}{\alpha} > Q_e^{neg,j}$, then $Q_e^{neg,j} > Q_e^{sl,j}$ and $a^{neg,j} > a^{sl,j}$.[10]

Hence the result.

iii) Given that $a_e^{sl,neg}(v_e) > a_e^{neg,sl}(v_e)$ holds, and that according to Proposition 2: $a > \frac{q(x_{sw})}{q(x_{sw})+q(y_{sw})} \Rightarrow Q_e^{sl,neg}(a) > Q_e^{neg,sl}(a)$, then we obtain $a > \frac{q(x_{sw})}{q(x_{sw})+q(y_{sw})} \Rightarrow Q_e^{sl,neg} > Q_e^{neg,sl}$ and $a_e^{sl,neg} > a_e^{neg,sl}$. Hence the result. ■

[10] In turn, we show that it is not possible for strict liability in tort law to yield a higher output and a lower R&D investment levels than negligence, i.e. $Q_e^{neg,j} < Q_e^{sl,j}$ and $a_e^{neg,j} > a_e^{sl,j}$ never hold where $j \in \{sl, neg\}$.

Figure 2 represents the graph of conditions (ME) and (7), illustrating how the equilibrium levels of output and R&D investment are affected by the combinations of liability laws. It is easy to verify that the slope (in absolute value) of (ME) is lower than the slope of (7) as a result of SOC: $-\left(\frac{dv_e}{da}\right)_{(ME)} < -\left(\frac{dv_e}{da}\right)_{(7)} = \frac{\gamma''(a)}{(1-F(v_e))+f(v_e)(1-v_e+C_H(x_{sw};\rho h)-C_M(y_{sw};\sigma h))}$.



$a)$ (neg, j) to (sl, j) $b)$ (i, neg) to (i, sl)

Figure 2 – Equilibrium output and R&D investments and liability costs

Let us consider simple changes in the legal design – such as a switch from negligence to strict liability in tort law (i.e., $C_H(x_{sw}; \rho h)$ increases) or in product liability (i.e., $C_M(y_{sw}; \sigma h)$ increases) – cf. Parts i) and ii) of Proposition 3.

On the one hand, switching from negligence to strict liability in tort law moves the curve (EM) upwards because this pushes downwards the demand-price curve as previously explained; hence, the level of output decreases for a given level of R&D investment; the curve (7) also shifts upwards because the difference in liability burdens between robots users and the manufacturer increases: investing more in R&D is beneficial for a given level of output, but this effect is dampened by the decrease in output. Overall, the net effects on the output and R&D at equilibrium are indeterminate (Figure 2a)).

On the other hand, when product liability switches from negligence to strict liability, the curve (EM) shifts upwards, since the supply-price curve is pushed upwards; hence the level of output decreases at any level of R&D investment; in turn, the curve (7) now shifts downwards because the difference in liability burdens between robots users and the manufacturer decreases: investing less in R&D reduces the impact on the manufacturer's liability burden at any level of output. Hence, the net effects on the output and R&D at equilibrium are negative (Figure 2b)).[11]

When the liability regime changes simultaneously in both areas of law, this introduces much more complexity, and the impact on the equilibrium is often indeterminate. Part ii) of Proposition 3 suggests – still considering technology maturity – clear effects arise in more specific circumstances : at the most advanced stages (i.e., a is much larger than $\frac{1}{2}$), strict liability in tort law combined with negligence in product liability allows for more output and greater investment in R&D than negligence in tort law combined with strict liability in product liability. Nevertheless, let us emphasize here that the condition $a > \frac{q(x_{sw})}{q(x_{sw})+q(y_{sw})}$ raises a practical issue. It is reasonable to anticipate that at very advanced stages of the technology, the probability of accident in the autonomous mode, $q(y_{sw})$, is also much lower than at preliminary stages – it is expected that in the future, not only will the autonomous mode be much safer than the human-driven mode, but also accidents will be very rare. The consequence for the discussion here, is that it may become less likely that the inequality $a > \frac{q(x_{sw})}{q(x_{sw})+q(y_{sw})}$ is satisfied.[12]

This said, Proposition 3 implies that it is generally not possible to have a complet ranking of the different legal designs with respect to their effects on the output and on R&D investments. Nonetheless, Proposition 3 has clear implications regarding the potential effects of changes in liability regimes compared to the current legal framework (negligence in tort law, strict liability in product liability): Shifting to strict liability in tort law in the very short run (when the market is weakly developed) would entail negative effects for the robot market, with a decrease in both the output and the R&D investment levels. In the long run, however (whenever the market is highly expanded), shifting to strict liability would have positive effects for the robot market, with a increase both in the output and the R&D investment levels. In contrast, our results suggest that adopting negligence also in product liability would allow an increase both of the output and the R&D investment levels, in the

[11] Remind that a is the probability of using the autonomous mode; therefore, the higher a is; the more likely the manufacturer is to be liable.

[12] As $q(y_{sw})$ decreases, then $\frac{q(x_{sw})}{q(x_{sw})+q(y_{sw})}$ is closer to 1.

short run as well as in the long run.

4 Issues in welfare analysis

We discuss here different issues related to the effect of legal design on social welfare. We first consider the differences in terms of monopoly distortion and social welfare between the alternative combinations of liability regimes. Then, we discuss the implications of the coordination (failure) of output regulation and liability laws.

4.1 Monopoly distortion, social welfare, and the legal design

We first compare the equilibrium outcome of the different legal designs with the social optimum.

Proposition 4. Compared with the first best solution: i) When strict liability is implemented both in tort law and product liability, then the monopoly provides not enough robots and underinvests in R&D if the condition $v_e^{sl,sl}(a) > E(v|v \geq v_{sw})$ is satisfied. ii) In the case where $v_e^{neg,j}(a) > v_{sw}(a)$ holds when negligence is implemented in tort law (regardless of the product liability regime), then the monopoly provides not enough robots and underinvests in R&D if the condition $v_e^{neg,j} > E(v|v \geq v_{sw})$ is satisfied. iii) In the case where $v_e^{sl,neg}(a) < v_{sw}(a)$ holds when strict liability is implemented in tort law while negligence is applied in product liability, then the monopoly provides too many robots and overinvests in R&D.

Proof. A) Consider conditions (ME) and (PO). We first show that conditional on any given level of R&D investment, the stage 2-output level under the combination (sl, sl) is smaller than the first best output level; in contrast, for any combination of liability rules, it can be either higher or lower. It is obvious that:

– Under the combination (sl, sl), since both the robot manufacturer and a robot users face a total liability cost which is equal to the social cost of the accident, the RHS in

(ME) is larger than the RHS in (PO) (because of the term $\left(\frac{1-F}{f}\right)_{|v_e(a)}$), implying thus $v_e^{sl,sl}(a) > v_{sw}(a) \Leftrightarrow Q_e^{sl,sl}(a) < Q_{sw}(a)$.

