OPTIMAL LIABILITY RULES FOR COMBINED HUMAN-AI HEALTH CARE DECISIONS

Documents de travail AFED
The French Law & Economics Association Working Papers Series

BERTRAND CHOPARD
OLIVIER MUSY

AFED WP No. 24-04
https://econpapers.repec.org/paper/afdwpaper

Optimal Liability Rules for Combined Human-AI Health Care Decisions

Bertrand Chopard*and Olivier Musy†

AFED Working Paper No. 2024–04

Abstract

The integration of AI for healthcare redefines medical liability, converting decision-making into a collaborative process involving a technological tool and its user. When a harm is caused, both users and manufacturers of AI may be responsible. The judicial system has yet to address claims of this nature. We build a model with bilateral care to study which combinations of liability rules are socially efficient. Both agents could face strict liability, be subject to negligence rules or face hybrid regimes: one agent faces a fault liability regime, while the other operates under strict liability. We highlight two crucial elements: (i) the sharing scheme of the payment of compensation between users and producers (ii) the nature of their cares (complements or substitutes). The latest AI Liability Directive from the European Parliament in the field of medicine advocates the implementation of a strict liability regime for producers and a fault liability regime for users. We show that this regime is socially efficient. A novel framework is not necessary.

JEL Codes: I11, L13, K13, K41.

Keywords: Medical Liability, Joint Liability, Artificial Intelligence, Bilateral Care, European Regulation.

*Université Paris Cité, LIRAES, F-75006 Paris, France. Mail: bertrand.chopard@u-paris.fr

†Université Paris Cité, LIRAES, F-75006 Paris, France. Mail: olivier.musy@u-paris.fr

1 Introduction

Recent technological developments in robotics and artificial intelligence (AI) in the medical area will change the nature of experts' decisions because a part of the final decision or action is delegated to an 'autonomous' product. This integration of AI into healthcare protocols has the potential to improve the well-being of patients since "combining AI methods and systems with an irreplaceable human clinician can advance the delivery of care in a way that outperforms what either can do alone" (Griffin, 2021). However, such utilization also carries the inherent risk of causing harm, thereby establishing a basis for potential legal claims, given that the judicial system has yet to address any claims of this nature. We anticipate that both healthcare practitioners employing these AIs and the manufacturers of the AIs may be deemed responsible. Given that it is not feasible to ascribe any fault directly to the AI, being a non-legal entity, the patient-victim naturally will turn to the healthcare professional for accountability. Subsequently, the practitioners will, in turn, seek redress from their employer, the healthcare institution. Finally, the healthcare institution may then pursue legal action against the supplier or the manufacturer of the AI.

In accordance with the latest European regulations pertaining to AI in the healthcare sector, two already established liability frameworks may facilitate the resolution of future disputes arising from the use of AI in the field of medicine[1]. Put differently, there would be no need to establish a new liability framework specifically for these disputes. These two existing liability regimes are (i) the strict liability regime for defective products (such as algorithm malfunctions) and (ii) the fault liability regime for healthcare professionals (e.g., improper use of AI). Let us examine these two liability regimes successively.

As outlined in the AI Liability Directive of the European Parliament, dated September 28, 2022, the manufacturer of AI will assume strict liability for its AI's defectiveness. The defectiveness of AI will be established upon the identification of a malfunction in the algorithm code. The confirmation of AI defectiveness poses challenges and incurs substantial costs, necessitating private and judicial expert assessments and counter-expertise to conclusively determine its flaws. Indeed, as stated by Price (2018) and Price *et al.* (2019, 2023), one major difference with traditional medicine is that the AI may communicate recommendations without communicating

[1]See for instance, the Proposal for a Directive of the European Parliament and of the Council on the adaptation of rules on non-contractual liability in the realm of AI, dated September 28, 2022. This directive is commonly referred to as the AI Liability Directive, accessible at: https://ec.europa.eu)

the reasons of the recommendation, leading to black box medicine.

Those AI manufacturers could then absolve themselves of this liability by invoking the so-called 'development risk' provision, particularly applicable to learning algorithms. Essentially, the manufacturer could demonstrate that, at the time of introducing the product to the market, the prevailing scientific and technical knowledge did not enable the detection of the identified defect. This assertion holds validity, especially in the context of deep learning algorithms, where computer codes are designed at a specific moment for autonomous learning. These algorithms generate developments independently through inferential reasoning based on accumulated data. Consequently, if defects arise in the self-generated developments post-initial programming, the manufacturer is absolved of liability, as anticipating such defects at the time of market release was not feasible.

The latest European regulation on AI use in healthcare could also modify the scope of the liability of healthcare professionals as AI users. The main idea is that the regime of fault liability for the healthcare professional must continue to apply, for example when the fault committed results from the improper use of AI by the physician. At the same time, physicians must also continuously monitor the AIs they use. They must continue to take responsibility for their choices, even when assisted by AI (Higgins, 2022). The most intricate scenario is as follows. If not following the recommendations of algorithmic systems constitutes a fault, considering the performance of certain systems, medical professionals may feel almost compelled to follow the recommendations of AIs for fear of having their liability invoked in case of non-compliance.

Viewed through the lens of joint liability, such health care decisions arising from the collaboration between humans and AI present a unique challenge. In this scenario, distinct agents find themselves subject to diverse liability frameworks. While physicians adhere to the traditional fault-based tort law, any proven defects in the AI system should trigger the application of standard product liability rules grounded in strict liability to the AI producers. Enhancing patient care and mitigating overall harm becomes a shared responsibility between the physician and the AI producer, each capable of individually reducing risks at a private cost. Notably, the literature has yet to explore the incentive implications for potential injurers in this context.

In this paper, we explore the care incentives provided by liability rules in an joint decision process in which potential harm to the patient depend on a joint process combining a physician (or the AI user) and an AI device produced by a private firm, called the AI producer. To do this, we extend the model of tort liability in a "buyer-seller" relationship with bilateral observable care developed by Daughety and Reinganum (2013, 2014). The first modification to this model comes from the specific context of medical malpractice (see Danzon, 1985a, 1985b, 1997). We assume that the patient's harm depends on the levels of care exercised by the AI producer and the AI user, while the final consumer, here the patient, remains passive. The patient has no direct influence on both levels of care and does not build any strategy. However the patient's welfare matters as the physician will take care about patient's health status, which is a standard assumption in health economics[2]. The second change lies in the assumption that the decision-making process is joint between the physician and the AI product[3] (see Agrawal *et al.*, 2019). We assume that both the AI user and the AI producer can cause harm to patient and are then subject to liability. Consequently, we also introduce in our study how liability should be assigned and shared between agents. Each agent can face a specific liability regime, which creates specific incentives in terms of care. We show that an important element of the analysis is how these different levels of care interact with each other, i.e. whether they are complements or substitutes.

The specific design of the model is such that the AI producer chooses its individually optimal level of care first, for instance during the development phase and/or the production phase of the AI system. One can interpret this level of care as the quality of AI system, i.e. the quality of the prediction or the ease-of-use of the algorithm. Next, the AI user, here the physician, chooses his level of care. Since the final decision is a joint process between the AI user and the AI, the higher the care of both agents, the lower the potential harm of the patient. Individuals levels of care can be substitutes or complements. Between these two decisions about the levels of care, there is a price-mediated relationship between the AI producer and the AI user. The AI user determines the quantity of AI system to demand, anticipating the amount of care that she will choose when using the AI system. It allows to calculate the AI user's marginal willingness-

[2]Standard theories of health providers suggest that physicians are motivated by both their own personal interest and by an altruistic motive relative to their patients' health benefits or losses (Brosig-Koch, Hennig-Schmidt, Kairies-Schwarz and Wiesen, 2016, 2017; Hennig-Schmidt and Wiesen, 2014). This medical altruism leads to the inclusion of the patient's health status in the physician's utility function. The literature generally uses a linear and separable form of the physician's utility function (Chalkley and Malcomson, 1998; Godager and Wiesen, 2013; Jack, 2005; Makris and Siciliani, 2013), which distinguishes between the physician's revenue, and costs due to the patient's health status.

[3]Our analysis can also apply to the study of other technology devices such as medical robots that notably improve precision in surgical procedures. For the conciseness of the presentation, we focus our discussion on AI technology devices.

to-pay for AI system. The probability of a medical accident, as well as the magnitude of the patient's damage (and compensation), depends on these two levels of care, taken together.

After calculating the socially efficient levels of care, we study which liability rules may induce, or not, socially efficient levels of care from both AI user and AI producer. We study various combinations of liability rules. Both the AI producer and the AI user can have strict liability. They may also be each of them subject to fault based liability rules. But one can also apply a different liability rule to each party. We thus study configurations in which one agent faces a fault liability, while the other operates under strict liability. When both parties have to compensate the patient's harm they cause, we consider a general sharing scheme of the payment of compensation between the AI producer and the AI user. Our main result is that the latest proposal of the European Parliament, that is a combination of strict liability for the AI producer and fault-based liability for the AI user, is socially desirable as it would induce socially efficient levels of care from both AI user and AI producer. Thus, the future concern lies not in determining the applicable liability regime for the involved parties but, rather, in exploring how these two liability frameworks can be subtly adjusted to adeptly address the challenges posed by the integration of AI in the field of medicine.

The paper is organized as follows. The literature review is given in Section 2. We present the model and our assumptions in Section 3. Then we successively study different liability regimes: strict liability in Section 4, and fault-based liability in Section 5. In these sections, we assume that both the AI producer and the AI user are subject to the same liability rules. In section 6, we move away from this assumption and we apply a different regime of liability to each party. In Section 7, we discuss and compare our overall outcomes. Section 8 concludes.

2 Related literature in tort law and economics

Most of the standard law and economics literature on products liability developed notably by Shavell (1980, 2007), Landes and Posner (1980) and Daughety and Reinganum (2013, 2014) assume that only the producer may be subject to liability rules, not the user of the product, or a final consumer such as a medical patient. When care is bilateral, the literature explains that the regimes of no liability and negligence generate socially efficient care by the two players (the producer and the consumer), while strict liability does not. Thus, strict liability must be augmented with a defense of contributory negligence on the part of consumer to restore incentives for socially efficient care. It means that the producer will be strictly liable if the consumer takes due care, while the producer has no liability if the consumer has failed to take due care. By analysing the specific market relation between the AI producer and the AI user, we go further in this direction at several levels.

Some papers have already studied optimal liability rules for the use of AI systems, notably Shavell (2020). He explains that none of the liability rules, mainly strict liability and fault-based liability, would perform well to reduce the accident risk. Then, he describes an alternative liability rule, so called the strict liability with damages paid to the state, that would be able to regulate the accident risk. The main change is that this liability requires to make a payment to the State (for the damage) rather than to the party who has suffered harm. It differs from our case since these AI systems, typically autonomous motor vehicles, are fully autonomous, which changes radically the analysis. Still in connection with autonomous motor vehicles, David and Muelheuser (2022) examine how liability regime may influence the innovation phase of these new technologies. It could serve as a source of inspiration as we do not study here the impact of liability rules on previous steps of analysis such as AI producer's market entry or innovation.

The combinations of liability rules we analyse make sense in settings were decisions are jointly done by AI-based technology and humans, which represents a mix of the previous cases. This type of decisions making has been studied by Agrawal et al. (2019). They make the distinction between two fundamental tasks: prediction and judgment. These two tasks are inputs into decision-making. Using this model, Obidzinski and Oytana (2022) study optimal liability law according to behavioral biases of AI users, using simple sharing rules to compensate victims. As in other papers on joint and shared liability between multiple injurers, they consider symmetric agents facing the same liability rules. In a duopoly market with vertical differentiation, Chopard and Musy (2023) study the effects of liability sharing between the AI producer and the AI user on the price of AI products and their levels of quality. They show that the transfer of liability from AI users to firms is priced into AI products, which makes this neutral in the end.

