NBER WORKING PAPER SERIES

ON THE REDESIGN OF ACCIDENT LIABILITY FOR THE WORLD OF AUTONOMOUS VEHICLES

Steven Shavell

Working Paper 26220
http://www.nber.org/papers/w26220

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
September 2019

I thank Kenneth Abraham, Mark Geistfeld, Louis Kaplow, A. Mitchell Polinsky, Robert Rabin, and Abraham Wickelgren for comments and Henry Lu and Shenghao Wang for able research assistance. The views expressed herein are those of the author and do not necessarily reflect the views of the National Bureau of Economic Research.

On the Redesign of Accident Liability for the World of Autonomous Vehicles
Steven Shavell
NBER Working Paper No. 26220
September 2019
JEL No. K13,K2,K32

ABSTRACT

This article studies a model of liability for automobile accidents in the coming world in which automobiles will be autonomous. In that world, travelers will not be drivers, rendering liability premised on driver fault irrelevant as a means of reducing accident dangers. Moreover, no other conventional principle of individual or of manufacturer liability would serve well to do so. Indeed, in the model considered, strict manufacturer liability, recommended by many commentators, would actually tend to leave accident risks unchanged from their levels in the absence of liability. However, a new form of strict liability—the hallmark of which is that damages would be paid to the state—would be superior to conventional rules of liability in alleviating accident risks and would be easy to administer.

Steven Shavell
Harvard Law School
1575 Massachusetts Avenue
Hauser Hall 508
Cambridge, MA 02138
and NBER
shavell@law.harvard.edu

On the Redesign of Accident Liability for the World of Autonomous Vehicles

Steven Shavell*

This article studies a model of liability for automobile accidents in the coming world in which automobiles will be autonomous. In that world, travelers will not be drivers, rendering liability premised on driver fault irrelevant as a means of reducing accident dangers. Moreover, no other conventional principle of individual or of manufacturer liability would serve well to do so. Indeed, in the model considered, strict manufacturer liability, recommended by many commentators, would actually tend to leave accident risks unchanged from their levels in the absence of liability. However, a new form of strict liability—the hallmark of which is that damages would be paid to the state—would be superior to conventional rules of liability in alleviating accident risks and would be easy to administer.

1. Introduction

The objective of this article is to identify a liability regime that will be well suited to limit accident risks in the future world in which motor vehicles on the roads will generally be fully autonomous.[1] In that world travelers in automobiles will not be drivers who control their vehicles—they will be passengers.

Although accident risks will be low in comparison to those today, accidents will still occur. The reasons include mechanical failure of vehicles, unsafe road conditions, pedestrian behavior, and software design error.[2] Given that the volume of traffic for the transport of

---

* Samuel R. Rosenthal Professor of Law and Economics, Harvard Law School. I thank Kenneth Abraham, Mark Geistfeld, Louis Kaplow, A. Mitchell Polinsky, Robert Rabin, and Abraham Wickelgren for comments and Henry Lu and Shenghao Wang for able research assistance.

[1] By fully autonomous vehicles, I refer to vehicles not requiring human intervention for their operation. Estimates of the time at which most vehicles on the roads will be fully autonomous vary, ranging from the late 2030s to the 2050s or 2060s. See, for example, Bansal and Kockelman (2017), Bierstedt, et al. (2014), and Litman (2018). The path to that juncture in the use of autonomous vehicles is frequently described by five stages in their development; see Society of Automotive Engineers (2014). In the analysis below, I make the assumption that all vehicles are fully autonomous, but I comment on the accident problem when this is not so in the final section of the article.

[2] Road accidents presently account for approximately 1.25 million deaths a year globally as well as substantial non–fatal injuries and economic harm; see WHO (2015), pp. x, xi, 1–16. A large fraction of these accidents are attributed to human error; see, for example, Singh (2018), reporting that 94% of accidents are caused by faulty driver behavior. Because driver mistake would be eliminated in an era of autonomous vehicles, it is often predicted that the accident toll would fall dramatically. For two reasons, however, a number of commentators have expressed some skepticism about the level to which accidents will decline. First, some accidents that are ascribed to driver fault were in fact caused by dangerous road conditions, weather events, and pedestrian mistake and therefore might still occur when vehicles are autonomous. Second, the advent of autonomous vehicles will be accompanied not only by the risk of software design problems, but also by other significant new hazards associated with their

individuals and goods will be great, society will have a continuing interest in constraining whatever are the residual vehicle accident risks through the use of the liability system.[3]

How will the present form of liability for vehicle accidents function in a world of autonomous vehicles? The major rule of tort liability that we apply today to reduce road accidents concerns, of course, fault of the driver. But this rule will be irrelevant when there are no drivers in active command of their vehicles. If an autonomous vehicle speeds or strays from its lane, the explanation for that event will not involve a driver.

That driver fault will be made moot in the contemplated world might suggest that society should turn its attention toward liability of manufacturers of vehicles. Liability of manufacturers could be strict—requiring them to pay damages for any accident caused by their vehicles—or grounded on a notion of fault, notably about manufacturing defect or deficient design. Liability could also be imposed on owners of vehicles, perhaps on a strict basis.

What will be explained in this article is that none of the conventional rules of liability would perform well to regulate accident risks. But a new alternative liability rule, *strict liability with damages paid to the state*,[4] would work efficiently to curb accident risks and is easy to administer.

The logic of the argument to be developed is straightforward and is made in Sections 2 and 3 below employing a basic stylized model of autonomous vehicle accidents and liability.[5] The social objective in the model is assumed to be the usual utilitarian goal associated with economic analysis of accidents, namely the maximization of the benefits parties obtain from their activities, here traveling in their vehicles, less the costs of achieving safety together with those of accidents themselves. Two variables affect this goal: vehicle safety and miles traveled. The optimal level of safety is that which best trades off the cost of enhancing safety against the benefits of accident risk reduction; and the optimal level of mileage is that which best trades off the cost of accidents generated by going extra miles against the benefits from so doing. In determining these optimal levels, it is crucial to recognize that the social cost of an accident between two vehicles is the *sum* of the loss to each. If persons A and B each suffer a $25,000 loss in an accident, then the social cost of the accident is $50,000.

Now let me sketch the problems with the conventional rules of liability and the advantage of the new rule introduced here. I will begin with the simplest conventional rule, strict liability of individuals for the harm that they cause in accidents. Consider individuals A and B just

---

technologies, such as interruption of vehicle communications and cyber intrusion. See, for example, International Transport Forum (2018), pp. 5-7, 12-14, and Noy, et al. (2018).

[3] The liability system instigates risk–reduction by threatening those who cause harm with having to pay damages. On this role of liability, see generally Calabresi (1970), Landes and Posner (1987), and Shavell (1987). Society will also wish to employ regulation to govern risk, and that will be true of autonomous vehicle dangers, but the focus in this article is on the instrument of liability.

[4] In identifying this scheme as a form of liability, I am committing a minor abuse of language because "liability" is usually interpreted as a requirement to make a payment to a party who has suffered harm rather than to the state.

[5] In Section 2 I analyze the model informally and in Section 3 I do so formally.

mentioned. Under strict liability, A would compensate B for her $25,000 loss, and B would likewise compensate A for his $25,000 loss. Hence, after making and receiving damage payments, A would bear $25,000 in expenses and so would B. Each therefore would bear less than, and in fact only half of, the $50,000 social cost of an accident. Accordingly, their incentives to demand safe vehicles from manufacturers will be socially inadequate and the mileage levels that they will choose to travel will be socially excessive. In particular, it is clear that their subpar behavior is exactly the behavior that they would display in the absence of any liability—strict liability will make no difference to outcomes.[6]

In contrast, consider strict liability to the state, in which each individual pays damages for harm done, but in which the damage payments are made to the state rather than to the victims of harm. Under this rule, person A would pay $25,000 to the state and also bear his own $25,000 loss—since he would not be compensated by B. Consequently, A would incur $50,000 of expenses from his involvement in an accident, giving him proper incentives to demand safety in his vehicle purchases and to moderate his mileage. The same would be true in regard to B. Therefore, strict liability to the state instills perfect incentives to reduce accident risks and dominates conventional strict liability.

Next, consider strict liability of manufacturers of vehicles rather than of individuals who own them. Under that rule, in an accident involving persons A and B, the manufacturer of A's vehicle would be liable for the $25,000 of harm caused for B and would pay B that amount in damages; and the manufacturer of B's vehicle would be liable for the $25,000 of harm caused for A.[7] In this situation, manufacturers would have to impound their liability payments into their charges to buyers of vehicles in order to cover their expenses. Under natural assumptions, a manufacturer will be led to charge buyers not only for their vehicles but also a mileage fee equal to the manufacturer's expected liability payments per mile. The mileage fee paid by individuals like A and B would therefore reflect their manufacturers' damage amount of $25,000 per accident. Neither A nor B would bear their own accident losses, however, because they will be compensated for their losses by manufacturers. Therefore, A and B would regard accidents as costing them only $25,000 each, embedded in the mileage fee that each pays. It follows that the situation of individuals would be essentially identical to what it was under strict liability of

[6] In the absence of liability, A would bear his own losses of $25,000 in accidents, and B would bear her own losses of $25,000. Because this is the dollar amount that each would bear under strict liability, their behavior would be the same as under strict liability. As will be discussed later, the reason for the conclusion that strict liability makes no difference to outcomes is that in the context of automobile accidents, strict liability means not only that parties must pay for harms that they cause for others, but *also* that they will be compensated for harms that they themselves experience. This conclusion of no difference depends on the assumption that the loss of each person in an accident is of the same magnitude. But even if that is not the case, strict liability still makes little difference to incentives relative to a regime of no liability, as is noted in the comment following Proposition 3.

[7] As will be clear, I define strict liability of the manufacturer to be true strict liability and, in particular, not to be premised on a finding that the manufacturer's product was defective. Thus, manufacturer strict liability here is different from manufacturer liability under our present law; see Restatement (Third) of Torts: Products Liability § 1 (1998) ("One engaged in the business of selling or otherwise distributing products who sells or distributes a defective product is subject to liability for harm to persons or property caused by the defect."); and see also Restatement (Second) of Torts § 402A (1965). I also define damages under manufacturer strict liability to reflect only harm to third parties, whereas under our present law, manufacturer strict liability may extend to harm to buyers. I treat manufacturer strict liability for harm to buyers as a separate rule, under the heading of product liability.

individuals, which is to say, individuals and manufacturers would display the same deficient incentives to curb accident risks that would be observed in the absence of any accident liability at all.[8]

But if manufacturers are strictly liable to the state, then incentives would be ideal. The reason is that in this case, an individual would not only pay a mileage fee based on the manufacturer's damage payment of $25,000 per accident, the individual would also bear his or her own $25,000 loss in an accident. Thus, individuals would effectively consider the cost of an accident to be $50,000 and would demand optimal safety in their vehicles and choose socially correct mileage.