– Under any other combination, either the robot manufacturer or robot users face a liability burden which is smaller than the social cost of the accident; thus (despite the influence of the term $\left(\frac{1-F}{f}\right)_{|v_e(a)}$), the RHS in (ME) may be lower as well as larger than the RHS in (PO). Hence, the monopoly may provide either not enough or too much output.

B) Consider now conditions (7) and (3).

– Under the combination (sl, sl), given that we have $Q_e^{sl,sl}(a) < Q_{sw}(a)$, a condition sufficient for the RHS in (7) to be lower than the RHS in (3) is given by $1 - v_e^{sl,sl}(a) \leq 1 - E(v|v \geq v_{sw}(a))$.[13] Assuming this restriction holds, then we obtain that $a_e^{sl,sl}(v_e) < a_{sw}(v)$, implying (given the feedback influence of R&D investments on the output level) that under (sl, sl), the monopoly provides not enough output $Q_e^{sl,sl} < Q_{sw}$ and not enough R&D investments $a_e^{sl,sl} < a_{sw}$. Otherwise, the comparison is indeterminate.

– Assume now that negligence is used under tort law (regardless of the liability rule under product liability). Since under the combination (neg, j) (for any $j \in \{sl, neg\}$) we have (see the inequalities in (8)):

$$C_H(x_{sw}; 0) - C_M(y_{sw}; \sigma h) < C_H(x_{sw}; 0) - C_M(y_{sw}; h),$$

then a sufficient condition for the RHS in (7) to be smaller than the RHS in (3) is that both $v_e^{neg,j}(a) < v_{sw}(a)$ and $1 - v_e^{sl,sl}(a) < 1 - v_{sw}(a)$ are satisfied – implying thus that $a_e^{neg,j}(v_e) < a_{sw}(v)$, and finally, at equilibrium: $Q_e^{neg,j} < Q_{sw}$ and $a_e^{neg,j} < a_{sw}$. Otherwise, the comparison is indeterminate.

– Finally, assume that strict liability is implemented in tort law, while negligence is applied in product liability. Under the combination (neg, sl), we have (see again the inequalities in (8):

[13] This condition is generally more restrictive than the condition $v_e^{sl,sl}(a) \geq v_{sw}(a)$, and depends on the characteristics of the probability distribution function for v, such as its concentration or asymmetry. It is satisfied when $v \in [0, 1]$ is distributed according to the uniform law, since then: $E(v|v \geq v_{sw}) = \frac{1+v_{sw}}{2}$. For the broader family of triangular PDF, we have checked that additional restrictions related to the mode are required .

$$C_H(x_{sw};h) - C_M(y_{sw};0) > C_H(x_{sw};0) - C_M(y_{sw};h).$$

Thus, a sufficient condition for the RHS in (7) to be smaller than the RHS in (3) is that $v_e^{sl,neg}(a) > v_{sw}(a)$ is satisfied (hence $1 - v_e^{sl,neg}(a) > 1 - v_{sw}(a)$ also holds) – implying thus that $a_e^{sl,neg}(v_e) > a_{sw}(v)$, and finally, at equilibrium: $Q_e^{sl,neg} > Q_{sw}$ and $a_e^{sl,neg} > a_{sw}$. Otherwise, the comparison is indeterminate. ■

Proposition 4 illustrates that regardless of the legal design, the monopoly provides both an inefficient number of robots and an inefficient level of R&D investments. However, it is generally not possible to determine whether the distortions correspond to under- or over-provision of output and/or investment. Additional restrictions are required to compare the market outcome with the social optimum, which depends on the combination of liability regimes in both areas of liability laws; these restrictions are required to qualify the "selection effect", and/or the output effect, that drive the incentives for R&D investments.

We illustrate now that the welfare comparisons between the different legal designs are generally also indeterminate, unfortunatly. Given the prevailing legal design, social welfare at an equilibrium $(v_e^{i,j}, a_e^{i,j})$ that satisfies (ME)-(7), is given by (see (1)):

$$W = \int_{v_e^{i,j}}^{1} \left(a_e^{i,j}.(1 - C_M(y_{sw};h)) + (1 - a_e^{i,j}).(v - C_H(x_{sw};h)) \right) dF(v) - \gamma(a_e^{i,j}). \qquad (9)$$

A switch from negligence to strict liability under tort law (resp. product liability) translates into an increase in robot users' liability costs (the manufacturer's liability cost) such that $dC_H(y_{sw}, \rho h) = q(x_{sw})h.d\rho$ (resp. $dC_M(y_{sw}; \sigma h) = q(y_{sw})h.d\sigma$). According to (9), by definition social welfare doesn't depend on injurers' liability burden, but is related to the full cost of accidentss (costs of care and maintenance, plus the harm to victims); hence, the effect of a change in the liability regime on social welfare arrives only through the indirect effects on output and R&D investments (in turn, there is no direct effect). Taking the total derivative of (9) yields:

$$dW = -dv_e^{i,j} \times f(v_e^{i,j}) \left(a_e^{i,j}.(1 - C_M(y_{sw}; h)) + (1 - a_e^{i,j}).(v_e^{i,j} - C_H(x_{sw}; h))\right)$$
$$+da_e^{i,j} \times \left(\int_{v_e^{i,j}}^{1} \left(1 - v_e^{i,j} + C_H(x_{sw}; h) - C_M(y_{sw}; h) - \frac{\gamma'(a_e^{i,j})}{1 - F(v_e^{i,j})} \right) dF(v) \right).$$

The signs of the term in brackets in the first line and second lines (multiplicand of $dv_e^{i,j}$ and $da_e^{i,j}$ respectively) is generally indeterminate. The multiplicand of $dv_e^{i,j}$ is the social expected (net) benefit of alloting a robot to the marginal user $v_e^{i,j}$ – where $v_e^{i,j} - C_H(x_{sw}; h)$ is always positive when $v_e^{i,j} > v_{sw}$ (as under under (sl, sl)), but it may take different signs under other combinations of liability regimes. The multiplicand of $da_e^{i,j}$ is the total (net) benefit of an additional unit of R&D; given that any equilibrium is inefficient, its sign is also indeterminate generally. Substituting with (7) in the second line, and rearranging yields:

$$dW = -dv_e^{i,j}.f(v_e^{i,j}) \left(a_e^{i,j}.(1 - C_M(y_{sw}; h)) + (1 - a_e^{i,j}).(v_e^{i,j} - C_H(x_{sw}; h))\right)$$
$$-da_e^{i,j}.\left(1 - F(v_e^{i,j})\right).\begin{pmatrix} (E(v|v \geq v_e^{i,j}) - v_e^{i,j}) + (C_M(y_{sw}; h) - C_M(y_{sw}; \sigma h)) \\ -(C_H(x_{sw}; h) - C_H(x_{sw}; \rho h)) \end{pmatrix},$$

where $E(v|v \geq v_e^{i,j}) \equiv \int_{v_e^{i,j}}^{1} v \frac{f(v)}{1-F(v_{sw})} dF(v) (> v_e^{i,j})$ denotes the (conditional) average type of users who buy a robot at equilibrium. On the one hand, we have by construction: $E(v|v \geq v_e^{i,j}) - v_e^{i,j} > 0$; however, the size/sign of the difference between the two other terms depends on the change in liability laws we consider. On the other hand, as shown in the Appendix (see the comparative statics analysis; see also Proposition 3), it comes that either $dv_e^{i,j}$ and $da_e^{i,j}$ have opposite signs, or the sign of $da_e^{i,j}$ is indeterminate. Hence, the sign of dW is generally indeterminate.

