International Review of Law and Economics 68 (2021) 106022

Contents lists available at ScienceDirect

International Review of Law and Economics

Car accidents in the age of robots

Alessandro De Chiara[a,*], Idoia Elizalde[b], Ester Manna[a,1], Adrian Segura-Moreiras[b]

[a] Department of Economics, Universitat de Barcelona, and Barcelona Economic Analysis Team (BEAT), Avinguda Diagonal 696, 08034 Barcelona, Spain
[b] Department of Law, Universitat Pompeu Fabra, Ramon Trias Fargas, 25-27, 08005 Barcelona, Spain

ARTICLE INFO

Article history:
Received 7 July 2020
Received in revised form 24 June 2021
Accepted 5 September 2021
Available online 16 September 2021

JEL classification:
K2
L1

Keywords:
Autonomous vehicles
Investment problem
Liability rules
Torts

ABSTRACT

In this paper, we compare liability rules in a world where human-driven and fully-autonomous cars coexist. We develop a model where a manufacturer can invest to improve the safety of autonomous cars. Human drivers may decide to purchase a fully-autonomous car to save precaution costs to avoid road accidents and shift liability to the car manufacturer. As compared to the negligence rule, a strict liability regime on both human drivers and car manufacturers is proved to be a superior policy. In particular, strict liability leads to more efficient R&D investments to enhance the benefits of the technology and favors the adoption of fully-autonomous cars. We also recommend that users of fully-autonomous cars make a technology-dependent payment to a third-party if there is an accident to discipline their activity levels.

1. Introduction

The use of automated machines in different industries is growing rapidly. Between 2013 and 2018 the yearly worldwide investment in the robot industry grew at an average rate of 19%. The main drivers of this rate were the automotive industry and the electronics industry, which together accounted for more than 55% of the total growth.[1] For the particular case of the vehicle industry, before the Covid-19 outbreak, it was expected that over the period 2020–2021 an additional 10% of the current US mobility manufacturers would adopt some type of autonomous or semi-autonomous systems.[2] In this paper, we conduct a normative study of how civil liability rules should adapt to a world where humans interact with automated machines. At the onset of the automation era, the technology has limitations and these machines are not perfect. Therefore, humans are in charge of their control. At a later stage, technological improvements will allow the machines to operate in a fully-autonomous way. We precisely focus our attention on this latter scenario, where machines have achieved a substantial level of development and function without any human intervention.

The robot industry is very broad, therefore in order to contextualize it in the right way, this study focuses on the autonomous vehicle industry as a particular type of automated machine. The reason why the automotive industry is chosen as the object of the study is twofold: First, the World Health Organization estimates that approximately 1.35 million people die each year as a result of road traffic and those accidents represent a cost of almost the 3% of the domestic product of each country.[3] Second, there are current policy debates discussing the benefits and costs of Artificial Intelligence, robotics, and the Internet of Things concerning this specific industry.[4]

The introduction of autonomous vehicles into the market will entail many benefits including a better use of travel time, a reduction in traffic congestion and, consequently, pollution, and primarily, a reduction in road accidents. This last benefit is due to the fact that 95% of the current US road accidents are due to driver

* Corresponding author.
E-mail addresses: aledechiara@ub.edu (A. De Chiara), idoia.elizalde@upf.edu (I. Elizalde), estermanna@ub.edu (E. Manna), adrian.segura01@estudiant.upf.edu (A. Segura-Moreiras).
1 Ester Manna is Professora Lectora Serra Húnter.

1 See the Executive Summary World 2019 Industrial Robots, *International Federation of Robotics*.

2 Information gathered in a survey of 128 large and miz-size firms by PwC (https://www.pwc.com/us/en/industries/industrial-products/library/industrial-mobility.html).

3 Global status report on road safety (2018), *World Health Organization*.

4 See the Commission Report on safety and liability implications of AI, the Internet of Things and Robotics, *European Commission*, 19 February 2020.

https://doi.org/10.1016/j.irle.2021.106022

behavior.[5] On the other hand, it will also imply some disadvantages on the assumption that accidents will still occur. As mentioned before, it seems plausible that the amount of road accidents will decrease but it is unclear whether the magnitude of damage will be bigger or smaller than that of human-driven cars accidents. Since the first automated vehicle accident until the present day, at least five fatal automated vehicle accidents have occurred. Four of them in the United States[6] and one in China. Many other automated vehicle accidents occurred around the world but without casualties. Regarding the mortal accidents there are a lot of open investigations and research but not a judicial decision yet. So, even though there are some probable causes identified, there is not any definitive conclusion.

Our work approaches this issue by presenting a model where the production choice between an autonomous or a traditional vehicle follows from an industrial investment problem. The main difference in terms of precaution between the autonomous and the traditional technology must be seen from the manufacturer's perspective. On the one hand, manufacturers of traditional vehicles put their efforts on mechanisms that require essentially only *ex ante* investments (i.e., airbags, anti-lock breaking systems, etc.). On the other hand, autonomous producers are mainly focused on *ex post* precautionary systems which require constant updates and improvements even after the initial investment has been undertaken to ensure their proper functioning (i.e., software operating systems).

We believe, as it is usually the case for most technological innovations, that the introduction of fully automated vehicles in the market will be gradual. Fully autonomous vehicles will eventually substitute the traditional ones if and only if it is rational for consumers to purchase them. In essence, market forces and consumers' preferences will determine the fraction of each type of vehicle that will be produced and sold. Our model builds on the idea that substantial R&D investments are needed to develop the autonomous-car technology, and such investments would only be profitable if enough drivers shift from the alternative provided by traditional, human-driven cars to the autonomous ones.

In this regard, we assume that fully autonomous cars are developed and sold by a monopolistic manufacturer, whereas human-driven cars are sold in a regime of perfect competition. This simplification is meant to capture the higher market power that producers of autonomous cars will enjoy. We also assume that the manufacturer of autonomous cars can invest to improve the technology of autonomous cars. The R&D investment decision is private and, if taken, reduces the cost of preventing road accidents. As drivers can still purchase traditional cars, the monopolistic car manufacturer may decide not to make this investment if only few drivers would be willing to pay a higher price to buy the autonomous car. In this environment, we examine how different liability rules affect investment, pricing, autonomous-car adoption, and precaution choices.

Purchasing an autonomous car allows a driver to save the precaution cost and to avoid liability, as this will be shifted to the car manufacturer. We find that imposing strict liability on both human drivers and the car manufacturer is the constrained optimum, even though first-best efficiency will not be achieved. While both the negligence rule and the strict liability rule can be designed to ensure that first-best care decisions are taken, only the latter allows the manufacturer to fully enjoy the benefit that its R&D investment generates, which is associated with the reduction in the probability that a road accident occurs. Intuitively, with the negligence rule, a human driver could escape liability by exercising the required level of care. Then, only the reduction in the precaution cost (but not the avoidance of the damages following an accident) can prompt a driver to purchase a fully autonomous car. This leads to an inefficiently low level of adoption and, consequently, of investment. Furthermore, we point out that when also victims' care is important to prevent road accidents, then we have a more balanced picture because the negligence rule can induce victims to take precautions. Our conclusion on the superiority of the strict liability rule does not rest on the inability of the monopolist autonomous-car manufacturer to price discriminate, and the associated market inefficiency. What appears to play a more prominent role is the inability to condition liability on the firm's initial investment in the quality of the technology.

We also examine mixed rules, under which different liability applies to human drivers and the autonomous-car manufacturer. We show that applying strict liability to human drivers and adopting a negligence rule for the autonomous car manufacturer could favor adoption and may stimulate investment.

Lastly, we also find that, when usage also affects the probability of an accident, users of fully-autonomous cars would choose the highest possible level of activity, irrespective of the quality of the technology. This is because they would disregard the effect of activity on the expected cost of taking precaution (and on the expected liability) which would be borne by the manufacturer. However, it would be possible to induce efficient activity levels by requiring the driver of a fully-autonomous car to make a payment to a third party (e.g., the state) when their car causes an accident. Notably, this payment would not exonerate the manufacturer from liability and, in order to obtain first-best, should be tied to the quality of the technology of the autonomous car.

The paper is organized as follows: Section 2 discusses the related literature. Section 3 presents the formal model. Section 4 explores the implications of the model under the main liability rules. Section 5 analyzes mixed rules, whereas Section 6 provides some extensions. Section 7 concludes.

2. Related literature

This paper intends to answer the topical question of which civil liability rule should be chosen to optimize social costs in a world where there are fully-autonomous vehicles. The topic of how the level of care chosen depends on the civil liability rule applicable has been studied by legal scholars since very early. The care taking literature (e.g., Calabresi, 1970; Brown, 1973; Shavell, 1980; and more recently, Shavell, 2003) looks into the incentives that agents have to prevent or to reduce accident risks. Those works approached this issue by studying the effects of liability systems on the compensation of victims and on the allocation of risks between parties. Some authors have compared the pros and cons of liability systems with respect to others methods of controlling harmful activities, such as corrective taxation (e.g., see Kaplow and Shavell, 2002; Shavell, 2011) or public regulation (e.g., see Shavell, 1984; Hiriart et al., 2004; Immordino et al., 2011; Schwartzstein and Shleifer, 2013).

How liability rules should adjust to accommodate the arrival of autonomous cars has attracted media and scholarly attention. Some recent studies have looked into this issue in-depth. Shavell (2020) models encounters between drivers of fully autonomous cars that can hurt each other. The author finds that a new alternative liability rule can be suitable for shifting the curve of accidents to lower levels. Specifically, a liability system that induces payments to the state, rather than to the harmed party, proves to be a superior rule.

[5] These estimates were collected at the National Motor Vehicle Crash Causation Survey (NMVCCS) by the U.S. Department of Transportation.

[6] Two men were found killed in Texas in April 2021 after a fatal crash of the autonomous Tesla with which they were taking a drive. Even though the causes were unclear at the time the accident occurred, this fateful example shows us that autonomous vehicles do not perfectly protect us from having accidents.

In a different model, where there is a clear separation between victims and injurers, we also show that a payment to a third party can help align private and social incentives, but it is not meant to replace damages paid to the harmed party. The scope of the paper is different in that we aim to study which liability regime best prompts adoption of autonomous cars and investments that improve their safety.

Friedman and Talley (2019) extend the standard multilateral precaution framework to analyze the interactions between algorithmic and human decision makers. The specific aim of this work is to investigate what sorts of liability structures fare best in distributing the risks during a transition period (where human and autonomous vehicles may coexist). Their results show that some negligence-based rules are capable of achieving an efficient outcome. Di et al. (2019) model the uncertainty in the behavior of human actors and in the impact of autonomous vehicles manufacturers through a unified game-theoretical model. They find that human drivers could develop moral hazard if they perceive that the road environment has become safer and, as a result, the design of an optimal liability rules turns out to be crucial. Despite dealing with some of the questions covered in the aforementioned contemporaneous studies, our paper takes a different approach. We endogenize the level of adoption of the autonomous car technology as well as its level of sophistication.

