ISSN 1936-5349 (print)
ISSN 1936-5357 (online)

HARVARD

JOHN M. OLIN CENTER FOR LAW, ECONOMICS, AND BUSINESS

PRODUCT SAFETY, CONTRACTS, AND LIABILITY

Xinyu Hua
Kathryn E. Spier

Forthcoming in *RAND Journal of Economics*

Discussion Paper No. 1012

08/2019

Harvard Law School
Cambridge, MA 02138

Product Safety, Contracts, and Liability*

Xinyu Hua† Kathryn E. Spier‡
HKUST Harvard University

August 4, 2019

Abstract

A firm sells a dangerous product to heterogeneous consumers. Higher consumer types suffer accidents more often but may enjoy higher gross benefits. The firm invests resources to reduce the frequency of accidents. When the consumer's net benefit function (gross benefits minus expected harms) is decreasing in consumer type, the firm contractually accepts liability for accident losses and invests efficiently. When the consumer's net benefit function is increasing in consumer type, the firm contractually disclaims liability and underinvests. Legal interventions, including products liability and limits on contractual waivers and disclaimers, are necessary to raise the level of product safety.

*We would like to thank Ying Chen, Eric Claeys, James Dana, Oliver Hart, Louis Kaplow, Susan Norton, Marco Ottaviani, Ron Siegel, Holger Spamann, the Editor and anonymous referees, participants in the 7th Annual Law and Economics Theory Conference at the Rice University (2017), the Law and Economics Conference at the Hong Kong University of Science and Technology (2017), the 28th Annual Meeting of the American Law and Economics Association (2018), the REStud Tour Reunion Conference at the University of Copenhagen (2018), the Workshop on Industrial Organization and Competition Policy at UIBE (2018), and seminar audiences at George Mason University (2018), Harvard Law School (2018), Harvard Economics Department (2018), National University of Singapore (2018), Vanderbilt University (2019), UniBG (2019), and University of Mannheim (2019).

†Hong Kong University of Science and Technology. xyhua@ust.hk.
‡Harvard Law School and NBER. kspier@law.harvard.edu.

1 Introduction

■ Should the manufacturer of a dangerous product be held liable when consumers, while using the product, suffer injury or harm? Although one's intuition might suggest that firms should be held responsible — especially when the product malfunctioned in some way — the case for products liability is not as strong as it first appears. After all, sophisticated, forward-looking consumers would be willing to pay a premium for safer products that malfunction less often. In a well-functioning market, firms can turn a handsome profit by giving consumers the types of goods and services that they desire and by designing private contracts to assure performance and allocate risk. Thus, on reflection, the logical case for products liability must hinge on the failure of private markets and private contracts.

In this article, the divergence between the private and social incentives to design safer products stems from the exercise of market power. A firm with market power will make cost-justified improvements in product safety to suit the needs and preferences of the marginal consumer who is just indifferent between purchasing the good and not. Problems may arise when the marginal consumer's needs and preferences are not representative of the average buyer of the product (Spence, 1975). This is true in many economic settings involving dangerous products, and poses a particular problem for the firm and for society as a whole when consumers who place a higher gross value on the product also suffer accidents more frequently. We argue that the problem of inadequate product safety is pernicious and cannot be overcome by private contracts in a free market.

The theoretical insights of this article are empirically relevant. Many dangerous products are supplied by imperfectly-competitive markets. According to the Consumer Product Safety Commission (CSPC), all-terrain vehicles (ATVs) have been associated with approximately 100 thousand emergency-department treated injuries and more than 600 fatalities each year.[1] The ATV industry is dominated by several large competitors, with Polaris and Honda jointly accounting for more than 60% of sales.[2] According to the Centers for Disease Control and Prevention (CDC), nearly half a million people die in the United States each year from smoking-related diseases, and 16 million more suffer from smoking-related illnesses.[3] With 41% of the market, Philip Morris' Marlboro brand has more sales

[1]Many of the victims are children. See 2016 Annual Report of ATV Related Deaths and Injuries available at www.cpsc.gov. These statistics do not include dune buggies or golf carts.

[2]In 2017, worldwide market shares were Polaris (36%), Honda (28%), Yamaha (13%) and Can-Am (13%). See https://www.statista.com/statistics/438085/global-all-terrain-vehicle-market-share/.

[3]See https://www.cdc.gov/tobacco/about/osh/index.htm .

than the next eight cigarette brands combined.[4] Many medical device makers and pharmaceutical companies have considerable market power as well. Before Merck pulled Vioxx from the market in 2004, the blockbuster arthritis drug enjoyed a large market share.

We begin with a simple benchmark model where a monopolist sells a potentially dangerous product to a population of heterogeneous, risk-neutral consumers. Product safety is fully observable to consumers at the time of sale. By investing more resources, the firm can reduce the likelihood of accidents. Taking the level of sales as fixed, the socially-optimal investment in product safety would minimize the aggregate production costs plus the aggregate harms to the consumers. A consumer's type affects both the consumer's propensity to suffer harm and the gross benefit of consumption: Higher consumer types suffer accidents more frequently but may enjoy larger gross benefits. Note that a consumer's net benefit from consuming the product – the gross benefit of consumption minus the expected harms – may be either increasing or decreasing in the consumer's type.[5] The firm cannot observe consumer types and therefore cannot price discriminate directly.

If the firm bears no financial responsibility for accidents, then the private and social incentives to invest in product safety diverge. When the consumer's net benefit function is decreasing in the consumer's type, then (absent liability) the firm over-invests in product safety. The reason for this result is that the consumer type who is just indifferent between purchasing the good and not purchasing it, the "marginal consumer," is someone who places a relatively high value on product safety. The firm chooses a high level of product safety to match the needs of this marginal consumer. Conversely, when the consumer's net benefit is an increasing function of the consumer's type, then the marginal consumer places a relatively low value on product safety. In this case, the firm chooses a low level of product safety to suit the needs of the low-type consumer. Thus, absent products liability or financial responsibility for consumer harms, product safety may be either too high or too low.