A numerical application would be useful to assess W directly in different legal regimes and to illustrate the different influences; however, it is beyond the scope of the current work. The properties of the probability distribution for users' type may play an important role (at least through the size of $v_e^{i,j}$ and $E(v|v \geq v_e^{i,j})$), aside from institutional and technological characteristics.

4.2 Combining market regulation and liability laws

4.2.1 Efficient regulation and price discrimination

Assume that a public regulator sets the monopoly's output at the socially optimal level, Q_{sw}; and also allows the manufacturer to use perfect price discrimination. Then, each unit of the efficient output – i.e., for any $v \in [v_{sw}, 1]$ – is sold at a specific price defined according to (using (4)):

$$P(v) = a + (1-a)v - (1-a)C_H(x_{sw}; \rho h). \tag{10}$$

Thus, equilibrium R&D investment under the combination (i, j), denoted $\tilde{a}_e^{i,j}$, maximizes the stage 1-profit $\Pi(a) = \pi(P(v), Q_{sw}) - \gamma(a)$. After substituting (10), this can be rewritten as $\Pi(a) = W(a) + Q_{sw}.\Lambda(a; h)$, where $\Lambda(a; h) \equiv SC(a; h) - SC(a; \rho h, \sigma h) \geq 0$ is the difference between the expected cost of accidents at the optimum, and the total expected liability burden under the prevailing liability regimes.[14] The following table shows the expression of $\Lambda(a; h)$, and how it varies with a in the different the legal designs.

	(sl, sl)	(sl, neg)	(neg, sl)	(neg, neg)
$\Lambda(a; h)$	0	$aq(y_{sw})h$	$(1-a)q(x_{sw})h$	$(1-a)q(x_{sw})h + aq(y_{sw})h$
$\frac{d\Lambda}{da}(a; h)$	$n.d.$	$q(y_{sw})h$	$-q(x_{sw})h$	$-(q(x_{sw}) - q(y_{sw}))h$

Table 2 – Absolute and marginal benefit of price discrimination

This useful to arrive at the following results:

Proposition 5. Assume that according to public regulation, the monopoly delivers the efficient output level and is allowed to use perfect price discrimination. Then: i) When strict liability is implemented both in tort law and product liability, the monopoly chooses the first best level of R&D investment (i.e., $\tilde{a}_e^{sl,sl} = a_{sw}$). ii) When strict liability is used in tort law, while negligence is used in product liability, the monopoly overinvests in R&D (i.e., $\tilde{a}_e^{sl,neg} > a_{sw}$). iii) When negligence is used under tort law (regardless of the liability rule

[14] Developping, we have: $SC(a; \rho h, \sigma h) = aC_M(y_{sw}; \sigma h) + (1-a)C_H(x_{sw}; \rho h)$
$= a\mathbb{C}(y_{sw}) + (1-a)c(x_{sw}) + (1-\sigma)p(a)q(y_{sw})h + (1-\rho)(1-p(a))q(x_{sw})h.$

under product liability), the monopoly may overinvest as well as underinvest in R&D (i.e., $\tilde{a}_e^{neg,j} \lesseqgtr a_{sw}$ for $j \in \{sl, neg\}$).

Proof. Part i) is straightforward. For the three other combinations, the FOC $\Pi'(a) = 0$ can be written as:

$$W'(a) + \Lambda(a;h).f(v_{sw})\left(-\frac{dv_{sw}}{da}\right)_{(PO)} + Q_{sw}.\frac{d\Lambda}{da}(a;h) = 0. \tag{11}$$

According to Table 2: the second LHS term in (11) is always positive; the third LHS term is positive only for the combination (sl, neg). Hence result ii). The third LHS term is negative when negligence is used at least in tort law; hence result iii). ■

The situation where strict liability is used in both areas of law mirrors the standard result in textbooks: the monopoly maximizes social welfare. The rationale here is that each kind of injurer faces the full social cost of robots associated with its actions (i.e., the social cost of accidents). The result no longer holds when negligence is used in one area of law at least: when it is used in product liability (resp. tort law), the monopoly produces (resp. sell) each unit of the efficient output at lower operating costs (resp. for a higher price), which generates a monopoly rent above social welfare. As a result, it becomes profitable to depart from the first best level of R&D investment.

This rent increases with R&D investments when robot users face strict liability whereas the monopoly faces negligence: raising R&D above the first best level is beneficial to the monopoly because this way, it becomes more likely that the autonomous mode is activated such that each user accepts to pay a higher price (since their expected liability burden decreases), while each unit is produced at a lower operative cost. In contrast, the rent may decrease with R&D investments when robot users face negligence because the price effect is annulled; hence it may be rational for the monopoly to cut R&D investments below the first best level.

Proposition 5 illustrates the coordination problem that exists between the legal design and the tools available to regulators. Efficient regulation solves the output distortion, and neutralizes the selection effect at the same time. However, for the monopoly equilibrium to coincide with the social optimum, the strict liability rule should be implemented in both areas of law, alongside allowing perfect price discrimination. Conversely, either using the negligence rule in one or both areas of law (despite perfect price discrimination), or forbidding perfect price discrimination (despite strict liability is adopted both in tort law and in product liability) would lead the monopoly to choose an inefficient level of R&D investments.[15]

4.2.2 Full market coverage and price discrimination

We show here that when the monopoly finds it profitable to cover the market, then perfect price discrimination generally fails to reach the first best level of investment, regardless of the design of liability laws.

When the monopoly fully covers the market ($Q_e = 1 > Q_{sw}$), and uses perfect price discrimination, then the price is given by (9) for each $v \in [0,1]$. Thus, stage 1-profit can be written as $\Pi(a) = \int_0^1 (B(a;v) - SC(a;\rho h, \sigma h))dF(v) - \gamma(a)$. For an interior solution, the equilibrium level of R&D expenditures $\bar{a}_e^{i,j}$ under liability regimes (i,j), satisfies now the following FOC (assuming $\Pi(\bar{a}_e^{i,j}) \geq 0$)

$$1 - E(v) + C_H(x_{sw}; \rho h) - C_M(y_{sw}; \sigma h) = \gamma'(a_e) \tag{12}$$

where $E(v) = \int_0^1 v dF(v)$ denotes the (unconditional) average type. The interpretation of (12) is very similar to (7). The main difference between (12) and (7) (apart of the output effect), relates to the selection effect (first term in the LHS); indeed, when the monopoly covers the market, no selection effect occurs by construction (all human-types are served and buy a robot). It is straightforward that the equivalent of Proposition 4 is obtained here:

Proposition 6. Assume that the monopoly fully covers the market and uses perfect price discrimination; then: i) The monopoly may overinvest as well as underinvest in R&D regardless of the combination of liability regimes in tort law and product liability law. ii)

[15] In the Appendix, we show that if the monopoly is constrained to sell all robots at a unique price such that it breaks even, then it does not choose the first best R&D investment level.