Besides the context related to the study of a medical relationship, our analysis differ from the previous ones since we consider a joint decision between two parties that can face specific liability regimes. We study the sharing of the compensation payment between the two parties. For example, going back to the first example, that is strict liability with a defense of contributory negligence on the part of the AI user. Usually, if the AI user is negligent, then she will solely pay the entire amount of compensation. In our context, that is questionable. One may suggest that the negligent AI user could share the payment of compensation with the AI producer as, under strict liability, the AI producer is legally responsible even in the absence of fault or negligence. The payment of compensation between the two parties can be divided, equally or not. Such a division of compensation may also arise when both the AI producer and the AI user are subject together to either strict liability or negligence rule (when both parties are negligent). Due to this division of payment of compensation between the two parties, we show that the type of levels of care, either substitute or complement, becomes important. The two cases must be considered. If they are substitutes (resp. complements), then the AI user will increase (resp. decrease) her amount of care when the amount of care chosen by the AI producer decreases (resp. increases). We then highlights which liability rules and payment sharing configurations lead to an optimal level of care. We highlight several mechanisms at play and we show which combinations are needed to obtain en efficient situations. We notably show that imposing strict liability to the AI producer and a pure defense of contributory negligence on the AI user, which resembles to the existing situation, is efficient.

3 The Model

3.1 Assumptions

To represent the market for artificial intelligence in healthcare, we use of a sequential model *à la* Daughety and Reinganum (2014) with bilateral care and proportional harm. We assume a market with a single AI producer and many identical AI users or physicians. In the whole document, we employ the terms 'physician' and 'AI user' interchangeably. The AI producer is a monopolist and chooses q the quantity of AI system to produce. The variable q also represents the intensity of use of AI tools in health care by a representative physician. For sketch of simplicity, we assume that a physician needs one unit of AI system to treat one patient.

For each patient treated with AI (or for each unit of AI system), the expected patient's harm depends on amounts of care chosen by both the AI producer and the AI user. This is a sequential care-taking, wherein the AI producer takes care in developing and manufacturing the AI system, and the physician takes care in using the AI system to treat q patients. Both levels of care reduce the expected harm of each patient treated with AI, denoted as $h(x, y)$. In case of compensation of the harm due to a medical error, this expected patient's harm equals the expected compensation payed to the patient, which is then equals to $h(x, y)$.

The AI producer's production cost, including care, is constant per unit of AI system and equal to xq. The physician's cost of care is also constant per unit of AI system and equal to yq. These are crucial assumptions. If (i) total production cost and cost of care on part of the AI producer, xq, is constant per unit of AI system (or per patient treated with AI), q, and (ii) total effort cost for taking care on the part of the AI user, yq, is also constant per unit of AI system, q, then both welfare and AI producer's profit can be expressed on a per AI user basis. In this paper, we thus examine the interaction between one monopolist with a single representative AI user as all physicians are assumed to be identical.

The game runs as follows.

- Step 1: The AI producer simultaneously chooses a level of care, x, and the quantity of AI system to produce, q. The quantity of AI system, the level of care and the liability regime imply a price p for each unit of AI system.

- Step 2: The AI user observing p and x, acquires an amount q of the AI system, anticipating the amount of care y she will choose to employ during the use of each unit of AI system.

- Step 3: The AI user chooses her level of care y per patient treated with AI (or equivalently per unit of AI system) subject to the constraints reflecting the liability rules and the the sharing of the expected compensation to patient harm between the AI user and the AI producer once a medical error occurs.

We assume that the AI producer is able to anticipate correctly the AI user's marginal willingness to pay for the AI system, as well as the care the AI user will subsequently take when using AI system. We denote $u(q)$ the AI user's utility when she uses q units of AI system to treat q patients. The higher the use of AI system, the higher the utility of the AI user. This captures the fact that combining AI system with clinician's decision improves the patient welfare. Any increase in patient's welfare appears in our altruistic physician's utility. We have $u'(q) > 0$ and $u''(q) < 0$.

The expected patient's harm (or compensation) h per unit of q can be reduced by choosing a level of care x for the AI producer and y for the AI user, with the total expected harm $h(x, y)q > 0$ for each AI user. Here, we clearly see that the patient's expected harm or compensation depends on a combined human - AI decision, the level of care of the later being decided by the product supplier. We assume that $h_x < 0$, $h_y < 0$, $h_{xx} > 0$ and $h_{yy} > 0$. It means that for each player, raising its level of care reduces expected harm but at a diminishing rate. We assume that $h_{xx}h_{yy} - h_{xy}^2$ for our problem to be concave. The sign of the cross derivatives h_{xy} and h_{yx} can be positive or negative, according to whether we consider that levels of care are either complements or substitutes. We denote $\pi(q, x, y)$ the firm's profit for any given choice of (q, x) and a given expected AI user's level of care y.

3.2 Social Welfare

The social objective is to maximize the net social benefits from the use of AI in the healthcare sector. The social planner would maximize the utility $(u(q))$ derived from the use of AI system in healthcare, minus (i) the expected patient's harm $(h(x, y)q)$ caused by negligence, misuse or product defect when an AI system is applied to enhance the patient care, and (ii) the sum of the cost of care (yq) borne by the AI user and the cost of care (xq) borne by the AI producer. Therefore, the social optimization problem is given by :

$$\max_{q,x,y} W(q, x, y) = u(q) - h(x, y)q - xq - yq \tag{1}$$

The socially efficient outcomes x^{SE}, y^{SE} and q^{SE} are respectively given by the following first-order conditions:

$$-h_x(x^{SE}, y^{SE}) = 1 \tag{2}$$

$$-h_y(x^{SE}, y^{SE}) = 1 \tag{3}$$

$$u'(q^{SE}) = h(x^{SE}, y^{SE}) + x^{SE} + y^{SE} \tag{4}$$

From equations (2) and (3), we obtain x^{SE} and y^{SE} such that the marginal benefits associated to patient's harm reduction are equal, because the AI user's and the AI producer's marginal costs of care are equal. Given those values, equation (4) gives the socially efficient quantity of AI. This quantity is such that the marginal gain from applying an additional unit of AI to take care of patients equals the marginal social cost of using this unit of AI, which includes both the expected patient's harm in case of medical error and the amounts of care chosen by AI user and AI producer.

Below, we explore how the liability regime either based on fault, or no fault principle, affects the choice of amounts of care and indirectly the patient's welfare. We especially study the influence of the way that the liability rule shares the payment of compensation for patient's harm between the AI user and the AI producer. We put aside the analysis of the quantity of AI: the AI producer, acting as a monopolist, is producing too little output.

4 Strict liability

Under a regime of no liability, it is straightforward to show that both the AI user and the AI producer do not take any care: $x^* = y^* = 0$. Under a strict liability regime, the patient harm has to be compensated, whatever the level of care of agents. This payment can be borne by a single agent, being either the AI user or the AI producer. But it can also be shared between the two agents as explained below.

Assumption 1 *Under strict liability, once a medical accident occurs, the compensation for patient's harm is shared out between the AI user and the AI producer. We denote $\beta \in [0, 1]$ the share of the compensation borne by the AI user. Thus $\beta h(x, y)$ is borne by the AI user while $(1 - \beta)h(x, y)$ is borne by the AI producer.*

The share of the compensation paid by the AI user or the AI producer (that is β) is not a function of the amounts of care x and y. Solely, the expected patient's harm $h(x, y)$ (or the compensation) depends on the levels of care. In other words, we do not consider the rule of comparative causation such that both parties' shares of compensation are functions of levels of care.

Assume that Assumption 1 holds. We solve our game by backward induction. At step 3, for given values x and q, the AI user solves the program:

$$\min_{y} q\beta h(x, y) + yq \tag{5}$$

The first order condition gives:

$$-\beta h_y(x, y^*) = 1 \tag{6}$$

where $y^*(x)$ is the AI user's best choice of care. This choice depends on (i) the share of compensation borne by the AI user (β), and (ii) the amount of care x chosen by the AI producer at step 1. If the payment of compensation is borne by the single AI user (that is $\beta = 1$), then we have $y^*(x) = y^{SE}(x)$ for all x. As a consequence, the AI user's choice of care for any given level of investment in care made by the AI producer is the efficient amount. However, it does not ensure that, at equilibrium, we have $y^* = y^{SE}$. It will depend on the AI producer's choice of care at step 1. When the share of compensation borne by the AI user β decreases, it reduces the his incentive to invest in care for any given level of investment in care made by the AI producer. If $\beta < 1$ then $y^*(x) < y^{SE}(x)$ for all x. If the payment of compensation is borne by the single AI producer ($\beta = 0$), then the AI user has no incentives to invest in care for any given level of investment in care made by the AI producer, that is $y^*(x) = 0$ for all x.

At step 2, expecting the choice $y^*(x)$ found in step 3, the AI user's choice of quantity is solution of the program:

$$\max_q u(q) - \beta h(x, y^*(x))q - qy^*(x) - pq \tag{7}$$

where p is the unit price of AI system paid by the AI user. The first-order condition yields the following inverse demand function for AI system:

$$p(q, x|y^*(x)) = u'(q) - \beta h(x, y^*(x)) - y^*(x) \tag{8}$$

The share of compensation borne by the AI user (β) has a direct and indirect effect on the AI user's willingness-to-pay for AI system. Consider first the direct effect. When β rises, the legal pressure on the AI user increases. The use of AI system becomes more costly due to the higher payment of compensation in case of medical accident. Consequently, the AI user's demand for AI system diminishes. Hence, there is a potential trend of healthcare professionals scaling back their involvement in patient care with AI. The apprehension of potential liabilities may lead these professionals to reconsider their use of AI, despite its significant potential for progress in medical detection, prevention, and treatment of diseases.

Next, two indirect effects are present because the share of compensation borne by the AI user will also influence the amount of care chosen by the AI user at step 3. If legal pressure on AI user increases (higher β), then the AI user chooses higher amount of care. The first indirect effect is that the AI user's cost of care increases (see the last term, $y^*(x)$, in Equation (8)). As the use of AI system becomes more costly, the demand for AI system diminishes, as well as the AI user's willingness to pay for. The second indirect effect is that a higher amount of care decreases the expected patient's harm or the magnitude of compensation ($h(x, y^*(x))$). The liability cost diminishes, which makes the AI system more attractive. The AI user's willingness to pay for AI system increases. In the end, the two indirect effects run in opposite directions. Note that when the level of care from the AI producer increases, expected harm diminishes, which raises the demand for the AI system.

Finally, at step 1, expecting the amount of care $y^*(x)$ chosen by the AI user at step 3, as well as the AI user's willingness to pay for AI system set at step 2, the AI producer chooses an amount of care that solves:

$$\max_{x} \pi(q, x|y^*(x)) = [u'(q) - \beta h(x, y^*(x)) - y^*(x)]\, q - (1-\beta)h(x, y^*(x))q - qx \tag{9}$$

The first order condition writes:

$$-h_x(x, y^*(x)) - \left[(1-\beta)h_y(x, y^*(x))\frac{\partial y^*(x)}{\partial x}\right] = 1 \tag{10}$$

Comparing Equation (14) with Equation (2), it appears an additional term at the left hand side, under the condition that $\beta < 1$. This term results from the sharing of the payment of patient's compensation between the AI producer and the AI user. This new term may be either positive or negative, thus being either a marginal benefit or cost. The sign of this term depends on the sign of $\frac{\partial y^*(x)}{\partial x}$. Using equation (6) and the implicit functions theorem, we find that[4]:

$$\frac{\partial y^*(x)}{\partial x} = -\frac{h_{yx}}{h_{yy}} \tag{11}$$

Since we have assumed that $h_{yy} > 0$, then the sign of the cross term h_{yx} (which indicates how x and y substitute so as to provide any constant level of $h(x, y)$) determines the sign of the slope of the AI user's reaction function $y^*(x)$ at step 3. Whether $y^*(x)$ is downward or upward-sloping indicates whether both amounts of care are either substitutes or complements at equilibrium. More precisely, if $h_{yx} > 0$ then x and y behave like substitutes. As the AI producer is moving first in our game, the AI user will increase (resp. decrease) her amount of care if the amount of care chosen by the AI producer decreases (resp. increases). At the opposite, if $h_{yx} < 0$ then x and y behave like complements at equilibrium. Both amounts of care increase or they both decrease at equilibrium.