Now suppose instead that the sales contract includes a stipulated-damage clause that entitles consumers to financial compensation for their accident losses. Consumers pay a (possibly higher) price up front when they purchase the product, but then receive stipulated-damage payments from the firm whenever accidents arise. Note that because high-type consumers suffer accidents more frequently than their low-type counterparts, they will collect the damage payments more

[4] https://www.statista.com/statistics/603940/market-share-leading-cigarette-brands-us/ .

[5] Many products fit this description. Consumers of durable goods such as cars, table saws, and pressure cookers often vary in their frequency of product use. Intensive users of Facebook or Amazon enjoy more benefits but are more likely to suffer harm from data leakage.

frequently too. In other words, when the firm is strictly liable for consumer harm, the effective price paid by a consumer (the upfront purchase price minus the expected future damage payments) is a decreasing function of the consumer's type. Hence, stipulated damage payments subsidize the purchases of the high-type consumers.

Would the firm want to include a stipulated-damage clause in the sales contract? Stipulated-damage payments may or may not be aligned with the firm's quest for price discrimination. When the consumer's net benefit is a decreasing function of the consumer's type, then the firm will want to include a stipulated-damage clause. Stipulated-damage payments are a way for the firm to price discriminate and give steeper discounts to the higher consumer types. Because damage payments allow the firm to extract a greater share of consumer surplus, the firm's incentive to invest in product safety are brought into alignment with the incentives of a social planner. So, when the net benefit function is decreasing in the consumer's type, private contracting solves the problem of excessive product safety. There is no need for legal intervention in this case.

When the consumer's net benefit is an increasing function of the consumer's type, the firm will not include a stipulated-damage clause in the sales contract. In this case, higher consumer types are willing to pay more for the product, not less. If the firm could perfectly price discriminate, it would charge higher prices to consumers with higher types. Stipulated-damage payments do exactly the opposite, because with damage compensation the higher-type consumers pay lower effective prices. In these settings, the firm will eschew stipulated-damage payments and other forms of contractual liability and the under-investment problem remains. The private market will produce products that are insufficiently safe, and accidents will occur too frequently.

Although our benchmark model is based on strong assumptions, its insights and implications are robust to various extensions. In Section 4 we show that the justification for legal intervention remains when product safety is not observed by consumers at the time of sale, when there is less-than-full market coverage, when the firm can sell many versions of the product with a menu of contracts, and when the consumers' net benefits are non-linear in type. As in our benchmark model, social welfare rises when products liability is mandated by law.

The insights provided by our article are relevant for public policy. In the United States, products liability law is a mixture of tort (i.e., accident) law and contract law. In practice, consumers who suffer product-related injuries may bring suit for liability in tort, breach of contract, or both.[6] According to the

[6] For example, in Buford et al. v. Toys R' Us, Inc. 217 Ga. App. 565 (1995), a child suffered serious injuries when his bicycle broke because of a defective weld. The lawsuit was brought under breach of implied warranty of merchantability and strict liability and negligence in tort.

Restatement (Third) of Torts (1998), manufacturers face strict liability for consumer harms arising from manufacturing defects, even when "all possible care was exercised in the preparation and marketing of the product."[7] Article 2 of the Uniform Commercial Code (UCC) covers the law of sales, including the obligations of manufacturers or sellers to stand behind their products and correct problems if their products fail.[8] Although the UCC allows for compensation of physical injuries and economic losses associated with a breach of warranty, private contracts limit the damage compensation to the repair and replacement of the product itself.[9] The issue of whether contractual disclaimers and other limitations of liability should be enforced has been debated in legislatures and the courts.[10]

Our article is organized as follows. Section 2 discusses the literature. Section 3 presents our benchmark model and the main results, describing the circumstances under which private contracts will fail to assure adequate safety and the optimal public policy responses. Section 4 considers several relevant extensions, including firm moral hazard, quantity distortions, versioning, and non-linear net benefits for consumers. Section 5 concludes by discussing several policy implications and directions for future research. The proofs and technical details are in the Appendix.

2 Literature

■ Our article contributes to several strands of literature. The basic distortion highlighted here – that a firm with market power may either over or under supply product safety depending on the identity of the marginal consumer – is familiar from the literature on product quality. In a classic article, Spence (1975) showed that if the marginal consumer places a lower than average value on a product's quality, then a monopolist will under-invest in quality relative to the social optimum. Spence (1975) did not consider the use of private contracts (e.g., damage payments and warranties) for price discrimination beyond a uniform price, and he did not investigate the policy interventions discussed here. In contrast, we

[7]Tort liability also arises for design defects and for failure to adequately warn consumers of product risks.

[8]The types of warranty are express warranty (§2-313), the implied warranty of merchantability (§2-314), and the implied warranty of fitness (§2-315).

[9]See UCC §2-715. Incidental damages refer to the time and effort of the injured party in dealing with the breach; consequential damages refer to injuries to people or property.

[10]The Magnuson-Moss Federal Warranty Act (1975) restricts a seller's ability to avoid warranty responsibility for consumer goods. For waivers in tort law, see Tunkl v. Regents, 383 P.2d 441 (Cal. 1963) and Dalury v. S-K-I Ltd., 670 A.2d 795 (Vt 1995).

consider the monopolist's quest for price discrimination in designing private contracts, and the need for products liability law when private contracts fail to assure adequate product safety.

Our article also contributes to the literature on products liability. Hamada (1976) was one of the first scholars to argue that products liability is unnecessary if product safety is observable to consumers at the time of sale.[11] Products liability can be socially desirable when consumers underestimate product risks (Spence, 1977; Epple and Raviv, 1978; Polinsky and Rogerson, 1983), or when product safety is not observed by consumers at the time of sale (Simon, 1981; Daughety and Reinganum, 1995, 1997, 2006, 2008a and b).[12] These articles do not consider private warranties or other private contractual mechanisms to mitigate moral hazard problems and align incentives. For the most part, the literature has evaluated products liability law under the (implicit or explicit) assumption that private contracts are incomplete.

Several articles describe economic settings where products liability outperforms private contracting. Ordover (1979) argues that products liability, by bundling insurance with the sale of a product, can mitigate the inefficiencies of insurance markets. Wickelgren (2006) argues that products liability is a valuable commitment device, preventing ex post renegotiation in settings where renegotiation will compromise ex ante incentives. See also Arlen and MacLeod (2003). Choi and Spier (2014) argue that competitive firms will set contractual damage payments too low in an attempt to "cream skim" low-risk consumers, compromising product safety.[13] Our article argues that monopolists will tend to under-invest in product safety, even when safety is observable to consumers at the time of sale and when private contractual liability is feasible. The fundamental divergence between the private and social incentives to supply safe products is not solved by the market.