When strict liability is used in tort law (regardless of the liability rule under product liability), the monopoly overinvests in R&D if $E(v) < E(v|v \geq v_{sw})$.

Proof. Using an argument similar to the one in Proposition 4, given the inequalities in (8), it is straightforward to show that $E(v) < E(v|v \geq v_{sw}) \Rightarrow \bar{a}_e^{sl,j} > a_{sw}$ for $j \in \{sl, neg\}$. ■

A first implication of proposition 6 is as follows: Whether covering the market represents a significant market expansion justifying the monopoly to expand its R&D investments above the socially optimal level, depends on the (statistical) properties of the distribution of human type (for example, whether it is left or right skewed). Part ii) of Proposition 6 illustrates that when the perceived market expansion is large enough (in the sense that $1 - E(v) > 1 - E(v|v \geq v_{sw})$) and robot users face strict liability (irrespective of the liability regime for the monopoly) , then the monopoly have incentives to invest more in R&D than what is socially optimal: price discrimination among all human types together with robot users facing strict liability enables a large marginal benefit for R&D investment, combined with a lower marginal cost for R&D expenditures. Nevertheless, the market expansion effect is not guaranteed – it occurs only under specific conditions regarding the distribution of robot users' types.

In turn, whenever robot users face negligence, the comparison with the social optimum becomes indeterminate again, since the second term on the LHS in (12) (which captures the combined increase in market price and manufacturer liability cost) is lower than in (3) (the cut in the social cost of accidents at optimum).

5 Conclusion

In this paper, we extend "the unilateral care model of accidents" to deal with some specific features of advanced robot-based technologies (possibility of two mode of motions, fully autonomous vs human-driven robots), and analyze how the legal design adopted for liability laws may affect R&D investments and the market for robots. We have shown that even in a case where any legal design (any combination of liability rules between tort law and product liability) provides efficient incentives in the prevention of accidents (i.e., all combinations are equivalent, and arrive at the social goal in terms of prevention), each combination of tort law and product liability yields a different outcome in terms of output and R&D investment. The reason is that the different combinations are not equivalent regarding the allocation of the cost of accidents between robot users, the manufacturer and victims – each legal design implements a specific cost allocation which is decisive for the market for robots and the incentives to invest in the development of the technology.

The current setting is already of broad application, since it may be understood as a application to the market for algorithms broadly speaking, or more restrictive the market for autonomous vehicles, or for medical robots. Remark also in passing that despite our work is focused on a world with highly advanced robots having two modes of motion, it can be easily extended to fit a world where humans may choose between, on the one hand, autonomous robots (having two alternative modes of motion), and on the other hand non autonomous robots (a unique mode of motion), both coexisting, such that we can interpret throughout the paper all humans types $v \in [0, v_e)$ as using a non autonomous robot.[16]

This said we provide several important highlights regarding the influence of liability laws on the incentives to invest in algorithms/robots development, in a set-up where a priori liability regimes are efficient. In our framework, the first-stage level of R&D investments mainly depends on the properties of the second stage market equilibrium, this latter reflecting the design of liability laws: the monopoly price adjustment is driven by the liability

[16] Suppose that upon the monopoly market for autonomous robots considered in the paper, a competitive market for standard/non-autonomous robots exists where robots are produced at zero marginal cost (thus the equilibrium competitive price is null). Let V be the expected benefit (net of liability cost) for owners of such robots. Since V appears as a constant on the LHS of (4), we can set $V = 0$ wlog. Hence v_e is the equilibrium number of non-autonomous robots. As a result, our framework encompasses the analysis of the market for Autonomous Vehicles by de Chiara et al. [2021] who assume two-type of drivers (high/low type). In de Chiara et al. [2021] the equilibrium output cannot take but two values: the monopoly provides the efficient output (only one type is served), or fully covers the market (the two types are served) – hence the "composition effect" and the "output effect" are always determinate and go in the same direction, in contrast to the continuous-type model introduced here. Thus our analysis provides a broader set of predictions, with also more general results consisting in more various market outcomes than in de Chiara et al. [2021], except our Proposition 6 which is similar to Remark 4.1 in de Chiara et al. [2021].

regime that consumers face (tort law, with a demand price lower under strict liability than negligence), while the monopoly cost is driven by court (with a liability burden higher under strict liability than negligence), implying differentiated effects on the level of output and on the selection of consumers' type. In all, the incentives to invest in R&D for robots are channeled by three influences: a first one is through the monopoly output, a second one is related to the characteristics/composition of the population of robots buyers, and a third one comes from the size of the difference in total private expenditures for preventing accidental harms to third-party victims between the autonomous and the human-driven modes for robots. We have shown that different combinations of liability regimes in court and tort law entail different effects on each of these three components.

In particular, the general use of strict liability recently promoted by the European Commission, has no clear advantages over the negligence rule neither in terms of output level furnished to the market, nor in terms of R&D investments level attained. One exception discussed here is when three conditions are met: the monopoly uses perfect price discrimination and faces strict liability, the output is regulated, and simultaneously consumers also face strict liability – in this case the monopoly market equilibrium replicates the social optimal. But to the least, this raises a coordination problem between different public authorities. Extending the analysis to alternative market structures will be useful in this perspective, since different market power conditions will attain a different (consumers) selection effect, as well as a different range of output expansion. But it can be anticipated that the properties of equilibria will be very dependent on behavioral determinants (pricing strategies), the existence of a market intervention, and the characteristics of the legal system. This will be the topic of future research.

References

ABRAHAM K. et RABIN R. [2019], ‹‹ Roboted Vehicles and Manufacturer Responsibility for Accidents: a New Legal Regime for a New Era ››, *Virginia law Review*, 105, p. 127-171.

BUITEN M. [2024], Product liability for defective IA, *European Journal of Law and Economics*, 57, p. 239-273.

CALABRESI G. [1970], *The cost of accidentss - A Legal and Economic Analysis*, Yale University Press.

CHARREIRE M. et LANGLAIS E [2021], « Should environment be a concern for competition policy when firms face environmental liability ? », *International Review of Law & Economics*, 67, p. 1-16.

CHOPARD B. et MUZY O. [2023], ‹‹ Market for artificial intelligence in health care and compensation for medical errors ››, *International Review of Law and Economics*, 75, Article 106153.