With regard to fault regimes, suppose that manufacturers are liable for accident harm suffered by other vehicles if their vehicles are not properly safe. Then it will be shown that manufacturers will be induced to obey the required standards of safety by the threat of liability for non–compliance and thus will not be held liable for harm in accidents. As a consequence, individuals would bear their own $25,000 of losses when accidents occur. Because this expense is less than $50,000, individuals would choose an excessive level of mileage. Hence, although the fault regime under consideration would lead to desirable safety levels, it would still be inferior to strict liability to the state. Under product liability, by which I mean a regime under which manufacturers are liable to buyers of their vehicles if their vehicles are not properly safe, it will be shown that manufacturers will not in general be induced to obey the required safety standards and also that mileage levels will be excessive.

After I develop these arguments in Sections 2 and 3, I turn in Section 4 to a summary and interpretation of the conclusions of the analysis of the model and a discussion of a number of issues that were not addressed in it. Among these are the ability to implement the rule of strict liability to the state; latent advantages of that rule in stimulating risk–reduction by manufacturers (such as their research and development efforts); the view that the rule is problematic because it would not furnish compensation to accident victims; the role of liability insurance under the rule; whether the rule is applicable more broadly than to autonomous vehicle accidents; and why, if the rule is as attractive as is claimed, it has not already been recognized as such.

Before proceeding, let me comment on economic writing related to this article. An early effort to analyze automobile accidents and the law from an economic perspective is Vickrey (1968), who emphasized that for incentives for risk reduction and mileage to be appropriate,[9] each driver in an accident should bear an expense equal to the entirety of the harms sustained by the drivers involved in the accident. Vickrey, however, did not examine legal rules in any detail and did not present a formal model of automobile accidents and liability. The first such model to my knowledge is developed by Diamond (1974), who studied single activity accidents, exemplified by accidents caused by the activity of driving.[10] He focused on levels of driver care

---

[8] See the last paragraph in Section 2.3(d) below for further explanation of this perhaps counterintuitive conclusion.

[9] See pp. 466–467.

[10] Brown (1973) earlier introduced a formal model of two activity accidents, involving injurers who are engaged in one activity and who cause harm, and victims who are engaged in another activity and who suffer harm

and the fault rule; he did not examine levels of activity (mileage in the case of driving) or strict liability.[11] The model of the present article obviously differs from Diamond's: here, as the reader knows, driver care and driver fault are irrelevant because vehicles are autonomous; instead, vehicle safety, determined by manufacturers, and mileage are relevant; moreover, strict liability rules and liability based on vehicle fault are considered. Of additional note is Friedman and Talley (2019), who investigate a model of accidents involving autonomous and other vehicles as injurers and parties such as pedestrians and bicyclists as victims.[12] Of note as well is Edlin and Karaca-Mandic (2006), who provided empirical estimates of the magnitude of the driving accident externality (reflecting the difference between the total harm in an accident and that suffered by a particular driver).

A burgeoning literature in law reviews and articles in the popular press specifically devoted to autonomous vehicles and tort liability is also of interest. This writing includes Abraham and Rabin (2019), Ben–Shahar (2016), Crane et al. (2017), Evas (2018), Geistfeld (2017), Logue (2019), and Vladek (2014). It mainly considers variations of product liability and strict liability of manufacturers (so-called enterprise liability), sometimes paying attention to incentives to reduce accident risks, but when so in ways that are incomplete. For example, most of the articles just cited fail to consider the effect of liability on mileage, and none recognizes the basic problem that the effect of liability on safety will be subpar unless the amount paid by each party involved in an accident reflects the total harm sustained by the concerned vehicles. Thus, none notices that strict manufacturer liability will have no tendency to improve accident risk reduction over a setting in the absence of any liability. Although I am unaware of discussion of the regime of strict liability to the state in the legal literature on autonomous vehicle accidents, that is not unexpected given the departure of this regime from the norm. Nevertheless, the appeal of the regime should not come as a complete surprise to readers familiar with writing on economic analysis of tort law, for the general virtue of having injurers pay damages to the state has occasionally been mentioned there.[13]

(for example, injurers could be drivers and victims could be pedestrians). In two activity accidents, injurers and victims, being distinct, can be subject to different rules of liability (for instance, injurers can face strict liability and victims can be held to a duty not to be contributorily negligent). In single activity accidents, it is natural to assume that all parties are subject to the same liability rule and that both may cause and also suffer harms in accidents. Most accident situations are described by the two activity model.

[11] See also Arlen (2009), who added to Diamond (1974), mainly by considering risk–aversion.

[12] In their two activity model, Friedman and Talley place stress on the information that victims possess about whether a vehicle that might injure them is or is not autonomous. The issues of concern to them are thus different from those involved where accidents are between vehicles and would typically result in harm to each other.

[13] See, for example, Green (1974), who remarks in the introduction to his analysis of liability rules that such a system would result in optimal care–taking, and Shavell (1987, pp. 29–30, 46), where I observe that any liability system with damage payments made by injurers to victims cannot optimally control their levels of activity (such as miles driven by motorists and miles walked by pedestrians), whereas a system in which injurers make damage payments to the state can accomplish that.

2. A Model of Autonomous Vehicle Accidents and Liability: Informal Analysis

2.1 Basic assumptions

I consider here a model in which a population of risk–neutral[14] individuals travel[15] in autonomous vehicles and all accidents involve exactly two vehicles. The questions to be addressed concern the effects and the social desirability of the use of liability rules addressing the harms caused in accidents. The major assumptions of the model are as follows.

To begin with, individuals derive benefits from travel that increase with miles driven. However, travel leads to the possibility of accidents. In particular, when vehicles travel, they will have *encounters* with other vehicles—events in which two vehicles come into close enough proximity that an accident could occur. The number of encounters that a vehicle has rises with its mileage: the frequency of encounters between any two specified vehicles, say that of person A and that of person B, will be proportional to the mileages of each.[16] Suppose that the number of encounters between A and B[17] will be 50 if A drives 10,000 miles a year and B drives 15,000 miles.[18] Then if A increases his mileage by 10% to 11,000, the number of encounters with B will rise by 10% to 55; likewise, if A doubles his mileage, the number of encounters with B will rise to 100; if B increases her mileage from 15,000 to 20,000, the number of encounters with A will rise by one–third to 66.67; and so forth. The total number of encounters between A and all other individuals will thus reflect A's mileage and the mileages of each of the others,[19] and similarly for every other person in the population.

---

[14] That is, individuals treat an uncertain prospect as equivalent to its probability–discounted amount. For example, an individual would consider a situation in which he or she would suffer a loss of $10,000 with a 10% probability as equivalent to a certain loss of 10% × $10,000 = $1,000. An interpretation of $1,000 is the amount that would be lost on average were the individual to repeatedly face the 10% risk of loss of $10,000. The assumption of risk–neutrality is made for simplicity when, as here, the object is to capture both the likelihood of events and their significance amounts in an analytically convenient manner. As was noted, risk aversion will be mentioned in Section 4.

[15] I will use "travel" and "drive" and their cognates interchangeably, with the understanding that individuals are not in fact driving vehicles in the usual sense of that word because the vehicles are assumed to be fully autonomous.

[16] In other words, the formula for the number of encounters between A and B is of the form ky_Ay_B, where k is a positive constant, y_A is the mileage of A, and y_B is the mileage of B. The constant k might depend on who A and B are because, for example, they live close to each other (raising k). The number of encounters should be interpreted as per unit of time, I will say annually for concreteness, and similarly with the expected number of accidents and other variables that are discussed below.

[17] I will use "A" to mean either the person traveling in a vehicle or the vehicle itself.

[18] This would be true if the constant k in note 16 is .000000333, for .000000333 × 10,000 × 15,000 = 50.

[19] If, for example, there are four individuals in the population, A, B, C, and D, then the number of encounters of A with others is the sum of those with B, those with C, and those with D. To calculate each of these numbers, we clearly need to know the mileages of each and the appropriate constant k to apply for each possible pair.

The *probability* of an accident in an encounter will reflect the *safety levels* of the two involved vehicles. The safety level of a vehicle is determined by such characteristics as its sensing devices, guidance software, and crashworthiness and will be identified with its cost of manufacture.[20] The higher the safety level of a vehicle, the lower will be the accident probability in any encounter. Suppose that if A's vehicle costs $30,000 and B's vehicle costs $35,000, the accident probability would be .3%. Then if A increases his safety level to $31,000,[21] the probability of an accident in an encounter would be lower—let us say it falls to .25%; and similarly if B increases his safety level.

The *number of accidents* that two named individuals will have is their frequency of encounters multiplied by their accident probability per encounter. If A and B will have 50 encounters and the probability of an accident per encounter is .3%, then they will have 50 × .003 = .15 accidents. Hence, the mileages of A and B and their safety levels allow us to calculate the number of accidents; for their mileages determine the number of encounters and their safety levels give us the accident probability per encounter. In this manner, we can in principle calculate the number of accidents between any two individuals and hence the total number across the population.

If an accident occurs, each individual will suffer the same level of *harm*, which for illustration will be taken to be $25,000.[22] Thus, the total harm sustained by the two vehicles involved in an accident will be $50,000. Accordingly, if A and B would be involved in .15 accidents, the harm caused by those accidents would be .15 × $50,000 = $7,500.[23]

2.2 Social welfare and socially optimal safety and mileage

In order to answer normative questions about outcomes under liability regimes, we must employ a measure of social welfare, and I will use a standard utilitarian objective: the sum of the benefits persons obtain from miles driven in vehicles, less the costs of safety—impounded in the costs of manufacture of the vehicles—and less the expected harms caused by accidents.[24]

---

[20] In reality, the cost of a vehicle would of course depend not only on its safety characteristics but also on its comfort and accessories. For my purposes, consideration of this point would be a distraction.

[21] Increasing the safety level can be interpreted as A obtaining a different vehicle or as adding a safety feature to the $30,000 vehicle.

[22] The harm in an accident may be interpreted as harm to a vehicle or its passengers. In reality, harm may vary due to many factors, such as the speed of the vehicles, their points of impact, and how many passengers they carry. That harm may vary should not, however, alter the main qualitative conclusions that will be drawn about liability rules.

[23] The accident probabilities and expected losses per pair of vehicles per year would undoubtedly be much lower in fact than in this illustrative example.

[24] Under this definition of the social objective, the goal of compensation of accident victims is absent. That goal is addressed in the remarks made in Section 4 on risk aversion and insurance, as I have mentioned.

Accordingly, two variables govern social welfare: the safety levels of vehicles and miles traveled.[25] Social welfare will be ideal if these two variables are chosen in the best possible manner for each individual in the population.