With the advent of autonomous cars, the borders between different branches of liability law have blurred. As a result, we also draw on other contributions, like those dealing with product liability. Polinsky (2003) examines some of the principles of liability when harm is caused by an agent who is under the supervision of a principal. In a similar line of research, Polinsky and Shavell (2009) review the main benefits of product liability. This is a legal term that stands for the liability that a manufacturer incurs for producing or trading a certain product. In particular, in this previous work it is argued that product liability may prove to be a useful tool when market forces and regulation of a given product are weak. In Daughety and Reinganum (2013) a vast literature on product liability is reviewed. There are some works dealing with the effects of product liability when firms offer products with privately-observed safety characteristics (Daughety and Reinganum, 2008). In Ganuza et al. (2016) it is discussed how firms' reputation interact with different product liability rules. Choi and Spier (2014) shows that imposing tort liability on manufacturers for uncovered accident losses and prohibiting private parties from waiving that liability can improve social welfare when the likelihood of an accident depends on the unobserved precautions taken by the manufacturer. Recently, Hua and Spier (2020) also considers a monopolist selling to heterogeneous consumers and investigates the role of contract clauses as well as legal interventions, such as product liability, in inducing the firm to invest in product safety.

Our paper is also related to the analyses of optimal investment incentives in the presence of care choices and the issues inherent to market structure with regard to product liability. There is a line of research shared by some authors exploring the relationship between competition and product liability. In Chen and Hua (2017), the authors studied how these two concepts interact.[7] Some authors explore how innovations alters the *status quo* of liability systems. Following this vein of research we should highlight the work by Galasso and Luo (2018) which provides empirical evidence from medical implants industry showing that an increase in liability risk has a relatively large and negative impact on downstream innovation, but has no relevant effect on upstream innovation. Dari-Mattiacci and Franzoni (2014) present a model that shows the interplay of the adoption of a new technology with the standards of a negligence system. The authors compare alternative negligence rules when injurers can adopt harm-reducing technologies highlighting the tension between incentives to exert efficient care and efficient adoption. As compared to these papers, our model highlights how the strict liability rule can better induce users to adopt a superior technology and stimulates firm's investment.

3. The model

We develop a standard law & economics precaution model where injurers, that we assume to be car drivers, can exercise care to reduce the probability of an accident. Our innovation is that injurers can purchase a fully-autonomous vehicle, in which case they do not need to take precautions to avoid accidents. In our analysis, victims are those human beings who can suffer injuries $H > 0$ in a car accident. Drivers derive utility $v > 0$ from using a car, that we assume large enough so that all drivers are willing to get one. Drivers decide whether to buy an autonomous car or a traditional vehicle. In the latter case, they also choose the level of care while driving. We assume that there is a unit mass of human drivers are heterogeneous with respect to the marginal cost of taking precautions. For simplicity, there are just two types of drivers: high and low marginal cost, with $c_h > c_l > 0$. It is known that there are $\alpha \in (0, 1)$ high-cost drivers and $1 - \alpha$ low-cost ones in the population. The probability of an accident is $p(x) \in (0, 1)$, that is decreasing in x and convex, i.e., $p_x(x) < 0$ and $p_{xx}(x) > 0$.

We assume that traditional vehicles are sold in a perfectly-competitive market at a price equal to marginal cost, that we normalize to zero.[8] Conversely, fully-autonomous cars are produced and sold by a monopolistic manufacturer, that is interested in maximizing its profits. The manufacturer decides (i) the investment, (ii) the price of the autonomous cars, and (iii) the level of precaution for its cars. More details on all the manufacturer's choices are provided in the next subsection.

A welfare-maximizing social planner decides the liability rule. We restrict attention to (i) strict liability, (ii) negligence, (iii) no liability rule. All players are risk neutral and welfare is given by the sum of the manufacturer's profit, the drivers' and the victims' utility.

The sequence of events is the following:

1. The social planner sets the liability rule.
2. The manufacturer may observe and take an investment opportunity to improve the technology of fully-autonomous cars.
3. After observing the result of the investment, the manufacturer chooses the price for the automated cars.
4. Consumers decide which car to buy (human-driven or automated).
5. Care is exercised: by humans if the car is human-driven; by the manufacturer if the car is autonomous.

We solve the game by backward induction and we employ as equilibrium concept subgame perfection. All technical proofs are reported in Appendix A.

[7] Theirs findings suggest that the kind of relationship is dependent on the causes that motivate a change in competition. In particular, given a particular market structure, when there is an increase in competition due to a lower level of product differentiation the social optimal product liability increases. However, if the increase in competition is motivated by an increase in the number of competitors the optimal liability varies non-monotonically, first decreasing and then increasing.

[8] Otherwise, one could interpret v as the gross benefit of using a car, that is inclusive of its marginal cost of production.

3.1. Precautions, price, and investment

In the absence of autonomous cars, each type of human driver takes precautions to minimize the following objective function:

$$\min_{x_i} c_i x_i + p(x_i)D,$$

for $i = l, h$, where D are the damages that depend on the liability rule.[9] Solving the minimization problem, we obtain $c_i = -p_x(x_i^*)D$ for $i = l, h$.

Suppose now that fully autonomous cars are also available. Buying an autonomous car allows a human driver to shift liability for accidents to the manufacturer. Let $d \in [0, 1]$ be the fraction of human drivers who buy an autonomous car. The manufacturer makes a single care decision, denoted z, at cost δ, that affects the probability of accidents of each driver that has purchased the autonomous car:

$$\min_{z} \delta z + dp(z)D.$$

First-order condition yields $\delta = -dp_z(z^*)D$, which is increasing in adoption. Intuitively, the higher the number of autonomous cars that have been sold the more socially desirable the manufacturer's precautions to avoid accidents. We assume that z has the same effect on the probability of an accident as x. Thus, the difference between human-driven and autonomous cars lies in the marginal cost. We assume a baseline technology for the fully-autonomous cars $\delta_h < \alpha c_h$. This means that, with this technology, it would be socially desirable if all the high-cost drivers switched to autonomous cars. It may or may not be desirable that low-cost human drivers purchase an autonomous car. Throughout the analysis, we also assume that $c_l < \alpha c_h$.

The manufacturer sets a purchasing price P for the autonomous car. A driver i would buy the autonomous car if and only if:

$$v - P \geq v - c_i x_i^* - p(x_i^*)D.$$

For an injurer's perspective, the advantage of buying an autonomous car is twofold: s/he saves the cost of taking precautions while driving and s/he will not incur in liability if an accident occurs. We have not considered any specific benefit in terms of utility associated with the purchase of an autonomous car. That is to say that some of the potential benefits that an autonomous car will have on the activity level of the vehicle (i.e., miles travelled) are disregarded.

In stage 2, we suppose that the firm *privately* observes an opportunity to improve the precaution technology by bearing a fixed investment cost $I > 0$. We can interpret this improvement as a reduction in the marginal cost of exerting care. That is, this reduces the marginal cost to $\delta_l < \delta_h$. If the investment is made, the new technology is publicly observable, that is everyone observes if $\delta = \delta_l$ or $\delta = \delta_h$. The manufacturer will invest whenever this is profitable.

3.2. Discussion of the assumptions

In this subsection, we comment on some of the assumptions of the baseline model.

First, we have assumed that there is a monopolistic manufacturer of the autonomous car. As only few key players are currently developing this technology, and this is known to require substantial R&D investments, we believe that focusing on a monopolistic market may be a fairly good approximation of the type of market structure it would emerge, at least initially, in this industry. Monopolies may give rise to inefficiencies that might drive our findings. Therefore, in Section 6.3 we discuss the robustness of our results when the monopolistic manufacturer can engage in first-degree price discrimination, as this typically restores allocative efficiency.

In our model, the manufacturer exerts care after the adoption of autonomous cars by humans, although it can initially make an investment that affects the marginal cost of taking precautions. With its care decision the manufacturer ensures the proper functioning of the autonomous vehicles, constantly update and upgrade algorithms to avoid failures, and prevent hacks to its software. All of this is essential to preventing car accidents. Many commentators have stressed that autonomous cars will require constant care *after* they are sold to consumers. As cars' operating systems become more and more central to their correct functioning, automakers are already reliant on sending updates via satellite, Wi-Fi or cellular signal. According to Richard Wallace, a director at the Center for Automotive Research Software, "upgrades will be almost mandatory once we move up to higher forms of autonomous driving... The artificial intelligence underpinning self-driving will require constant upgrading to deal with novel situations."[10] For the need to receive constant security updates so as to prevent hacking see Filiz (2020).

The opportunity to make an investment that improves the precaution technology for autonomous vehicles is privately observed by the manufacturer. That a legislator cannot observe which investment opportunities are available to a firm ex-ante but can verify which investments are effectively undertaken, or their results, seem plausible.[11] The role played by this assumption will be clarified in Section 6.4 where it is relaxed.

Lastly, in the baseline model, we assume that precautions can only be taken by injurers and we have not included the activity level that may affect both the benefits of driving and the probability of accidents. We discuss how are results change when we allow for bilateral care and activity level in Sections 6.1 and 6.2, respectively.

4. Comparison of liability rules

In this section, we describe and compare the different liability regimes with respect to their precaution and adoption incentives. As a benchmark, we describe the efficient solution. First-best (denoted FB) care for low-cost and high-cost drivers, respectively, would be:

$$\begin{cases} c_l = -p_x(x_l^{FB})H; \\ c_h = -p_x(x_h^{FB})H. \end{cases}$$

Similarly, for the manufacturer of fully-autonomous cars the first-best care decision is:

$$\delta_i = -dp_z(z_i^{FB})H,$$

for $i = l, h$. In all cases, the efficient level of care is decreasing in the cost. That is, as the marginal cost of taking precaution decreases, the efficient level of precaution increases and the expected harm

[9] For ease of exposition, we are assuming away any direct, monetary or psychological, cost that the driver would suffer from the accident.

[10] See *Your Car's New Software Is Ready. Update Now?* (New York Times, September 8, 2016).

[11] Berkovitch and Israel (2004), Armstrong and Vickers (2010), and De Chiara and Iossa (2019), among others, study models where a principal (e.g., a regulator) can verify the characteristics of the project or investment selected by an agent, but cannot observe which other projects or investments were available to that agent.

goes down. To see this, note that the optimal level of precaution decreases as the marginal cost of taking precaution goes up:[12]

$$\frac{\partial x^{FB}}{\partial c} = -\frac{1}{p_{xx}(x^{FB})} < 0.$$

As a result, $p(x_h^{FB}) > p(x_l^{FB})$. It would be socially desirable that all drivers switch to an autonomous vehicle if $\delta < c_l$ and only high-cost drivers if $\delta \in [c_l, \alpha c_h)$.