DAVOLA A. [2018], ‹‹ A Model for Tort Liability in a World of Driverless Cars: Establishing a Framework for the Upcoming Technology ››, *Idaho Law Review*, 54, p. 592-614.

DE CHIARA A., ELIZALDE I., MANNA E., et SEGURA-MOREIRAS A. [2021], ‹‹ Car accidents in the age of robots ››, *International Review of Law and Economics*, 68, Article 106022.

DI X., CHEN X., et TALLEY E. [2020], ‹‹ Liability design for autonomous vehicles and human-driven vehicles: a hierarchical game-theoretic approach ››, *Transportation Research Part C*, 118, p. 1-36.

EUROPEAN COMMISSION [2022a], *Directive of the European Parliament and of the Council, on liability for defective products*, COM(2022) 395 final-2022/0302 COD.

EUROPEAN COMMISSION [2022b], *Directive of the European Parliament and of the Council, on adapting non-contractual liability rules to artificial intelligence*, COM(2022) 396 final-2022/03032 COD.

GUERRA A., PARISI F., et PI D. [2022a], ‹‹ Liability for robots I: legal challenges ››, *Journal of Institutional Economics*, 18, p. 331–343.

GUERRA A., PARISI F., et PI D. [2022b], ‹‹ Liability for robots II: an economic analysis ››, *Journal of Institutional Economics*, 18, p. 353-568.

HAMILTON S. [1998], ‹‹ Taxation, Fines, and Producer Liability Rules: Efficiency and Market Structure Implications ››, *Southern Economic Journal*, 65, p. 140-150.

HAMILTON S., RIDLAND H., et SUNDING D. [2023], ‹‹ Imperfect Competition and the Optimal Deterrence of Environmental Accidents ››, Available at SSRN: https://ssrn.com/abstract=443546

KIM J-Y [2024], ‹‹ Law and Economics of Artificial Intelligence: optimal liability rule for accident losses caused by fully autonomous vehicles ››, *The Korean Economic Review*, 40 (1), p. 49-75.

LEMLEY M. et CASEY B. [2019], ‹‹ Remedies for robots ››, *The University of Chicago Law Review*, 86, p. 1311-1396.

OBIDZINSI M. et OYTANA Y. [2022], ‹‹ Advisory algorithms and liability rules ››, *Working Papers 2022-04*, CRESE.

SHAVELL S. [2020], ‹‹ On the redesign of accident liability for the world of Autonomous Vehicles ››, *Journal of Legal Studies*, 49, p. 243-285.

TALLEY E. [2019], ‹‹ Robotorts: How should accident law adapt to Autonomous Vehicles? Lessons from Law and Economics ››, Standford University, Hoover IP2 Working Paper Series No.19002.

Appendix

Interpretation of Proposition 2 in terms of equilibrium outcome in the robot market. Let us denote $v_e^{i,j}(a)$ and $P_e^{i,j}(a)$ as the solution to the system (4)-(6) when tort law is associated with the liability regime $i \in \{sl, neg\}$, while product liability is associated with liability regime $j \in \{sl, neg\}$.

On the one hand, according to the definition of $v_e^{i,j}(a)$ we have the following relationship between $v_e^{i,j}$ and $P_e^{i,j}(a)$ after rewriting (4) as

$$P_e^{i;j}(a) = B(a; v_e^{i,j}(a)) - (1-a)c(x_{sw}) - \begin{cases} (1-a)q(x_{sw})h, & \text{under strict liability in tort law} \\ 0, & \text{under negligence in tort law} \end{cases}, \tag{4bis}$$

corresponding to the upward sloping curve in Figure 1, with a slope which equals to $\left(\frac{dP_e^{i,j}}{dv_e^{i,j}}\right)_{|(4bis)} = 1 - a > 0$. Thus according to (4bis) all else equal the monopoly price is higher at any level of v_e when the negligence rule is used in tort law, than when strict liability is used. On the other hand, according to (6), the stage 2-equilibrium monopoly price, a second relationship between $v_e^{i,j}$ and $P_e^{i,j}(a)$ can be written as

$$P_e^{i,j}(a) = (1-a).\left(\frac{1-F}{f}\right)_{|v_e^{i,j}} + a\mathbb{C}(y_{sw}) + \begin{cases} aq(y_{sw})h, & \text{under strict liability in product liability} \\ 0, & \text{under negligence in product liability} \end{cases}, \tag{6bis}$$

corresponding to the downward sloping curve in Figure 1, with a slope which equals to $\left(\frac{dP_e^{i,j}}{dv_e^{i,j}}\right)_{|(6bis)} = (1-a)\left(\frac{1-F}{f}\right)'_{|v_e^{i,j}} < 0$. Thus according to (6bis), all else equal, the monopoly price is lower when the negligence rule is used in product liability than when strict liability is used.

As a result, the combination (sl,sl) leads to the highest threshold v_e and (neg,neg) to the lowest threshold v_e, while the other two combinations result in values that fall in between: $v_e^{sl,sl}(a) > \sup\left\{v_e^{sl,neg}(a), v_e^{neg,sl}(a)\right\} > \inf\left\{v_e^{sl,neg}(a), v_e^{neg,sl}(a)\right\} > v_e^{neg,neg}(a)$. The comparison of (neg,sl) and (sl,neg) yields an ambiguous result in general. Hence the result:

$$\begin{aligned} 1-F(v_e^{sl,sl}(a)) &< \inf\left\{1-F(v_e^{sl,neg}(a)),1-F(v_e^{neg,sl}(a))\right\} \\ &< \sup\left\{1-F(v_e^{sl,neg}(a)),1-F(v_e^{neg,sl}(a))\right\} < 1-F(v_e^{neg,neg}(a)). \end{aligned}$$

As explained in the text, if $(1-a)q(x_{sw})h > aq(y_{sw})h$, then we obtain $v_e^{sl,neg}(a) > v_e^{neg,sl}(a)$, and thus $1-F(v_e^{sl,neg}(a)) < 1-F(v_e^{neg,sl}(a))$.

■

Comparative statics analysis (Proposition 3).

Differentiating (ME)-(7), it can be checked that we obtain

$$\begin{aligned} \alpha dv_e + \beta da &= dC_H + \frac{a}{1-a}dC_M \\ \lambda dv_e + \gamma''(a)da &= (1-F(v_e))\left(dC_H - dC_M\right) \end{aligned}$$

where C_H and C_M set for $C_H(x_{sw};\rho h)$ and $C_M(y_{sw};\sigma h)$ respectively, and

$$\begin{aligned} \alpha &= 1-\left(\frac{1-F}{f}\right)'_{|v_e} > 0; \beta = \frac{1-C_M(y_{sw};\sigma h)}{(1-a)^2} > 0; \\ \lambda &= f(v_e)\left(1-v_e+C_H(x_{sw};\rho h)-C_M(y_{sw};\sigma h)+\left(\frac{1-F}{f}\right)_{|v_e}\right) > 0. \end{aligned}$$

Solving yields

$$\begin{aligned} dv_e &= \frac{\gamma''(a)\left(dC_H+\frac{a}{1-a}dC_M\right)-\beta(1-F(v_e))\left(dC_H-dC_M\right)}{\Delta} \\ da &= \frac{\alpha(1-F(v_e))\left(dC_H-dC_M\right)-\lambda\left(dC_H+\frac{a}{1-a}dC_M\right)}{\Delta} \end{aligned}$$

with $\Delta \equiv \alpha\gamma''(a) - \beta\lambda > 0$ according to SOC. Hence, the sign of dv_e and da are given by the sign of their numerators.