Let me first describe the socially optimal level of safety for a person, taking as given the person's mileage choice and the safety and mileage choices of others.[26] A possible expenditure on safety by a person will be socially worthwhile if its amount is less than the social benefit it yields, which is the reduction in expected accident harms it engenders. In the case of A and B, we had said that if A's vehicle costs \$30,000 and B's costs \$35,000, the accident probability per encounter would be .3%, and that if A increased safety spending to \$31,000, the accident probability per encounter would fall to .25%. Would this extra \$1,000 expenditure on safety by A be socially worthwhile? Given the mileages of A and B, they will have 50 encounters, and since the probability of an accident per encounter would fall by .05%, the expected number of accidents would fall by $.0005 \times 50 = .025$. Because the harm per accident is \$50,000, the reduction in expected harm would be $.025 \times \$50{,}000 = \$1{,}250$, and since this reduction exceeds the \$1,000 cost of greater safety, the expenditure would be socially worthwhile. The optimal level of safety is found by increasing safety expenditures up to the point at which a further increase of a dollar would no longer yield a reduction in expected harm exceeding a dollar.

Note that the optimal level of safety implicitly depends on a person's mileage, for the accident reduction benefit from having a safer vehicle depends on how many encounters a person expects to have, which is a direct function of his mileage.

Note also that in the preceding calculation of the social value of greater safety, I presumed that B was the only other person in the population. More generally, of course, with a large population, the determination of the social value of increased safety on the part of A would take into account the lower likelihood of accidents in encounters with not only B but with all other individuals in the population. For purposes of illustration, however, I will generally continue to assume below that A and B are the only individuals in the population.

Let me now make a calculation regarding a person's mileage, taking as given the person's safety level and the safety levels and mileage choices of others. An increase in miles driven by a person will be socially worthwhile if its benefit exceeds its social cost, which is the additional accident harm it engenders. Suppose that if A increases his mileage from 10,000 miles to 11,000 miles, his benefits would rise by \$1,500. We know that if A raises his mileage by 10%, the number of encounters will increase by 10% and thus so too will the number of accidents, which we had calculated to be .15. Thus the number of accidents will rise by .015, meaning that the expected harm will increase by $.015 \times \$50{,}000 = \$750$. Because this increase is less than the value to A of \$1,500 of driving the extra 1,000 miles, the mileage increase would be socially desirable. The optimal number of miles for A to travel is ascertained by elevating mileage until

---

[25] It is clear that the ingredients of social welfare depend on these two variables: the benefits of travel are a function of miles traveled; the costs of safety depend on the safety levels; and the harms due to accidents are determined by the safety levels and miles traveled (because miles traveled determine encounters, and safety levels determine the probability of an accident in an encounter).

[26] A necessary condition for maximizing social welfare is that each individual must have chosen his or her level of safety in order to maximize social welfare holding all else constant.

the point is reached at which the benefits of another mile to A no longer exceed the increase in accident losses per mile.

2.3 Behavior under liability rules

Having discussed socially ideal behavior, let us now analyze behavior under various liability rules. To this end, observe that because individuals are risk–neutral, they will seek to maximize their expected utility, which is the benefits they obtain from driving less the price they pay for their vehicles, less expected harms that they sustain, less expected liability payments that they must make, but plus expected liability payments that they receive.

Individuals are assumed to purchase their vehicles from manufacturers in a competitive industry.[27] Hence, the price of a vehicle will be equal its production cost plus liability expenses per vehicle borne by a manufacturer. (The manufacturer must collect a total amount that covers not only its production costs but also its liability expenses in order to break even.) Additionally, individuals are assumed to understand the safety levels of vehicles sold by manufacturers.[28]

Given these assumptions, we can now ascertain how individuals will behave, and whether outcomes will be socially desirable, in the absence of liability and under various liability rules.

(a) *No liability.* Let us first examine a regime of no liability for vehicle accidents. In this regime, a person's safety choice will be socially inadequate. Consider person A and the calculation we made showing that it would be socially desirable for A to increase his safety level from $30,000 to $31,000. We reasoned in Section 2.2 that this $1,000 expenditure would reduce the number of accidents by .025, and that meant that the reduction in accident harm would be .025 × $50,000 or $1,250 because total harm in each accident is $50,000. However, in the absence of liability, the loss to A in any accident would be only his own loss of $25,000. Therefore, the expected reduction in his personal losses due to a $1,000 increase in the safety level would be only .025 × $25,000 or $625. Accordingly, A would not find it worthwhile to spend $1,000 more on safety when deciding on the kind of vehicle to purchase. This illustrates the general point that, in the absence of liability, the level of safety will be systematically low from a social standpoint.[29] Individuals will make decisions to purchase vehicles with a socially

---

[27] I do not consider rental of vehicles in the model.

[28] This assumption is apt if vehicles of any given type are sold in large numbers and are evaluated by independent entities, and perhaps by government agencies.

[29] The reader might wonder why this point is described as general, for in a modification of the example, A might have a sufficient incentive to increase the safety level to $31,000. (Suppose, for example, that the reduction in the number of accidents was .06. Then A would save .06 × $25,000 = $1,500, giving him a sufficient motive to spend the extra $1,000 on safety.) However, if A is able to change safety expenditures in a continuous manner (not just in a discrete increment like $1,000), then it is true that he would always choose too low a level of safety. In essence, the reason is that he will increase safety up to the point that an extra dollar spent on safety just equals one dollar in savings of his own expected accident losses. But when his own expected accident savings are one dollar, the social savings will be two dollars (because his loss of $25,000 in an accident is one half of the social loss of $50,000). This implies that, from a social standpoint, he should always increase his spending on safety beyond the point that he would find personally best. This shown in the formal analysis of the next section. Similar remarks will apply to what is asserted about mileage below.

inadequate level of safety because they will bear only part of, and in the model only half of, the social losses in vehicle accidents.

A person's decision about mileage will also be socially improper and excessive in the absence of liability, again because he will not bear the full social costs of accidents. We found that if person A traveled an extra 1,000 miles, the expected number of accidents would increase by .015 and thus that the social harm would rise by .015 × $50,000 = $750. But in the absence of liability, the losses that A would sustain in accidents from driving another 1,000 miles would be only .015 × $25,000 = $375, half the social harm. For that reason, it is clear that A might choose to drive the extra 1,000 miles even though that would be socially undesirable; this would be true if the benefit he obtained from traveling the 1,000 miles is between $375 and $750. More generally, this argument implies that the number of miles that A would find best to drive would exceed the socially optimal number of miles.

(b) *Strict liability.* Now we will consider a succession of liability rules, beginning with strict liability, by which I mean that an individual is strictly liable for any harm caused to another individual in a vehicle accident.[30] This rule can immediately be seen to be equivalent to the regime with no liability. Consider that in an accident between persons A and B, each would suffer $25,000 in harm and that each would be strictly liable to the other. Thus A would have to compensate B for her losses and B would have to compensate A for his losses. Hence, in the end A would incur $25,000 in expenses—his liability payment to B. The position of A, that he would bear an expense of $25,000, would thus be what it was in the no liability regime. The same would be true for B.[31]

Because strict liability of individuals to each other devolves into a no liability regime, the outcomes under the strict liability regime must be the same as in the absence of liability, that is, the levels of safety will be socially subpar and mileages will be socially excessive.

It may be facially puzzling that a rule of strict liability is equivalent to a regime of no liability. How can it be that imposing absolute liability on parties for causing harm is no different from leaving them free of liability? A way of understanding this conclusion is that because there is only *one* kind of actor in the model—an individual who both causes accident losses and suffers these losses—the imposition of liability means not only that individuals who injure others must pay for harms caused, it *also* means that they will *receive* compensation for harms that they sustain. Since the amount paid and the amount received are the same in any accident, $25,000, liability has no impact. (If we had allowed for the two vehicles to suffer different levels of loss, however, then strict liability would not be equivalent to a regime of no liability, but strict liability still could not lead to optimality because individuals' incentives would be compromised to the

---

[30] It is assumed for simplicity that the only issue of causation required for liability is that each vehicle asserted to have been involved in an accident was in fact involved—the presence of each was a but–for cause of the accident. It is also assumed that there is no defense of fault—because there are no drivers who could have acted negligently. (See, however, subsections (g) and (h) below, where I mention the possibility of fault in the sense that a vehicle might have a sub–par safety level.)

[31] This observation was made by Diamond (1974), p. 117, and Arlen (1990), p. 77.

extent that they received liability payments—and strict liability would still lead to results resembling those in the absence of liability.[32])

(c) *Strict liability to the state.* As has just been explained, the root of the problem with the strict liability rule is that the receipt of liability payments dilutes the incentives of individuals to take the $50,000 of social losses caused by each accident into proper account in deciding on safety levels and mileage. This invites us to ask how the receipt of damages could be avoided in a regime under which parties are held strictly liable for the harm that they cause. An answer is to employ the rule of strict liability to the state. Under this rule, individuals who cause harm in accidents must pay for that harm, but their payments are made *to the state*, not to accident victims. Thus, neither party involved in an accident receives damage payments and, as noted in the introduction, each party will bear the entirety of the losses caused by it. If A is involved in an accident with B, A will bear his own loss of $25,000 and also pay $25,000 to the state for the loss caused to B, so that the personal cost to A of involvement in the accident will be $50,000; and similarly for B.

It is apparent that because a person will bear the harm sustained by the two vehicles involved in an accident, not just his own harm, his safety choices will be socially desirable. We calculated in Section 2.2 that it was socially desirable for A to increase his safety level by $1,000 because that would lower the number of accidents by .025, implying a reduction in accident harm of $.025 \times \$50,000 = \$1,250$, exceeding the $1,000 cost. Since when A faces strict liability to the state he would also bear $50,000 in costs per accident, he would make the socially desirable decision to increase his safety level, in contrast to the situation in the absence of liability or under strict liability of individuals to each other.

Likewise, the mileage decision of an individual will be socially desirable. It was observed that if person A traveled an extra 1,000 miles, the expected number of accidents would increase by .015 and thus that the social harm would rise by $.015 \times \$50,000 = \$750$. Under strict liability to the state, A would make the same calculation about the cost to himself of traveling another 1,000 miles, so would make the socially correct decision about increasing mileage.

The reader may wonder why individuals will behave socially desirably when the sum of the amounts that they bear in an accident is *twice* the harm in the accident. A and B each bear $50,000 under the regime of strict liability to the state, so together they incur $100,000 in expenses in an accident involving total social harm of only $50,000. Intuition might suggest that their having to pay twice the social harm would lead to excessive safety and/or would unduly curb mileage. This intuition would be correct if the parties made their safety and mileage decisions jointly, *as a unit*. Suppose that A and B were the only individuals in the world and decided together, ex ante, on their levels of safety and of mileage. They would then tend to find

---

[32] Suppose for example that one type of person always suffers losses of $25,000 in an accident and another type of person always suffers $20,000. Then in any accident involving the two different types, the first will lose $25,000, receive that amount in damages, and thus bear $20,000 in liability expenses. The second type will be compensated for his $20,000 of losses and thus in the end will bear $25,000 in liability expenses. Neither would face the true social loss of an accident of $45,000. Thus, each would choose inadequate levels of safety, but different levels. Moreover, on average, each would bear $22,500 in expense, which is the average expense that each would bear in a regime of no liability. Hence, their behavior would not be substantially different in that regime from under strict liability. See also the remark after Proposition 3 in Section 3.

it in their joint interest to invest excessively in safety and to travel too few miles. But the assumption made in the model is that individuals who might become involved in accidents are strangers to one another and choose their levels of vehicle safety and mileage independently.