Consider now the socially optimal choice of investment. Suppose first that $\delta_h \in (c_l, \alpha c_h)$. If $\delta_l \in [c_l, \delta_h)$, then human drivers should continue to drive their car themselves if they are low cost. The investment is socially desirable if its ex-post benefits outweigh the fixed cost[13] :

$$[\delta_h z_h^{FB}(\alpha) + \alpha p(z_h^{FB}(\alpha))H] - [(\delta_l z_l^{FB}(\alpha) + \alpha p(z_l^{FB}(\alpha))H)] \geq I.$$

If the reduction in the cost of care would be such that $\delta_l < c_l$, investment would be socially desirable if:

$$\alpha\left[\frac{\delta_h z_h^{FB}(\alpha)}{\alpha} + p(z_h^{FB}(\alpha))H - (\delta_l z_l^{FB} + p(z_l^{FB})H)\right] + (1-\alpha)[c_l x_l^{FB} + p(x_l^{FB})H - (\delta_l z_l^{FB} + p(z_l^{FB}))H)] \geq I.$$

This is because also the low-cost human drivers should get a fully-autonomous car if their technology has been enhanced.

If $\delta_h \leq c_l$, even low-cost human drivers should use fully-autonomous cars with the baseline technology. Thus, the condition under which an improvement in technology is socially desirable is the following:

$$[\delta_h z_h^{FB} + p(_h^{FB})H - (\delta_l z_l^{FB} + p(z_l^{FB})H)] \geq I.$$

Another benchmark is provided by the no liability rule (NL), under which the victim is not reimbursed if there is an accident, that is $D = 0$. It follows that neither human drivers nor manufacturers would take any precaution, that is $x_i^{NL} = z_i^{NL} = 0$ for $i = l, h$. No human driver would be willing to purchase the autonomous car.[14] The manufacturer would not have any incentive to invest. Interestingly, our model highlights how the lack of a liability regime may not prompt investment and adoption of (precaution) cost-saving technologies.

We examine strict liability and negligence rules in detail in the next two subsections. Importantly, as we focus on liability rules, we do not contemplate a no-fault insurance as an alternative since it dispenses, precisely, with liability rules.[15]

4.1. Strict liability rule

Under the strict liability rule (SL), should an accident occur, the injurer would have to reimburse the victim for the suffered harm, i.e., $D = H$. The equilibrium levels of care are:

$$c_i = -p_x(x_i^{SL})H,$$

for $i = l, h$, and

$$\delta_i = -dp_z(z_i^{SL}(d))H,$$

for $i = l, h$. Thus, strict liability induces the first-best level of care for the injurer: $x_i^{SL} = x_i^{FB}$ and $z_i^{SL}(d) = z_i^{FB}(d)$ for all i and demand levels. The firm would choose the price that maximizes expected profits[16] :

$$\max_{P \geq 0} \quad Pd(P) - [\delta z_i^{FB}(d(P)) + d(P)p(z_i^{FB}(d(P)))H]. \tag{1}$$

The human drivers' participation constraints are type-dependent:

$$PC_i^{SL}$$

$$v - P \geq v - \left(c_i x_i^{FB} + p(x_i^{FB})H\right).$$

The market demand function for autonomous cars is decreasing in the price and equal to:

$$d^{SL}(P) = \begin{cases} 0, & \text{if } P > c_h x_h^{FB} + p(x_h^{FB})H; \\ \alpha, & \text{if } P \in [c_l x_l^{FB} + p(x_l^{FB})H, c_h x_h^{FB} + p(x_h^{FB})H]; \\ 1, & \text{if } P < c_l x_l^{FB} + p(x_l^{FB})H. \end{cases}$$

To see why there exist prices for which only high-cost drivers would purchase autonomous cars, consider the following inequalities:

$$c_h x_h^{FB} + p(x_h^{FB})H > c_l x_h^{FB} + p(x_h^{FB})H \geq c_l x_l^{FB} + p(x_l^{FB})H,$$

where the first strict inequality owes to $c_l < c_h$ and the second inequality owes to $x_l^{FB} \in \arg\min c_l x + p(x)H$.

Depending on the parameter values, there are two possible equilibrium prices, as shown in the following lemma.

Lemma 1. *If $\delta \in [c_l, \alpha c_h)$, the manufacturer sets:*

$$P^{SL} = c_h x_h^{FB} + p(x_h^{FB})H.$$

If $\delta < c_l$, there exists $\tilde{\alpha} \in (0, 1)$ such that the manufacturer sets

$$P^{SL} = \begin{cases} c_h x_h^{FB} + p(x_h^{FB})H, & \text{if } \alpha \geq \tilde{\alpha} \\ c_l x_l^{FB} + p(x_l^{FB})H, & \text{otherwise.} \end{cases}$$

If the quality of the technology is intermediate, i.e., $\delta \in [c_l, \alpha c_h)$, the manufacturer will serve only high-cost drivers. If so, it will charge a price that makes these drivers indifferent between purchasing a traditional car and an autonomous one.

If the quality is high, that is $\delta < c_l$, the firm may be willing to sell also to low-cost drivers, provided that their fraction in the population is large enough. Intuitively, the manufacturer faces the typical monopolist trade-off between volume and margin: selling to low-cost drivers increases sales volume but lowers the margin on all units sold. It is profitable if the increase in volume is large enough as to more than compensate for the reduction in the margin. Henceforth, we say that the market expands when autonomous cars would be sold to low-cost drivers only if there is a superior technology for automated cars, that is, if $\delta_l < \delta_h$.

Consider now the investment incentives that this liability regime provides. When the reduction in the marginal cost would not generate a demand expansion effect, e.g., when $\delta_l \in [c_l, \alpha c_h)$, then it is easy to see that the condition for the investment to be

[12] This is shown for human drivers by using the Implicit Function Theorem. The same result can be found for the autonomous-car manufacturer and by using other techniques: e.g., monotone comparative statics (see, e.g., Milgrom and Shannon, 1994) showing that the function $-c\,x - p(x)H$ has increasing differences in $(x, -c)$ and, as a result, x^* is decreasing in c.

[13] As z_i^{FB} is a function of the manufacturer's demand, we now write $z_i^{FB}(\alpha)$ if only high-cost drivers buy the autonomous cars and simply z_i^{FB} if all human drivers purchase the autonomous cars. We will use a similar notation for the other solutions analyzed in the paper.

[14] If the injurers bear some positive but small costs when an accident occurs, some care would be exerted by both human drivers and car manufacturers. Then, some drivers might be willing to switch to the autonomous vehicle.

[15] See Schellekens (2018) and Engelhard and de Bruin (2018) for a no-fault insurance model as a solution for autonomous vehicles accidents.

[16] In the baseline model, we maintain the assumption that the manufacturer sets only one price, implicitly assuming that it cannot engage in price discrimination.

undertaken coincides with the first-best one. For instance, if $\delta_l \in [c_l, \alpha c_h)$, the investment is made only if[17] :

$$\pi^{SL}(\delta_l) - \pi^{SL}(\delta_h) \geq I$$

$$\Leftrightarrow [\delta_h z_h^{FB}(\alpha) + \alpha p(z_h^{FB}(\alpha))H] - (\delta_l z_l^{FB}(\alpha) + \alpha p(z_l^{FB}(\alpha))H)] \geq I.$$

Suppose instead that the investment would lead to a demand-expansion effect. This is the case if $\delta_h \in [c_l, \alpha c_h)$ and $\delta_l < c_l$ and $\alpha < \tilde{\alpha}$. The condition under which the firm will make the investment is:

$$\pi^{SL}(\delta_l) - \pi^{SL}(\delta_h) \geq I$$

$$\Leftrightarrow [(c_l x_l^{FB} + p(x_l^{FB})H) - (\delta_l z_l^{FB} + p(z_l^{FB})H)]$$

$$-\alpha \left[(c_h x_h^{FB} + p(x_h^{FB})H - \left(\frac{\delta_h z_h^{FB}(\alpha)}{\alpha} + p(z_h^{FB}(\alpha))H \right) \right] \geq I.$$

Compared to the efficient level, it is possible to see that this is more unlikely to hold.[18] We summarize this result in the following proposition.

Proposition 1. *Under strict liability, investment incentives are aligned with first-best unless there is a market-expansion effect, in which case there is underinvestment.*

Strict liability tends to provide the manufacturer with efficient incentives to undertake investments that increase the efficiency of autonomous cars technology. The only exception arises when the investment would lead to market expansion and would disappear if the monopolist were able to price discriminate, as we show in 6.3.

4.2. Negligence rule

Under the negligence rule (N), the social planner would impose that human driver (respectively, the manufacturer) $i = l, h$ would have to choose $x_i \geq x_i^{FB}$ (resp., $z \geq z_i^{FB}(d)$) to avoid having to pay $D = H$ if there is an accident. The underlying assumption is that the social planner can identify the different marginal costs of human drivers and that of the manufacturer. This would induce optimal care by both human drivers and the manufacturer.

Thus, under negligence, the injurers, irrespective of their type and nature, can make sure that they do not pay damages by selecting the minimum required level of care. This is exactly what they would do in equilibrium. Let us now determine the price the manufacturer would set under negligence:

$$\max_{P \geq 0} \quad Pd(P) - \delta z_i(d(P)). \tag{2}$$

The drivers' participation constraints are:

PC_i^N

$$v - P \geq v - c_i x_i^{FB},$$

for $i = l, h$. In stark contrast to the case of strict liability, under negligence it may be the case that the firm does not find it profitable to sell the autonomous car even to high-cost drivers when $\delta_i < c_h$. To see why, notice that the maximum price the high-cost drivers would be willing to pay for the autonomous car is $c_h x_h^{FB}$, namely the equilibrium precaution cost they would save by purchasing the autonomous car. However, the manufacturer may not turn a positive profit with this price even if $\delta_i < \alpha c_h$: if $z^{FB}(\delta_i) > x_h^{FB}$, it may be the case that $\alpha c_h x_h^{fb} < \delta_i z^{FB}(\delta_i)$. Hence, negligence may entail an inefficiently low adoption of autonomous cars.

Remark 1. Under negligence, the manufacturer may decide not to sell fully-autonomous cars even if $\delta_h < \alpha c_h$.

Further note that this inefficiency would not be solved even if the manufacturer could price discriminate. Therefore, while price discrimination would lead to efficient adoption of autonomous cars by human drivers under the strict liability rule, this would not necessarily be the case under the negligence rule. We impose a restriction, which ensures that the equilibrium precaution cost is increasing in the marginal cost. When this assumption holds, this downside of the negligence rule does not occur.

Assumption 1. We impose that $\frac{\partial cx(c)}{\partial c} > 0$.

Then, with this restriction, the manufacturer's demand is:

$$d^N(P) = \begin{cases} 0, & \text{if} \quad P > c_h x_h^{FB}; \\ \alpha, & \text{if} \quad P \in [c_l x_l^{FB}, c_h x_h^{FB}]; \\ 1, & \text{if} \quad P < c_l x_l^{FB}. \end{cases}$$

We can now proceed with characterizing the two possible equilibrium prices in a negligence regime.