Next, we replace $\frac{\partial y^*(x)}{\partial x} = -\frac{h_{yx}}{h_{yy}}$ in equation (10). The equilibrium AI producer's choice of care is solution to:

$$-h_x(x, y^*(x)) + \frac{(1-\beta)h_y(x, y^*(x))h_{yx}}{h_{yy}} = 1 \tag{12}$$

where $y^*(x) < y^{SE}(x)$. Below, we compare Equation (16) with Equation (2) for two (sets of) values of the share of the compensation borne by the AI user: $\beta = 1$ and $\beta < 1$.

First, assume that the payment of compensation is borne by the single AI user, that is $\beta = 1$. The AI producer has no liability while the AI user faces strict liability. According to Equation (6) and Equation (12), we find that $x^* = x^{SE}$ and $y^* = y^{SE}$. It results in the socially efficient choices of level of care. As the AI user solely bears the payment of compensation, she will choose

[4] For easier exposition of the section, we present the second derivatives without references to the variables.

the socially efficient level of care for any given level of investment made by the AI producer. We find that the AI producer will choose his level of care x so that the AI user meets the due-care standard by choosing $y^* = y^{SE}$. Put differently, the AI producer also chooses the socially efficient level of care at equilibrium $x^* = x^{SE}$ even if he has no liability. The reason is that from the AI producer's viewpoint, investment in care (x) becomes profitable because it transfers into a higher willingness-to-pay of the AI user who bears the entire payment of compensation in case of medical accident. This creates an incentive to invest $x^* = x^{SE}$ at equilibrium. Then, the AI user responds via her best-response function to the AI producer's choice of care of x^{SE}. It gives $y^* = y^{SE}$. In the end, it is the AI user who pays for the entire care, directly for (y) and indirectly for (x).

Now, we study the implication of the assumption $\beta < 1$. In this case, the way in which the quantities of care of each agent interact plays an important role. Two cases must be considered. First, we assume that amounts of care are substitutes. Next, we assume that amounts of care are complements.

Care levels as substitutes

Assume that $h_{yx} > 0$. Both levels of care behave like substitutes. Next, remind that the AI producer, moving first, can influence the AI user's care level. Look at the first term in left-hand side of Equation (12), that is $-h_x(x, y^*(x))$. This term is the AI producer's direct marginal benefit of care. Increasing investments in care reduces the expected payment of compensation in case of medical accident. It does not depend on the liability cost sharing rule (β). This results from the market relationship between the AI user and the AI producer. More precisely, the share of compensation borne by the AI user, β, reflects in an higher willingness-to-pay on the part of the AI user. Consequently, when choosing x, the AI producer takes into account not only his share of compensation $(1 - \beta)$ but also the AI user's share (β) because this becomes valuable to the AI producer. A second point is important here. As a decrease in the AI producer's care level causes an increase in the AI user's care level (as they are substitutes), then the AI producer could shift some of the costs of care to the AI user who is moving second in our game. However, if the AI producer ever did that, this would be reflected in a lower marginal willingness-to-pay on the part of the AI user for AI system. Thus, such a strategy is not valuable for the AI producer. Finally, as $h_{yx} > 0$ (or $h_{xy} > 0$), $h_x < 0$ and $y^*(x) < y^{SE}(x)$ under the condition $\beta < 1$ at step 3, we have $h_x(x, y^{SE}(x)) > h_x(x, y^*(x))$. Thus, $-h_x(x, y^*(x)) > -h_x(x, y^{SE}(x))$. That means the AI producer has higher incentive (in comparison with the socially efficient case) to invest in care because levels of care are substitutes.

Next, look in detail the second term in left-hand side of Equation (16), that is $\frac{(1-\beta)h_y(x,y^*(x))h_{yx}}{h_{yy}}$. As $h_{yx} > 0$, this term is negative. Remember that an AI producer's increased investment in care reduces the AI user's level of care as both are substitutes at equilibrium. Consequently, the AI producer's payment of compensation tends to increase due to the AI user's lower investment in care (for any given increase of the AI producer's level of care). This marginal cost is proportional to the share of compensation borne by the AI producer, $1-\beta$. This marginal cost does not appear in Equation (2) determining the socially efficient level of care from the AI producer. Thus, this additional marginal cost reduces the AI producer's incentive for investment in care. The two effects depicted above run in opposite direction to each other. Thus, under *Property* 1, if both levels of care are substitutes, the AI producer generally does not choose the socially efficient level of care. The AI producer can over-invest in care ($x^* > x^{SE}$). This induces the AI user to under-invest in care ($y^* < y^{SE}$) at equilibrium. But the AI producer may also under-invest in care ($x^* < x^{SE}$). In such a situation, the AI user will therefore over-invest in care ($y^* > y^{SE}$) or under-invest in care ($y^* < y^{SE}$) at equilibrium.

Care levels as complements

Assume that $h_{yx} < 0$. Both levels of care behave like complements.The AI producer's direct marginal benefit of care in Equation (12) still equals $-h_x(x, y^*(x))$. When the AI producer invests in care, he reduces the expected amount of compensation payed to the patient in case of medical accident. As $h_{yx} < 0$ (or $h_{xy} < 0$), $h_x < 0$ and $y^* < y^{SE}$ under the condition $\beta < 1$ at step 3, we have $h_x(x, y^{SE}(x)) < h_x(x, y^*(x))$. Thus, $-h_x(x, y^*(x)) < -h_x(x, y^{SE}(x))$. The AI producer has therefore a lower incentive (in comparison with the socially efficient case) to invest in care. Next, study the second term in left-hand side of Equation (16), that is $\frac{(1-\beta)h_y(x,y^*(x))h_{yx}}{h_{yy}}$. As $h_{yx} < 0$, this term is positive. An AI producer's increased investment in care increases the AI user's level of care as both are complements at equilibrium. Consequently, the AI producer's payment of compensation tends to decrease due to the AI user's higher investment in care (for any given increase of the AI producer's level of care) because both levels of care are complements. Consequently, the AI producer's payment of compensation tends to decrease due to the AI user's higher investment in care. This marginal benefit is proportional to the share of compensation borne by the AI producer, $(1-\beta)$. Notice that this marginal benefit does not appear in Equation (2). Consequently, this additional marginal benefit increases the AI producer's incentive for investment in care. The two effects depicted above run in opposite direction. If both levels of care are complements and *Property* 1 holds, then the AI producer generally does not choose the socially efficient level of care. If the AI producer over-invests in care ($x^* > x^{SE}$), then the AI user will respond at equilibrium by either under-investing in care ($y^* < y^{SE}$) or over-investing in care ($y^* > y^{SE}$). But the AI producer may also under-invest in care at equilibrium ($x^* > x^{SE}$). This induces the AI user to under-invest in care too ($y^* < y^{SE}$) at equilibrium.

We summarize our main results in proposition below. Let us recall that x^* and y^* denote the levels of care chosen by the AI producer and the AI user at equilibrium when both players are subject to strict liability.

Proposition 1 *When the AI user faces a regime of strict liability while the AI producer has no liability, we have $x^* = x^{SE}$ and $y^* = y^{SE}$, socially efficient levels of care are chosen by both agents. When both the AI producer and the AI user face a regime of strict liability and Assumption 1 holds, then two cases must be considered, none of them being efficient:*

(i) Assume that levels of care are substitutes. If $x^* > x^{SE}$ then $y^* < y^{SE}$. But, if $x^* < x^{SE}$ then $y^* < y^{SE}$ or $y^* > y^{SE}$.

(ii) Assume that levels of care are complements. If $x^* > x^{SE}$ then $y^* < y^{SE}$ or $y^* > y^{SE}$. But, if $x^* < x^{SE}$ then $y^* < y^{SE}$.

When both the AI producer and the AI user have strict liability, one has to think about the division of payment of compensation between the two parties in case of medical accident. The key point is that as soon as one reduces the share of compensation borne by the AI user (and simultaneously transfers this part of compensation to the AI producer), this reduces the AI user's incentive to invest in care, which may be socially inefficient. Due to this division of compensation between the two parties, it also appears that the type of care matters. More precisely, the type of care, complement or substitute, is helping to determine the AI producer's choice of care at first step as well as the AI user's reaction at last step of the game. If levels of care are substitutes then as the AI producer is moving first, the AI user will increase (resp. decrease) her amount of care if the amount of care chosen by the AI producer decreases (resp. increases). At the opposite, if levels of care behave like complements then both amounts of care increase or they both decrease at equilibrium.

First, assume that levels of care are substitutes. As the AI user tends to reduce her investment in care (her liability exposure is reduced by the division of compensation), the marginal product of AI producer's investment in care increases. The AI producer is prompted to invest more in care. But, at the same time, any AI producer's increased investment in care reduces the AI user's level of care as both are substitutes. This second effect is higher the more the share of compensation borne by the AI producer is large. As the division on compensation between the two parties encourages the AI user to reduce her investment in care, it is expected that the AI user will under-invest in care while the AI producer will over-invest in care as both levels are substitutes. This outcome is plausible. But, due to the above effects, another outcome may occur. There could be under investment in care by the AI producer at equilibrium. Under this scenario, the AI user's choice of care is unclear.

Suppose the AI producer and the AI user are still subject to strict liability, but their levels of care are complements. Contrary to the previous case, the marginal product of AI producer's investment in care decreases (in comparison with the socially efficient case). This reduces the AI producer's incentive to invest in care. But a second effect also works in reverse. An AI producer's increased investment in care increases the AI user's level of care, as both are now complements, further reducing the AI producer's expected payment of compensation. In the end, as the division on compensation between the two parties still encourages the AI user to reduce her investment in care, one plausible outcome is that both the AI user and the AI producer under-invest in care. But, due to the effects depicted above, the AI producer could also over-invest in care at equilibrium in which case the AI user's reaction is unclear.

5 Negligence rule

In this section, we explore the negligence rule. Under this liability rule, a solely negligent party bears the entire compensation payed to the injured patient. The compensation is shared between the AI producer and the AI user, only when both parties are negligent. There is no comparative causation, meaning that parties do not bear a compensation proportional to their causal contribution to the accident loss (for instance, proportional to their amounts of care)[5]. Instead, the sharing of the compensation between the AI user and the AI producer is

[5]Other negligence based liability rules exist (Parisi and Fon, 2004). These rules are based on comparative causation, the only tort regime that allows injurers and victims to share accident losses. Their setting is different because we rather study the compensation sharing between two injurers, the AI user and the AI producer. Moreover, the accident results from the interaction of two parties who are strangers to each other. Thus, contrary to our game, these parties are unable to allocate the risk of accident between them prior an

exogenously determined.

Assumption 2: *A non-negligent party has no liability, regardless the amount of care chosen by the other party. The entire compensation for accident loss is borne by the negligent party when the other party happens to be diligent. When both parties are negligent, then the compensation is exogenously shared between them. We denote β the share of compensation payed by the negligent AI user. The share $1-\beta$ is born by the negligent AI producer.*

The game between the AI user and the AI producer is still played in three steps.

5.1 The AI user's choice of level of care

At step 3, under Assumption 2, after having purchased q units of AI, the AI user must decide on the level of care to take which is solution to:

$$\min_y \left\{ \begin{array}{c} yq \text{ if } y \geq y^{SE} \text{ and regardless of the value of } x \\ yq + \beta h(x,y)q \text{ if } y < y^{SE} \text{ and } x < x^{SE} \\ yq + h(x,y)q \text{ if } y < y^{SE} \text{ and } x \geq x^{SE} \end{array} \right\} \tag{13}$$

where x^{SE} and y^{SE} are the socially efficient amounts of care taken by the AI producer and the AI user respectively. The expression of the expected harm depends on both the AI user's and the AI producer's care decisions, while the cost of taking care remains the same in the three cases. When the negligence rule prevails, the AI user's expected harm depends on whether she, as well as the AI producer, will choose the socially efficient amount of care. We successively study each line of Equation (13).