(d) *Strict liability of manufacturers.* We now turn to examine the liability of manufacturers. By definition of strict liability of manufacturers, a manufacturer is held responsible for the harms that its vehicles cause to other vehicles. Thus, in an accident involving person A and person B, the manufacturer of A's vehicle would pay damages of $25,000 to B because this was the harm that a vehicle produced by A's manufacturer caused for B. Similarly, the manufacturer of B's vehicle would pay $25,000 in damages to A. Hence, under this rule individuals will be fully compensated for the harm they suffer in accidents and will not bear liability themselves.

However, individuals will turn out to have to pay a *mileage fee* to manufacturers to cover manufacturer liability expenses. In particular, when a manufacturer sells a vehicle to a person, the manufacturer will bear a liability expense that depends on the buyer's mileage and on the safety level of the vehicle. This liability expense of the manufacturer must be recovered from the buyer for the manufacturer to break even, as was mentioned at the beginning of this section. The manufacturer can accomplish the recovery of its expected liability expenses by charging the buyer a mileage fee (say, annually) equal to miles traveled multiplied by the expected liability expense per mile. In the case of person A, we know that driving an extra 1,000 miles would cause .015 accidents a year given the safety level of A; thus the mileage fee per 1,000 miles would be .015 × $25,000 = $375. A mileage fee should be readily implementable because mileage is assumed to be observable to the manufacturer.[33]

It should be remarked that there is another way for a manufacturer to recover its liability expenses. Suppose that the manufacturer adds to the vehicle price an enhancement amount at the time of purchase equal to the predicted liability expense of the manufacturer. This scheme is inferior for the buyer than the mileage fee. The reason is that under the price enhancement scheme, the buyer would have no motive to moderate his mileage to reflect the fact that traveling more miles will contribute to the manufacturer's liability expenses. Consequently, the buyer would drive an extra 1,000 miles whenever doing so would yield him any benefit, such as $100, even though these miles would cost the manufacturer $375 and thus cost the buyer an extra $375 in the enhancement price. Using this logic, it can be shown that the buyer would turn out to pay more in the enhanced price scheme than under the mileage fee arrangement and would be worse off.[34] Therefore, the mileage fee arrangement would characterize the agreement between buyers of vehicles and manufacturers.

Given that manufacturers will charge a mileage fee to cover their liability costs, we can see that an individual's situation reduces to the one that obtains in the absence of liability. The mileage fee covers the expected accident losses for one vehicle, based on miles driven, the safety level of the vehicle, and an expense of $25,000 per accident. This is the same expected loss that

---

[33] In a world of autonomous vehicles that will be electronically linked and mileage will be verifiably recorded.

[34] See the proof of Proposition 5 in Section 3.

an individual would face in the absence of liability. Hence, we infer that when manufacturers are strictly liable, individuals will behave just as they would in the absence of liability, choosing too low a level of safety and a socially excessive mileage. More directly, we can see that if an individual increases the safety level of his vehicle, his mileage fee will fall to a degree reflecting an accident cost of only $25,000 rather than the full social harm of $50,000, so that his incentive to increase safety will be too low;[35] and if an individual increases his mileage, his mileage fee will increase to a degree reflecting again only the $25,000 amount, so the number of miles he will choose to travel will be excessive.[36]

We have therefore arrived at the conclusion that the outcome under strict liability of manufacturers is the same as the outcome under strict liability of individuals for harm caused—and both are equivalent to a regime of no liability. The underlying explanation is that strict liability imposes liability but also compensates individuals involved in accidents in equal measure; and when manufacturers are the bearers of liability, their liability burden is imparted to buyers through market forces and contractual mileage fee arrangements, so that in the end it is as if individuals are strictly liable.

One more comment should be added. A reader might be led to think that when a manufacturer faces strict liability and must pay $25,000 for each accident that its vehicle causes, then "of course" the manufacturer will be induced to choose a higher safety level than if the manufacturer faces no liability. How can this intuition be reconciled with the different conclusion that we have drawn? The flaw in the intuition is that when a manufacturer does not face any liability, the manufacturer still has a motive to produce vehicles with a distinctly positive level of safety. That motive is that then *buyers* will want vehicles with a safety level to reflect the $25,000 of losses that they will sustain per accident in the absence of liability. This level of safety is exactly that which manufacturers would wish to adopt when they face strict liability.

(e) *Strict liability of manufacturers to the state*. We saw earlier in subsection (c) that when individuals faced the rule of strict liability to the state, outcomes were optimal, the explanation being that then an individual involved in an accident had to bear the entirety of the harm from it, $50,000. The same conclusion holds when manufacturers are subject to strict liability to the state, by which we mean that the manufacturer must pay for the $25,000 harm

---

[35] The mileage fee per 1,000 miles is, we said, .015 × $25,000 = $375 when the safety level of A's vehicle is $30,000, so when A drives 10,000 miles, the total fee will be $3,750. If A contemplates raising the safety level to $31,000, we assumed that this would reduce the accident probability per encounter from .3% to .25%. Since there are five encounters for every 1,000 miles, the number of accidents per 1,000 miles would fall to .0125, implying that the mileage fee would become .0125 × $25,000 = $312.50. Hence for 10,000 miles the fee would fall to $3,125. The savings in the mileage fee would therefore be only $625, so spending the extra $1,000 on safety would not be worthwhile for A.

[36] Because the mileage fee per 1,000 miles is $375 given the safety level of A's vehicle, and A contemplates increasing his mileage from 10,000 to 11,000, he would have to pay an additional $375 to do so. But the social cost of increasing his mileage by 1,000 miles is .015 × $50,000 = $750. Hence, if the value he places on driving the 1,000 additional miles is between $375 and $750, he would increase his mileage to 11,000 even though that would not be socially worthwhile.

caused by one of its vehicles when it is involved in an accident with another vehicle, but where the damages payment is made to the state rather than to the other party involved in the accident.

In particular, we know by essentially the argument of the last subsection that when manufacturers are strictly liable for harm caused to other vehicles, the manufacturers will charge a mileage fee to buyers based on their expected liability. Hence, as we observed, person A will pay a mileage fee per 1,000 miles of .015 × $25,000 = $375 to his manufacturer, given that the safety level of his vehicle is $30,000. However, A will *also* bear his own losses of $25,000 in any accident because he will not receive any damage payments from the manufacturer of B's vehicle. Accordingly, given his safety level, A will bear expected expenses of .015 × $50,000 = $750, which are the social costs per 1,000 miles.

It follows that individuals will choose socially correct mileage levels. If A is driving 10,000 miles and contemplates increasing his mileage to 11,000, he will view the cost of so doing as $750, which is the social cost, and increase his mileage only if his personal benefit from the extra mileage exceeds $750.

Analogously, individuals will choose socially correct safety levels for their vehicles. If A plans to drive 10,000 miles a year and contemplates increasing his safety level from $30,000 to $31,000, he will be charged less in his mileage fee premised on the $25,000 liability of the manufacturer and also will bear less in his own accident losses of $25,000 per accident. This will give him the incentive to increase his safety level by $1,000, as is socially desirable.[37]

We have thus explained the rationale for the claimed conclusion that socially correct decisions will be made both with regard to mileage levels and the safety of vehicles. Again, the logic involves the point that manufacturer liability is transferred to individuals through the mileage fee. When this transferred liability amount is added to the bearing of individuals' own losses—which occurs because liability payments are made to the state—individuals' incentives become the social incentives corresponding to the harm suffered by both vehicles in an accident.

(f) *Strict liability of manufacturers to buyers.* Another form of manufacturer strict liability is the rule under which a manufacturer would have to compensate a buyer of its vehicle for harm sustained in an accident (rather than the other vehicle that suffered harm). Under this rule, if A is involved in an accident with B, the manufacturer of A's vehicle would pay A damages of $25,000 because this was the harm that A suffered. Hence, A would not be out–of–pocket if an accident occurred. However, because manufacturers would bear liability expenses, they would impose mileage fees for essentially the reasons we discussed in subsection (d). Therefore, A would pay a mileage fee equal to the expected accident losses he suffers. This implies that individuals would be in the same position as they would be in the absence of liability, when they bear only their own losses. Accordingly, the outcome will be the same, and individuals will choose socially inadequate levels of safety and socially excessive mileage.

(g) *Fault–based manufacturer liability*. Let us next consider *fault–based manufacturer liability*. By fault, or negligence in this context, I refer to the manufacture of a vehicle with a less than the socially desirable safety level. According to the fault–based manufacturer liability rule,

---

[37] In note 35 above, we observed that if A raises the safety level of his vehicle from $30,000 to $31,000, his mileage fees for driving 10,000 miles will fall by $625. But now his own expected accident expenses will also fall by $625, so that his total savings will be $1,250, exceeding the $1,000 extra expense on safety.

a manufacturer will be held liable for the harm caused by its vehicle to another vehicle (not to the buyer's) in an accident if and only if its vehicle's safety level was subpar.

It can be shown that under this rule, manufacturers would be led to produce and sell vehicles with socially optimal levels of safety. The spirit of the demonstration proceeds from the point that if a person bought a vehicle that did not satisfy the required socially optimal safety level, the manufacturer would be liable for accidents. That would imply that the buyer would have to pay a mileage fee based on expected liability payments. But it will be argued that the buyer would be better off purchasing a more expensive vehicle satisfying the optimal safety level in order to avoid the liability–related mileage fee. Hence, a manufacturer selling a vehicle with a less than desired safety level could not survive in business.

To illustrate, recall that it would be socially desirable for A to have a vehicle with a safety level of $31,000 rather than $30,000 because the reduction in the risk of accidents would be worth $1,250, exceeding $1,000. Hence, the manufacturer of A's vehicle would be found negligent for an accident if the safety level of A's vehicle is $30,000 but would be free of liability if the safety level is $31,000. Now consider which choice of safety level person A would make. If he chooses the $30,000 safety level, the manufacturer will be held liable in accidents, meaning that it will impose a mileage fee of $375 per 1,000 miles or $3,750 for the 10,000 miles A drives. Additionally, A will bear his own accident losses of $3,750,[38] implying that A will have expenses of $7,500, apart from his $30,000 cost for his vehicle. If, however, A increases the safety level to $31,000, the manufacturer will not be at fault, so that it will not charge a mileage fee. Yet A will still bear his own accident losses, which will amount to $3,125 since the accident risk will be lower. Thus, A's total expenses above $30,000 if he spends the $1,000 more on safety will be $1,000 plus $3,125 in his accident losses, adding to $4,125. This amount is less than the extra $7,500 he would bear if he obtains the $30,000 vehicle. Accordingly, A will choose to purchase a vehicle with the optimal level of safety.