Lemma 2. *Suppose Assumption* 1 *holds. If* $\delta \in [c_l, \alpha c_h)$*, the manufacturer sets:*

$$P^N = c_h x_h^{FB}.$$

If $\delta < c_l$*, there exists* $\hat{\alpha} \in (0, 1)$ *such that the manufacturer sets*

$$P^N = \begin{cases} c_h x_h^{FB}, & \text{if} \quad \alpha \geq \hat{\alpha}, \\ c_l x_l^{FB}, & \text{otherwise.} \end{cases}$$

When Assumption 1 holds, we obtain a result which is reminiscent of that illustrated for strict liability. If the quality of the technology is intermediate, i.e., $\delta \in [c_l, \alpha c_h)$, the manufacturer will serve only high-cost drivers. It will set a price for the fully-autonomous car that makes them exactly indifferent between purchasing either type of vehicle, thereby leaving them with no rent. If the quality is high, that is, $\delta < c_l$, the firm will also sell to low-cost drivers, provided that their fraction in the population is large enough.

Consider now investment incentives. By improving the technology, the ex-post requirements would be increased, accordingly. The firm will escape liability for accidents if $z \geq z_l^{FB}(d)$. If the investment could not increase demand ($\delta_l \geq c_l$). Investment would be made if:

$$\pi^N(\delta_l) - \pi^N(\delta_h) \geq I,$$

$$\delta_h z_h^{FB}(\alpha) - \delta_l z_l^{FB}(\alpha) \geq I.$$

While $\delta_h > \delta_l$, the requirement would be stricter: $z_h^{FB}(\alpha) < z_l^{FB}(\alpha)$. As a result, total prevention cost may increase, and the profit the firm may obtain which is closely related to it, may actually decrease. So even if I were very small, the firm may be reluctant to make the investment. If Assumption 1 holds, profits do increase. Even so, as $p_h(z_h^{FB}(\alpha))H - p(z_l^{FB}(\alpha))H > 0$, there would always be under-investment. Likewise, if $\delta_h < c_l$ and $\alpha < \hat{\alpha}$, investment would be made if

$$\pi^N(\delta_l) - \pi^N(\delta_h) \geq I,$$

$$\delta_h z_h^{FB} - \delta_l z_l^{FB} \geq I.$$

Intuitively, with the negligence rule there would be under-investment because the firm is not made to internalize the benefits of a reduction in the probability of causing accidents.

[17] If $\delta_l < \delta_h < c_l$ and $\alpha < \tilde{\alpha}$, we obtain the same condition with the only difference that α is replaced by 1.

[18] The left-hand side of the above inequality can be rewritten as the left-hand side of the efficient condition plus a negative term.

Under-investment is observed also in the case of market expansion as formally shown in the proof of the following proposition.

Proposition 2. *The negligence rule always leads to under-investment.*

Comparing the two liability regimes, we have highlighted that strict liability may entail under-investment and too little adoption of autonomous cars. Under negligence, the adoption problem is magnified and the under-investment problem always arises. Therefore, considering the market structure and investment incentives is critical to understanding the pros and cons of alternative liability regimes.

5. Mixed rules

As illustrated in the previous section, the downside of the negligence rule is that the firm does not capture the benefit of investment on the reduction in the probability of causing accidents. In fact, once the technology improves, the firm is subject to stricter requirements. Other solutions may generate other inefficiencies. For instance, while keeping the negligence rule, the planner could impose laxer requirements on the firm. This can increase adoption of autonomous vehicles, thereby boosting the firm's expected profits, and stimulating investment. Yet, an ex-post inefficiency would emerge as too little care would be taken. Alternatively, the requirements on the firm can be made stricter over time, irrespective of the actual technology in use. This would prompt the manufacturer to undertake investment just to keep up with the more stringent standards imposed by the planner. This approach could backfire if the firm does not have the opportunity or simply fails to improve the technology. While the issue of entry has not been considered in the model, in reality, a car manufacturer could decide against producing autonomous cars, anticipating that it will be subject to overly strict ex-post requirements in terms of car safety.

More in general. one could argue that human drivers and manufacturers may well be subject to different liability regimes, whereas we have so far restricted the social planner to use the same rule regardless of the identity of the injurer. We now explore to what extent our results vary in these mixed regimes.

First, suppose that human drivers are subject to the negligence rule, whereas the manufacturer is subject to strict liability. We henceforth denote this mixed regime with the superscript *NSL*. In a sense, when an autonomous car is involved in a road accident, its manufacturer may be liable for the damages that their product has caused. As Friedman and Talley (2019) document, while the negligence standard is typically invoked for conventional vehicular accidents, product liability rules normally impose strict liability on injurers, at least in the U.S. As emphasized in the previous subsections, with unilateral care, both negligence and strict liability can induce the optimal level of care in stage 5. As compared to strict liability on both human drivers and manufacturers, this mixed regime would lead to too little adoption in stage 4 and its effect on the manufacturer's investment in stage 2 is ambiguous.

Remark 2. Under strict liability for the autonomous-car manufacturer and negligence for human drivers,

(a) adoption incentives are lower than under negligence;
(b) investment incentives may be higher or lower than under strict liability.

To understand the dampened effect of this mixed rule on adoption, consider that human drivers can escape liability by taking the required precautions. Hence, their only incentive to purchase an autonomous car lies in the possibility of offloading the precaution costs to the manufacturer. By contrast, the manufacturer will find it profitable to serve the market only if the price can cover the cost of precautions as well as the damages it will have to pay in the case of an accident. Thus, the price at which autonomous cars can be sold is low, whereas the liability-related costs that the manufacturer would have to incur are high.

This mixed regime would strengthen incentives to invest, as compared to the negligence one. This is because the firm would internalize the benefits of a reduction in the probability of causing accidents. Yet, as its profits are lower than with strict liability on both categories of injurers, the firm may not have a strong enough incentive to invest resources to improve the technology.

Alternatively, human drivers may be subject to strict liability, whereas the manufacturer may be subject to negligence. We henceforth refer to this mixed rule with the superscript *SLN*. Although this rule may entail inefficiencies, it may favor adoption and investment.

Remark 3. Under negligence for the autonomous-car manufacturer and strict liability for human drivers,

(a) there are excessive incentives for autonomous-car adoption;
(b) investment incentives may be higher or lower than under strict liability.

To understand why there might be over-adoption, notice that human drivers would be willing to purchase autonomous cars to avoid having to pay damages and save the precaution cost. As the manufacturer can always avoid damages by selecting the efficient level of care, human drivers can be easily convinced to purchase autonomous cars. The effect on investment incentives is ambiguous. On the one hand, being subject to negligence, the manufacturer would not directly enjoy the benefits associated with a reduction in the probability of causing accidents. On the other hand, thanks to a regime which is especially burdensome to human drivers, the firm can obtain higher profits and, thus, its incentive to invest is stronger.

Very similar results to *SLN* could be obtained under a negligence rule imposed on both human drivers and the autonomous car manufacturer where the determination of negligence also includes the choice of the vehicle. Specifically, suppose that a human driver who causes an accident could be deemed negligent regardless of her precaution choice if she should have switched to autonomous cars for efficiency reasons (i.e., if her cost of taking precautions were higher than the manufacturer's). This modification would favor adoption, since a driver's participation constraint would effectively be (PC_i^{SL}) rather than (PC_i^{N}) whenever she should optimally switch to autonomous cars. Akin to *SLN*, under this modified negligence rule the manufacturer could charge higher prices and, as a result, it may have stronger incentives to invest in R&D.

To conclude this section, we observe that our analysis finds that mixed rules may be desirable under some circumstances. In particular, applying the negligence rule to the car manufacturer, whereas human drivers are subject to strict liability may especially be desirable if the social planner's primary aim is to promote the manufacturer's investment in a superior technology. By contrast, if strict liability applies to the car manufacturer whereas negligence applies to human drivers, autonomous car adoption will be discouraged and, as a consequence, investment could be depressed.

6. Discussion and extensions

In this section, we discuss the results of our model, we comment on some of the assumptions, and we present some extensions.

6.1. Bilateral care

Up to now, we have assumed that only injurers can take precautions that lower the probability of an accident. In reality, victims also can exercise care that reduces the probability of an accident. The literature (Shavell, 2003, 2007) has highlighted that the negligence rule can induce both victims and injurers to take the desired precautions, by appropriately setting the standards. By contrast, imposing strict liability on the injurers can prompt victims to take little or no precautions, anticipating that the injurers will have to make them whole if an accident occurs.

Thus, without formally developing a model, but building on the framework developed and analyzed in the previous sections, we can illustrate a trade-off between incentives for precautions and incentives for investment and adoption of superior technologies. The negligence rule proves (weakly) superior to the strict liability rule for what concerns the provision of incentives to take precautions. The strict liability rule performs better with respect to the provision of incentives for undertaking R&D investments and the adoption of the autonomous cars by the drivers.

The more important the effort that victims can take to avoid accidents, the more likely it is that the negligence rule dominates. Conversely, strict liability becomes more desirable when incentivizing investments, as well as ensuring prompt adoption by the users, is more important.

In light of this, a policy prescription that our model yields is that of relying on strict liability as the technology is still developing and moving to a regime of negligence when the technology reaches maturity.

6.2. Liability rules and activity levels

In this extension, we augment the baseline model by including the *activity level*, namely, the drivers' level of usage of the cars. Following the standard approach in the literature (e.g., Shavell, 1980, 2009), we assume that the activity level, denoted $a \in [0, \overline{a}]$, affects the benefits that drivers derive from using the cars and the probability of an accident. In particular, we now impose that the benefits of driving a car are given by the twice continuously differentiable function $v(a)$, which is strictly increasing and concave, i.e., $v_a(a) > 0$, and $v_{aa}(a) < 0$. Furthermore, we assume that the probability of an accident is linearly increasing in the level of the activity: $a\, p(x)$. Notably, we assume that the cost of exerting care is independent of the activity level, irrespective of the type of car that is used. Arguably, this assumption is more plausible for autonomous cars than for conventional, human-driven cars. However, we have opted for this modeling choice not to bias the adoption in favor of self-driving cars.[19]

At first best, activity and precaution levels for each type of human driver are derived from the maximization of the following expression:

$$U(a, x) \equiv v(a) - cx - ap(x)H.$$

Optimal levels of precaution and activity are thus given by the first-order conditions:

$$\begin{cases} c_i = -a_i^{FB} p_x(x_i^{FB})H; \\ v_a(a_i^{FB}) = p(x_i^{FB})H, \end{cases} \tag{3}$$

for $i = l, h$. Likewise, for autonomous cars, the corresponding first-order conditions for the optimal levels of precaution and activity are given by:

$$\begin{cases} \delta_i = -da_i^{FB} p_z(z_i^{FB}(d, a_i^{FB}))H; \\ v_a(a_i^{FB}) = dp(z_i^{FB}(d, a_i^{FB}))H, \end{cases}$$

for $i = l, h$. The first optimality condition in (3) tells us that, at the social optimum, the marginal precaution cost per unit of activity must be equal to the marginal benefit of precautions, which is associated with the reduction in the expected accident harm. The second optimality condition says that the level of activity should be chosen in such a way that the marginal benefit of using the car be equal to the expected total cost per unit of activity. The latter is simply the expected harm.