The large literature on product warranties focuses primarily on the private motivations for warranties or the design of private contracts for profit maximization. Product warranties, for example, allow for more efficient risk sharing between firms and consumers (Priest, 1981).[14] Second, warranties can signal firms' pri-

[11]Others, including Polinsky and Shavell (2010), argue that market forces and regulations should suffice to assure product safety. Chen and Hua (2017), and Baker and Choi (2018) explore the interaction between liability and market forces such as reputation and competition.

[12]Chen and Hua (2012) consider the commitment effect of liability when product safety is not observed by firms at the time of sale. Hay and Spier (2005) show that products liability is necessary when victims are bystanders and injurers (the consumers) are judgment proof.

[13]Their insights build on Rothschild and Stiglitz (1976). In Choi and Spier (2014), firms are competitive and the probability of an accident is additively separable in effort, so all consumer types have the same incremental willingness-to-pay for higher safety.

[14]Che (1986) studies the use of return guarantees for risk sharing, and Boom (1988) argues

vate information about product quality (Grossman, 1981; Lutz, 1989; Moorthy and Srinivasan, 1995; Shieh, 1996). Third, warranties can address moral hazard problems and motivate firms to invest in product quality (Cooper and Ross, 1985; Emons, 1988; Mann and Wissink, 1988). Finally, warranties can screen heterogeneous consumers and facilitate price discrimination (Matthews and Moore, 1987; Padmanabhan, 1995; and Lutz and Padmanabhan, 1998).[15]

Our article differs from the literature on product warranties in several ways. First, in contrast to the prior literature, our analysis is distinctly normative. We study the divergence between the private and social incentives to stipulate damage payments and the optimal public policy response. Additionally, the literature presumes the harm level equal to the consumer benefit without heterogeneity or with only discrete types. In contrast, we consider continuous types of consumers and allows the net benefit function to be either increasing or decreasing in consumer type (depending on the firm's investment in product safety).

More broadly, our article is related to the mechanism-design literature regarding how to minimize information rents (Hansen, 1985; Cremer and McLean, 1988; Riordan and Sappington, 1988; DeMarzo, Kremer and Skrzypacz, 2005; Che and Kim, 2010). As shown in this literature, when there is an ex-post signal about an agent's private information, the optimal mechanism for rent extraction should use the signal in a way to "flatten" the sensitivity of the agent's payoff to the private information. Consistent with this insight, we show that in various scenarios, stipulated damage payments may or may not help a firm to flatten the sensitivity of consumers' net benefit to their private information. Different from this literature, we take a normative approach and explore how the incentive to reduce information rents affects the firm's safety investments and the need for products liability law.

3 The Benchmark Model

■ There is a unit mass of risk-neutral consumers, each of whom buys at most one unit of a product from a firm. The product has constant unit cost $c(\pi)$ where $\pi \in [0,1]$ affects the probability of an accident or product failure (as described below). Products with lower values of π are more expensive to produce, and we assume that $c'(\pi) < 0$, $c''(\pi) > 0$, $\lim_{\pi\to 0} c'(\pi) = -\infty$, and $\lim_{\pi\to 1} c'(\pi) = 0$ to assure an interior solution. For now, we assume that product safety π is observed

that replacement warranty is more efficient than money-back warranty for risk sharing. See also Heal (1977).

[15] Courty and Li (2000) study the usage of refund contracts for sequential screening in a context where consumers learn their types after purchasing products.

by consumers at the time of purchase. We relax this assumption in Section 4.

Consumers are heterogeneous and privately observe their types, x, distributed according to density $f(x) > 0$ on the interval $[\underline{x}, \overline{x}]$ where $1 \geq \overline{x} > \underline{x} > 0$.[16] A type x consumer enjoys a gross benefit $b_0 + bx > 0$ from product use but suffers harm $h > 0$ with probability πx where b_0, b, and h are constants. We generalize the analysis to non-linear settings in Section 4. If $b = 0$ all consumer types have the same gross benefit, b_0, but have different propensities for harm. More generally, the coefficient b may be either positive or negative, so a consumer who is more likely to suffer harm (i.e., a higher type) may enjoy higher or lower gross benefits from using the product.[17] We assume that the benefit b_0 is sufficiently high so that the firm chooses its price to sell to all consumer types $x \in [\underline{x}, \overline{x}]$ (so there are no quantity distortions).[18]

Note that the harm h could include a loss of the product's usage value (fully or partially) when the product fails to work, in addition to physical injuries and economic losses. So our model captures the firm's quality investment to ensure product functioning as well.

This general specification is aligned with a variety of economic settings. First, consumers often vary in their frequency of product use. Some buyers of table saws use their saws infrequently (only for occasional projects) whereas other buyers use their table saws daily. Similarly, some car buyers only use the car to do local errands (low intensity of use) whereas other buyers use their cars to commute long distances (high intensity consumers). High intensity consumers derive higher total benefits because they use the product more frequently, but may also suffer accidents more frequently. One can interpret x as the frequency of use, $b > 0$ as the incremental benefit per use, and πh as the expected harm per use.[19]

Although consumers who use products with greater intensity may have proportionally greater benefits and risks associated with product use, this is not always true. High intensity users may develop greater skills and expertise while using the product, thereby lowering their risk exposure. A professional chef is less likely to lose a finger when dicing an onion than a novice; an experienced hunter is less likely to shoot himself in the foot than someone who rarely uses a

[16] The normalization $\overline{x} \leq 1$ ensures the probability of consumers being harmed is no larger than 1. If $\overline{x} > 1$, the results remain the same as long as $\pi\overline{x} \leq 1$.

[17] Alternatively – and equivalently – one could assume that consumers are indexed by their gross benefits of product use, v, and the probability of harm takes the form $\alpha_0 + \alpha v$ where α may be either positive or negative.