However, the mileage choices of individuals would be socially excessive. The reason is that because manufacturers would not be found liable, individuals would bear only their own losses from accidents. If A increases his mileage by 1,000 miles, his extra expense would be $312.50, not the social cost of $625. Hence, A would have a socially inadequate incentive to moderate his mileage.[39]

(h) *Fault–based manufacturer liability to buyers (product liability*[40]*)*. This rule is like the last one, except that when the manufacturer is held liable in an accident because its vehicle's

---

[38] This would be true if B is not at fault, an assumption I make because the claim I am establishing is that A will be induced to choose a vehicle with optimal safety given that all others do that. For details, see the analysis of the fault–based rule in the next section.

[39] That mileage would be socially excessive is a reflection of the fundamental problem of the negligence rule that parties subject to it will be led to engage in potentially harmful activities to a socially excessive extent—for if a party behaves in a non–negligent manner, the party does not pay for accident losses that still occur. See originally my article Shavell (1980) developing this point.

[40] This rule may be interpreted as product liability because that form of liability is predicated on types of fault: defective design; manufacturing defect; and failure to warn. See Restatement (Third) of Torts: Products Liability §§ 1–2 (1998).

safety level is below the required level, it must compensate the buyer of its vehicle for harm rather than the other party. That the *buyer* is the party compensated for vehicle fault turns out to imply that manufacturers will *not* be motivated to sell vehicles obeying the safety level requirement. The reason, in essence, is that if a manufacturer sells vehicles with subpar safety and thus will bear liability for accidents, the resulting mileage fee will not represent a true cost to the buyer—for he will be receiving equivalent damage payments. Accordingly, buyers will wish to purchase vehicles with a safety level that they find best ignoring considerations of liability, and that will reflect only the $25,000 loss that a buyer suffers in an accident.

To illustrate the point, let us reconsider the numerical example from the previous subsection and see how the argument about person A's best choice between a $30,000 safety level and a $31,000 safety level changes. If A chooses the $30,000 safety level, the manufacturer will as explained be at fault in accidents and will charge a mileage fee of $3,750 for the 10,000 miles that A drives. However, A will also *receive* in expected terms exactly this amount. Hence, A's total costs will be $30,000 for the vehicle and $3,750 of his own accident losses, adding to $33,750. If in the alternative A purchases a vehicle with a $31,000 safety level, his total costs were explained to be $31,000 plus his own expected accident losses of $3,125, which add to $34,125. Evidently, then, A is better off purchasing a vehicle with only a $30,000 safety level.

This example suggests what will be shown in the analysis of the next section, that under fault–based manufacturer liability to buyers, safety level will be inadequate, guided by only the $25,000 loss suffered by the buyer in an accident, and mileage levels will be socially excessive.

(i) *Fault–based liability.* Finally, let us consider fault–based liability of the individual, by which I mean that the individual will be held liable for the harm caused by his vehicle if its safety level was below the socially optimal level. In other words, liability is again premised on what was described subsection (g) in regard to subpar vehicle safety, but here the individual is responsible for paying the victim, not the manufacturer.

Under this rule, by an argument similar to that provided in subsection (g), it can be established that individuals will choose to purchase vehicles that meet optimal safety standards, so will be non-negligent, but they will drive an excessive number of miles.

3. The Model: Formal Analysis

3.1 Basic assumptions[41]

A population of size 1 of identical risk–neutral individuals possess autonomous vehicles. The individuals are indexed by the variable t (interpreted as a particular person or as the person's vehicle), which is uniformly distributed on [0, 1]. Let

$y(t)$ = miles traveled by t;

$b(y(t))$ = benefits to t from traveling $y(t)$ miles; $b(0) = 0$; $b'(y) > 0$; and $b''(y) < 0$.

The occurrence of accidents, harmful events involving two vehicles, is determined by two factors. The first is the number of encounters that vehicles have with each other, where an encounter involves two vehicles coming into sufficiently close proximity for an accident to

[41] The assumptions and framework of analysis are as set out in the informal analysis in Section 2, so that I will be brief in this section.

occur. The second is the probability of an accident given an encounter. For any two persons t and t', the rate of encounters is proportional to the mileages of each, namely,

$ky(t)y(t')$ = expected rate of encounters involving t and t'; $k > 0$.

The probability of an accident in an encounter depends only on the safety levels of the two involved vehicles. Let

$x(t)$ = safety level of t,

which also corresponds to the production cost of the vehicle (production cost that is unrelated to safety is ignored for simplicity). The accident probability falls as the safety of either vehicle increases; let

$q(x(t), x(t'))$ = probability of an accident in an encounter involving t and t';

$q_1(x(t), x(t')) < 0$; $q_2(x(t), x(t')) < 0$.

It is assumed that q is a convex function and that it is symmetric in its arguments: $q(x(t), x(t')) = q(x(t'), x(t))$.

If an accident occurs, each of the two persons sustains the same level of harm (interpreted as harm to the vehicle or personal injury); let

h = harm suffered by each person in an accident; $h > 0$,

where h is of fixed magnitude.

Social welfare is the sum of welfare over all individuals t. The welfare of a person t equals the benefit he derives from travel, less the production cost of his vehicle, less the expected accident losses suffered by him. Thus, the welfare of a person t is

$$(1)\quad W(t) = b(y(t)) - x(t) - \int_0^1 ky(t)y(z)q(x(t), x(z))hdz.$$

The integral term is the expected losses suffered by t from accidents, since he will have $ky(t)y(z)$ encounters with each person z, where z ranges over all possible individuals, and since in any encounter, the accident probability will be $q(x(t), x(z))$. Hence, social welfare is

$$(2)\quad W = \int_0^1 W(t)dt = \int_0^1 [b(y(t)) - x(t) - \int_0^1 ky(t)y(z)q(x(t), x(z))hdz]dt.$$

The problem of social welfare maximization is to choose the safety level function $x(t)$ and the mileage function $y(t)$ to maximize (2). I will restrict attention to socially optimal functions in which all individuals behave identically;[42] these optimal functions will be denoted by $x^*(t) = x^*$ and $y^*(t) = y^*$, and x^* and y^* will be assumed to be positive.

3.2 Socially optimal behavior

A necessary condition for social welfare maximization is that for any t, $x(t)$ and $y(t)$ maximize social welfare W given the choices of all other individuals. To characterize this

[42] This is a natural restriction, for all individuals have identical utility functions and are governed by identical functions determining encounters, accident probabilities, and harm from accidents.

necessary condition, we need to state the effect of a person's $x(t)$ and $y(t)$ on social welfare W; this effect will be designated $R(t)$. It is claimed that

$$(3) \quad R(t) = b(y(t)) - x(t) - 2\int_0^1 ky(t)y(z)q(x(t), x(z))h\,dz;$$

in other words, the assertion is that $R(t)$ differs from $W(t)$ in that the expected accident loss integral in $W(t)$ is multiplied by 2. The explanation for the factor of 2 is the inherent externality in an accident. Namely, whenever person t is involved in an accident and suffers a loss of h, the other person involved in that accident also suffers a loss of h. This externality can be seen formally from (1) and (2). Consider a person t' different from t. Observe that t' will be reflected in $W(t)$, for the accident integrand $ky(t)y(z)q(x(t), x(z))h$ will equal $ky(t)y(t')q(x(t), x(t'))h$ when $z = t'$. But this value of the integrand will *also* appear in $W(t')$. For in $W(t')$, the accident integrand will be $ky(t')y(z)q(x(t'), x(z))h$; and when $z = t$, this integrand will equal $ky(t')y(t)q(x(t'), x(t))h = ky(t)y(t')q(x(t), x(t'))h$ (by the assumption that q is symmetric in its arguments). Thus, each value of the accident integrand in $W(t)$ has a twin in $W(z)$ for some z.

Accordingly, a necessary condition for optimality of $x(t)$ is found by maximizing $R(t)$ with respect to that variable. This leads to the first–order condition

$$(4) \quad 1 = -2\int_0^1 ky(t)y(z)q_1(x(t), x(z))h\,dz.$$

In other words, the marginal cost 1 of raising the safety level equals the marginal reduction in expected accident losses. Likewise, a necessary condition for optimality of $y(t)$ is found by maximizing $R(t)$ with respect to it, namely,

$$(5) \quad b'(y(t)) = 2\int_0^1 ky(z)q(x(t), x(z))h\,dz;$$

the marginal utility from traveling another mile equals the associated marginal increase in expected accident losses.

If we assume that, aside from person t, all other individuals have common mileage $\hat{y}$ and safety level $\hat{x}$, then (1)–(5) become

$$(1') \quad W(t) = b(y(t)) - x(t) - ky(t)\hat{y}q(x(t), \hat{x})h,$$

$$(2') \quad W = \int_0^1 W(t)dt = \int_0^1 [b(y(t)) - x(t) - ky(t)\hat{y}q(x(t), \hat{x})h]dt.$$

$$(3') \quad R(t) = b(y(t)) - x(t) - 2ky(t)\hat{y}q(x(t), \hat{x})h,$$

(4') $1 = -ky(t)\hat{y}q_1(x(t), \hat{x}))2h,$

(5') $b'(y(t)) = k\hat{y}q(x(t), \hat{x}))2h.$

Given the assumption that at the social optimum all individuals behave identically, (4') and (5') imply that x^* and y^* are determined by

(6) $1 = -ky^{*2}q_1(x^*, x^*))2h,$

(7) $b'(y^*) = ky^*q(x^*, x^*))2h.$

In other words, the marginal cost 1 of raising safety equals the marginal benefit in expected accident cost reduction, taking into account that $2h$ is the cost of an accident, and the marginal benefit of increasing mileage equals the marginal cost of driving another mile, taking into account that $2h$ is the cost of an accident. In other words, we have

Proposition 1. The socially optimal levels of safety x^* and mileage y^* for individuals are determined by (6) and (7).

If we relax the assumption that the harm h is fixed, the determination of the social welfare optimum would be straightforward and essentially unchanged. Notably, if h is probabilistic and independently distributed for each individual in an accident, then the mean $E(h)$ would replace h in all the expressions above; and if h is a function $h(x(t), x(z))$ of safety levels, the first–order conditions would change in obvious ways.

3.3 Behavior under liability regimes

Let us now consider outcomes under different accident liability regimes involving individuals and manufacturers. It will be assumed that individuals purchase autonomous vehicles from manufacturers in a competitive industry. The price of a vehicle is therefore taken to equal its production cost x, which is its level of safety, plus any liability expenses that the manufacturer faces. It will also be assumed that manufacturers observe a buyer's mileage.

We will suppose that individuals act identically in equilibrium, that is, that equilibria are uniform, and we will compare social welfare to socially optimal welfare given the assumption that individuals act identically at the social welfare optimum.