First consider the AI user's optimal decision when $x \geq x^{SE}$ at step 1. First-order condition yields that $y^*(x) = y^{SE}(x)$ for all x, when $y^*(x)$ is the optimal AI user's choice of care for any given amount of investment in care made by the AI producer. We find that the equilibrium AI user's choice of care (for any amount of investment in care made by the AI producer) is the efficient amount. However, this does not ensure that the equilibrium level of care under

accident occurs. We give two examples. First, under the *rule of comparative negligence*, a solely negligent party still bears the entire accident loss. However, the accident loss is now proportionally shared between the injurer and the victim when both are negligent. Negligent parties bear accident loss in shares that are proportional to their causal contribution to the accident loss, that is to their relative amounts of care. One may also consider the *rule of comparative causation under negligence*. Anew, a solely negligent party bears the entire accident loss. However, the accident loss is proportionally shared between the two parties, when both are negligent *or* when both are non-negligent.

negligence rule is the same as y^{SE}. Indeed, assume that at step 1, the AI producer chooses x^{SE}. Then at step 3, we will have $y^* = y^{SE}$. But, assume that $x > x^{SE}$ at step 1. At step 3, the optimal AI user's choice of care will depend on the sign of the cross derivative h_{xy}. If $h_{xy} > 0$ then x and y are substitutes. As a consequence, if $x > x^{SE}$ then the optimal AI user's response gives $y < y^{SE}$ according to her equilibrium best response function $y^*(x) = y^{SE}(x)$. When $h_{xy} < 0$: x and y are complements. If $x > x^{SE}$, one might believe that $y > y^{SE}$ given that $y^*(x) = y^{SE}(x)$. But, according to Equation (17), the AI user's cost function of care becomes yq if $y > y^{SE}$. Thus, minimizing the cost of care leads the AI user to choose $y^* = y^{SE}$ at equilibrium under two conditions: (i) $x > x^{SE}$ and (ii) x and y are complements. More generally, we find that the AI user has no incentives to choose $y > y^{SE}$ whatever the level of investment in care made by the AI producer at step 1.

Finally, we consider the AI user's optimal decision when the AI producer chooses $x < x^{SE}$ at step 1. As explained before, the AI user will not choose at equilibrium $y > y^{SE}$. Next, for every $x < x^*$, the AI user's best response is either y^* or one value function for y that solves $1 = -\beta h_y(x, y)$. This first-order condition gives the AI user's optimal choice of care denoted as $y^*(x) < y^{SE}(x)$ for $\beta < 1$ under the condition that for x the AI user's cost of care is lower if she does not meet the socially efficient amount of care y^{SE}. In this case, the optimal AI user's choice of care for any given amount of care made by the AI producer such as $x < x^{SE}$ is strictly inferior to the efficient amount. It is the result of the sharing of the compensation between the two parties. For $\beta = 1$ (we exclude this limit case here[6]), we have $y^*(x) = y^{SE}(x)$. Thus, if x and y are substitutes or complements, then the optimal AI user's response is $y^*(x) < y^{SE}(x)$. This could lead at equilibrium that $y^* \leq y^{SE}$.

For summary, if x and y are complements then the equilibrium AI user's choice of care is $y^* = y^{SE}$ if $x \geq x^{SE}$. If the AI producer chooses $x < x^{SE}$ at step 1, then the equilibrium AI user's choice of care is $y^*(x) \leq y^{SE}$ where $y^*(x)$ is solution to $1 = -\beta h_y(x, y)$. If x and y are substitutes, then the AI user's best response function is $y^*(x) = y^{SE}(x) < y^{SE}$ if $x \geq x^{SE}$ at step 1. If the AI producer chooses $x < x^{SE}$ at step 1, then the AI user's best response function is $y^*(x) \leq y^{SE}$ where $y^*(x)$ solves $1 = -\beta h_y(x, y)$.

[6]When $\beta = 1$ then the payment of compensation is solely borne by the AI user, thus meaning that the AI producer has no liability even if he is negligent. This is not on line with Assumption 2.

5.2 The AI user's choice of AI quantity to demand

We next turn to finding the AI user's choice of quantity to demand at step 2. The AI user's marginal willingness-to-pay for AI comes from the maximization problem for the AI user, where she rationally anticipates using her best choice-of-care at step 3. We have shown that the AI user's best choice-of-care depends on whether x and y are substitutes or complements. We consider each case successively.

Care levels as complements

Assume that x and y behave like complements. The AI user's best choice-of-care rule is either y^{SE} if $x \geq x^{SE}$, or $y^*(x) \leq y^{SE}$ that solves $1 = -\beta h_y(x,y)$, if $x < x^{SE}$ as explained above. Thus, the AI user's program at step 2 is:

$$\max_q \left\{ \begin{array}{c} u(q) - pq - y^{SE}q \text{ if } x \geq x^{SE} \\ u(q) - \beta h(x, y^*(x)) - pq - y^*(x)q \text{ if } x < x^{SE} \end{array} \right\} \tag{14}$$

where p is the price paid by the AI user for each unit of AI. The first order conditions yield the inverse demand functions:

$$p(q) = u'(q) - y^{SE} \text{ if } x \geq x^{SE}$$
$$p(q, x, y^*(x)) = u'(q) - y^*(x) - \beta h(x, y^*(x)) \text{ if } x < x^{SE}.$$

Care levels as substitutes

Assume that x and y behave like substitutes. The AI user's best choice-of-care rule is either $y^{SE}(x) \leq y^{SE}$ if $x \geq x^{SE}$, or $y^*(x) \leq y^{SE}$ if $x < x^{SE}$. Thus, the AI user's program at step 2 is:

$$\max_q \left\{ \begin{array}{c} u(q) - h(x, y^{SE}(x)) - pq - y^{SE}(x)q \text{ if } x \geq x^{SE} \\ u(q) - \beta h(x, y^*(x)) - pq - y^*(x)q \text{ if } x < x^{SE} \end{array} \right\} \tag{15}$$

where p is the price paid by the AI user for each unit of AI. The first order conditions yield the inverse demand functions:

$$p(q, x, y^{SE}(x)) = u'(q) - y^{SE}(x) - h(x, y^{SE}(x)) \text{ if } x \geq x^{SE}$$
$$p(q, x, y^*(x)) = u'(q) - y^*(x) - \beta h(x, y^*(x)) \text{ if } x < x^{SE}.$$

5.3 The AI producer's choice of care

We turn to the step 1 of the game. We again study two cases: first, x and y are complements, second, x and y are substitutes.

Care levels as complements

When negligence rule prevails, the AI producer's program at step 1 is:

$$\max_x \left\{ \begin{array}{c} \pi(q,x,y^{SE}) = (u'(q) - y^{SE})q - xq \text{ if } x \geq x^{SE} \\ \pi(q,x,y^*(x)) = (u'(q) - y^*(x) - \beta h(x,y^*(x)))q - xq - (1-\beta)h(x,y^*(x))q \text{ if } x < x^{SE} \end{array} \right\} \tag{16}$$

We do not need the details of the AI producer's choice of x. Notice that $\max_x \{(u'(q) - y^{SE})q - xq\}$ yields the AI producer to choose $x = x^{SE}$. Thus, for $x = x^{SE}$, the profit equals $(u'(q) - y^{SE})q - x^{SE}q$. The AI producer will choose $x < x^{SE}$ if and only if there does exist a level of care x such that:

$$(u'(q) - y^*(x) - \beta h(x,y^*(x)))q - xq - (1-\beta)h(x,y^*(x))q > (u'(q) - y^{SE})q - x^{SE}q.$$

where $y^*(x)$ is solution to $1 = -\beta h_y(x,y)$. We can rewrite this condition as:

$$(x^{SE} - x) > (y^*(x) - y^{SE}) + h(x,y^*(x)). \tag{17}$$

The AI producer will under-invest in care, that is $x < x^{SE}$, if the gain associated to the reduction of the cost of care per unit of AI, as measured by $(x^{SE} - x)$, is superior to the expected harm per unit of AI (as the AI producer is now liable because he chooses $x < x^{SE}$) plus an additional negative term, $(y^*(x) - y^{SE}) < 0$. We remark that the overall expected harm (or compensation) matters, not only the expected compensation payed by the AI producer. The AI producer takes into account her part of the expected compensation payed to the patient plus the portion paid by the AI user. This effect comes from the inverse-demand function. If the AI producer does not choose the socially efficient amount of care, then the AI user will not meet the socially efficient amount of care too. Due to this liability cost under negligence rule, the AI user will consequently reduce her demand for AI. That reduces the AI producer's incentive to increase her level of care in order to meet the standard of care. Further, we find that the higher is the share of compensation borne by the AI user (that is higher β), the more the AI user invests in care thereby decreasing the expected compensation payed in case of medical accident. Therefore more likely that the AI producer will under-invest in care at equilibrium. Finally, we see in Equation (17) that another element creates pressure on AI producer to reduce investment in care. When the AI producer anticipates that the AI user will respond to her decision to reduce care by under-investing in care (that is $y^*(x) - y^{SE} < 0$), it reinforces the AI producer's incentive to under-invest in care too.

Care levels as substitutes

Assume that x and y are substitutes. The AI producer's program at step 1 is:

$$\max_x \left\{ \begin{array}{c} \pi(q,x,y^{SE}(x)) = (u'(q) - y^{SE}(x) - h(x,y^{SE}(x)))q - xq \text{ if } x \geq x^{SE} \\ \pi(q,x,y^*(x)) = (u'(q) - y^*(x) - \beta h(x,y^*(x)))q - xq - (1-\beta)h(x,y^*(x))q \text{ if } x < x^{SE} \end{array} \right\} \tag{18}$$

We remark that $\max_x \{(u'(q) - y^{SE}(x) - h(x,y^{SE}(x)))q - xq\}$ yields the AI producer to choose $x = x^{SE}$. Thus, for $x = x^{SE}$, the AI producer's profit equals $(u'(q) - y^{SE} - h(x^{SE}, y^{SE}))q - x^{SE}q$. The AI producer will choose $x < x^{SE}$ at equilibrium if and only if there does exist a level of care x such that:

$$(u'(q) - y^*(x) - \beta h(x,y^*(x)))q - xq - (1-\beta)h(x,y^*(x))q > (u'(q) - y^{SE} - h(x^{SE},y^{SE}))q - x^{SE}q.$$

where $y^*(x)$ is solution to $1 = -\beta h_y(x,y)$. We can rewrite this condition as:

$$(x^{SE} - x) > (y^*(x) - y^{SE}) + h(x,y^*(x)) - h(x^{SE},y^{SE}) \tag{19}$$

The AI producer will under-invest in care ($x < x^{SE}$) if the gain associated to the reduction of the cost of care per unit of AI, measured by $(x^{SE} - x)$, is superior to the right-hand side of inequality (19). There are two terms at the right-hand side in (19). The first term, $y^*(x) - y^{SE}$, is negative. This element creates pressure on AI producer to reduce her amount of care. As explained above, if the AI producer anticipates that the AI user will respond to her decision to reduce care by under-investing in care, then the AI producer under-invests in care too. The second term at the right-hand side of (19), $h(x, y_\beta^{NR}(x)) - h(x^*, y^*)$, is positive. It means that the lower is the share of compensation payed by the AI user, the less the AI user invests in care. As a consequence, the expected compensation borne by the AI user increases, thereby making less likely that the AI producer under-invests in care. We summarize the main results in proposition below. Anew, let us recall that x^* and y^* denote the levels of care chosen by the AI producer and the AI user at equilibrium when both players are subject to negligence rule.

Proposition 2 *Assume that both AI user's liability and AI producer's liability are based on the principle of fault liability, and that Assumption 2 holds. We have $x^* \leq x^{SE}$ and $y^* \leq y^{SE}$. More precisely, if $x^* = x^{SE}$ then $y^* = y^{SE}$. But, if $x^* < x^{SE}$ then $y^* < y^{SE}$.*

When (i) both AI user's liability and AI producer's liability are based on the principle of fault liability and (ii) the payment of compensation is shared out between them if they are negligent, then neither the AI producer, nor the AI user will over-invest in care at equilibrium. If the AI producer chooses the efficient amount of care, then the AI user will also choose the efficient level of care. If the AI producer under-invests in care, then the AI user will under-invest in care at equilibrium too. These results hold whether both levels of care are complements or substitutes.

6 Different liability rules for AI user and AI producer

In this section, the arm caused to the patient is still compensated by the AI user and/or the AI producer. However, since the nature of their activities are different, we assume that they can face different liability regimes. For example, in many legal systems, AI user's liability is based on the principle of fault liability, while industrial producers face strict liability.