Importantly, when the marginal cost of taking precautions is lower, the equilibrium level of care is higher. To see this, consider that a human driver's objective function $U(x, a; c) = v(a) - cx - ap(x)H$ exhibits increasing differences in $(x, -c)$. Then,

$$\frac{\partial U^2}{\partial a \partial c} = -p_x(x_i^{FB})H \frac{\partial x_i^{FB}}{\partial c} < 0,$$

that is, $U(x, a\,; c)$ also exhibits increasing differences in $(a\,; -c)$ and, consequently, the lower c, the higher the precaution level and the higher the equilibrium level of activity. Stated differently, a reduction in the marginal cost of care will lead to a greater usage of the car.[20] The same logic applies for autonomous cars, i.e., the lower δ, the higher z and the higher a.

With respect to the baseline model, first-best adoption and investment choices are different in that the drivers would also enjoy the benefits associated with a greater usage of the vehicle. The specific conditions for the socially optimal investment choices are reported in the appendix alongside other technicalities related to this extension.

Consider now the different liability rules and assume that drivers perfectly predict future benefits they derive from the car when they decide which car to purchase.[21] Under the *no liability rule*, the driver would always choose $a^{NL} = \overline{a}$ and $x^{NL} = z^{NL} = 0$. Therefore, there would be excessive usage and zero precautions which, taken together, would result in an excessive number of accident and harm. Therefore, with the no liability rule, drivers would have no incentive to purchase a fully-autonomous car, which means that there would be no investment to develop the technology.

The *strict liability rule* gives rise to excessive activity with a fully-autonomous car, but also induces the optimal level of care, conditional on over-usage. To understand why, consider that activity and precaution choices are made by different players when the car is fully autonomous: the manufacturer makes the precaution choice, whereas the driver makes the activity choice. As a result, the driver would disregard the positive effect of usage on the probability of an accident and choose $a = \overline{a}$. The manufacturer would choose z_i to maximize $-\delta_i z_i - d\overline{a}p(z_i)H$ which leads to $z_i = z_i^{FB}(d, \overline{a})$. Importantly, drivers would overuse the fully-autonomous car, regardless of the quality of its technology. As we show in the appendix, this leads to generally non-optimal adoption and investment choices. One the one hand, the investment does not affect the driver's benefit from using the vehicle which discourages investment. On the other hand, the manufacturer is more willing to invest when drivers overly use the fully-autonomous cars: as over-usage increases the probability of accidents, there is a greater return from investing in a

[19] See Nussim and Tabbach (2009) for a more general model that allows for interaction between care and activity level.

[20] Note that second-order conditions for a maximum are satisfied at the optimum if $-v_{aa}(a)p_{xx}(x)H - 2p_x(x)H^2 > 0$ at $a = a^{FB}, x = x^{FB}$. See that $v_{aa}(a) < 0$ and $-ap_{xx}(x)H < 0$.

[21] See Baniak and Grajzl (2017) for a model where consumers mispredict future usage.

technology that reduces the occurrence of accidents for any activity level.

The strict liability rule can (and should) be amended to restore its desirable property of inducing optimal investment and adoption by inducing users of fully-autonomous cars to choose the first-best activity level. While the activity level may not be verifiable, the social planner could set a fine F that the users of fully-autonomous cars should pay in the case of an accident. The fine should be set in such a way that a user finds it beneficial to choose the optimal activity level when she purchases a fully-autonomous car. Thus, the fine would be a function of the technology of the car and the market demand. Specifically,

$$F = d(P)H,$$

so that, when choosing the activity level, the driver of a fully-autonomous car would maximize:

$$\max_{a\in[0,\bar{a}]} \quad v(a) - dap(z_i)F \Leftrightarrow \max_{a\in[0,\bar{a}]} \quad v(a) - dap(z_i)H,$$

exactly as in first-best. Note that this co-payment F is not meant to reduce the liability on the car manufacturer, that should continue to pay H to the injured party. Therefore, the payment F could be made to the state. In this respect, this *hybrid rule* is similar to that proposed by Shavell (2020), whereby the payments are made to the state rather than to the victims. However, the two proposals differ in that our alternative rule entails the manufacturer paying the damages to the victim – so that there are first-best incentives to exert care – and the user of the fully-autonomous car making a payment to the state – to discipline her usage incentives.[22]

Consider now the *negligence rule*. Following the literature, we assume that the usage level is not included in the negligence standard, since it would be prohibitively costly to obtain information about the activity level. Yet, in setting the standard, the social planner would take into account the consumers' incentives to choose the activity. When the car user drives the car herself and there is a clear separation between the identity of injurers and victims, it is well established that the negligence rule can be set in such a way that the driver is willing to choose the optimal level of care. Specifically, this can be achieved by imposing damages $D = H$ if it turns out that $x_i < x_i^{FB}$ for $i = l, h$ and $D = 0$, otherwise. Then, the driver would exercise the first-best level of care, but usage level would also be above first best. Similarly, if the user of the car drives a fully-autonomous car, she would not factor in the increase in the expected total cost of precautions that her activity entails. Therefore, the user of a fully-autonomous car would choose $a = \bar{a}$.

Following a similar approach to the one described above for the strict-liability rule, it is possible to devise a *hybrid rule* that can overcome the inefficiency of the negligence rule: specifically, the user of the fully-autonomous car can pay a fine to the state when an accident occurs. This solution can restore incentives to choose the optimal usage level. Yet, although this negligence rule augmented with this co-payment to the state can induce $a = a^{FB}$, adoption and investment continue to be inefficient. This is because a human driver would not consider the benefits associated with the avoidance of the damages following an accident, but only the reduction in the precaution cost, when making the purchase decision.

6.3. Price discrimination

In this subsection, we examine to what extent the results of the baseline model depend on the assumption that the monopolistic car manufacturer sets a uniform price for the autonomous cars. As is well known, an inefficient outcome typically arises under a uniform pricing policy because the monopolist may find it profitable not to serve all consumers with a valuation greater than the marginal production cost. We now suppose that the monopolistic firm can engage in first-degree price discriminate and, accordingly, sets two distinct prices: P_h and P_l for high- and low-cost human drivers, respectively.

Under both liability regimes, the car manufacturer will set a price that extracts the entire drivers' willingness to pay for the good. That is, $P_i^{SL} = c_i x_i^{FB} + p(x_i^{FB})H$, for $i = l, h$ under strict liability, and $P_i^N = c_i x_i^{FB}$, for $i = l, h$, under negligence.[23]

Notably, optimal investment is achieved under strict liability: in the baseline model, under-investment occurs because the manufacturer is unable to fully extract the drivers' increased surplus associated with the development of a superior technology. It follows that, if the manufacturer were able to price discriminate, it would always make the first-best investment. By contrast, price discrimination does not solve the under-investment problem under the negligence rule, which ultimately owes to the too little increase in surplus the manufacturer could appropriate following an investment. The following remark summarizes this result.

Remark 4. If the monopolist can perfectly price discriminate,

1. investment incentives are always aligned with first best under strict liability;
2. there is always under-investment under negligence.

6.4. Public observability of the investment

In contrast to the baseline model, we now suppose that the opportunity to make an investment that improves the precaution technology is *publicly* observable. Then, in designing the liability rule, the social planner could take into account whether the firm invested or not, imposing more stringent requirements if the manufacturer fails to make the desired technological upgrade. While the analysis of the strict liability rule is unaffected by this amendment to the model, the outcome that can be achieved with the negligence rule can be greatly affected. Specifically, the precaution requirement on the firm might consist of (i) an efficient ex-post precaution decision and (ii) an efficient ex-ante investment. If the firm fails in at least one dimension, it will have to pay damages $D = H$ when an accident takes place.

As a result, the adoption choices will be unaltered, whereas the under-investment problem can be at least partially solved. Specifically, it is solved unless the investment would lead to a market-expansion effect, in which case it may fail to provide enough incentives. Notably, the manufacturer's ability to price discriminate between the two types of drivers would not lead to an efficient investment. When the market expands, the feature of a negligence rule that an injurer can avoid paying damages if it takes the efficient level of precaution inevitably dampens the injurer's incentives to invest in a better precaution technology. The following remark summarizes this result.

[22] It is also important to point out that Shavell (2020) studies a different scenario in which the two parties to an accident harm each other, whereas we focus on the case in which only one party is harmed. Furthermore, Shavell (2020) also allows for a mileage fee that the manufacturer can charge to the buyer.

[23] We focus on whether the ability to engage in first-degree price discrimination can lead to investment efficiency. How it affects the manufacturer's willingness to serve both types of drivers or only high-cost ones is not reported here for brevity, but its analysis be provided under request.

Remark 5. Under the negligence rule and public observability of the manufacturer's investment opportunity, investment incentives are aligned with first-best unless there is a market expansion effect. The manufacturer's ability to perfectly price discriminates does not overcome the investment inefficiency.

We show below how investment efficiency can be achieved under negligence if the availability of the investment is publicly observable and there is no market expansion. Suppose first that $\delta_h > \delta_l \geq c_l$ and investment is socially desirable. If the firm invests, it will later choose $z_l = z_l^{FB}(d)$ as this is the minimum level of care that complies with the standard. If the firm does not invest, it will choose z_h to minimize the liability cost: $\delta_h z_h + d\,p(z_h)H$, which coincides with that of strict liability and gives $z_h = z_h^{FB}(d)$. Intuitively, if the firm does not invest when it is socially desirable, it will be found negligent in the case of an accident no matter the precaution taken. Now consider the firm's investment incentives:

$$\pi^N(\delta_l) - \pi^N(\delta_h) \geq I,$$

$$\delta_h z_h^{FB}(\alpha) + \alpha p(z_h^{FB}(\alpha))H - \delta_l z_l^{FB}(\alpha) \geq I.$$

When investing is socially beneficial, this inequality is always satisfied. This owes to the damages that the manufacturer would have to pay if there is an accident and it did not invest to improve the precaution technology.

Consider now the case in which $\delta_l < \delta_h < c_l$ and $\alpha < \hat{\alpha}$. Once again, if the firm decides against making a socially desirable investment, it will always be found negligent for an accident. Therefore, in choosing z_h the manufacturer would minimize $\delta_h z_h + p(z_h)H$ and first-best precautions would be taken. But then the firm would always find it profitable to invest. To see this note that:

$$\pi^N(\delta_l) - \pi^N(\delta_h) \geq I,$$

$$[\delta_h z_h^{FB} + p(z_h^{FB})H - \delta_l z_l^{FB}] \geq I,$$

which is always satisfied when the investment is socially desirable. The case in which the investment engenders a market expansion is dealt with in the appendix, under both uniform and personalized pricing.