Figure 3 provides a graphical representation of the equilibrium. In the space (a, v_e), the condition (ME) defines a curve with a slope equal to

$$\left(\frac{dv_e}{da}\right)_{(ME)} = \frac{C_M(y_{sw};h)-1}{(1-a)^2}\left(1-\left(\frac{1-F}{f}\right)'_{|v_e}\right)^{-1} < 0,$$

while condition (7) is represented by a curve having a slope equal to

$$\left(\frac{dv_e}{da}\right)_{(7)} = -\frac{\gamma''(a)}{f(v_e)\left(1-v_e+C_H(x_{sw};\rho h)-C_M(y_{sw};\sigma h)+\left(\frac{1-F}{f}\right)_{|v_e}\right)} < 0$$

since the numerator is negative by the SOC. Using the SOC, it comes that

$$\gamma''(a_e) > -\left(\frac{dv_e}{da}\right)_{(ME)} f(v_e)\left(1-v_e+C_H(x_{sw};\rho h)-C_M(y_{sw};\sigma h)+\left(\frac{1-F}{f}\right)_{|v_e}\right)$$

implying thus $-\left(\frac{dv_e}{da}\right)_{(7)} > -\left(\frac{dv_e}{da}\right)_{(ME)}$, meaning that, in the space (v_e, a), the curve corresponding to (7) has a slope (in absolute terms) which is always greater than the slope of the curve representing (ME). This leads to Figure 3 (regardless of the combination of liability regimes in both areas of law):

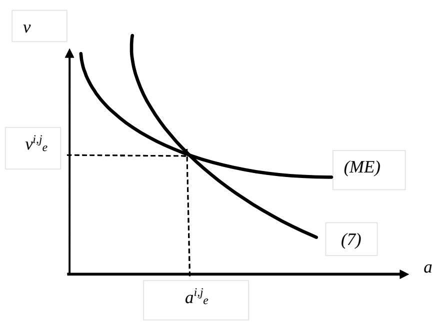


Figure 3 – The SPNE and liability regimes

In turn, it is easy to verify that the liability costs of both parties scale along these curves. An increase in $C_H(x_{sw};\rho h)$ (such as when the tort law switches from negligence to strict liability) or $C_M(y_{sw};\sigma h)$ (such as when the product liability law switches from negligence to strict liability) shifts the graph of the curve (ME) up and to the right. An increase in $C_H(x_{sw};\rho h)$ or a decrease in $C_M(y_{sw};\sigma h)$ shifts the graph of the curve (7) up and to the right.

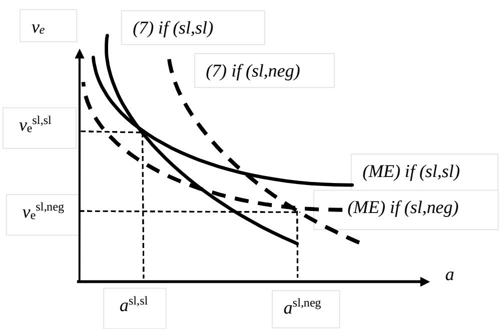
A/ Let us first consider "simple" cases (changing the liability rule in one area, keeping the other area stable). We have that:

a) Comparison of (i, neg) **and** (i, sl): regardless of the liability regime in tort law, an increase in C_M (e.g., when the liability regime in product liability switches from negligence $\rho = 0$ to strict liability $\rho = 1$) yields an increase in v_e (i.e., $v_e^{i,neg} < v_e^{i,sl}$ where $i \in \{sl, neg\}$), but a decrease in a (i.e., $a_e^{i,neg} > a_e^{i,sl}$ where $i \in \{sl, neg\}$), since we have:

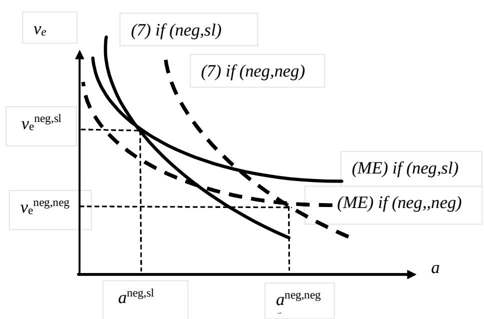
$$\frac{dv_e}{dC_M} = \frac{\frac{a}{1-a}\gamma''(a) + \beta(1 - F(v_e))}{\Delta} > 0; \frac{da_e}{dC_M} = -\frac{\alpha(1 - F(v_e)) + \lambda\frac{a}{1-a}}{\Delta} < 0$$

Figure 4.1 illustrates the outcome of the comparison of the combinations (sl, sl) (strict liability in both tort law and product liability) and (sl, neg) (strict liability in tort law but negligence in product liability). When strict liability is implemented in tort law, a switch in product liability from negligence to strict liability (i.e., the liability cost to the manufacturer increases) yields a shift in condition (ME) upwards and to the right, while condition (7) moves downwards and to the left. As a result, the levels of output and R&D investment both decrease.

4.1: (sl, sl) versus (sl, neg)



4.2: (neg, sl) versus (neg, neg)



Figures 4 – Comparison of (i, sl) ands (i, neg) where $i \in \{sl, neg\}$

Similarly, Figure 4.2 compares the combinations (neg, sl) (negligence in tort law and strict liability in product liability) and (neg, neg) (negligence both in tort law and in product liability). When negligence is implemented in tort law, a switch in product liability from negligence to strict liability (i.e., again the liability cost to the manufacturer increases) yields a shift in condition (ME) upwards and to the right, while condition (7) moves downwards and to the left. As a result, the levels of output and R&D investment both increase again.

b) Comparison of (neg, j) **and** (sl, j)**:** regardless of the liability regime in product liability, the effects of an increase in C_H (e.g., when the liability regime in tort law switches from negligence $\rho = 0$ to strict liability $\rho = 1$) on v_e and a is generally indeterminate (i.e., $v_e^{neg,j} \lesseqgtr v_e^{sl,j}$ and $a^{neg,j} \lesseqgtr a^{sl,j}$ where $j \in \{sl, neg\}$), since we have:

$$\frac{dv_e}{dC_H} = \frac{\gamma''(a) - \beta(1 - F(v_e))}{\Delta} \lesseqgtr 0; \frac{da_e}{dC_H} = \frac{\alpha(1 - F(v_e)) - \lambda}{\Delta} \lesseqgtr 0$$