6.1 AI producer's strict liability and AI user's negligence rule

We assume that AI user's face a regime of fault liability, while a regime of strict liability is applied to AI producers. In this context, two situations are discussed.

In the first one, if the AI user chooses the socially efficient level of care, then the AI producer is strictly liable, while if the AI user fails to take socially efficient care, the payment of compensation will be shared out between the two parties. We call this issue an AI producer's strict liability with defense of contributory negligence on part of AI user.

In the second case, the AI producer is still strictly liable if the AI user takes due care. But the AI producer will no longer be liable for any harm (independent of the level of care chosen by the AI producer) if the AI user does not meet the due care standard (that is the socially efficient amount of care). In such a case, the AI user is responsible for all damages caused by the use of AI system. We call this issue an AI producer's strict liability with *pure* defense of contributory negligence on part of AI user.

6.1.1 AI producer's strict liability with defense of contributory negligence on part of AI user

In this section, we explore the implications of the following assumption, bearing in mind that the AI user operates under negligence rule while we apply strict liability to AI producer.

Assumption 3 *Under negligence rule, a non-negligent AI user has no liability, regardless the amount of care chosen by the AI producer subject to strict liability. Once a medical accident occurs, the AI producer will solely bear the payment of compensation to the patient if the AI user is non-negligent. Finally, a negligent AI user will have to share the payment of compensation with the AI producer as follows. The share β of compensation, with $0 < \beta < 1$, is borne by the negligent AI user. The share $1 - \beta$ of compensation is borne by the AI producer.*

The AI user's choice of level of care

At step 3, after having purchased q units of AI, the AI user must decide on the level of care to take which is solution to:

$$\min_y \left\{ \begin{array}{c} yq \text{ if } y \geq y^{SE} \text{ and regardless of the value of } x \\ yq + \beta h(x, y)q \text{ if } y < y^{SE} \text{ and regardless the value of } x \end{array} \right\} \tag{20}$$

where y^{SE} is the socially efficient level of care taken by the AI user. We successively study each line of Equation (20). We remark that the AI user has no incentive to choose at equilibrium $y > y^{SE}$. Next, regarding the second line of (20), we find that, whatever the amount of care chosen by the AI producer, the AI user's best response is either y^{SE} or one value function for y that solves $1 = -\beta h_y(x, y)$. This first-order condition gives the AI user's equilibrium choice of care denoted as $y^*(x)$. We have: $y^*(x) < y^{SE}(x)$ for $\beta < 1$. In words, the AI user is requested to reduce her investment in care as she shares the payment of compensation with the AI producer. Thus, the equilibrium AI user's choice of care is given $\min \left\{ y^*(x), y^{SE} \right\}$ where $y^*(x) < y^{SE}(x)$ is solution to $1 = -\beta h_y(x, y)$. The central point is that the AI user may have an incentive to under-invest in care, that is to choose $y \leq y^{SE}$. However, the AI user will never over-invest in care at equilibrium.

The AI user's choice of AI quantity to demand

We next turn to finding the AI user's choice of quantity to demand at step 2. The AI user's marginal willingness-to-pay for AI comes from the maximization problem for the AI user, where she rationally anticipates using her best choice-of-care at step 3. The AI user's program at step 2 is:

$$\max_q \left\{ \begin{array}{c} u(q) - \beta h(x, y^*(x))q - pq - y^*(x)q \text{ if } y^*(x) < y^{SE} \\ u(q) - pq - y^{SE}q \text{ if } y^*(x) \geq y^{SE} \end{array} \right\} \tag{21}$$

where p is the price paid by the AI user for each unit of AI. The first-order conditions yield the inverse demand functions:

$$p(q,x,y^*(x)) = u'(q) - y^*(x) - \beta h(x,y^*(x)) \text{ if } x \text{ such that } y^*(x) < y^{SE}$$
$$p(q,x) = u'(q) - y^{SE} \text{ if } x \text{ such that } y^*(x) \geq y^{SE}$$

where $y^*(x)$ is solution to $1 = -\beta h_y(x,y)$.

The AI producer's choice of care

We turn to the step 1 of the game. Under strict liability (except for the AI user subject to negligence rule), the AI producer's program at step 1 is:

$$\max_x \left\{ \begin{array}{c} \pi(q,x,y^*(x)) = u'(q) - y^*(x)q - xq - h(x,y^*(x))q \text{ if } x \text{ such that } y^*(x) < y^{SE} \\ \pi(q,x) = u'(q) - y^{SE}q - xq - h(x,y^{SE})q \text{ if } x \text{ such that } y^*(x) \geq y^{SE} \end{array} \right\} \quad (22)$$

Look at the second line in (22): the AI producer anticipates that the AI user will never over-invest in care when using her AI system. When comparing both lines in equation (22), we find that over or under-investment in care by the AI producer may lead the AI user to under-invest in care, that is $y^*(x) < y^{SE}$. Therefore, the AI producer may either go up or down. However, notice that by raising the investment in care, the AI producer can increase expected profit as the AI user's willingness-to-pay increases due to the reduction of expected harm induced by this increased investment in care. Moreover, even if the payment of compensation may be divided between the AI producer and the AI user under some condition, we find at step 1 that the market interaction between the two parties leads the AI producer to bear the full payment of compensation. As a consequence, this should not lead to a too lower investment in care by the AI user as the AI producer wants to limit the amount of compensation to be paid to injured patients. We summarize our results in Proposition 4 below.[7]

Proposition 3 *Assume that AI user's liability is based on the principle of fault liability while we apply strict liability to AI producer. Also assume that assumption 3 holds. Then the AI user*

[7]Look at the first line in Equation (22). The first-order condition is:

$$-h_x(x,y^*_x(x)) - y^*_x(x)(1 + h_y(x,y^*(x)) = 1 \quad (23)$$

Using the first-order condition coming from step 3 and the implicit functions theorem, we find that the sign of $y^*_x(x)$ is given by the sign of h_{yx}. More precisely, if x and y are complements (that is $h_{yx} < 0$) then $y^*_x(x) > 0$. If x and y are substitutes (that is $h_{yx} > 0$) then $y^*_x(x) < 0$. According to Equation (23), we find that $x^*(y) > x^{SE}(y)$ if levels of care are complements. If levels of care are substitutes then $x^*(y) < x^{SE}(y)$.

will never over-invest in care at equilibrium. The AI producer may either over-invest in care, under-invest in care, or choose the efficient amount of care. This results holds whether both levels of care are either complements or substitutes.

6.1.2 AI producer's strict liability with *pure* defense of contributory negligence on part of AI user

In this section, the AI user's liability is still based on the principle of fault liability, while we apply strict liability for AI producers. In comparison with subsection 5.1.1, we study the implications of a slightly different liability scheme outlined below (notice that Assumption 4 is equivalent to Assumption 3 when $\beta = 1$).

Assumption 4 *Under negligence rule, a non-negligent AI user has no liability, regardless the level of care chosen by the AI producer subject to strict liability. Thus, once a medical accident occurs, the AI producer will solely bear the payment of compensation if the AI user is non-negligent. A negligent AI user will have to pay the entire compensation in case of medical error. As a consequence, the AI producer has no liability if the AI user is negligent.*

The AI user's choice of level of care

At step 3, after having purchased q units of AI, the AI user must decide on the level of care to take which is solution to:

$$\min_y \left\{ \begin{array}{c} yq \text{ if } y \geq y^{SE} \text{ and regardless of the value of } x \\ yq + h(x,y)q \text{ if } y < y^{SE} \text{ and regardless the value of } x \end{array} \right\} \tag{24}$$

where y^{SE} is the socially efficient amount of care taken by the AI user.

Look at the first line of Equation (26). The AI user will never choose at equilibrium $y > y^{SE}$. Regarding the second line of Equation (26), we find that, whatever the amount of care chosen by the AI producer at step 1, the equilibrium AI user's choice of care is $y^*(x) = y^{SE}(x) \leq y^{SE}$. If the AI producer invests the socially efficient amount of care then the AI user will do the same.

If the AI producer over-invests in care (or chooses $x > x^{SE}$) then the AI user will choose $y^*(x) = y^{SE}(x) < y^{SE}$ if levels of care are substitutes. If they are complements then the AI user will invest the socially efficient amount of care.

Assume that the AI producer chooses to under-invest in care. Then the AI user will choose $y^*(x) = y^{SE}(x) < y^{SE}$ if levels of care are complements. If they are substitutes then the AI user will invest the socially efficient amount of care.

The AI user's choice of AI quantity to demand

We next turn to finding the AI user's choice of quantity to demand at step 2. The AI user's marginal willingness-to-pay for AI comes from the maximization problem for the AI user, where she rationally anticipates using her best choice-of-care.

When care levels are complements, the AI user's program at step 2 is:

$$\max_q \left\{ \begin{array}{c} u(q) - pq - y^{SE}q \text{ if } x = x^{SE} \\ u(q) - pq - y^{SE}q \text{ if } x > x^{SE} \\ u(q) - h(x, y^{SE}(x)) - pq - y^{SE}(x)q \text{ if } x < x^{SE} \end{array} \right\} \tag{25}$$

where p is the price paid by the AI user for each unit of AI. When care levels are substitutes, the AI user's program becomes:

$$\max_q \left\{ \begin{array}{c} u(q) - pq - y^{SE}q \text{ if } x = x^{SE} \\ u(q) - pq - y^{SE}q \text{ if } x < x^{SE} \\ u(q) - h(x, y^{SE}(x)) - pq - y^{SE}(x)q \text{ if } x > x^{SE} \end{array} \right\} \tag{26}$$

Let us do a comparison between (25) and (26). When the AI producer invests the socially efficient amount of care then the type of care, substitute or complement, does not matter because the AI user is induced to choose the socially efficient amount of care too. But, when the AI producer deviates from this path (and no longer invests the socially efficient amount of care), then the AI user (subject to negligence liability rule) will naturally continue to choose the socially efficient amount of care in only two of those cases: (i) under-investment in care by AI user with substitutable players' amounts of care and (ii) over-investment in care by AI user with complementary players' amounts of care. The AI user eventually under-invested in care if the AI producer chooses to over-invest (under-invest) in care while care levels are substitutes (resp. complements). On this basis, the first order conditions yield the following inverse demand functions. When care levels are substitutes, we have:

$$\begin{aligned} p(q, x) &= u'(q) - y^{SE} \text{ if } x = x^{SE} \\ p(q, x) &= u'(q) - y^{SE} \text{ if } x < x^{SE} \\ p(q, x, y^{SE}(x)) &= u'(q) - y^{SE}(x) - h(x, y^{SE}(x)) \text{ if } x > x^{SE} \end{aligned}$$

When care levels are complements, we have:

$$p(q,x) = u'(q) - y^{SE} \text{ if } x = x^{SE}$$
$$p(q,x) = u'(q) - y^{SE} \text{ if } x > x^{SE}$$
$$p(q,x,y^{SE}(x)) = u'(q) - y^{SE}(x) - h(x, y^{SE}(x)) \text{ if } x < x^{SE}$$

The AI producer's choice of care

We turn to the step 1 of the game. We have to distinguish between two cases depending on whether levels of care are either complements or substitutes.

Care levels as complements

Assume that levels of care are complements. Under strict liability (except for the AI user subject to negligence rule), the AI producer's program at step 1 is:

$$\max_x \left\{ \begin{array}{c} \pi(q,x) = (u'(q) - y^{SE})q - xq - h(x,y^{SE})q \text{ if } x \geq x^{SE} \\ \pi(q,x,y^{SE}(x)) = (u'(q) - y^{SE}(x) - h(x,y^{SE}(x)))q - xq \text{ if } x < x^{SE} \end{array} \right\} \quad (27)$$

If $x \geq x^{SE}$ then the first-order condition equals:

$$-h_x(x, y^{SE}(x)) = 1. \quad (28)$$

As a consequence, the AI producer will choose the efficient amount x^{SE}.