6.5. Non-monotonically increasing precaution cost

In studying the negligence rule in Section 4.2, we have made the assumption that the precaution cost is monotonically increasing in the marginal cost. This has ensured that the manufacturer can find it profitable to sell autonomous cars to more inefficient human drivers. If Assumption 1 does not hold, it may well happen that the precaution cost is hump shaped and, for $P = c_l x_l$, only low-cost drivers would purchase autonomous cars, whereas high-cost drivers would not switch. Notably, if the precaution cost were monotonically decreasing, that is, if $\frac{\partial cx(c)}{\partial c} < 0$, the negligence rule might lead to no adoption of autonomous cars. To understand why, notice that not only would this opposite monotonicity condition imply that $c_l x_l > c_h x_h$, but also that $\delta_l z_l > c_l x_l$. Then, if the fraction of the low-cost drivers in the population is not too low, it may happen that the manufacturer would find it unprofitable to sell autonomous cars to either group of drivers.[24]

To summarize, the negligence rule might hinder the adoption of safer autonomous cars, unless a monotonicity condition is satisfied. This problem is not due to uniform pricing but solely to the limited savings that a human driver could realize by switching to fully-autonomous cars.

[24] Recall that the manufacturer takes a single precaution decision for all cars that have been sold. If only low-cost drivers purchased autonomous cars, $\delta_l = -(1-\alpha)p_z(z_l^{FB}(1-\alpha))H$.

7. Conclusion

To the best of our knowledge, this is the first paper that studies how alternative liability rules can affect incentives to invest and adopt fully-autonomous cars. The main take-away of the model is that strict liability provides better incentives for investment and adoption of this technology and, as a result, dominates the negligence rule in so far as victims' precautions are not very important to avoid road accidents. Unlike the negligence rule, strict liability allows the manufacturer of fully-autonomous cars to fully enjoy the benefits of its R&D investments that are aimed at improving car safety. What is more, strict liability encourages adoption of fully-autonomous cars by higher-cost drivers, who would have more to gain from shifting liability to car manufacturers. Intuitively, under the negligence rule, human drivers can only save precaution costs by purchasing the fully-autonomous car whereas, with strict liability, they would avoid road-accident related damages too. We have also put forward a variant to the strict-liability and negligence rules which involves a payment made by the driver to a third party in the case of an accident. This payment would kick in only when the driver uses a fully-autonomous car and is meant to discipline incentives to choose activity levels, which would otherwise be excessively high as the driver could shift liability to the manufacturer.

Whilst the model considers a one-period investment problem, its results can also provide some insight on how we can expect different liability rules to influence the speed of adoption of autonomous cars. In light of the above, strict liability appears to be better positioned to stimulate investment. Therefore, this rule should be favored, at least as long as the technology is in a development phase. Clearly, our stylized model has limitations and should only be viewed as a first step in the direction of incorporating investment and adoption decisions in the design of liability rules for fully-autonomous vehicles. In particular, in order to simplify the analysis, we have distinguished between two polar market structures: traditional cars are sold in a regime of perfect competition, whereas there is a monopolistic seller of fully-autonomous cars. While this does not exactly reflect the real market structures, it is meant to capture the higher market power that manufacturers of fully-autonomous cars are likely to enjoy. In any case, the manufacturer of fully-autonomous cars must enjoy some market power to be willing to undertake some quality-enhancing investment.

Acknowledgements

The authors would like to thank the two anonymous referees for their insightful and valuable comments to this paper. We also express our sincere gratitude for previous revisions of this manuscript and for the discussion of some features of the work to: Antonio Cabrales, Carlos Gómez, Pablo Salvador, Francisco Marcos, Fernando Gómez, Rosa Ferrer, Gideon Parchomovsky, Toni Gómez, and all participants at the 2019 Spanish Association of Law & Economics Annual Meeting in Oporto, and at the 2019 European Association of Law & Economics Annual Meeting in Tel Aviv. Special thanks to Nuno Garoupa for his guidance throughout this project. All errors are our solely responsibility.

Appendix A

A.1 Proof of Lemma 1

If $\delta \in [c_l, \alpha c_h)$, then the best the manufacturer can do is to set $P^{SL} = c_h x_h^{FB} + p(x_h^{FB})H$ so as to induce high-cost drivers to purchase the autonomous car, without leaving them any surplus. The firm would obtain profit:

$$\alpha\left[\left(c_h x_h^{FB}+p(x_h^{FB})H\right)-\left(\frac{\delta z^{FB}(\alpha)}{\alpha}+p(z^{FB}(\alpha))H\right)\right],$$

which is strictly positive because

$$c_h x_h^{FB}+p(x_h^{FB})H>\frac{\delta}{\alpha}x_h^{FB}+p(x_h^{FB})H\geq\frac{\delta}{\alpha}z^{FB}(\alpha)+p(z^{FB}(\alpha))H,$$

where the first inequality owes to $\delta<\alpha c_h$ and the second inequality owes to $z^{FB}(\alpha)\in\arg\min_x c_h x+\alpha p(x)H$. Following a similar argument, for $\delta\geq c_l$ the firm could not set a price that is simultaneously acceptable to low-cost drivers and is profitable.

If $\delta<c_l$, the manufacturer can also set a price that can induce low-cost drivers to purchase the autonomous car and is profitable. The best such price is $c_l x_l^{FB}+p(x_l^{FB})H$ which exactly satisfies the low-cost drivers' participation constraint. Yet, as the firm cannot engage in price discrimination, it would have to lower the price to high-cost drivers. By comparing the profits, it is possible to determine the threshold value of α below which the firm sells to both drivers' types. In particular, let the intermediate profits with a low and a high price be, respectively:

$$\pi_l^{SL}\equiv\left(c_l x_l^{FB}+p(x_l^{FB})H\right)-\left(\delta z^{FB}+p(z^{FB})H\right)$$

and

$$\pi_h^{SL}(\alpha)\equiv\alpha\left(c_h x_h^{FB}+p(x_h^{FB})H\right)-\left(\delta z^{FB}(\alpha)+\alpha p(z^{FB}(\alpha))H\right).$$

To prove the existence of $\tilde{\alpha}\in(0,1)$ such that $\pi_l^{SL}=\pi_h^{SL}(\tilde{\alpha})$, note that π_l^{SL} is independent of α whereas $\pi_h^{SL}(\alpha)$ is continuously increasing in α. To see that, note that for the Envelope Theorem:

$$\frac{\partial\pi_h^{SL}(\alpha)}{\partial\alpha}=\left(c_h x_h^{FB}+p(x_h^{FB})H\right)-p(z^{FB}(\alpha))H,$$

and this is always positive because $p(x_h^{FB})>p(z^{FB}(\alpha))$. When $\alpha\to1$, $\pi_h^{SL}(\alpha)>\pi_l^{SL}$ because $\left(c_l x_l^{FB}+p(x_l^{FB})H\right)<\left(c_h x_h^{FB}+p(x_h^{FB})H\right)$, and when $\alpha\to0$, $\pi_h^{SL}(\alpha)=0<\pi_l^{SL}$. Therefore, there must exist $\tilde{\alpha}\in(0,1)$ for which $\pi_l^{SL}=\pi_h^{SL}(\tilde{\alpha})$. □

A.2 Proof of Lemma 2

If $\delta\in[c_l,\,\alpha c_h)$, then the best the manufacturer can do is to set $P^N=c_h x_h^{FB}$ so as to induce high-cost drivers to purchase the autonomous car, without leaving them any surplus. The firm would obtain profit:

$$\alpha c_h x_h^{FB}-\delta z_h^{FB}(\alpha),$$

which is strictly positive. To see why, note that $x_h^{FB}=x^{FB}(c_h)$ and $z_h^{FB}(\alpha)=x^{FB}(\frac{\delta}{\alpha})$ and

$$\begin{aligned}&\alpha c_h x^{FB}(c_h)>\delta x^{FB}\left(\frac{\delta}{\alpha}\right)\\&\Leftrightarrow\left(c_h-\frac{\delta}{\alpha}\right)x^{FB}(c_h)+\frac{\delta}{\alpha}x^{FB}(c_h)-\frac{\delta}{\alpha}x^{FB}\left(\frac{\delta}{\alpha}\right)>0\\&\Leftrightarrow x^{FB}(c_h)+\frac{\delta}{\alpha}\frac{x^{FB}(c_h)-x^{FB}(\frac{\delta}{\alpha})}{c_h-\frac{\delta}{\alpha}}>0.\end{aligned}$$

Taking the limit for c_h which goes to $\frac{\delta}{\alpha}$ we can rewrite the last inequality as:

$$x^{FB}\left(\frac{\delta}{\alpha}\right)+\frac{\delta}{\alpha}x'^{FB}\left(\frac{\delta}{\alpha}\right)>0,$$

which is always satisfied if Assumption 1 holds.

If $\delta<c_l$, the manufacturer can also set a price that can induce low-cost drivers to purchase the autonomous car and is profitable. The best such price is $c_l x_l^{FB}$ which exactly satisfies the low-cost drivers' participation constraint. Yet, as the firm cannot engage in price discrimination, it would have to lower the price to high-cost drivers, giving up a rent to these drivers. By comparing the profits, it is possible to determine the threshold value of α below which the firm sells to both drivers' types. To determine the cutoff $\hat{\alpha}$, let the intermediate profits with a low and a high price be, respectively,

$$\pi_l^N\equiv c_l x_l^{FB}-\delta z^{FB}$$

and

$$\pi_h^N(\alpha)\equiv\alpha c_h x_h^{FB}-\delta z^{FB}(\alpha).$$

Note that $\pi_h^N(\alpha)$ is continuous in α and, as $\alpha\to1$, $\pi_h^N(\alpha)>\pi_l^N$ because $c_h x_h^{FB}>c_l x_l^{FB}$ due to Assumption 1, and, as $\alpha\to0$, $\pi_h^N(\alpha)\to0<\pi_l^N$. Yet, we need to impose some condition to guarantee that there exists a unique $\hat{\alpha}\in(0,1)$ such that $\pi_h^N(\hat{\alpha})=\pi_l^N$. Specifically, we impose that $\pi_h^N(\alpha)$ be strictly increasing in α which requires that $c_h x_h^{FB}>\frac{(p_z(z^{FB}(\alpha)))^2H}{p_{zz}(z^{FB}(\alpha))}$. To see why, consider that the first order derivative of $\pi_h^N(\alpha)$ with respect to α is $c_h x_h^{FB}-\delta z_\alpha^{FB}(\alpha)$. For the Implicit Function Theorem,

$$z_\alpha^{FB}(\alpha)=-\frac{p_z(z^{FB}(\alpha))H}{\alpha p_{zz}(z^{FB}(\alpha))H},$$

whereas from the first order condition, $\delta=-\alpha p_z(z^{FB}(\alpha))H$. Hence, $\frac{\partial\pi_h^N(\alpha)}{\partial\alpha}=c_h x_h^{FB}-\frac{(p_z(z^{FB}(\alpha)))^2H}{p_{zz}(z^{FB}(\alpha))}$. □