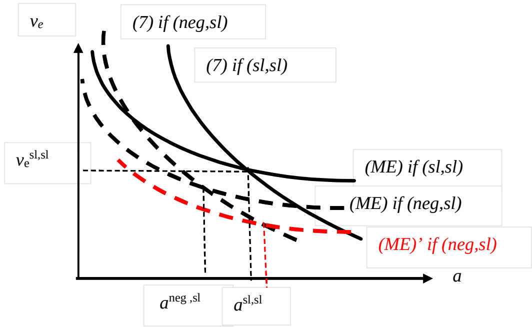
By the SOC, the inequality $\frac{\gamma''(a)}{\beta} > \frac{\lambda}{\alpha}$ always holds; thus if $\frac{da_e}{dC_H} < 0$ holds, it cannot be that $\frac{dv_e}{dC_H} < 0$, since this would require that $\frac{\gamma''(a)}{\beta} < 1 - F(v_e) < \frac{\lambda}{\alpha}$, contradicting the SOC (thus it is not possible to obtain $a^{neg,j} > a^{sl,j}$ and $Q_e^{neg,j} < Q_e^{sl,j}$). In turn, any of the following configurations may occur, where $j \in \{sl, neg\}$:

– if $1 - F(v_e) > \frac{\gamma''(a)}{\beta}$, then $\frac{dv_e}{dC_H} < 0$ and $\frac{da_e}{dC_H} > 0$; hence: $Q_e^{neg,j} < Q_e^{sl,j}$ and $a_e^{neg,j} < a_e^{sl,j}$;

– if $\frac{\lambda}{\alpha} > 1 - F(v_e)$, then $\frac{da_e}{dC_H} < 0$ and $\frac{dv_e}{dC_H} > 0$; hence: $Q_e^{neg,j} > Q_e^{sl,j}$ and $a_e^{neg,j} > a_e^{sl,j}$;

– if $\frac{\gamma''(a)}{\beta} > 1 - F(v_e) > \frac{\lambda}{\alpha}$, then $\frac{dv_e}{dC_H} > 0$ and $\frac{da_e}{dC_H} > 0$; hence: $Q_e^{neg,j} > Q_e^{sl,j}$ and $a_e^{neg,j} < a_e^{sl,j}$.

Figure 5 illustrates the comparison of combinations (sl, sl) and (neg, sl). When strict liability is implemented in product liability, a switch in tort law from negligence to strict liability (i.e., the liability cost to the users increases), then both condition (7) and condition (ME) moves upwards and to the right. Figure 5 shows that, depending on the magnitude of the shift in (ME), the decrease in the output level may be accompagnied by either an increase (dashed (ME) in black) or a decrease (dashed (ME)' in red) in the level of R&D.



Figures 5 – Comparison of (sl, sl) versus (neg, sl)

B/ In contrast, when the liability rule changes at the same time in both areas of law, the analysis is much more complex, since we have:

$$dv_e = \underbrace{\frac{(\gamma''(a) - \beta(1 - F(v_e)))}{\Delta}}_{\gtreqless 0}.dC_H + \underbrace{\frac{\left(\gamma''(a)\frac{a}{1-a} + \beta(1 - F(v_e)\right)}{\Delta}}_{>0}.dC_M$$

$$da_e = \underbrace{\frac{(\alpha(1 - F(v_e)) - \lambda)}{\Delta}}_{\gtreqless 0}dC_H - \underbrace{\frac{\left(\alpha(1 - F(v_e)) + \lambda\frac{a}{1-a}\right)}{\Delta}}_{>0}dC_M$$

c) Comparison of (neg, neg) **and** (sl, sl)**:** When both tort law and product liability switch from negligence to strict liability, there is an increase in v_e (i.e., $v_e^{neg,neg} < v_e^{sl,sl}$ but an ambiguous effect on a (i.e., $a^{neg,neg} \lesseqgtr a^{sl,sl}$), since assuming $d\rho = d\sigma$, $dC_H = q(x_{sw})h.d\rho$, and $dC_M = q(y_{sw})h.d\rho$, we have:

$$\frac{dv_e}{d\rho} = \frac{(\gamma''(a) - \beta(1 - F(v_e)))}{\Delta}.q(x_{sw})h + \frac{\left(\gamma''(a)\frac{a}{1-a} + \beta(1 - F(v_e)\right)}{\Delta}.q(y_{sw})h \lesseqgtr 0$$

$$\frac{da_e}{d\rho} = \frac{(\alpha(1 - F(v_e)) - \lambda)}{\Delta}.q(x_{sw})h - \frac{\left(\alpha(1 - F(v_e)) + \lambda\frac{a}{1-a}\right)}{\Delta}.q(y_{sw})h \lesseqgtr 0$$

Hence, $\gamma''(a) - \beta(1 - F(v_e)) > 0 > \alpha(1 - F(v_e)) - \lambda$ is sufficient to obtain $\frac{dv_e}{d\rho} > 0$ and $\frac{da_e}{d\rho} < 0$ (i.e., $v_e^{neg,neg} > v_e^{sl,sl}$ and $a_e^{neg,neg} < a_e^{sl,sl}$); cf. b) above.

d) Comparison of (neg, sl) **and** (sl, neg)**:** When tort law switches from negligence to strict liability, and simultaneously product liability also switches from negligence to strict liability, the effect on v_e and a are ambiguous (i.e., $v_e^{neg,neg} \lessgtr v_e^{sl,sl}$ and. $a^{neg,neg} \lessgtr a^{sl,sl}$), since assuming now $d\sigma = -d\rho$, $dC_H = q(x_{sw})h.d\rho$, and $dC_M = -q(y_{sw})h.d\rho$, we have:

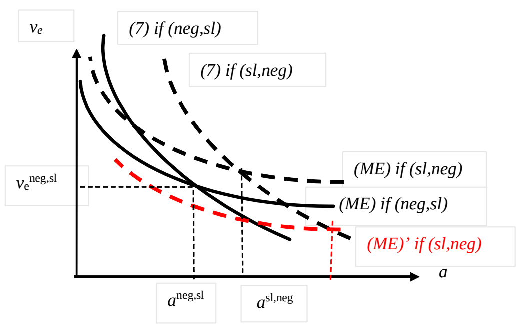
$$\begin{aligned}
\frac{dv_e}{d\rho} &= \frac{(\gamma''(a) - \beta(1 - F(v_e)))}{\Delta}.q(x_{sw})h - \frac{\left(\gamma''(a)\frac{a}{1-a} + \beta(1 - F(v_e)\right)}{\Delta}.q(y_{sw})h \\
&= \frac{h}{\Delta}\left(\gamma''(a)\left(q(x_{sw}) - \frac{a}{1-a}q(y_{sw})\right) - \beta(1 - F(v_e))\left(q(x_{sw}) + q(y_{sw})\right)\right) \\
\frac{da_e}{d\rho} &= \frac{(\alpha(1 - F(v_e)) - \lambda)}{\Delta}.q(x_{sw})h + \frac{\left(\alpha(1 - F(v_e)) + \lambda\frac{a}{1-a}\right)}{\Delta}.q(y_{sw})h \\
&= \frac{h}{\Delta}\left(\alpha(1 - F(v_e))\left(q(x_{sw}) + q(y_{sw})\right) - \lambda\left(q(x_{sw}) - \frac{a}{1-a}q(y_{sw})\right)\right)
\end{aligned}$$