[18] Our assumption that b_0 is sufficiently high and $f(x) > 0$ for $x \in [\underline{x}, \overline{x}]$ implies full market coverage. Our main insights do not hinge on full market coverage, however. See section 4.

[19] The frequency of use x is the consumer's type, not a choice variable. Our results hold qualitatively if consumers have different and private values of the coefficient b and they choose the level of product use x. The analysis is available upon request.

gun. Our framework accommodates such settings. Letting $b < 0$, consumers who experience higher expected harms have lower gross benefits.[20]

Our model also captures settings where consumers are privately informed about the likelihood of needing the product in the future. For example, when buying their very first car, a young couple may not know for sure whether a child will later occupy the back seat. Formally, suppose there are two states of the world, low and high, and consumers privately observe the probability $x \in [0, 1]$ of the high state. When a consumer's needs are in the low state, the consumer derives a benefit b_0 and suffers no harm; when the consumer's needs are in the high state, the consumer derives a higher benefit $b_0 + b$ and suffers expected harm πh.

Finally, our model applies to settings with just two consumer types, $x \in \{\underline{x}, \overline{x}\}$, where the high-harm type $\overline{x}$ enjoys a gross benefit $\overline{b}$ and the low-harm type $\underline{x}$ enjoys a gross benefit $\underline{b}$.[21] With just two types, the density function has two mass points where a proportion θ of the consumer population has type $\overline{x}$ and a proportion $1 - \theta$ has type $\underline{x}$. This two-type model can of course be easily extended to a continuous setting where consumers have private information about θ which is drawn from a density on $[0, 1]$. Then, using our earlier notation, the consumer's type is $x = (1 - \theta)\underline{x} + \theta\overline{x}$ and we are back to our original framework.[22]

We allow the firm to offer the product for sale with a menu of contracts. Using the revelation principle, we consider direct revelation mechanisms of the form $(\pi, \rho(x), \omega(x))$ where π is the product's safety, $\rho(x)$ is the price, and $\omega(x)$ is the stipulated-damage payment. A consumer of type x pays $\rho(x)$ at the time of purchase and receives a damage payment of $\omega(x) \in [0, h]$ whenever he or she suffers harm h.[23] The "effective price" paid by a consumer of type x is therefore $\rho(x) - \pi x \omega(x)$. Note that we are assuming that the monopolist produces a single standard version of the product rather than a menu of specialized versions, an assumption that we relax in section 4.[24] Menu costs and other fixed costs of production may limit the variety of products sold.

[20] When $b < 0$, type $\underline{x}$ has higher gross benefits and suffers smaller accident losses than type $\overline{x}$. So the professional is type $\underline{x}$ and the novice is type $\overline{x}$.

[21] Just as above, the harm suffered by type x is $\pi h x$ and the gross benefit enjoyed by type x is $b_0 + bx$ where the coefficients are given by $b_0 = \frac{\overline{x}\underline{b} - \underline{x}\overline{b}}{\overline{x} - \underline{x}}$ and $b = \frac{\overline{b} - \underline{b}}{\overline{x} - \underline{x}}$. It is straightforward to verify that $b_0 + b\overline{x} = \overline{b}$ and $b_0 + b\underline{x} = \underline{b}$.

[22] Suppose that $H(\theta)$ is the cumulative density of θ where $H(0) = 0$ and $H(1) = 1$. Then $F(x) = H(\frac{x - \underline{x}}{\overline{x} - \underline{x}})$ is the cumulative density of x.

[23] If consumers are risk averse, then $\omega(x) \in [0, h]$ has risk-sharing benefits, too. However, given the firm's quest for price discrimination, our insights about the firm's suboptimal incentive to offer damage payments and to invest in product safety would still be valid.

[24] In practice firms can and do sell different versions of products with different safety and quality levels to appeal to different market segments.

Our assumption that the stipulated-damage payment $\omega(x) \in [0, h]$ is empirically relevant. In practice, supra-compensatory damage payments with $\omega(x) > h$ might induce consumers to take actions to raise accident probabilities,[25] and negative damage payments with $\omega(x) < 0$ would create an incentive for consumers to hide accidents from the firm. Formally, we assume that the accident state is contractible but can be hidden by consumers, which allows for payments to consumers when accidents occur but precludes net rebates to consumers when accidents do not occur.[26]

Note that our core analysis assumes that private contracts are fully enforceable: consumers who suffer harm are not entitled to compensation beyond what is stipulated in the contract. In other words, firms are free to waive or disclaim financial responsibility for consumer harms. In the real world, not all contracts are fully enforceable: consumers who suffer harm may be legally entitled to damage compensation, regardless of what is stipulated in the contract. We will also explore settings where a benevolent social planner may regulate the damage compensation, thus overriding the level stipulated in the contract.[27]

Finally, we establish a social-welfare benchmark. Because it is socially efficient for all consumer types to purchase the product, social welfare is given by

$$W(\pi) = \int_{\underline{x}}^{\overline{x}} \left[b_0 + bx - \pi hx - c(\pi) \right] f(x) dx. \tag{1}$$

Differentiating, the socially-optimal safety level, π^{**}, is the implicit solution to

$$W'(\pi^{**}) = -hE(x) - c'(\pi^{**}) = 0. \tag{2}$$

where $E(x)$ is the average value of x in the consumer population. Given our earlier assumptions, $\pi^{**} \in (0, 1)$ is uniquely defined.[28] We let $b^{**} = \pi^{**} h$.

□ **Preliminary Analysis.** The firm chooses a direct revelation mechanism

[25] In our model, the firm would choose $\omega(x) > h$ and earn higher profits when $b < 0$, but the firm's choice of product safety would not change. Our assumption is consistent with the courts' reluctance to enforce stipulated damages that exceed actual harm (the "penalty doctrine").

[26] Consider an extended mechanism $(\pi, \rho(x), \omega(x), r(x))$ where $r(x)$ is a "rebate" when accidents do not occur. This is equivalent to a mechanism without a rebate: $\widetilde{\rho}(x) = \rho(x) - r(x)$, $\widetilde{\omega}(x) = \omega(x) - r(x)$, and $\widetilde{r}(x) = 0$. Formally, we assume that the net rebate $r(x) - \omega(x) \leq 0$.