It is convenient to examine first outcomes in the absence of liability for vehicle accidents. Under this *no–liability* regime the expected utility of a person t is $W(t)$ given by (1'), for he would obtain $b(y(t))$ from his mileage, would pay the manufacturer $x(t)$ since the manufacturer would have no liability expenses, and he would bear only his own accident losses. The first–order condition for the choice of $x(t)$ will therefore be

(8) $1 = -ky(t)\hat{y}q_1(x(t), \hat{x})h,$

where $\hat{x}$ and $\hat{y}$ will continue to denote (here and below) the common values of x and y that are presumed to hold in equilibrium. Comparing (8) to the socially optimal condition (4'), it is clear the solution $x(t)$ to (8) is lower than the socially optimal one because $q_{11}(x(t), \hat{x})) > 0$. In other words, because individuals bear only h, half of the losses of $2h$ that occur in an accident, they choose too low a level of safety, other things being equal. Similarly, the first–order condition for choice of $y(t)$ will be

(9) $b'(y(t)) = k\hat{y}q(x(t), \hat{x})h.$

The solution to this must exceed the solution to socially optimal condition (5'), since $b''(y(t)) < 0$. Because individuals bear only half of the accident losses that they generate, they travel too many miles.

Given the assumption of uniform equilibria, and denoting by x_N and y_N the common choices of safety and miles under the no liability regime, (8) and (9) imply that the following must hold.

(10) $1 = -ky_N{}^2q_1(x_N, x_N)h.$

(11) $b'(y_N) = ky_Nq(x_N, x_N)h.$

We thus have

Proposition 2. The socially optimal outcome does not occur in a regime of no liability for accidents. Individuals choose inadequate safety levels and socially excessive mileage given the behavior of others; their safety levels x_N and mileage y_N are determined by (10) and (11).

It should be noted that although we know that safety is socially inadequate and mileage is excessive given the behavior of others, we cannot conclude how safety and mileage relate to their socially ideal levels x^* and y^* (and similar comments will apply to other results below about less than socially optimal outcomes). For instance, it is not necessarily the case that $x_N < x^*$: if $y_N > y^*$, it is possible that $x_N > x^*$ because privately–chosen safety increases with the mileage of others.[43]

Let us now turn to liability regimes, beginning with *strict liability* of individuals. Under this rule, if person t is involved in an accident with another person t' and thus causes harm of h for t', then t must pay h to t'. Since the rule also applies to t', he must pay h to t. Hence, in the end, the regime of strict liability is equivalent to a regime of no liability, for each party bears h in an accident. We therefore have

Proposition 3. A regime of strict liability of individuals is equivalent to a regime of no liability. Therefore, the regime leads to suboptimal safety levels and excessive mileage, as described in Proposition 2.

---

[43] From (8), we see that an individual's choice of safety $x(t)$ will be determined by $1 = -k\hat{y}^2q_1(x, \hat{x})h$ (where $x(t)$ is written as x for simplicity). This expression implicitly determines x as a function of $\hat{y}$. If we differentiate the expression with respect to $\hat{y}$, we find that the sign of $x'(\hat{y})$ equals the sign of $q_{11}(x, \hat{x})$, which is positive because q was assumed to be convex.

It is worth observing that the inadequacy of incentives under strict liability—and its resemblance to a regime of no liability—would carry over in a substantial sense to a more general model in which harm suffered by individuals in accidents might vary across individuals and by chance. In particular, in any such model, suppose that person t suffers a loss of h_1 and person t' suffers a loss of h_2 in an accident. Then under strict liability, t would bear an expense of h_2 and t' would bear an expense of h_1 (since each fully compensates the other for that person's loss). Thus, in any accident each party bears an expense that is less than the social harm of $h_1 + h_2$ and on average is only $(h_1 + h_2)/2$, just 50% of the social harm, and which also equals the average harm actually suffered that would be borne in the absence of liability. Therefore, incentives for safety would remain substantially short of optimal and would resemble those in the absence of liability.

Next, let us consider the regime of *strict liability to the state*. Under this rule, if a person t is involved in an accident with another person t', then t must pay the state (not t') for the harm h that he caused. Because this rule also applies to t', t' must pay h to the state (not to t). Accordingly, t bears total expenses of $2h$ in an accident—his liability payment of h plus the loss of h that he personally sustains. The price of a vehicle will be $x(t)$ because the only expense of a manufacturer will be production cost. Thus the expected accident losses of t will be $2ky(t)\hat{y}q(x(t), \hat{y})h$. Therefore, the utility of t will be given by $R(t)$ in (3'). Because the goal of t is the social goal, this strict liability regime supports the social optimum.[44]

Proposition 4. The socially optimal outcome is supported by the rule of strict liability of individuals to the state.

We now consider the rule of *strict liability of manufacturers* for harms caused by their vehicles to *other* vehicles (as opposed to harms suffered by the buyers of their vehicles). Under this rule, if persons t and t' are involved in an accident, t's manufacturer must pay h to t'; and likewise, the manufacturer of t' must pay h to t. Thus, individuals do not bear their own losses h.

However, individuals will be charged a *mileage fee* by manufacturers equal to its expected liability per mile. To demonstrate this, note first that a manufacturer's costs will be $x(t)$ together with its expected liability expenses,

$$(12) \quad x(t) + ky(t)\hat{y}q(x(t), \hat{x})h,$$

which the manufacturer must collect from a buyer to break even. An individual's objective is therefore to maximize $b(x(t))$ minus (12), for he does not bear his own accident losses under the rule in question. Note that a person's objective is therefore to maximize $W(t)$ as given in (1'). If an individual is charged $x(t)$ for the purchase of his vehicle and a mileage fee of

[44] That is, the social optimum is a Nash equilibrium under the strict liability regime, for the goal of each individual is to maximize social welfare W given the behavior of others, namely (3').

$ky(t)\hat{y}q(x(t), \hat{x})h$ based on his mileage $y(t)$, then he will choose $x(t)$ and $y(t)$ in a personally optimal way, to maximize (1'). Thus, the mileage fee arrangement must be optimal for him, so I will assume that this is the arrangement. This fee arrangement is implementable because manufacturers will be able to observe mileage. Moreover, a manufacturer and an individual would not find any other contractual arrangement superior because the mileage fee permits the individual to maximize $W(t)$ over $x(t)$ and $y(t)$ without constraint. In this regard, it should be observed that another contractual arrangement that a manufacturer could make satisfying (12) is one in which the manufacturer charges a fixed amount at the time of purchase of a vehicle equal to $x(t) + \delta$, where δ is a constant calculated to equal the liability cost of the manufacturer based on the miles that the individual would be predicted to drive. If that arrangement were made, the individual would drive more miles than the $y(t)$ that maximizes $W(t)$: for after buying a vehicle, the individual would not bear any accident losses. In effect, then, his maximization of $W(t)$ would not be unconstrained and he would be worse off than under the mileage fee arrangement depending on actual miles driven.

Because we have concluded that individuals will not bear their own accident losses but will pay a mileage fee equal to the expected accident losses they cause, their situation will be identical to that under the no–liability regime; they will choose $x(t)$ and $y(t)$ to maximize (1'). Accordingly, we have

Proposition 5. A regime of strict liability of manufacturers to individuals harmed by their vehicles (but not to buyers of their vehicles) is equivalent to a regime of no liability. Therefore, the regime leads to suboptimal safety levels and excessive mileage, as described in Proposition 2. Additionally, under this regime, a manufacturer will charge a price equal to production cost and a mileage fee based on miles actually traveled by the buyer that will cover the manufacturer’s liability expense.

Next, we consider a rule of *strict liability of manufacturers to the state*. Under this rule, a manufacturer must pay the state h for the harm caused by one of its vehicles to the other vehicle in an accident. Therefore, individuals do not receive liability payments for harm suffered in an accident, so that an individual’s utility exclusive of what he pays a manufacturer will be

(13) $b(y(t)) - ky(t)\hat{y}q(x(t), \hat{x})h.$

A manufacturer’s costs will again be given by (12), which it must collect from a buyer to break even. Therefore, an individual’s objective function is expression (13) minus expression (12), which is to say $R(t)$ in (3'). The individual will therefore wish to maximize $R(t)$ over $x(t)$ and $y(t)$ and will do so if a manufacturer charges $x(t)$ for the purchase of the vehicle supplemented by a mileage fee of $ky(t)\hat{y}q(x(t), \hat{x})h.$ This mileage fee would in fact be charged by essentially the argument just given in demonstrating that the mileage fee would be charged under manufacturer liability to individuals harmed by their vehicles.

Because the individual’s goal is to maximize $R(t)$, the social goal, the strict liability regime under consideration supports the social optimum, and we have

Proposition 6. The socially optimal outcome is supported by the rule of strict liability of manufacturers to the state. Under this rule, a manufacturer will charge a price equal to production cost and a mileage fee per mile actually traveled by the buyer that will cover the manufacturer's liability expense.

We next consider what will be called *strict manufacturer liability to buyers*, by which is meant strict liability of manufacturers to the buyers of their vehicles for harms suffered in accidents. Thus, if person t is involved in an accident with person t', the manufacturer of t's vehicle compensates t for his loss; and likewise t' is compensated for his loss by his manufacturer.

The rule under consideration is equivalent to a rule of no liability. In particular, a manufacturer's costs will be given by (12), so that by an argument analogous to those given above, we know that manufacturers will impose a mileage fee equal to its expected liability per mile in addition to charging $x(t)$ for the vehicle. Hence, a person will choose $x(t)$ and $y(t)$ in order to maximize $W(t)$ as in (1'), which is what a person facing no liability would do. Accordingly, we conclude that

Proposition 7. A regime of strict manufacturer liability to buyers is equivalent to a regime of no liability. Therefore, strict liability of owners leads to suboptimal safety levels and excessive mileage, as described in Proposition 2. Additionally, under this regime, manufacturers will charge a price equal to production cost and a fee per mile actually traveled by the buyer of a vehicle.

Let us now consider a rule of *fault–based manufacturer liability*, where manufacturer fault means having produced a vehicle with a socially inadequate level of safety. By the latter is meant a safety level below a conditionally socially optimal level of safety. This socially optimal safety level is defined as that which minimizes expected accident losses plus safety costs *given* the mileage $y = y(t)$ of an individual t and also given the behavior $\hat{x}$ and $\hat{y}$ of others. Denoting this level of safety by $x^*(y)$ (and suppressing $\hat{x}$ and $\hat{y}$ in the notation), it is the x that minimizes

$$(14) \quad x + 2ky\hat{y}q(x, \hat{x})h.$$

From (14), we know that $x^*(y)$ is determined by the first–order condition

$$(15) \quad 1 = -2ky\hat{y}q_1(x, \hat{x}))h.$$

I assume that under the fault rule in consideration, the manufacturer will be liable for harm h that its vehicles cause to other vehicles (not to buyers) if and only if its safety level x is less than $x^*(y)$. This rule can be implemented because y has been assumed to be observable.

I claim that $x^*(y)$ is a Nash equilibrium under the rule. To show this, consider the choice of x of a person given his y. If he chooses $x < x^*(y)$, the manufacturer will be liable, implying that the person will have to pay the manufacturer x plus an amount equal to its expected liability. Moreover, the person will bear his own losses—because the hypothesis is that safety levels of others are $x^*(y)$, meaning that he would not receive damages from the manufacturers of other vehicles. Hence, the person would choose $x < x^*(y)$ to minimize (14). But if (14) is minimized without any constraint on x, the minimum is $x^*(y)$. Hence, (14) evaluated at $x^*(y)$ is lower than at any $x < x^*(y)$. Moreover, at $x^*(y)$ the person bears less than (14) because the manufacturer is no longer liable. Hence, the person will choose $x \geq x^*(y)$. For such x, the person bears

$$(16) \quad x + ky\hat{y}q(x, \hat{x}))h,$$

which he seeks to minimize. Because the derivative of (16) exceeds that of (14) at any x, (16) is minimized at $x^*(y)$.