Hence, a condition sufficient to obtain $\frac{dv_e}{d\rho} < 0$ and $\frac{da_e}{d\rho} > 0$ (i.e., $v_e^{neg,sl} > v_e^{sl,neg}$ and $a^{neg,sl} < a^{sl,neg}$) is:

– either $\gamma''(a) - \beta(1 - F(v_e) < 0 < \alpha(1 - F(v_e)) - \lambda$.

– or $q(x_{sw}) - \frac{a}{1-a}q(y_{sw}) < 0$.

Figure 6 compares the combinations (neg, sl) and (sl, neg) . Condition (7) under (neg, sl) is below and to the left of condition (7) under (sl, neg). However, condition (ME) under (neg, sl) may be either below and to the left, or above and to the right of condition (ME) under (sl, neg) (depending on the size of $\frac{a}{1-a}$). As a consequence, the level of R&D investment under (sl, neg) is larger than under (neg, sl); in contrast, the impact on the level of output is indeterminate.



Figures 6 – Comparison of (neg, sl) and (sl, neg)

Hence the results. ■

Equilibrium versus Optimum (Proposition 4). Condition (PO) defines a curve in the space (a, v_{sw}) with a slope equal to $\left(\frac{dv_{sw}}{da}\right)_{(PO)} = \frac{1}{(1-a)^2}\left(C_M(y_{sw};h) - 1\right) < 0$, while condition (3) is represented by a curve with a slope equal to

$$\left(\frac{dv_{sw}}{da}\right)_{(3)} = -\frac{\gamma''(a)}{f(v_{sw})(1-v_{sw}+C_H(x_{sw};\rho h)-C_M(y_{sw};\sigma h))} < 0$$

(the numerator is positive according to the SOC). By the SOC once again, we obtain that $-\left(\frac{dv_e}{da}\right)_{(3)} > -\left(\frac{dv_e}{da}\right)_{(PO)}$.

The comparison with the case where strict liability is used both under tort law and product liability is the most straightforward. On the one hand, the graph of (PO) is always below the graph of (ME) under (sl, sl) because of the term $\left(\frac{1-F}{f}\right)_{|v_e}$ appearing on the RHS term of (ME) (moreover, $\left(-\frac{dv_{sw}}{da}\right)_{(PO)} > \left(-\frac{dv_{sw}}{da}\right)_{(ME)}$); hence, for any level of R&D investment, the stage-2 output level is below the socially optimal one (i.e., for any a, then $v_e > v_{sw}$). On the other hand, comparing (3) and (7) for (sl, sl), it comes that the first term in the expression of the total marginal benefit of R&D in (3) depends on $E(v|v \geq v_{sw})$, while it depends on v_e in (7) (the second term of (3) and (7) being identical given the combination of rules (sl, sl)). Thus a sufficient condition for the graph of (3) to be above the graph of (7), as in Figure 7, is given by $E(v|v \geq v_{sw}) < v_e$:

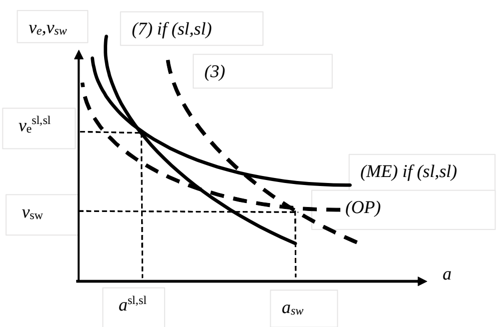


Figure 7 – Comparison of the Optimum and the Equilibrium under (sl, sl) when $E(v|v \geq v_{sw}) < v_e$

As a consequence, it comes that $v_e^{sl,sl} > v_{sw} \Leftrightarrow 1 - F(v_e^{sl,sl}) < 1 - F(v_{sw})$ and $a_e^{sl,sl} < a_{sw}$.

Complement to Section 4 – paragraph 4.1.

Assume that according to a public regulation, the monopoly affords the efficient output, Q_{sw}, at the (unique) best price sufficient to break even. Then, the robot manufacturer chooses the level of R&D investment $\hat{a}_e^{i,j}$ that maximizes the stage 1-expected profit given the liability regimes, defined by

$$\Pi(a) = (P - aC_M(y_{sw}; \sigma h))\, Q_{sw} - \gamma(a)$$

under (PO), and chooses the price $\hat{P}_e^{i,j}$ such that $\Pi(\hat{a}_e^{i,j}) = 0$, where v_{sw} is set according to (PO). It is easy to verify that the derivative w.r.t. $\hat{a}_e^{i,j}$ is given by:

$$\Pi'(a) = -C_M(y_{sw}; \sigma h).Q_{sw} + (P - aC_M(y_{sw}; \sigma h)) \,.f(v_{sw}) \left(\frac{dv_{sw}}{da}\right)_{(PO)} - \gamma'(a) \qquad (13)$$

with an equilibrium price that satisfies: $\hat{P}_e^{i,j} = \hat{a}_e^{i,j} C_M(y_{sw}; \sigma h) + \frac{\gamma(\hat{a}_e^{i,j})}{Q_{sw}}$. It is clear that this expression differs from the expressions we obtain in the different cases studied formerly (that lead to the different FOCs (3), or (7) or (11)), resulting thus in a different R&D investment level. Specifically, compared to the regulated monopoly that uses perfect price discrimination, it is straightforward that the first term in (13) is smaller than in (11) – therefore, for a given liability regime, the regulated monopoly that sets a unique price achieves a lower level of R&D investment compared to when it uses perfect price discrimination; i.e., $\hat{a}_e^{i,j} < \tilde{a}_e^{i,j}$ for any identical liability regime (i, j).

Furthermore, after substituting $\hat{P}_e^{i,j}$ and rearranging, we obtain:

$$\frac{\gamma(\hat{a}_e^{i,j})}{Q_{sw}} . \left(\frac{1 - C_M(y_{sw}; \sigma h)}{(1 - \hat{a}_e^{i,j})^2}\right) .f(v_{sw}) - C_M(y_{sw}; \sigma h).Q_{sw} = \gamma'(\hat{a}_e^{i,j})$$

Clearly, this latter condition is also distinct form condition (3), implying that generally $\hat{a}_e^{i,j} \neq a_{sw}$.