[27] Although we model legal intervention as a regulatory regime, note that regulations may be enforced through private litigation. A consumer who suffers harm may file suit to collect any residual damage compensation that they are entitled to by law.

[28] π^{**} is the first-best level of safety for the average consumer in the population. In a first-best world with different versions of the product, a consumer with $x < E(x)$ ($x > E(x)$) would buy a product with lower (higher) safety features, $\pi > \pi^{**}$ ($\pi < \pi^{**}$).

$(\pi, \rho(x), \omega(x))$ to maximize its profits

$$\int_{\underline{x}}^{\overline{x}} [\rho(x) - \pi\omega(x)x - c(\pi)] f(x)dx \tag{3}$$

subject to

$$\rho(x) - \pi\omega(x)x \leq \rho(\widehat{x}) - \pi\omega(\widehat{x})x \quad \forall x, \widehat{x} \tag{4}$$

$$\rho(x) - \pi\omega(x)x \leq b_0 + bx - \pi hx \quad \forall x \tag{5}$$

$$\omega(x) \in [0, h] \quad \forall x, \tag{6}$$

where $\rho(x) - \pi\omega(x)x$ is the effective price paid by a type-x consumer, (4) is the incentive compatibility (IC) constraint, and (5) is the individual rationality (IR) constraint.

We can rewrite the IC constraint in (4) as

$$\pi[\omega(x) - \omega(\widehat{x})]\widehat{x} \leq \rho(x) - \rho(\widehat{x}) \leq \pi[\omega(x) - \omega(\widehat{x})]x \quad \forall x, \widehat{x}. \tag{7}$$

This inequality implies that $\omega'(x) \geq 0$ and $\rho'(x) \geq 0$.[29] Higher consumer types receive (weakly) higher stipulated-damage payments and pay (weakly) higher prices. Using standard techniques,[30] one can show that the effective price paid by a type-x consumer, $\rho(x) - \pi\omega(x)x$, is:

$$\rho(\underline{x}) - \pi\omega(\underline{x})\underline{x} - \pi \int_{\underline{x}}^{x} \omega(y)dy. \tag{8}$$

Taking the derivative of this expression reveals that the effective price is weakly decreasing and concave in the consumer's type, x. First, the slope of (8), $-\pi\omega(x)$, is negative because by assumption $\omega(x) \geq 0$. Intuitively, because higher consumer types experience accidents more frequently, and collect a damage payment from the firm every time an accident occurs, higher consumer types pay lower effective prices. The second derivative of (8), $-\pi\omega'(x)$, is negative as well because as shown above $\omega'(x) \geq 0$.

We now present an important result. Lemma 1 tells us that we may restrict attention to simple contracts (π, p, w) that specify a single up-front price, p, and a single stipulated-damage payment to be paid in the event of an accident, w. A proof may be found in the Appendix.

Lemma 1. *Consider any individually rational and incentive compatible direct-revelation mechanism, $(\pi, \rho(x), \omega(x))$. There exists a simple contract (π, p, w) that gives the firm weakly higher profits.*

[29]Suppose $x > \widehat{x} > 0$. Because $\pi[\omega(x) - \omega(\widehat{x})]\widehat{x} \leq \pi[\omega(x) - \omega(\widehat{x})]x$ we have $\omega(x) - \omega(\widehat{x}) \geq 0$, and also that $\rho(x) - \rho(\widehat{x}) \geq 0$.

[30]See the proof of Lemma 1 in the Appendix.

This result follows from the individual rationality and incentive compatibility constraints. Recall that the IR constraint in (5) implies that a type x consumer is willing to purchase the product when the effective price, $\rho(x)-\pi\omega(x)x$, is less than or equal to the benefit of consumption, $b_0+bx-\pi hx$. Notice that the consumer's benefit of consumption is a linear function of x. However, as shown above, the IC constraint in (4) implies that the effective price is a weakly decreasing and concave function of x. If the firm chooses a simple contract, (π, p, w), then the effective price, $p-\pi wx$, is a linear function of x as well. Thus, the simple contract is the most effective way for the firm to price discriminate.

With a simple contract, (π, p, w), the firm's cost of supplying one unit of the product to a consumer of type x is

$$\pi wx + c(\pi). \tag{9}$$

Note that this is an increasing function of x. Consumers with higher types x experience accidents more frequently and are therefore more expensive for the firm to serve. A consumer's net benefit from the contract also depends on x,

$$b_0 + bx - \pi(h-w)x - p. \tag{10}$$

Importantly, the consumer's net benefit may be either increasing or decreasing in the consumer's type, x, depending on the values b, h, and w. Define $\widehat{\pi}(w)$ to be the safety level for which the consumer's net benefit or surplus in (10) is independent of the consumer's type:

$$\widehat{\pi}(w) = \frac{b}{h-w}. \tag{11}$$

When $b > 0$, for all $w \in [0, h)$ we have $\widehat{\pi}'(w) > 0$, $\widehat{\pi}''(w) > 0$, and $\lim_{w\to h}\widehat{\pi}(w) = \infty$. When $b < 0$, $\widehat{\pi}(w) < 0$ for all $w \in [0, h)$.

When maximizing its profits, the monopolist sets the price to extract all surplus from the marginal consumer. With full market coverage, the marginal consumer may be the consumer with $x = \underline{x}$ or with $x = \overline{x}$, depending on whether the consumer's net benefits in (10) are increasing or decreasing in x.

Suppose that $\pi < \widehat{\pi}(w)$ where $\widehat{\pi}$ is defined in (11) so the product is not too risky.[31] In this case, $b - \pi(h-w) > 0$ so the consumer's net benefit in (10) is an increasing function of x. If a consumer of type x' chooses to buy the product, then all consumers with $x > x'$ will buy the product as well. So, with full market coverage, type $\underline{x}$ is the marginal consumer and the monopoly price would be $p = b_0 + b\underline{x} - \pi(h-w)\underline{x}$. Using (10), an infra-marginal consumer with type $x > \underline{x}$

[31] For all $w \in [0, h]$, $\pi < \widehat{\pi}(w)$ implies that $b > 0$.

will receive rents of $(x - \underline{x})(b - \pi(h - w)) > 0$. The firm's profit function when $\pi < \widehat{\pi}(w)$ is

$$\underline{S}(\pi, w) = W(\pi) - (E(x) - \underline{x})(b - \pi(h - w)). \tag{12}$$

In other words, the firm's profits are equal to social welfare $W(\pi)$ minus the total rents that are paid to consumers, $(E(x) - \underline{x})(b - \pi(h - w))$.[32] Note that the $\underline{S}(\pi, w)$ is a well-defined function for all values π and w, although it represents the firm's actual profit function only when $\pi < \widehat{\pi}(w)$.