If $x < x^{SE}$ then the first-order condition, after some manipulation, still equals:

$$-h_x(x, y^{SE}(x)) = 1. \quad (29)$$

This result comes from the strict liability regime combined with the expected AI user's equilibrium best response. Indeed, if $x < x^{SE}$ the first-order condition equals:

$$-h_x(x, y^{SE}(x)) - y_x^{SE}(x)(1 + h_y(x, y^{SE}(x))) = 1 \quad (30)$$

But, using the first-order condition coming from step 3, we have $1 + h_y(x, y^{SE}(x)) = 0$. After replacing in (30), we find that the AI producer will choose the efficient amount x^{SE}.

Care levels as substitutes

Finally, assume that levels of care are substitutes. The AI producer's program at step 1 is:

$$\max_x \left\{ \begin{array}{c} \pi(q,x) = (u'(q) - y^{SE})q - xq - h(x,y^{SE})q \text{ if } x \leq x^{SE} \\ \pi(q,x,y^{SE}(x)) = (u'(q) - y^{SE}(x) - h(x,y^{SE}(x)))q - xq \text{ if } x > x^{SE} \end{array} \right\} \quad (31)$$

The same reasoning applies. There is no under or over-investment in care at equilibrium. For instance, when levels of care are substitutes, by over-investing in care, the AI producer could avoid paying the patient's compensation in case of medical error by encouraging the AI user to under-invest in care, thus becoming liable and paying compensation instead of the AI producer. But such a gain would be compensated by the AI producer's extra cost of investment in care. Thus, both the AI producer and the AI user will invest the efficient amount of care when both amounts of care are substitutes. We summarize our results in Proposition 5 below.

Proposition 4 *Assume that AI user's liability is based on the principle of fault liability while we apply strict liability to AI producer. Under assumption 4, we have:* $x^* = x^{SE}$ *and* $y^* = y^{SE}$.

One liability scheme will lead both the AI user and the AI producer to invest the socially efficient level of care, regardless of the type of levels of care being either complement or substitute. A non-negligent AI user will have no liability, regardless the level of care chosen by the AI producer subject to strict liability. Once a medical accident occurs, the AI producer solely pays the compensation if the AI user is non-negligent. But a negligent AI user will have to pay the entire compensation in case of medical error.

6.2 AI producer's negligence rule with AI user's strict liability

In this section, we apply strict liability for AI users while AI producer's liability is based on the principle of fault liability. We study two situations.

In the first case, if the AI producer chooses the socially efficient level of care, then the AI user is strictly liable, while if the AI producer fails to take socially efficient care, the payment of compensation will be shared out between the two parties. We call this AI user's strict liability with defense of contributory negligence on part of AI producer.

In the second case, the AI user is strictly liable if the AI producer takes due care. But the AI user will no longer be liable for any harm (independent of the level of care chosen by the AI user) if the AI producer under-invests in care. In such a case, the AI producer is responsible for all damages caused by the use of AI system. We call this AI user's strict liability with *pure* defense of contributory negligence on part of AI producer.

6.2.1 AI user's strict liability with defense of contributory negligence on part of AI producer

In this section, we explore the implications of the following assumption.

Assumption 5 *Under negligence rule, a non-negligent AI producer has no liability, regardless the amount of care chosen by the AI user subject to strict liability. Thus, once a medical accident occurs, the AI user will solely bear the payment of compensation if the AI producer is non-negligent. A negligent AI producer will share the payment of compensation with the AI user as follows. The share β of compensation is borne by the AI user. The share $1-\beta$ of compensation is borne by the negligent AI producer.*

The AI user's choice of level of care

At step 3, after having purchased q units of AI, the AI user must decide on the level of care to take which is solution to:

$$\min_{y} \left\{ \begin{array}{c} yq + h(x,y)q \text{ if } x \geq x^{SE} \\ yq + \beta h(x,y)q \text{ if } x < x^{SE} \end{array} \right\} \tag{32}$$

where x^{SE} is the socially efficient amount of care taken by the AI producer. Look at the first line of Equation (32). If $x \geq x^{SE}$ then the AI user's best response is $y^*(x) = y^{SE}(x)$. Regarding the second line of Equation (32), we find that the AI user's best response is now $y^*(x) < y^{SE}(x)$ where $y^*(x)$ is solution to $-\beta h_y(x,y) = 1$ under the condition that the AI producer chooses $x < x^{SE}$ at step 1.

The AI user's choice of AI quantity to demand

We next turn to finding the AI user's choice of quantity to demand at step 2. The AI user's marginal willingness-to-pay for AI comes from the maximization problem for the AI user, where she rationally anticipates using her best choice-of-care. The AI user's program at step 2 is:

$$\max_{q} \left\{ \begin{array}{c} u(q) - h(x, y^{SE}(x)) - pq - y^{SE}(x)q \text{ if } x \geq x^{SE} \\ u(q) - \beta h(x, y^*(x)) - pq - y^*(x)q \text{ if } x < x^{SE} \end{array} \right\} \tag{33}$$

where $y^*(x)$ is solution to $-\beta h_y(x,y) = 1$. We denote p the price paid by the AI user for each unit of AI. These two first-order conditions yield the inverse demand functions:

$$p(q, x, y^{SE}(x)) = u'(q) - y^{SE}(x) - h(x, y^{SE}(x)) \text{ if } x \geq x^{SE}$$
$$p(q, x, y^*(x)) = u'(q) - y^*(x) - \beta h(x, y^*(x)) \text{ if } x < x^{SE}$$

where $y^*(x)$ is solution to $-\beta h_y(x,y) = 1$.

The AI producer's choice of care

We turn to the step 1 of the game. Under negligence rule, the AI producer's program at step 1 is:

$$\max_x \left\{ \begin{array}{c} \pi(q,x,y^{SE}(x)) = (u'(q) - y^{SE}(x) - h(x,y^{SE}(x)))q - xq \text{ if } x \geq x^{SE} \\ \pi(q,x,y^*(x)) = (u'(q) - y^*(x) - \beta h(x,y^*(x)))q - xq - (1-\beta)h(x,y^*(x)))q \text{ if } x < x^{SE} \end{array} \right\} \tag{34}$$

If $x \geq x^{SE}$ the first-order condition equals:

$$-h_x(x,y^{SE}(x)) - y_x^{SE}(x)(1+h_y) = 1. \tag{35}$$

From analysis of step 3, we have: $(1+h_y) = 0$. Thus, $x^*(y) = x^{SE}(y)$. Further, we have previously shown that: $y^*(x) = y^{SE}(x)$. Consequently, If the AI producer chooses the efficient amount of care, then she anticipates that the AI user will also choose the efficient amount of care. Further, the AI producer has no incentive to over-invest in care.

If $x < x^{SE}$ we have shown that $y^*(x) < y^{SE}(x)$. Consequently, the AI user will choose $y^* < y^{SE}$ at equilibrium.

The AI producer can either choose the efficient amount of care, or under-invest in care. By comparing equilibrium profits in these situations, we find that the AI producer will under-invest in care and choose x (with AI producer choosing $y(x)$) at step 3 if and only if $h(x,y(x)) - h(x^{SE},y^{SE}) < (y^{SE} - y(x)) + (x^{SE} - x)$. In words, the AI producer will under-invest in care if the gain resulting from the reduction of the cost of care (for both AI user and AI producer) is superior to the increase of the amount of compensation that results from under-investing in care. If the above inequality does not hold, then both the AI producer and the AI user will choose the efficient levels of care. No-one over-invests in care at equilibrium. We summarize our results in Proposition 6 below.

Proposition 5 *Assume that AI producer's liability is based on the principle of fault liability while we apply strict liability to AI user. Under Assumption 5, we have: $x^* \leq x^{SE}$ and $y^* \leq y^{SE}$. More precisely, if $x^* = x^{SE}$ then $y^* = y^{SE}$. But, if $x^* < y^{SE}$ then $y^* < y^{SE}$.*

Consider that the AI producer operates under negligence rule. Thus, a non-negligent AI producer has no liability, regardless the amount of care chosen by the AI user, here subject to strict liability. Thus, once a medical accident occurs, the AI user will solely bear the payment of compensation if the AI producer is non-negligent. However, a negligent AI producer will share the payment of compensation with the AI user. Under this liability scheme, the AI producer will never over-invest in care at equilibrium due to negligence rule. If the AI producer invests the efficient amount of care, then the AI user will also choose the efficient amount of care. If the AI producer under-invests in care, then the AI user will under-invest in care too.

6.2.2 AI user's strict liability with *pure* defense of contributory negligence on part of AI producer

In this section, we keep the same liability regime as in previous section, and we explore the implications of Assumption 6 below.

Assumption 6 *Under negligence rule, a non-negligent AI producer has no liability, regardless the amount of care chosen by the AI user subject to strict liability. Once a medical accident occurs, the AI user will solely bear the payment of compensation if the AI producer is non-negligent. If the AI producer is proven to be negligent, then the payment of compensation is solely borne by the AI producer. The AI user has no liability, and there is no longer any sharing of payment of compensation between the AI producer and the AI user.*

The AI user's choice of level of care

At step 3, after having purchased q units of AI, the AI user must decide on the level of care to take which is solution to:

$$\min_y \left\{ \begin{array}{c} yq + h(x,y)q \text{ if } x \geq x^{SE} \\ yq \text{ if } x < x^{SE} \end{array} \right\} \tag{36}$$

where x^{SE} is the socially efficient amount of care taken by the AI producer. Look at the first line of Equation (36). If $x \geq x^{SE}$ then the AI user's best response is $y^*(x) = y^{SE}(x)$. Regarding the second line of (36), we find that the AI user's best response becomes $y^*(x) = 0$.

The AI user's choice of AI quantity to demand

We next turn to finding the AI user's choice of quantity to demand at step 2. The AI user's marginal willingness-to-pay for AI comes from the maximization problem for the AI user, where she rationally anticipates using her best choice-of-care. The AI user's program at step 2 is:

$$\max_q \left\{ \begin{array}{c} u(q) - h(x, y^{SE}(x)) - pq - y^{SE}(x)q \text{ if } x \geq x^{SE} \\ u(q) - pq \text{ if } x < x^{SE} \end{array} \right\} \tag{37}$$

where p is the price paid by the AI user for each unit of AI. These two first order conditions yield the inverse demand functions:

$$p(q, x, y^{SE}(x)) = u'(q) - y^{SE}(x) - h(x, y^{SE}(x)) \text{ if } x \geq x^{SE}$$
$$p(q, x) = u'(q) \text{ if } x < x^{SE}$$

The AI producer's choice of care

We turn to the step 1 of the game. Under negligence rule, the AI producer's program at step 1 is:

$$\max_x \left\{ \begin{array}{c} \pi(q, x, y^{SE}(x)) = (u'(q) - y^{SE}(x) - h(x, y^{SE}(x)))q - xq \text{ if } x \geq x^{SE} \\ \pi(q, x) = (u'(q))q - xq - h(x, 0)q \text{ if } x < x^{SE} \end{array} \right\} \tag{38}$$

If $x \geq x^{SE}$ the first order condition equals:

$$-h_x(x, y^{SE}(x)) - y_x^{SE}(1 + h_y) = 1. \tag{39}$$

From analysis of step 3 of the game, we have: $(1 + h_y) = 0$. Thus, using Equation (38), we find that: $x^*(y) = x^{SE}(y)$. We also know that $y^*(x) = y^{SE}(x)$. Consequently, if the AI producer chooses the efficient amount of care, then the AI user will choose the efficient level of care too. Notice also that the AI producer has no incentives to over-invest in care.

If $x < x^{SE}$ then the first-order condition equals:

$$-h_x(x, 0) = 1 \tag{40}$$

As a consequence, the AI producer's strategy is given by $x^*(0)$. The AI producer anticipates that the AI user will not invest in care at all if she under-invests in care. We also know that the AI user will choose the efficient amount of care if the AI producer chooses the efficient amount of care. By comparing equilibrium profits, we find that the AI producer will under-invest in care ($x < x^{SE}$) if and only if $h(x, 0) - h(x^{SE}, y^{SE}) < y^{SE} + (x^{SE} - x)$. The AI producer will under-invest in care if the gain resulting from the reduction of the cost of care (for both AI producer and AI user) is superior to increase of the amount of compensation that result from the under investment in care. We summarize our results in proposition below.