A.3 Proof of Proposition 2

To show that there is under-investment also in the case of market expansion, suppose $\delta_l<c_l$, $\delta_h\in[c_l,\alpha c_h)$, and $\alpha<\hat{\alpha}$. Investment is undertaken if:

$$\pi^N(\delta_l)-\pi^N(\delta_h)\geq I,$$

$$[c_l x_l^{FB}-\delta_l z_l^{FB}]-\alpha\left[c_h x_h^{FB}-\frac{\delta_h z_h^{FB}(\alpha)}{\alpha}\right]\geq I.$$

To see that there is under-investment, consider the following:

$$\begin{aligned}&\pi^N(\delta_l)-\pi^N(\delta_h)=[c_l x_l^{FB}-\delta_l z_l^{FB}]-\alpha\left[c_h x_h^{FB}-\frac{\delta_h z_h^{FB}(\alpha)}{\alpha}\right]\\&\Leftrightarrow\alpha[(c_l x_l^{FB}-\delta_l z_l^{FB})-(c_h x_h^{FB}-\frac{\delta_h z_h^{FB}(\alpha)}{\alpha})]\\&\quad+(1-\alpha)[c_l x_l^{FB}-\delta_l z_l^{FB}]\\&\quad\leq\alpha[\frac{\delta_h z_h^{FB}(\alpha)}{\alpha}+p(z_h^{FB})H-(\delta_l z_l^{FB}+p(z_l^{FB})H)]\\&\quad+(1-\alpha)[c_l x_l^{FB}+p(x_l^{FB})H-(\delta_l z_l^{FB}+p(z_l^{FB})H)].\end{aligned}$$

Consider first the terms multiplied by $(1-\alpha)$: since $p(x_l^{FB})H-p(z_l^{FB})H>0$, the term on the right-hand side of the inequality is larger. Consider now the terms multiplied by α. Note that

$$(p(z_h^{FB})H+c_h x_h^{FB})-(p(z_l^{FB})H+(c_l x_l^{FB}))>0$$

because $p(z_h^{FB})>p(z_l^{FB})$ and $c_h x_h^{FB}>c_l x_l^{FB}$ under Assumption 1. □

A.4 Proof of Remark 2

Point (a). The manufacturer's problem in stage 4 is given by (1), where the drivers' participation constraints are given by (PC_i^N) and consequently, under Assumption 1, the manufacturer's demand is $d^N(P)$. To see that the manufacturer has a reduced incentive to sell cars and, therefore, efficient exchanges may fail to take place, it suffices to compare the manufacturer's intermediate profits under this mixed rule with those under the negligence rule with a high and a low price:

$$\pi_h^{NSL}(\alpha) \equiv \alpha\left[c_h x_h^{FB} - \left(\frac{\delta z^{FB}(\alpha)}{\alpha} + p(z^{FB}(\alpha)H\right)\right] < \alpha c_h x_h^{FB} - \delta z^{FB}(\alpha) = \pi_h^N(\alpha),$$

and

$$\pi_l^{NSL} \equiv c_l x_l^{FB} - (\delta z^{FB} + p(z^{FB})H) < c_l x_l^{FB} - \delta z^{FB} = \pi_l^N.$$

Point (b). It is easy to see that the conditions for investing coincide with those under strict liability when there is no market expansion. In the case of market expansion, the firm would invest if

$$\begin{aligned}&\pi^{NSL}(\delta_l) - \pi^{NSL}(\delta_h) \geq I\\ &\Leftrightarrow [c_l x_l^{FB} - (\delta_l z_l^{FB} + p(z_l^{FB})H)]\\ &\quad - \alpha\left[c_h x_h^{FB} - \left(\frac{\delta_h z_h^{FB}(\alpha)}{\alpha} + p(z_h^{FB}(\alpha))H\right)\right] \geq I.\end{aligned}$$

Compared to the left-hand side of the investment condition under strict liability, the following terms are missing:

$$p(x_l^{FB})H - \alpha p(x_h^{FB})H,$$

which can be greater or lower than 0. □

A.5 Proof of Remark 3

Point (a). The manufacturer's problem in stage 4 is given by (2), where the drivers' participation constraints are given by (PC_i^{SL}) and consequently the manufacturer's demand is $d^{SL}(P)$. To see that there is an excessive incentive for adoption, note that if $\delta > c_l$, the firm could set a price $P^{SLN} = c_l x_l^{FB} + p(x_l^{FB})H$, which is acceptable to low-cost drivers, and gives positive profit to the firm:

$$c_l x_l^{FB} + p(x_l^{FB})H - \delta z > 0,$$

if $p(x_l^{FB})H > \delta z - c_l x_l^{FB}$. Note that the manufacturer will opt to sell to all drivers when:

$$c_l x_l^{FB} + p(x_l^{FB})H - \delta z > \alpha(c_h x_h^{FB} + p(x_h^{FB})H) - \delta z(\alpha),$$

which may or may not hold.

Point (b). It is easy to see that the conditions for investing coincide with those under negligence when there is no market expansion. Hence, there would be under-investment. In the case of market expansion, the firm would invest if

$$\begin{aligned}&\pi^{SLN}(\delta_l) - \pi^{SLN}(\delta_h) \geq I\\ &\Leftrightarrow [(c_l x_l^{FB} + p(x_l^{FB})H) - \delta_l z_l^{FB}]\\ &\quad - \alpha\left[(c_h x_h^{FB} + p(x_h^{FB})H) - \frac{\delta_h z_h^{FB}(\alpha)}{\alpha}\right] \geq I.\end{aligned}$$

Compared to the left-hand side of the investment condition under strict liability, the following terms are missing:

$$\alpha p(z_h^{FB}(\alpha))H - p(z_l^{FB})H),$$

which can be greater or lower than 0. □

A.6 Liability rules and activity levels

Consider *socially optimal investment choices*. If $\delta_h \in (c_l, \alpha c_h)$ and $\delta_l \in [c_l, \alpha\delta_h)$, investment is socially desirable if:

$$\left(v(a^{FB}(\delta_l)) - \frac{\delta_l a^{FB}(\delta_l)z^{FB}(\delta_l)}{\alpha} - a^{FB}(\delta_l)p(z^{FB}(\delta_l))H\right) - \left(v(a^{FB}(\delta_h)) - \frac{\delta_h a^{FB}(\delta_h)z^{FB}(\delta_h)}{\alpha} - a^{FB}(\delta_h)p(z^{FB}(\delta_h))H\right) \geq \frac{I}{\alpha}.$$

If $\delta_h \in (c_l, \alpha c_h)$ and $\delta_l < c_l$, investment is socially desirable if:

$$\begin{aligned}&\left(v(a^{FB}(\delta_l)) - \delta_l a^{FB}(\delta_l)z^{FB}(\delta_l) - a^{FB}(\delta_l)p(z^{FB}(\delta_l))H\right)\\ &- \alpha\left(v(a^{FB}(\delta_h)) - \frac{\delta_h a^{FB}(\delta_h)z^{FB}(\delta_h)}{\alpha} - a^{FB}(\delta_h)p(z^{FB}(\delta_h))H\right)\\ &- (1-\alpha)\left(v(a^{FB}(c_l)) - c_l a^{FB}(c_l)x^{FB}(c_l) - a^{FB}(c_l)p(x^{FB}(c_l))H\right) \geq I.\end{aligned}$$

If $\delta_h \leq c_l$, the condition for investment to be socially desirable is:

$$\begin{aligned}&\left(v(a^{FB}(\delta_l)) - \delta_l a^{FB}(\delta_l)z^{FB}(\delta_l) - a^{FB}(\delta_l)p(z^{FB}(\delta_l))H\right)\\ &- \left(v(a^{FB}(\delta_h)) - \delta_h a^{FB}(\delta_h)z^{FB}(\delta_h) - a^{FB}(\delta_h)p(z^{FB}(\delta_h))H\right) \geq I.\end{aligned}$$

Let us now show that *strict liability* does not induce optimal adoption and investment choices when activity matters for the probability of accidents and the cost of exerting precautions. We use the superscript *SLa* to denote the solutions, as we are also including the activity levels. The firm would choose the price that maximizes expected profits:

$$\max_{P \geq 0} \quad Pd(P) - [\delta z_i^{FB}(d, \bar{a}) + d(P)\bar{a}p(z_i^{FB}(d(P), \bar{a}))H].$$

The drivers' participation constraints are:

PC_i^{SLa}

$$v(\bar{a}) - P \geq v(a^{FB}(c_i)) - \left(c_i x_i^{FB} + a^{FB}(c_i)p(x_i^{FB})H\right).$$

With respect to the baseline model, drivers have an additional reason to purchase the fully-autonomous car: given that their usage choice would not be restrained from the fear of having to pay damages in the case of a car accident, they are more inclined to adopt the new technology.

Similarly to what shown in Lemma 1, depending on the value of δ and the fraction of high-cost drivers, we can have two equilibrium prices. Differently from Lemma 1, the threshold values of δ need not coincide with c_l and this can once again be ascribed to the choice of the activity level. Specifically, there exist two cutoff values of δ, $0 < \delta_1 < \delta_2$, such that for $\delta \in [\delta_1, \delta_2]$:

$$P^{SLa} = [v(\bar{a}) - v(a^{FB}(c_h))] + \left(c_h x_h^{FB} + a^{FB}(c_h)p(x_h^{FB})H\right).$$

At this price only high-cost drivers would purchase the autonomous car. The firm would obtain:

$$\alpha\left[\left(v(\bar{a}) - v(a^{FB}(c_h)) + c_h x_h^{FB} + a^{FB}(c_h)p(x_h^{FB})H\right) - \left(\frac{\delta z^{FB}(\alpha, \bar{a})}{\alpha} + \bar{a}p(z^{FB}(\alpha, \bar{a}))H\right)\right].$$

To see why the cutoff δ_1 need not coincide with c_l, note that

$$\left(c_h x_h^{FB} + a^{FB}(c_h)p(x_h^{FB})H\right) - \left(\frac{\delta z^{FB}(\alpha, \bar{a})}{\alpha} + \bar{a}p(z^{FB}(\alpha, \bar{a}))H\right),$$

may not be strictly positive because $\bar{a} > a^{FB}(c_h)$. At the same time, there is an additional term $v(\bar{a}) - v(a^{FB}(c_h)) > 0$. Thus, $\delta_1 \gtreqless c_l$.