Holding w fixed, the safety level that maximizes the function $\underline{S}(\pi, w)$ in (12), $\pi^H(w)$, is the implicit solution to

$$W'(\pi^H(w)) + (E(x) - \underline{x})(h - w) = 0. \tag{13}$$

This expression reveals some important properties. First, when the damage payment is fully compensatory, $w = h$, then the second term in (13) disappears. The safety level that maximizes firm profits in $\underline{S}(\pi, w)$ is the same one that maximizes social welfare, $\pi^H(h) = \pi^{**}$. When $w < h$, then the second term in (13) is strictly positive and so the firm would under-invest in product safety, $\pi^H(w) > \pi^{**}$.[33]

Now suppose instead that $\pi > \widehat{\pi}(w)$ so the product is relatively risky.[34] In this case, $b - \pi(h - w) < 0$ and so the consumer's net benefit in (10) is a decreasing function of x. Type $\overline{x}$ is the marginal consumer, and so the monopoly price is $p = b_0 + b\overline{x} - \pi(h - w)\overline{x}$. Note that an infra-marginal consumer with type $x < \overline{x}$ earns rents of $(x - \overline{x})(b - \pi(h - w)) > 0$, so the firm's profit function when $\pi > \widehat{\pi}(w)$ is:

$$\overline{S}(\pi, w) = W(\pi) - (E(x) - \overline{x})(b - \pi(h - w)). \tag{14}$$

Holding the damage payment w fixed, the safety level that maximizes the function $\overline{S}(\pi, w)$ in (14), $\pi^L(w)$, is the implicit solution to

$$W'(\pi^L(w)) + (E(x) - \overline{x})(h - w) = 0. \tag{15}$$

As before, when $w = h$ the second term in (15) disappears and so $\pi^L(w) = \pi^{**}$.[35] When $w < h$, then $(E(x) - \overline{x})(h - w) < 0$ and so $\pi^L(w) < \pi^{**}$. In this case, the

[32] These rents are of course non-negative because $b - \pi(h - w) > 0$ when $\pi < \widehat{\pi}(w)$. The marginal consumer with type $x = \underline{x}$ is the most profitable consumer for the firm (as this consumer is least likely to have accidents).

[33] Note that $\pi^H(w)$ does not depend on b, the consumer's incremental benefit of product use. And if punitive damage payments are feasible, then for any $w > h$ the firm would over-invest in product safety, $\pi^H(w) < \pi^{**}$.

[34] Note that, when $b < 0$, for all $w \in [0, h]$, we have $\pi \geq 0 > \widehat{\pi}(w)$.

[35] Note however that when $b \geq 0$ and $w = h$ then $b - \pi(h - w) \geq 0$ and so $\pi^L(w) < \widehat{\pi}(w)$. Although the function $\overline{S}(\pi, w)$ is well defined, it does not correspond with firm profits in this case.

firm over-invests in product safety.[36]

The properties of the functions $\pi^H(w)$ and $\pi^L(w)$ defined in (13) and (15) are summarized in the following lemma.

Lemma 2. *When $w \in [0, h)$, then $d\pi^L(w)/dw > 0$, $d\pi^H(w)/dw < 0$ and $0 < \pi^L(w) < \pi^{**} < \pi^H(w) < 1$. When $w = h$ then $\pi^L(h) = \pi^{**} = \pi^H(h)$.*

This lemma foreshadows the result that if the manufacturer offers a contract that is less-than-fully compensatory, $w < h$, then the safety level of the product may not be socially desirable. When the consumer with type $\underline{x}$ is the marginal consumer, and types $x > \underline{x}$ earn positive rents, product safety will be substandard. When the consumer with type $\overline{x}$ is the marginal consumer and types $x < \overline{x}$ earn rents, product safety will be excessive. This result is aligned with the analysis of Spence (1975) who shows that the quality choice of a monopolist will not reflect the needs of the infra-marginal consumers.

□ **Results.** We will now present a series of results. Before doing so, it is useful to rewrite the firm's profit function as

$$S(\pi, w) = W(\pi) - (E(x) - x^M)[b - \pi(h - w)] \tag{16}$$

where $x^M \in \{\overline{x}, \underline{x}\}$ is the marginal consumer. In other words, firm profits are equal to social welfare minus consumer surplus. As shown above, $x^M = \underline{x}$ is the marginal consumer when $b - \pi(h - w) > 0$, and $x^M = \overline{x}$ is the marginal consumer when $b - \pi(h - w) < 0$.[37]

The next proposition establishes that, in the absence of stipulated-damage payments and other types of liability, the firm may either over-invest or under-invest in product safety. This is a benchmark result, but would be empirically relevant in legal and institutional settings where the transaction costs of maintaining stipulated-damage payments or liability systems are prohibitive.

Proposition 1. *Suppose that there is no stipulated-damage payment, $w = 0$. Define b^{**} to be such that $b^{**} = \pi^{**}h$ where π^{**} is the socially-optimal safety level. If $b < b^{**}$ then the firm over-invests in product safety, $\pi^*(0) = \max\{\pi^L(0), \widehat{\pi}(0)\} < \pi^{**}$. If $b = b^{**}$ the firm invests efficiently in product safety, $\pi^*(0) = \pi^{**}$. If $b > b^{**}$ then the firm under-invests in product safety, $\pi^*(0) = \min\{\pi^H(0), \widehat{\pi}(0)\} > \pi^{**}$.*

[36] If punitive damage payments are feasible, for any $w > h$ the firm would under-invest in product safety, $\pi^L(w) > \pi^{**}$.

[37] In the former case, $S(\pi, w)$ is equivalent to $\underline{S}(\pi, w)$ in (12) and in the latter case it is equivalent to $\overline{S}(\pi, w)$ in (14).