Proposition 6 *Assume that AI producer's liability is based on the principle of fault liability while we apply strict liability to AI user. Under assumption 6, we have:* $x^* \leq x^{SE}$ *and* $y^* \leq y^{SE}$*. More precisely, if* $x^* = x^{SE}$ *then* $y^* = y^{SE}$*. But, if* $x^* < x^{SE}$ *then* $y^* = 0$*.*

Consider that the AI producer is subject to negligence rule. A non-negligent AI producer has no liability, and the AI user will solely bear the payment of compensation due to AI user's strict liability. If the AI producer is proven to be negligent, we assume that the payment of compensation is solely borne by the AI producer. The AI user has no liability here. So there is no longer any sharing of payment of compensation between the AI producer and the AI user. Under this liability scheme, the AI producer will never over-invest in care due to negligence rule. But, the same is true for the AI user. Moreover, the AI producer has two issues at equilibrium. First, the AI producer may invest the efficient amount of care. Then, the AI user responds by choosing the efficient amount of care too. Second, the AI producer can under-invest in care at equilibrium. Consequently, the AI user does not invest in care at all.

7 Discussion

In this section, we discuss the main findings of our paper. The results are summarized in Table 1 below. We also discuss an extension by exploring the possibility that the compensation sharing rule may also be based on the levels of care themselves. To begin with, we consider the simplest liability and compensation regime. The payment of compensation is borne by the single AI user while the AI producer has no liability. Due to the liability cost, the AI user will choose the socially efficient level of care. As a follow-on, even if the AI producer is not liable for any patient's damage, we show that he will also choose the socially efficient level of care at equilibrium. The reason is that the AI user is willing to pay more in return for higher AI producer's investment in care, that will reduce her expected payment of compensation. However, this relationship is completely unbalanced as the entire liability cost is borne by the AI user.

7.1 Strict liability for AI user and producer

When both the AI producer and the AI user are subject to strict liability, a factor may profoundly change the incentives of both parties. We think about the division of payment of compensation between the AI user and the AI producer in case of medical accident. The key point is that as soon as one reduces the share of compensation borne by the AI user (and simultaneously transfers this part of compensation to the AI producer), this reduces the AI user's incentive to invest in care, thus limiting the ability of physicians to provide the highest quality patient care. But, at the same time, we show that this reduction of potential liabilities could

Figure 1: Choice of levels of care by AI user and AI producer under different liability regimes

AI user (AI)		AI producer (AIP): Strict liability	AI producer (AIP): Negligence rule – Defense of contributory negligence	AI producer (AIP): Negligence rule – *Pure* defense of contributory negligence
Strict liability		If levels of care are *complements*, two possible outcomes: **(1)** Overinvestment in care by AIP and unclear reaction of AIU, or **(2)** under investment in care by AIP and AIU. If levels of care are *substitutes*, two possible outcomes: **(1)** overinvestment in care by AIP and underinvestment in care by AIU, or **(2)** under investment in care by AIP and unclear reaction by AIU.	No overinvestment in care by AIP and AIU Two possible outcomes: **(1)** socially efficient investments in care by AIP and AIU, or **(2)** underinvestment in care by AIP and AIU	No overinvestment in care by AIP and AIU Two possible outcomes: **(1)** socially efficient investments in care by AIP and AIU, or **(2)** underinvestment in care by AIP and no care investment by AIU
Negligence rule	Defense of contributory negligence	No overinvestment in care by AIU Unclear investment decisions of AIP and AIU	No overinvestment in care by AIP and AIU Two possible outcomes: **(1)** socially efficient investments in care by AIP and AIU, or **(2)** under investment in care by AIP and AIU	
Negligence rule	*Pure* defense of contributory negligence	**Socially efficient investments in care by AIP and AIU**		

lead physicians to increase their use of AI, which may be a significant progress in medical treatment in diseases. Further upstream, the AI producer will anticipate these changes. As the AI user tends to reduce her investment in care, the marginal product of AI producer's investment in care automatically increases. This effect associated with the increased demand for AI encourages the AI producer to invest more in care. This outcome is all the more plausible that both amounts of care behave like substitutes at equilibrium. It is then expected that the AI user will under-invest in care while the AI producer will over-invest in care. But, due to many direct and indirect effects, opposite or not, we find that another outcome may also occur. There could be under investment in care by the AI producer at equilibrium. Under this scenario, the AI user's choice of care is unclear. The most unfavorable scenario would arise if she also under-invests in healthcare. This liability framework could consequently result in a heightened incidence of medical incidents. Overall, there are few chances that both AI user and AI producer will choose the socially efficient levels of care. This is sufficient here to disqualify this liability framework.

7.2 Negligence rule for AI producer and/or AI user

Let us closely examine the implications of assigning liability to AI producer and/or AI user solely upon establishing negligence on their part. Notable changes are twofold. First, as soon as the AI producer or the AI user is subject to negligence rule, he or she will logically no longer over-invest in care. Second, as the division of payment of compensation becomes less central in our analysis, it appears that the type of care, either substitute or complement, will matter little or not at all.

The most common situation in health law and economics is one AI user's liability based on negligence rule while the AI producer is subject to strict liability. Thus, a non-negligent AI user has no liability anymore regardless the level of care chosen by the AI producer. If a medical accident occurs, the AI producer will solely pay for the compensation. However, if the AI user has proven to be negligent, then the AI producer has no liability and the AI user will have to pay for the entire compensation. We show that this combination is socially efficient. Consequently, our theoretical findings support the recent proposition put forth by the European Parliament regarding the AI Liability Directive in the field of medicine. This directive posits that there is no imperative need for a novel liability framework within this domain. Instead, it advocates the implementation of strict liability regime for AI producers in cases of defective products and a fault liability regime for AI users.

To go deeper in the detail of the analysis of this liability regime, we introduce one *minor* change in case the AI user is proven to be negligent. The AI user will now have to share the payment of compensation with the AI producer instead of solely bearing the payment of compensation. This assumption is plausible given the fact that the AI producer is legally responsible even in the absence of fault or negligence while the negligent AI user is also responsible. Under this assumption, we show that both parties will generally not invest the socially efficient levels of care. This ensures only that the AI user will never over-invest in care. The AI producer's decision is rather unclear. He may either over or under-invest in care. Hence, it is crucial to ensure that, within the framework of the negligence rule, the AI user experiences a sufficiently high level of legal pressure, thereby fostering the creation of proper care incentives.

Presently, we examine the potential outcomes that could be anticipated if both the AI producer and the AI user were subject to negligence rule? First of all, it may be worth noting that the final outcome is (practically) the same, no matter the AI user is subject to negligence rule or strict liability. Second, neither of AI producer or AI user will over-invest in care. Third, there is some chance that the AI producer invests the efficient amount of care, in which case the AI user will do the same. If we exclude this outcome, only one another outcome can emerge at equilibrium: both the AI user and the AI producer under-invest in care. These results hold true whether both levels of care are complements or substitutes.

Finally, consider the scenario where the AI producer is subject to negligence rule while the AI user has strict liability. Next, assume that the negligent AI producer must pay the entire compensation in case of medical accident. Then, the AI user has no liability in this case, but remains liable if the AI producer is non negligent. Consequently, if the AI producer under-invests in care, then the AI user will not invest in care at all. In this specific occurrence, liability hanging over the AI user clearly appears inadequate.

7.3 Sharing policy for payment of compensation, levels of care and negligence rule

In previous analysis, when required by the liability regime, the compensation is shared exogeneously between the AI producer and the AI user. But it could be assumed that these parties pay a share of the compensation that is proportional to their causal contribution to health loss, that is to their levels of care. The important thing is, firstly, that such a modification makes sense only under negligence rule. Under strict liability, we do not seek to evaluate ex post the amounts of care and even less to share the payment of compensation among parties. Accordingly, we consider in this section that both parties are subject to negligence rule. Consequently, a non-negligent party has no liability, regardless the level of care chosen by the other party. The entire payment of compensation is borne by the negligent party when the other one happens to have chosen the socially efficient level of care. When both parties are negligent, then the compensation is shared out between them according to their respective levels of care. The share of compensation borne by the AI user, β, now depends on both levels of care, x and y. We thus replace β in previous analysis by a function $\beta(x, y)$. The game is like the one described in Section 5. The question raised is whether there exists a function $\beta(x, y)$ (and, if so, what that function is) whereby the AI user will choose a level of care that maximizes her expected utility for any AI producer's level of care provided that the AI user now anticipates that her choice of level of care has a greater influence on the sharing policy for payment of compensation.

We focus on step 3 of the game. Using the game depicted in Section 5, the AI user solves this program:

$$\min_{y} q\beta(x, y)h(x, y) + qy \tag{41}$$

if the AI user does not meet the socially efficient amount of care, y^{SE}. The first order condition gives:

$$-\beta(x, y)h_y(x, y) - \beta_y(x, y)h(x, y) = 1. \tag{42}$$

Compared to the previous equilibrium described in the paper (when β is exogeneous), there is an additional marginal benefit: $-\beta_y(x, y)h(x, y)$. Consequently, not only the amount of compensation $h(x, y)$, but also the share of this compensation borne by the AI user $\beta(x, y)$ decreases with each additional level of care. Thus it logically creates more incentives to invest in care at step 3.

We seek to find the function $\beta(x, y)$ that can be used to verify the first order condition above. In Appendix 2, we consider the solution to this non homogeneous first order linear differential equation. We show that one solution is $\beta(x, y) = \frac{y^{SE}-y}{h(x,y)}$. This share of compensation (borne by the AI user) depends on both levels of care, the amount of compensation and the AI user's socially efficient level of care. It is logically increasing with the AI producer's level of care, x. The share of compensation borne by the AI user is decreasing with the AI user's level of care (y) under the sufficient condition that the elasticity of the harm/compensation to the AI user's level of care is superior to 1. The intuition of this result runs as follows. This share of compensation borne by the AI user depends on the amount of compensation, h. This amount of compensation in turn depends directly on the level of care, y. Thus, this amount of compensation must be adequately responsive to her level of care in order to ensure that any increase in AI user's level of care will reduce the share of compensation borne by the AI user.

8 Conclusion

The integration of AI into the realm of medicine is poised to redefine the very essence of medical decision-making, converting it into a collaborative process involving both the AI user and the technological tool. This transformation, however, gives rise to a paradigm shift in the landscape of liability, particularly in cases of medical malpractice. The pivotal questions that emerge relate to the determination of culpability and the appropriate liability framework to be applied. Despite the ongoing evolution of tort liability concerning AI, with very limited direct scrutiny in court cases, the responsibility in instances of AI-assisted health decision-making, where the ultimate decision rests with the AI user, warrants special attention (Price et al., 2023).

A comprehensive comprehension of the implications of these liability rules is imperative to harness the full potential of emerging technologies for the benefit of patients. Our research underscores a crucial finding: the introduction of these novel techniques and the consequent alterations in liability dynamics should not translate into a diminished liability for AI users as it could result in socially inefficient investments in care. It is noteworthy, however, that this stance doesn't advocate for the sole burden of medical error costs to be borne by AI users. Presently, the legal treatment of medical errors follows distinct paths: physicians are subject to a fault based regime, while AI producers are held under a strict liability regime. Our analysis demonstrates that maintaining this status quo aligns with the right incentives for care. The future trajectory of research in this domain should pivot away from determining the applicable liability regime for involved parties. Instead, it should focus on adeptly adjusting these liability frameworks to address the intricacies posed by the integration of AI in the field of medicine. A primary concern lies in the acquisition of information pertaining to AI user faults (such as instances of AI misuse) or AI product defects or documentation errors. This informational gap could heighten uncertainty regarding the legal standard of care required. Consequently, AI users and producers will be able to operate with less awareness of the implications of liability.