If $\delta < \delta_1$,

$$P^{SLa} = \begin{cases} [v(\bar{a}) - v(a^{FB}(c_h))] + \left(c_h x_h^{FB} + a^{FB}(c_h)p(x_h^{FB})H\right), & \text{if } \alpha \text{ is sufficiently high;} \\ [v(\bar{a}) - v(a^{FB}(c_l))] + \left(c_l x_l^{FB} + a^{FB}(c_l)p(x_l^{FB})H\right), & \text{otherwise.} \end{cases}$$

Consider now the investment incentives that this liability regime provides. Suppose first that the reduction in the marginal cost would not generate a demand expansion effect. For instance, if $\delta_l \in [\delta_1, \delta_h)$, the investment is made only if[25] :

$$\pi^{SLa}(\delta_l) - \pi^{SLa}(\delta_h) \geq I$$
$$\Leftrightarrow \alpha\left[\frac{\delta_h z_h^{FB}(\alpha, \bar{a})}{\alpha} + \bar{a}p(z_h^{FB}(\alpha, \bar{a}))H - \left(\frac{\delta_l z_l^{FB}(\alpha, \bar{a})}{\alpha} + \bar{a}p(z_l^{FB}(\alpha, \bar{a}))H\right)\right] \geq I.$$

Note that the investment will be generally different from first best, since injurers will overuse fully-autonomous cars. On the one hand, there is a positive term equal to $v(a^{FB}(\delta_l)) - v(a^{FB}(\delta_h))$ missing on the left-hand side of the inequality because the investment does not affect the benefits of the activity. The reason is that drivers would overuse the fully-autonomous car, regardless of the quality of its technology. On the other hand, as $\bar{a} > a^{FB}(\delta_l) > a^{FB}(\delta_h)$, the firm has an additional incentive to invest under the strict-liability rule: since the excessive use of fully-autonomous cars increases the probability of accidents, there is a greater return from investing in a technology that reduces their occurrence for any activity level.

Lastly, note that the investment would be different from first-best, even if it led to a market-expansion effect. In that case, the condition under which the firm will make the investment is:

$$\pi^{SLa}(\delta_l) - \pi^{SLa}(\delta_h) \geq I$$
$$\Leftrightarrow [v(\bar{a}) - v(a^{FB}(c_l) + (c_l x_l^{FB} + a^{FB}(c_l)p(x_l^{FB})H) - (\delta_l z_l^{FB} + \bar{a}p(z_l^{FB}(\bar{a}))H)]$$
$$- \alpha[v(\bar{a}) - v(a^{FB}(c_h) + c_h x_h^{FB} + a^{FB}(c_h)p(x_h^{FB})H - \left(\frac{\delta_h z_h^{FB}(\alpha, \bar{a})}{\alpha} + \bar{a}p(z_h^{FB}(\alpha, \bar{a}))H\right)] \geq I,$$

which does not coincide with the socially-optimal one.

A.7 Proof of Remark 4

Consider strict liability first. It is immediate to see that the conditions under which the manufacturer invests coincide with those in the baseline model when the adoption of a superior technology does not lead to market expansion. By contrast, when there is market expansion, investment is first-best. To see this, note that the firm invests if[26] :

$$\pi_{PD}^{SL}(\delta_l) - \pi_{PD}^{SL}(\delta_h) \geq I$$
$$\Leftrightarrow (1-\alpha)\left(c_l x_l^{FB} + p(x_l^{FB})H\right)$$
$$- \left[\left(\delta z^{FB} + p(z^{FB})H\right) - \alpha\left(\frac{\delta z^{FB}(\alpha)}{\alpha} + p(z^{FB}(\alpha))H\right)\right] \geq I,$$

which can be straightforwardly rewritten as the expression for the efficient condition for investment.

Consider the negligence rule. Without market expansion, there is no difference with the baseline model and negligence is thus associated with under-investment. Same occurs in the presence of market expansion. To see this, investment is undertaken if:

$$\pi_{PD}^{N}(\delta_l) - \pi_{PD}^{N}(\delta_h) \geq I$$
$$(1-\alpha)c_l x_l^{FB} - \delta_l z_l^{FB} + \delta_h z_h^{FB}(\alpha) \geq I$$
$$\Leftrightarrow \alpha\left[\frac{\delta_h z_h^{FB}(\alpha)}{\alpha} - \delta_l z_l^{FB}\right] + (1-\alpha)[c_l x_l^{FB} - \delta_l z_l^{FB}] \geq I.$$

Since $p(x_l^{FB})H - p(z_l^{FB})H > 0$ and $p(z_h^{FB}(\alpha))H - p(z_l^{FB})H > 0$, there will be under-investment. □

A.8 Proof of Remark 5

Suppose that the investment would engender a market expansion. To see that the negligence rule may not prompt the firm to make a socially desirable investment, suppose that $\delta_l < c_l < \delta_h$ and $\alpha < \hat{\alpha}$. Investment is undertaken if:

$$\pi^{N}(\delta_l) - \pi^{N}(\delta_h) \geq I$$
$$\Leftrightarrow [c_l x_l^{FB} - \delta_l z_l^{FB}] - \alpha[c_h x_h^{FB} - \delta_h \frac{z_h^{FB}(\alpha)}{\alpha} - p(z_h^{FB}(\alpha))H] \geq I$$
$$\Leftrightarrow \alpha[(c_l x_l^{FB} - \delta_l z_l^{FB}) - (c_h x_h^{FB} - \frac{\delta_h z_h^{FB}(\alpha)}{\alpha} - p(z_h^{FB}(\alpha))H)]$$
$$+ (1-\alpha)(c_l x_l^{FB} - \delta_l z_l^{FB}) \geq I.$$

Compared to the first-best investment condition, the terms multiplied by α may be higher or lower because $c_h x_h^{FB} > c_l x_l^{FB}$ under Assumption 1, but there is a term $-p(z_l^{FB})H$ missing. The terms multiplied by $1-\alpha$ are larger in the condition for the investment to be socially desirable as shown in the Proof of Proposition 2.

Suppose now that the manufacturer can perfectly price discriminate. An investment leading to market expansion would be undertaken if:

$$\pi_{PD}^{N}(\delta_l) - \pi_{PD}^{N}(\delta_h) \geq I,$$
$$(1-\alpha)[c_l x_l^{FB} - \delta_l z_l^{FB}] + \alpha[\delta_h \frac{z_h^{FB}(\alpha)}{\alpha} + p(z_h^{FB})H - \delta_l z_l^{FB}] \geq I.$$

Compared to the first-best investment condition, the term multiplied by α is lower because there is a term $-p(z_l^{FB})H$ missing; the terms multiplied by $1-\alpha$ are larger in the condition for the investment to be socially desirable as shown in the Proof of Proposition 2. □

References

Armstrong, M., Vickers, J., 2010. A model of delegated project choice. Econometrica 78 (1), 213–244.

Baniak, A., Grajzl, P., 2017. Optimal liability when consumers mispredict product usage. Am. Law Econ. Rev. 19 (1), 202–243.

Berkovitch, E., Israel, R., 2004. Why the npv criterion does not maximize npv. Rev. Financ. Stud. 17 (1), 239–255.

Brown, J.P., 1973. Toward an economic theory of liability. J. Legal Stud. 2 (2), 323–349.

Calabresi, G., 1970. The Cost of the Accidents, A Legal Economic Analysis. Yale UP.

Chen, Y., Hua, X., 2017. Competition, product safety, and product liability. J. Law Econ. Organ. 33 (2), 237–267.

Choi, A.H., Spier, K.E., 2014. Should consumers be permitted to waive products liability? Product safety, private contracts, and adverse selection. J. Law Econ. Organ. 30 (4), 734–766.

Dari-Mattiacci, G., Franzoni, L.A., 2014. Innovative negligence rules. Am. Law Econ. Rev. 16 (2), 333–365.

Daughety, A.F., Reinganum, J.F., 2008. Communicating quality: a unified model of disclosure and signalling. RAND J. Econ. 39 (4), 973–989.

Daughety, A.F., Reinganum, J.F., 2013. Cumulative harm, products liability, and bilateral care. Am. Law Econ. Rev. 15 (2), 409–442.

De Chiara, A., Iossa, E., 2019. Public Procurement As A Demand-Side Policy: Project Competition and Innovation Incentives. CEPR Discussion Paper No. DP13664.

[25] If $\delta_l < \delta_h < \delta$ and α is lower than the cutoff, we obtain the same condition with the only difference that α is replaced by 1.

[26] We use the subscript *PD* to denote the profit expressions when the manufacturer can price discriminate.

Di, X., Chen, X., Talley, E., 2019. Liability Design for Autonomous Vehicles and Human-Driven Vehicles: A Hierarchical Game-Theoretic Approach. arXiv:1911.02405.
Engelhard, E., de Bruin, R., 2018. Legal analysis of the EU common approach on the liability rules and insurance related to connected and autonomous vehicles. In: Evas, T., Rohr, C., Dunkerley, F., Howarth, D. (Eds.), A Common EU Approach to Liability Rules and Insurance for Connected and Autonomous Vehicles: European Added Value Assessment Accompanying the European Parliament's Legislative Own-Initiative Report. EU Publications, pp. 38–132.
Filiz, C., 2020. Can autonomous vehicles prevent traffic accidents? Accident Analysis and Prevention.
Friedman, E., Talley, E., 2019. Automatorts: How Should Accident Law Adopt to Autonomous Vehicles, Lessons from Law and Economics., pp. 3, Draft, April.
Galasso, A., Luo, H., 2018. When Does Product Liability Risk Chill Innovation? Evidence From Medical Implants. NBER Working Paper No. 25068.
Ganuza, J.J., Gomez, F., Robles, M., 2016. Product liability versus reputation. J. Law Econ. Organ. 32 (2), 213–241.
Hiriart, Y., Martimort, D., Pouyet, J., 2004. On the optimal use of ex ante regulation and ex post liability. Econ. Lett. 84 (2), 231–235.
Hua, X., Spier, K.E., 2020. Product safety, contracts, and liability. RAND J. Econ. 51 (1), 233–259.
Immordino, G., Pagano, M., Polo, M., 2011. Incentives to innovate and social harm: Laissez-faire, authorization or penalties? J. Public Econ. 95 (7–8), 864–876.
Kaplow, L., Shavell, S., 2002. On the superiority of corrective taxes to quantity regulation. Am. Law Econ. Rev. 4 (1), 1–17.
Milgrom, P., Shannon, C., 1994. Monotone comparative statics. Econometrica: J. Econ. Soc., 157–180.
Nussim, J., Tabbach, A.D., 2009. A revised model of unilateral accidents. Int. Rev. Law Econ. 29 (2), 169–177.
Polinsky, A.M., 2003. Principal-Agent Liability, Available at SSRN 413103.
Polinsky, A.M., Shavell, S., 2009. The uneasy case for product liability. Harv. Law Rev. 123, 1437.
Schellekens, M., 2018. No-fault compensation schemes for self-driving vehicles. Law Innov. Technol. 10 (2), 314–333.
Schwartzstein, J., Shleifer, A., 2013. An activity-generating theory of regulation. J. Law Econ. 56 (1), 1–38.
Shavell, S., 1980. Strict liability versus negligence. J. Legal Stud. 9 (1), 1–25.
Shavell, S., 1984. A model of the optimal use of liability and safety regulation. RAND J. Econ. 15 (2), 271–280.
Shavell, S., 2003. Economic Analysis of Accident Law. Technical Report. National Bureau of Economic Research.
Shavell, S., 2007. Liability for accidents. Handb. Law Econ. 1, 139–182.
Shavell, S., 2009. Economic Analysis of Accident Law. Harvard University Press.
Shavell, S., 2011. Corrective taxation versus liability. Am. Econ. Rev. 101 (3), 273–276.
Shavell, S., 2020. On the redesign of accident liability for the world of autonomous vehicles. J. Legal Stud. 49 (2), 243–285.