This result can be seen by examining the firm's profit function in (16). When $w = 0$ then the profit function becomes

$$S(\pi, 0) = W(\pi) - (E(x) - x^M)(b - \pi h). \tag{17}$$

Suppose by chance that $b = \pi^{**}h$ in (17). This is a knife-edged situation, and will serve as a benchmark. If the firm invests efficiently, so $\pi = \pi^{**}$, then the second term disappears and $S(\pi^{**}, 0) = W(\pi^{**})$. In this knife-edged case, the firm is able to perfectly price discriminate when it chooses the socially-efficient safety level. The consumers receive no rents at all, and the firm captures the entire maximized social value. Thus, in this benchmark, private and social incentives are aligned.

Next, suppose that $b > \pi^{**}h$ in (17). If the firm invests efficiently, $\pi = \pi^{**}$, then the second term $b - \pi^{**}h > 0$ and $x^M = \underline{x}$ would be the marginal type. The firm's profits would be $S(\pi^{**}, 0) = W(\pi^{**}) - (E(x) - \underline{x})(b - \pi^{**}h)$. The firm has an incentive to raise π above π^{**}. A small increase in π has a second-order effect on $W(\pi)$ (as social welfare is maximized at π^{**}), but has a first-order effect on the rents that are paid to consumers. Therefore when $w = 0$ and $b > \pi^{**}h$, the firm will under-invest in product safety.

Conversely, if $b < \pi^{**}h$ and the firm invests efficiently, $\pi = \pi^{**}$, then $b - \pi^{**}h < 0$ and so $x^M = \overline{x}$. The firm's profits would be $S(\pi^{**}, 0) = W(\pi^{**}) - (E(x) - \overline{x})(b - \pi^{**}h)$. A small reduction in π would have a second order effect on $W(\pi)$ but would reduce the consumer rents in a first-order sense. Therefore when $w = 0$ and $b < \pi^{**}h$, the firm will over-invest in product safety.

A higher benefit, b, does not necessarily increase the consumer rent or consumer surplus. If $b < b^L = \pi^L(0)h$, then $x^M = \overline{x}$ is the marginal consumer and $\pi^*(0) = \pi^L(0)$. In this case, the consumer rent, $(\overline{x} - E(x))(\pi^L(0)h - b)$, decreases in b. If $b > b^H = \pi^H(0)h$, then $x^M = \underline{x}$ is the marginal consumer and $\pi^*(0) = \pi^H(0)$. Then the consumer surplus, $(E(x) - \underline{x})(b - \pi^H(0)h)$, increases in b. If $b \in [b^L, b^H]$, then $\pi^*(0) = \widehat{\pi}(0)$, so that consumers receive no rent (perfect price discrimination).

Corollary 1. *Suppose that there is no stipulated-damage payment, $w = 0$. Then the consumer rent (or consumer surplus) decreases in b for $b < b^L$, equals zero for all $b \in [b^L, b^H]$, and increases in b for $b > b^H$.*

The results in Proposition 1 are illustrated in the following numerical example.

Numerical Example 1: *Assume that x follows the uniform distribution on $[1/2, 1]$ and $c(\pi) = a\pi + \frac{a}{\pi}$ for $\pi \in (0, 1]$ and normalize $h = 1$. Then the socially-optimal safety level is $\pi^{**} = (\frac{4a}{4a+3})^{1/2}$. It can also be verified that without any stipulated-damage payment, $w = 0$, the firm may over-invest or under-invest, in particular, $\pi^*(0) = \max\{(\frac{a}{a+1})^{1/2}, b\} < \pi^{**}$ for $b < \pi^{**}$, and $\pi^*(0) =$*

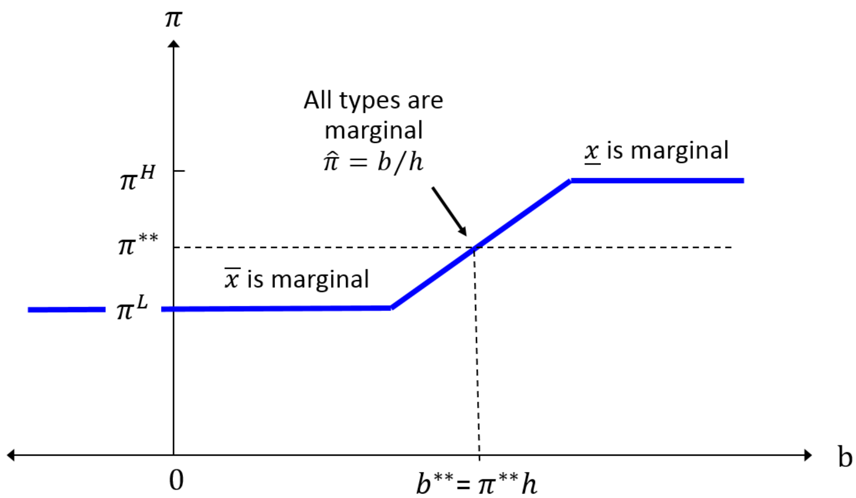

Figure 1: Product Safety Without Stipulated Damages ($w = 0$)

$\min\{b, (\frac{2a}{2a+1})^{1/2}\} > \pi^{**}$ *for* $b > \pi^{**}$. *Figure 1 depicts the relationship between* $\pi^*(0)$ *and* b*; Figure 2 shows how* b *affects the consumer rent given* $w = 0$.

We now allow both π and w to be choice variables for the firm. The firm obviously has an incentive to engage in price discrimination to extract the consumers' information rents. As shown in the mechanism-design literature, the optimal mechanism "flattens" the sensitivity of the agent's payoff to the agent's private information.[38] In our context, the firm has two instruments (product safety π and damage payment w) to flatten the sensitivity of consumers' net benefit with respect to type x. Distorting product safety π away from the socially-efficient level π^{**} reduces the total rents that are available for extraction, so the stipulated-damage payment w is a "less costly" instrument.

The next proposition states that when $b > 0$, a profit-maximizing firm will not choose to offer the fully-compensatory damage payment. With a lower stipulated-damage payment, $w < h$, the firm can better extract the rents that would otherwise be paid to the infra-marginal consumers. Moreover, the level of product safety will never be excessive, and may be insufficient from a social-welfare perspective.