Now that we have verified that $x^*(y)$ is a Nash equilibrium, let us determine the y that individuals will select. An individual will wish to choose y to maximize

$$(17) \quad b(y) - [x^*(y) + ky\hat{y}q(x^*(y), x^*(\hat{y}))h],$$

so that the first–order condition determining y will be

$$(18) \quad b'(y) = x^{*\prime}(y) + k\hat{y}q(x^*(y), x^*(\hat{y}))h + x^{*\prime}(y)ky\hat{y}q_1(x^*(y), x^*(\hat{y}))h.$$

However, the social problem for an individual who is constrained to choose $x^*(y)$ is to select y to maximize

$$(19) \quad b(y) - [x^*(y) + 2ky\hat{y}q(x^*(y), x^*(\hat{y}))h],$$

because the social losses from an accident are twice the individual's. Hence the first–order condition for the socially optimal y is

$$(20) \quad b'(y) = x^{*\prime}(y) + 2k\hat{y}q(x^*(y), x^*(\hat{y}))h + 2x^{*\prime}(y)ky\hat{y}q_1(x^*(y), x^*(\hat{y}))h.$$

Because the right–hand sides of (18) and (20) are different, the individual's choice of y will generally be socially suboptimal. Subtracting the right–hand side of (18) from the right–hand side of (20), we obtain

$$(21) \quad k \quad q(x^*(y), x^*(\hat{y}))h + x^{*\prime}(y)ky\hat{y}q_1(x^*(y), x^*(\hat{y}))h.$$

These terms correspond to the marginal change in expected accident losses suffered by other individuals when a person increases his mileage, a consideration that is overlooked in the private calculus under the fault rule. If (21) is positive, then the y satisfying (18) will be higher than that satisfying (20), meaning that the privately chosen y is excessive; if (21) is negative, then the privately chosen y will be too low; and if (21) is zero, the two levels of y will be equal.[45]

Let me summarize as follows.

---

[45] In standard models of accidents, the optimal level of care x^* does not change with the level of activity; see, for example, Shavell (1987, p. 41). Were that true, then (18) would become $b'(y) = k\hat{y}q(x^*(y), x^*(\hat{y}))h$ and (20) would become $b'(y) = 2k\hat{y}q(x^*(y), x^*(\hat{y}))h$, so that mileage chosen by the individual would be socially excessive.

Proposition 8. Under a regime of fault–based manufacturer liability to individuals harmed by a manufacturer's vehicles (but not to its buyers), where fault means socially inadequate safety, individuals will choose to purchase vehicles with adequate safety. Hence, manufacturers will be free of fault. However, individuals' mileage choices will generally be suboptimal.

Next, let us consider the rule of *fault–based liability* of individuals, where fault means having purchased a vehicle with inadequate safety, which is to say with $x < x^*(y)$ as defined in the proof of Proposition 8. Here, if an individual t is involved in an accident and is at fault, the individual would have to pay the other party t' the amount h to compensate him for his loss. By an argument analogous to that give for Proposition 8, we obtain

Proposition 9. Under a regime of fault–based liability of individuals, where fault means socially inadequate safety, individuals will choose to purchase vehicles with adequate safety. Hence, individuals will be free of fault. However, individuals' mileage choices will generally be suboptimal.

Last, let us consider a rule of *fault–based manufacturer liability to buyers (product liability).* Under this rule, I assume that there is some standard of safety $\tilde{x}(y)$ that must be satisfied and that may depend on y for the manufacturer to be free of liability; otherwise, if $x < \tilde{x}(y)$, then the manufacturer must pay h in damages to the buyer of its vehicle in any accident.[46] I claim that this rule is equivalent to a regime of no liability. Suppose first that $x < \tilde{x}(y)$. This means that the buyer will not suffer any loss in an accident because he will be compensated by the manufacturer. However, the price paid for the vehicle must equal $x + ky\hat{y}q(x, \hat{x}))h$ because the second term is the liability expense of the manufacturer. Suppose next that $x \geq \tilde{x}(y)$. Then the manufacturer will not be liable so will charge x for the vehicle and the buyer will bear his own accident losses, so the buyer's total expenses will again be $x + ky\hat{y}q(x, \hat{x}))h$. Hence, the buyer's utility will be $W(t)$ as in (1'), what it is in the absence of liability. Whether manufacturers will be found at fault will depend on whether $x_N < \tilde{x}(y_N)$. In sum, we have

Proposition 10. Under a regime of fault–based manufacturer liability to buyers (product liability), outcomes will be identical to those in a regime without liability; safety levels will be less than socially appropriate and mileage will be excessive.

[46] As will be seen, the definition of $\tilde{x}(y)$ does not matter. Under any definition, the same outcome will occur.

4. Discussion

In this section, I appraise the main conclusions reached in the analysis of the model and consider a number of issues not addressed in it.

(a) *Strict liability to the state possesses an as yet unmentioned advantage over fault: control of all dimensions of safety, not just those that can be well–policed by courts in a fault regime*. A general advantage of strict liability, and thus of strict liability to the state in the context of autonomous vehicle accidents, was not mentioned in the analysis. Namely, strict liability should motivate parties to make not just a single abstract precautionary decision in a socially correct fashion, but to make all manner of decisions about safety appropriately—to move any lever of influence over risk when the cost of so doing would be justified by the resulting reduction in expected accident losses. Hence, manufacturers of autonomous vehicles should be induced to make a great multitude of choices properly, including those concerning the development and testing of vehicle control software, communications systems and cyber security, braking mechanisms, steering apparatus, elements of design to enhance crashworthiness, and methods of quality assurance in production.

In contrast, the fault system will not induce parties to make many of the decisions confronting them in a socially appropriate manner. A primary reason is that courts will lack the information necessary to apply fault rules in a way that approximates the socially desirable. As was just emphasized, the safety of vehicles depends on a plethora of manufacturer decisions about the extremely complex device of an autonomous vehicle. It is implausible that courts would be able to formulate socially desirable standards for many of these decisions, let alone be able to identify many types of decisions that are relevant to risk. Consider, for instance, the ability of courts to formulate correct fault standards in regard to communications between vehicles to reduce accident dangers, whether sensors should be able to detect black ice, or the degree of testing that should be conducted into the structural integrity of a vehicle using a modified material. Because courts' ability to carry out such tasks will be limited owing to their lack of information, they will sometimes insist on fault standards that are too rigorous, other times permit standards that are lax, and other times simply omit types of precautions from their fault considerations. Moreover, even where courts can formulate proper fault standards, courts must be able to observe whether manufacturers lived up to the standards, and that could be problematic.

None of these informational difficulties afflicting the courts in applying the fault rule would arise under strict liability because, as I stressed, manufacturers will then find it in their financial self–interest to move all the levers under their control to properly reduce accident risks, whether or not courts would have knowledge of how these levers ought to be moved, of the existence of all the levers, and of how they were in fact moved.[47]

(b) *The advantage of strict liability to the state over fault in controlling mileage*. The main advantage of the rule of strict liability to the state over vehicle fault rules that was identified in the model concerned mileage choices: whereas these choices will be socially desirable under such strict liability because parties will have to pay for all harms caused, mileage choices will be socially excessive under vehicle fault rules because individuals will not have to

[47] See, for example, Shavell (1987, p. 9) on the general point made in these paragraphs.

pay for harms that they cause. This advantage of strict liability also has a broader interpretation when one recognizes that users of autonomous vehicles will not merely choose a single scalar, mileage, they will also do so when they know that the risk accident risk per mile will depend on where they travel, when they travel, and weather conditions. They will therefore be motivated to reduce mileage more in circumstances of higher risk.[48]

The advantages of strict liability to the state stressed in (a) and that concerning mileage just noted suggest that strict liability to the state should dominate vehicle fault as a means of controlling accident risks. But these advantages do not tell us precisely why strict liability to the state should be superior to conventional strict liability, and I now summarize the conclusions from the analysis about that issue.

(c) *Conventional strict liability engenders a fundamental dilution of incentives to reduce vehicle accident risks—and renders it equivalent to, or similar to, a regime without liability.* The problem with conventional strict liability observed in the model was that because any individual who was involved in an accident would by definition be *compensated* for his own losses by the counter party, the individual would bear only liability expenses. Thus, the party would bear only *half* the total harm caused by the accident. This self–diluting aspect of strict liability meant not only that individuals' incentives to reduce accident risks would be socially inadequate, it also implied that their incentives would be no different from their motives to reduce risk in the absence of liability. For in the absence of liability, individuals would also bear half of the total harm caused in an accident (their own harm). Moreover, even if the assumption of the model that each individual involved in an accident suffers the same loss is relaxed, strict liability still results in the individuals' bearing only half of the total harm on average, and thus the incentives that it creates are similar to those applying in the absence of liability.[49]

The claims of the foregoing paragraph were shown to apply to manufacturer strict liability as well as to individual strict liability. But that raised an apparent puzzle: How could it be true that imposing strict liability on manufacturers would not augment their incentives to produce safe cars over what they would be in the absence of liability? The answer (which was given in the last paragraph of Section 2.3(d)) is that in the absence of liability, individuals would bear their own losses. For that reason, manufacturers would face market pressure to provide safety. This market pressure would be equivalent to the liability pressure that manufacturers would experience if they are held strictly liable for harms caused by their vehicles in accidents.

(d) *Lack of compensation for losses under the rule of strict liability to the state would not be problematic for risk–averse individuals owing to the availability of first–party insurance.* The rule of strict liability is designed *not* to compensate individuals involved in accidents because this feature of liability is needed in order to produce desirable incentives to reduce risk. This lack of compensation would doubtlessly be considered socially detrimental by many legal scholars, judges, and citizens today because of the broadly held view that making victims of loss whole is a, if not the, primary purpose of tort law. However, first–party insurance for property losses and

[48] Consider for example the calculus of large commercial users of vehicles, like Amazon, FedEx, and UPS. Under a strict liability regime, these users would be expected factor into their calculus how risks vary with route choice, time of day, and weather.

[49] See the paragraph following Proposition 3.

personal injury is widespread in modern economies and constitutes a less costly mechanism for compensation than the tort system. On general functional grounds, therefore, that the proposed rule does not compensate risk–averse victims of accident losses should not be judged a disadvantage.

(e) *Liability insurance under the rule of strict liability to the state.* Under strict liability of persons to the state, individuals would be directly liable in the event of accidents, so, if risk averse, they would tend to purchase liability insurance policies. Ownership of the policies would not be likely to compromise incentives to reduce risk, however. The reason is that insurance premiums would depend on mileage and risk per mile, both of which would be readily observable by insurers. Hence, insureds would have a proper motive to control mileage and also to pay for enhanced safety when purchasing their vehicles.