Appendices

Appendix 1: Division of compensation under strict liability

We study the effect of the allocation of compensation between the AI producer and the AI user, on the optimal AI user's care decision when both players operate under strict liability rule. At step 3, the first order condition for the problem $\min_y \{\beta h(x,y) + qy\}$ is:

$$-\beta h_y(x,y) - 1 = 0 \tag{43}$$

Equation (43) defines the AI user's optimal care decision given the AI producer's care decision. We have previously calculated the condition characterizing the AI producer's optimal care decision at step 1. Once incorporating the AI user's optimal care decision, we find at equilibrium that:

$$-h_x + \frac{h_{yx}h_y}{h_{yy}}(1-\beta) - 1 = 0 \tag{44}$$

Let the solution to Equation (44) be denoted as x^{SL}. This solution depends on the parameter β. Using Implicit Function Theorem and Equation (43), one can derive the marginal effect of β on the AI user's care decision at equilibrium given that the AI producer's care level also depends on the parameter β. We find that the sign of $\frac{\partial y}{\partial \beta}$ is the sign of expression:

$$-h_y(x^{SL},y) - \beta h_{yx}(x^{SL},y)\frac{\partial x^{SL}}{\partial \beta} \tag{45}$$

The first term of Equation (45) is positive. The second term, $-\beta h_{yx}(x^{SL},y)\frac{\partial x^{SL}}{\partial \beta}$, is negative because $h_{yx}(x^{SL},y)$ and $\frac{\partial x^{SL}}{\partial \beta}$ have opposite sign. Indeed, we have shown that: $x^{SL} < x^*$ if $h_{yx} > 0$. As the socially efficient AI user's level of care corresponds to the case where $\beta = 1$, we have $\frac{\partial x^{SL}}{\partial \beta} > 0$ if $h_{yx} > 0$. At the opposite, one can readily show that $\frac{\partial x^{SL}}{\partial \beta} < 0$ if $h_{yx} < 0$.

As a consequence, there are two opposite effects of the parameter β on the AI user's optimal level of care: (i) a positive direct effect given by $-h_y(x^{SL},y) > 0$, and (ii) a negative indirect effect given by $-\beta h_{yx}(x^{SL},y)\frac{\partial x^{SL}}{\partial \beta} < 0$. The second one corresponds to the effect of the parameter β on AI user's level of care via the AI producer's care decision.

Using Equation (44) and Implicit Function Theorem, we have:

$$\frac{\partial x^{SL}}{\partial \beta} = \frac{h_{yx}h_y h_{yy}}{-h_{xx}h_{yy}^2 + (1-\beta)(h_{yxx}h_y h_{yy} + h_{yx}^2 h_{yy} - h_{yx}h_y h_{yyx})}. \tag{46}$$

Thus, $\frac{\partial x^{SL}}{\partial \beta} > 0$ if and only if

$$h_{xx}h_{yy}^2 h_y - h_{yx}^2 h_{yy} h_y < h_y^2(1-\beta)(h_{yxx}h_{yy} - h_{yx}h_{yyx}). \tag{47}$$

First, study the LHS of Equation (47). By convexity of h, we have $h_{xx}h_{yy} - h_{yx}^2 > 0$. Multiplying by $h_{yy}h_y < 0$, we find that the LHS of Equation (47) is negative. Next, study the right-hand side of Equation (47). As $h_y^2(1-\beta) > 0$, the sign of the RHS of Equation (47) is equal to the sign of expression $h_{yxx}h_{yy} - h_{yx}h_{yyx}$. By assumption, we have $h_{yxx} > 0$ and $h_{yyx} > 0$. As a consequence, the RHS of Equation (47) is positive if $h_{yx} < 0$. Thus, $\frac{\partial x^{SL}}{\partial \beta} > 0$ if $h_{yx} < 0$. As the socially efficient AI user's level of care y^* corresponds to the case $\beta = 1$, we find that $y^{SL} < y^*$ if $h_{yx} < 0$.

When $h_{yx} < 0$, the positivity of the RHS of Equation (47) requires a restriction on the magnitude of h_{yx}, that is $0 < h_{yx} < \frac{h_{yxx}h_{yy}}{h_{yyx}}$. Under this condition, we anew find that $y^{SL} < y^*$. But, when $h_{yx} > \frac{h_{yxx}h_{yy}}{h_{yyx}}$, the RHS of Equation (47) is negative. Thus, the AI user's optimal level of care may be either increasing or decreasing with the parameter β. As the socially efficient AI user's level of care y^* corresponds to the case $\beta = 1$, we find either over-supply or under-supply of care by the AI user when $h_{yx} > \frac{h_{yxx}h_{yy}}{h_{yyx}}$.

Appendix 2: Sharing of compensation and negligence rule

Consider the non homogeneous linear first order condition denoted as (E):

$$\beta(y)h_y(y) + \beta'(y)h(y) = -1. \tag{48}$$

For convenience, we omit the variable x in the formulas. We can find the solution $\beta(y)$ as follows. We use the method of variation of constant.

First, consider the associated homogeneous equation denoted as (E_H). It is:

$$\beta(y)h_y(y) + \beta'(y)h(y) = 0 \tag{49}$$

Next, we denote β_H the solutions to (E_H) and β_P a particular solution to (E). The solutions of (E) are $\beta(y) = \beta_H(y) + \beta_P(y)$.

We start with the resolution of the associated homogeneous equation (E_H). The solutions to (E_H) are $\beta_H(y) = ke^{-F(y)}$ for $k \in R$ where $F(y) = \int \frac{h_y(y)}{h(y)}dy = \ln h(y)$. Thus, $\beta_H(y) = ke^{-\ln h(y)} = \frac{k}{h(y)}$.

Next, let us consider the particular solution to (E). It is $\beta_P(y) = K(y)e^{-F(y)}$ where $F(y) = \ln h(y)$. We have to find the function $K(y)$ such that $\beta_P(y) = \frac{K(y)}{h(y)}$. To begin, we calculate $\beta_P'(y) = \frac{K'(y)h(y) - K(y)h'(y)}{h^2(y)}$. As $\beta_P(y)$ is solution to (E), we have $\beta_P(y)h'(y) + \beta_P'(y)h(y) = -1$.

Now substitute $\beta'_P(y)$ and $\beta_P(y)$ in this equation, we have after some manipulations, $K'(y) = -1$ or $K(y) = -y$. Finally, substitute $K(y) = -y$ in $\beta_P(y) = \frac{K(y)}{h(y)}$, we find that $\beta_P(y) = \frac{-y}{h(y)}$.

Put together, the solutions of (E) are $\beta(y) = \beta_H(y) + \beta_P(y) = \frac{k-y}{h(y)}$. Finally, one have to determine k. We use one initial condition, that is a point that the solution must satisfy. Here, this condition is $\beta(y^{SE}) = 0$. The reason is that the amount of compensation borne by the AI user equals zero if he chooses the socially efficient amount of care as negligence rule prevails. As $\beta(y) = \frac{k-y}{h(y)}$, we have $\beta(y^{SE}) = \frac{k-y^{SE}}{h(y^{SE})} = 0$ or $k = y^{SE}$. Thus the general solution to (E) is $\beta(y) = \frac{y^{SE}-y}{h(y)}$.

References

[1] Agrawal, A., Gans, J. S., and Goldfarb, A., 2019. Exploring the impact of artificial intelligence: Prediction versus judgment. *Information Economics and Policy.* 47, 1–6

[2] Brosig-Koch, J., Hennig-Schmidt, H., Kairies-Schwarz, N., and Wiesen, D., 2016. Using artefactual field and lab experiments to investigate how fee-for-service and capitation affect medical service provision. *Journal of Economic Behavior & Organization* 131, 17-23.

[3] Brosig-Koch, J., Hennig-Schmidt, H., Kairies-Schwarz, N., and Wiesen, D., 2017. The effects of introducing mixed payment systems for physicians: Experimental evidence. *Health Economics*, 26, 243–262.

[4] Chalkley M., and Malcomson J. M., 1998. Contracting for health services when patient demand does not reflect quality. *Journal of Health Economics*, 17(1), 1–19.

[5] Chopard, B., and Musy, O., 2023. Market for Artificial Intelligence in Health Care and Compensation for Medical Errors. *The International Review of Law and Economics.*

[6] Danzon, P.M., 1985a. Liability and Liability Insurance for Medical Malpractice. *The Journal of Health Economics.* 4, 309–331

[7] Danzon, P.M., 1985b. Medical Malpractice: Theory, Evidence, and Public Policy, Cambridge (Mass.), Harvard University Press.

[8] Danzon, P.M., 1997. Tort Liability: A Minefield for Managed Care. *The Journal of Legal Studies.* 26(2), 491–519.

[9] Daughety, A., and Reinganum, J., 2013. Economic Analysis of Products Liability: Theory. In *Research Handbook on the Economics of Torts.* Edward Elgar Publishing.

[10] Daughety, A., and Reinganum, J., 2014. Cumulative Harm and Resilient Liability Rules for Product Markets. *Journal of Law, Economics, and Organization* 30(2), 83–108.

[11] Dawid, H., and Muehlheusser, G., 2022. Smart Products: Liability, Investments in Product Safety, and the Timing of Market Introduction. *Journal of Economic, Dynamics and Control.* 134, 104288.

[12] Godager G., and Wiesen D., 2013. Profit or patients' health benefit? Exploring the heterogeneity in physician altruism. *Journal of Health Economics*, 32(6), 1105–16.

[13] Griffin F., 2021. Artificial Intelligence and Liability in Health Care, *Health Matrix*, 31, pp. 65-106.

[14] Guerra A., Parisi, F., Pi, D., 2021. Liability for Robots: an Economic Analysis. Available at SSRN 39399486.

[15] Hennig-Schmidt, H., and Wiesen, D., 2014. Other-regarding behavior and motivation in health care provision: An experiment with medical and non-medical students, *Social Science & Medicine*, 108, 156–165.

[16] Higgins B.W. 2022, Legal Elements of an AI Regulatory Permit Program, in *The Oxford Handbook of AI Governance*, edited by J. Bullock, Y-C. Chen, J. Himmelreich, V.M. Hudson, A. Korinek, M. Young, and B. Zhang.

[17] Jack W., 2005. Purchasing health care services from providers with unknown altruism. *Journal of Health Economics*, 24(1), 73–93.

[18] Landes, W.M., and Posner, R.A., 1980. Joint and Multiple Tortfeasors: An Economic Analysis. *The Journal of Legal Studies.* 9(3), 517–555.

[19] Makris, M., and Siciliani, L., 2013. Optimal Incentive Schemes for Altruistic Providers. *Journal of Public Economic Theory.* 5(5), 675–699.

[20] Obidzinski, M. and Oytana, Y. 2022. Prediction, human decision and liability rules. CRED Working paper No 2022-06.

[21] Parisi, F., and Fon, V., 2004. Comparative Causation, *American Law and Economics Review.* 6(2).

[22] Price, W. N., 2018, Medical Malpractice and Black-Box Medicine, in *Big Data, Health Law, and Bioethics*, ed. by I. Cohen, H. Lynch, E. Vayena, and U. Gasser, pp. 295-306, Cambridge University Press.

[23] Price, W.N., Gerke S. and Cohen I.G., 2019, Potential Liability for Physicians Using Artificial Intelligence. *Journal of the American Medical Association*,322(18), pp. 1765–1766.

[24] Price, W. N., Gerke, S., and Cohen, I. G. , 2023. Liability for Use of Artificial Intelligence in Medicine. in *Medicine, Research Handbook on Health, AI and the Law*, ed. by B. Solaiman and I. GG. Cohen, Edward Elgar Publishing Ltd.

[25] Shavell, S., 1980. Strict Liability versus Negligence. *The Journal of Legal Studies.* 9(1), 1–25.

[26] Shavell, S., 2007. Liability for Accidents. in *Handbook of Law and Economics, vol. 1.*, ed. by Polinsky, A. M., and Shavell, S., Elsevier. Chapter 2, 139–182.

[27] Shavell, S., 2020. On the Redesign of Accident Liability for the World of Autonomous Vehicles. *The Journal of Legal Studies.* 49(2), 243–285.

DOCUMENTS DE TRAVAIL *AFED* PARUS EN 2024
AFED Working Papers Released in 2024

24-01 Claudio Ceccarelli & Antonio Cappiello
OECD PMR Indicators on Professional Services: Top Performances or Outliers?

24-02 Benjamin Monnery & Alexandre Chirat
Trust in the Fight against Political Corruption: A Survey Experiment among Citizens and Experts

24-03 Stefano Dughera & Alain Marciano
Buchanan and the Social Contract: Coordination Failures and the Atrophy of Property Rights

24-04 Bertrand Chopard & Olivier Musy
Optimal Liability Rules for Combined Human-AI Health Care Decisions