Under strict liability of manufacturers to the state, individuals would have no reason to purchase liability insurance. They would, however, pay manufacturers mileage fees based on risk, and as was explained in Section 2, that would create socially proper incentives for risk–reduction and mileage. Manufacturers might be unlikely to purchase liability coverage because they might be viewed as approximately risk–neutral. If they did purchase such coverage, that would not tend to distort their incentives to take care, again because their premiums would reflect vehicle safety and mileage levels.

(f) *Application of the rule of strict liability to the state should be straightforward.* To apply the rule, the state will have to determine when accidents occur and then damages. (Liability should be unproblematic because it is strict, unaccompanied by any defenses.)

The state should be able to ascertain when accidents happen because vehicles will be linked electronically to systems that continuously record their movements. It is important to note, however, that the state will need to require reporting of accidents from the electronic systems because parties will have no affirmative reason to inform the state that they were involved in accidents. Indeed, the parties involved in an accident will actually have a perverse incentive to conceal accidents, because in that way they would jointly save an amount equal to their total losses.[50]

The determination of damages should not present issues out of the ordinary and could be routinized by the state to promote settlement. Moreover, efficiencies could be achieved if accident litigation was adjudicated by specialized courts.

(g) *Accidents involving semi–autonomous vehicles, conventional vehicles, motorcyclists, bicyclists, and pedestrians.* How might the rule of strict liability to the state apply in accidents involving an autonomous vehicle and either a different kind of automobile—semi–autonomous or conventional—or a motorcyclist, bicyclist, or pedestrian? These questions are of interest because it will be decades before the supermajority of automobiles are fully autonomous and because conventional vehicles may be driven for sport, and motorcyclists, bicyclists and pedestrians will often be at risk for an indeterminable future.

---

[50] In the example of Section 1 in which person A's loss is $30,000 and person B's loss is $20,000, under strict liability to the state, A bears his loss of $30,000 and pays B $20,000 in damages to the state, and B bears her loss of $20,000 and pays A $30,000 in damages to the state. If they conceal the accident, A avoids $20,000 in damage payments and B avoids $30,000—so their joint incentive to conceal is $50,000, the sum of the accident losses that the two sustain.

An answer to the questions is provided by a general observation about strict liability to the state. Namely, the argument supporting its use does not depend on the assumption that vehicles are autonomous—the rule should induce proper incentives to take care and to choose levels of activity such as mileage in *any* accident context.[51] Thus, if some individuals exercise direct control over their level of safety—say they are drivers of conventional automobiles—that does not imply that they should be subject to the fault rule. Indeed, they would continue to have socially proper incentives to reduce risk under strict liability to the state through the exercise of precautions and the choice of mileage levels.

It follows that strict liability to the state could in principle be employed for any combination of types of parties. If, however, mixed liability regimes were employed, some distortion of incentives from ideal ones would occur. For example, suppose that if an autonomous vehicle and a conventional vehicle are involved in an accident, liability of the conventional vehicle is governed by fault, whereas the autonomous vehicle owner would be held strictly liable and pay damages to the state. Then the incentives of the conventional owner to take care would be appropriate, but he would be led to drive too many miles; and the incentives of the autonomous vehicle owner would be proper.

(h) *If strict liability to the state instills socially desirable incentives to reduce accident risks in any accident context, should its use be more broadly recommended?* The answer is not clear. The principal reason is that, as mentioned in (f) but not considered in the analysis, a serious problem with the proposed rule is that it would motivate parties to hide their involvement in accidents in order to avoid making payments to the state. It happens that this problem does not afflict the rule in the context of autonomous vehicle accidents, for as also observed in (f), these accidents will be electronically recorded, furnishing the state ready access to information about their occurrence. That aspect of autonomous vehicle accidents makes them a natural candidate for the use of strict liability to the state, whereas the problem of reporting of accidents might be difficult to circumvent for other kinds of accidents.

Reinforcing the view that the proposed rule might not be advantageous for other kinds of accidents is the conjecture that its advantages in risk–reduction may be only modest (and thus not sufficient to outweigh the problem of concealment of accidents). In particular, in accident contexts involving distinct injurers, who suffer little or no harm themselves in accidents, and distinct victims, who suffer most or all of the harm, the rule of strict liability with a defense of contributory negligence may provide tolerably good incentives to reduce accident risks.[52] If so, the risk–reduction gain from use of the proposed rule would be circumscribed. In the setting of

[51] Regardless of whether the accident context involves one activity (like driving) or multiple activities (driving and bicycling, or driving and bicycling and walking), and regardless of how many variables each party controls (dimensions of care, dimensions of activity), it is clear that if *each* party is made to pay for the *sum* of all the losses caused in an accident in which the party is involved and that payment is made to the state, the party will act socially optimally given the behavior of all other parties. Thus, the social optimum will be supported by the rule of strict liability to the state.

[52] If accidents cause harm to victims only, then under strict liability with a defense of contributory negligence, injurers would in fact be paying for the sum of accident harms. That is because, as is well–known, this rule would lead victims to take optimal care. Therefore, the only problem of incentives under this rule would concern the activity level of victims.

automobile accidents, however, the gap between socially desirable incentives and those engendered by conventional liability rules could be significant, as emphasized in (c).

To summarize, it appears that in the context of autonomous vehicle accidents, there is both an unusual opportunity to implement the rule of strict liability to the state and a substantial incentive gain to be had from doing so. In contrast, in most accident contexts, the opportunity to employ the rule would be hindered by the parties' incentives not to report accidents, and the social benefit that would be obtained from use of the proposed rule might not be substantial.

(i) *If the virtues of strict liability to the state for autonomous vehicle accidents are as asserted, why has the rule not drawn attention before?* A reply is that the proposed rule represents a radical deviation from our present tort regime—chiefly because it requires damages to be paid to the state rather than to victims of harm, and secondarily because it is a form of strict liability, whereas fault is the predominant basis of liability in tort law. But anticipation of the advent of fully autonomous vehicles raises questions in a sharp way about the utility of the conventional rules of tort liability. And, as explained in this article, when one inquires about the dual problems of assuring appropriate vehicle safety and mileage levels in the future world of autonomous vehicles, the proposed rule of strict liability to the state emerges as a solution.

References

Abraham, Kenneth S., and Robert L. Rabin. 2019. Automated Vehicles and Manufacturer Responsibility for Accidents: A New Legal Regime for a New Era. *Virginia Law Review* 105:127-171.

Arlen, Jennifer H. 1990. Reconsidering Efficient Tort Rules: The Case of Single Activity Accidents. *William & Mary Law Review* 32:41–103.

Bansal, Prateek, and Kara M. Kockelman. 2017. Forecasting Americans' Long-term Adoption of Connected and Autonomous Vehicle Technologies. *Transportation Research Part A* 95:49-63.

Ben-Shahar, Omri. September 22, 2016. Should Carmakers Be Liable When A Self-Driving Car Crashes? *Forbes* https://www.forbes.com/sites/omribenshahar/2016/09/22/should-carmakers-be-liable-when-a-self-driving-car-crashes/2/#309606f74378

Bierstedt, Jane, Aaron Gooze, Chris Gray, Josh Peterman, Leon Raykin, and Jerry Walters. 2014. Effects of Next-Generation Vehicles on Travel Demand and Highway Capacity. *Fehr and Peers* https://orfe.princeton.edu/~alaink/Papers/FP_NextGenVehicleWhitePaper012414.pdf.

Brown, John Prather. 1973. Toward an Economic Theory of Liability. *Journal of Legal Studies* 2:323–349.

Calabresi, Guido. 1970. *The Costs of Accidents.* New Haven, Conn.: Yale University Press.

Crane, Daniel A., Kyle D. Logue; Bryce C. Pilz. 2017. A Survey of Legal Issues Arising from the Deployment of Autonomous and Connected Vehicles. *Michigan Telecommunications & Technology Law Review* 23:191–320.

Diamond, Peter A. 1974. Single Activity Accidents. *Journal of Legal Studies* 3:107–164.

Edlin, Aaron S., and Pinar Karaca-Mandic. 2006. The Accident Externality from Driving. *Journal of Political Economy* 114: 931–955.

Evas, Tatjana. 2018. A Common EU Approach to Liability Rules and Insurance for Connected and Autonomous Vehicles. *European Parliamentary Research Unit*. European Parliament, Brussels, Belgium.

Friedman, Ezra, and Eric Talley. 2019. Automatorts: How Should Accident Law Adopt to Autonomous Vehicles? Lessons from Law and Economics. Draft, April 3, 2019, available from the authors, ezra-friedman@law.northwestern.edu and etalley@law.columbia.edu.

Geistfeld, Mark A. 2017. A Roadmap for Autonomous Vehicles: State Tort Liability, Automobile Insurance, and Federal Safety Regulation. *California Law Review* 105:1611–1694.

Green, Jerry. 1976. On the Optimal Structure of Liability Laws. *The Bell Journal of Economics* 7:553–574.

International Transport Forum. 2018. Safer Roads with Automated Vehicles? *ITF/OECD*. https://www.itf-oecd.org/safer-roads-automated-vehicles-0

Landes, William M., and Richard A. Posner. 1987. *The Economic Structure of Tort Law*. Cambridge, Mass.: Harvard University Press.

Litman, Todd. 2018. *Autonomous Vehicle Implementation Predictions: Implications for Transport Planning*. Victoria Transport Policy Institute. http://vtpi.org/avip.pdf.

Logue, Kyle D. 2019. The Deterrence Case for Comprehensive Automaker Enterprise Liability. *Journal of Law and Mobility*. 2019:1-31.

Noy, Ian Y., David Shinar, and William J. Horre. 2018. Automated Driving: Safety Blind Spots. *Safety Science* 102:68–78.

Shavell, Steven. 1980. Strict Liability versus Negligence. *Journal of Legal Studies* 9:1–25.

Shavell, Steven. 1987. *Economic Analysis of Accident Law*. Cambridge, Mass.: Harvard University Press.

Singh, Santokh. 2018. Critical Reasons for Crashes Investigated in the National Motor Vehicle Crash Causation Survey. *Traffic Safety Facts Crash Stats*. Report No. DOT HS 812 506. Washington, DC: National Highway Traffic Safety Administration.

Society of Automotive Engineers. 2014. Taxonomy and Definitions for Terms Related to On-Road Motor Vehicle Automated Driving Systems. https://www.sae.org/misc/pdfs/automated_driving.pdf.

Vickrey, William. 1968. Automobile Accidents, Tort Law, Externalities, and Insurance: An Economist's Critique. *Law and Contemporary Problems* 33:464–487.

Vladeck, David C. 2014. Machines without Principals: Liability Rules and Artificial Intelligence. *Washington Law Review* 89:117–150.

WHO. 2015. Global status report on road safety 2015. World Health Organization, Geneva, Switzerland.