Proposition 2. *If* $b \leq b^{**} = \pi^{**}h$ *then the firm chooses a stipulated-damage payment* $w^* = \min\{h - b/\pi^{**}, h\}$ *and invests efficiently in product safety. If* $b > b^{**}$ *then the firm chooses* $w^* = 0$ *and under-invests in product safety.*

[38]See in particular Hansen (1985) Cremer and McLean (1988), Riordan and Sappington (1988), DeMarzo, Kremer and Skrzypacz (2005), and Che and Kim (2010).

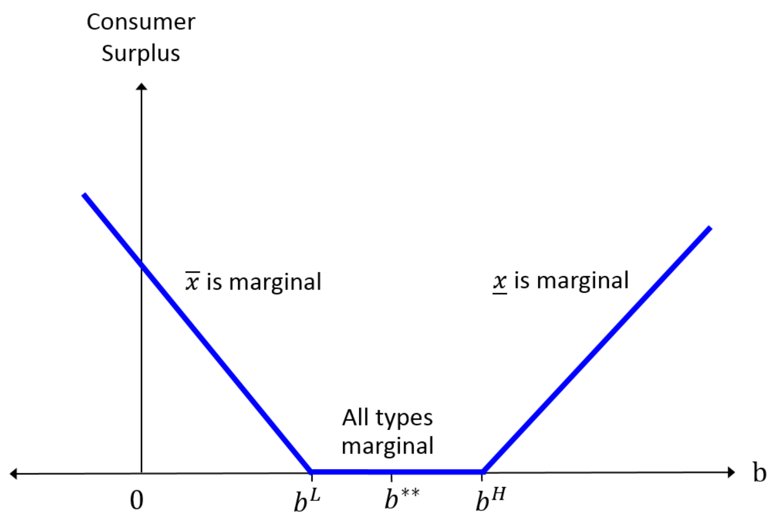

Figure 2: Consumer Rent Without Stipulated Damages ($w = 0$)

Suppose first that $0 < b \leq \pi^{**}h$. As shown earlier, if $w = 0$ then the firm over-invests in safety and the consumer's net benefit is decreasing in type x. A small increase in w (i.e., a positive stipulated-damage payment) offers steeper discounts to the higher consumer types and allows the firm to price discriminate (i.e. making consumers' net benefits less sensitive to their private types). Accordingly, the firm's incentives to invest in product safety are more aligned with the incentives of a social planner. In particular, suppose that the firm chooses $w^* = h - b/\pi^{**} \in [0, h)$. Plugging this into the firm's profit function in (16),

$$S(\pi, w^*) = W(\pi) - (E(x) - x^M)b[1 - \pi/\pi^{**}]. \tag{18}$$

If the firm invests efficiently and chooses $\pi = \pi^{**}$, then the second term disappears and the firm captures the entire maximized social welfare, $S(\pi^{**}, d^*) = W(\pi^{**})$. When $\pi = \pi^{**}$ and $w^* = h - b/\pi^{**} \in [0, h)$ then the consumer's net benefit in (10) does not depend on the consumer's type at all. The firm is engaging in perfect price discrimination, and extracting the full surplus from all consumer types.[39] In this way, the social optimum is obtained.

Note that when $b \leq \pi^{**}h$, the firm can use a partially-compensatory damage payment to fully extract consumer rents. The firm's optimal damage payment, $w^* = \min\{h - b/\pi^{**}, h\}$, is weakly decreasing in the consumer's incremental

[39]This result holds in non-linear settings, too. See Section 4. In particular, when b is sufficiently small, and the consumer's net benefit function is concave, the firm can implement perfect price discrimination by offering a menu of partially-compensatory damage payments.

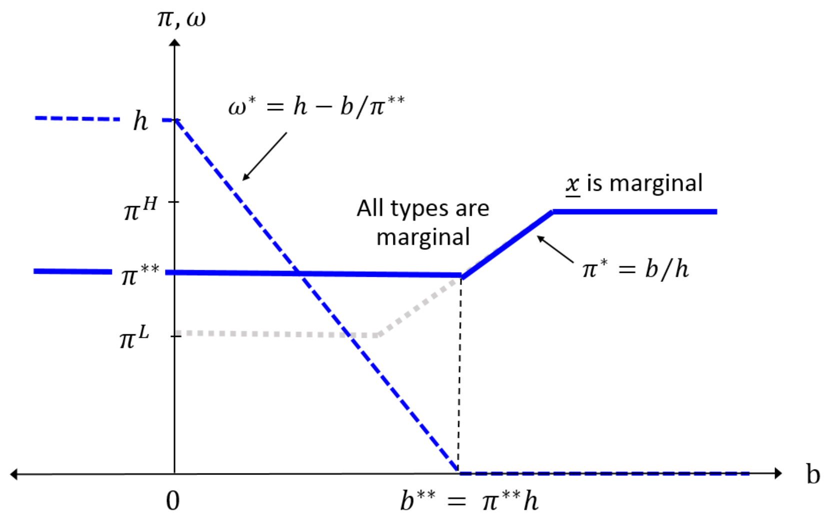

Figure 3: Firm's Optimal Stipulated Damages and Product Safety

gross benefit, b: An increase in the benefit b enlarges consumer heterogeneity; accordingly the firm reacts to reduce the damage payment to offset the increased heterogeneity.

Next, imagine that $b > \pi^{**}h$. If the firm could use a negative stipulated-damage payment, $w = h - b/\pi^{**} < 0$, then the firm would have an incentive to choose $\pi = \pi^{**}$ just as it did above. Negative stipulated-damage payments are not possible, however. As argued earlier, payments from consumers to the firm after accidents occur would encourage consumers to hide their accidents. When $b > \pi^{**}h$, the best the firm can do is to choose $w^* = 0$.[40] Offering a positive damage payment would be counter-productive for rent extraction, because the damage payment would be subsidizing the infra-marginal consumers with types $x > \underline{x}$ and making consumers' net benefits more sensitive to their types. Then, as in Proposition 1, the firm will under-invest in product safety, which reduces social welfare but helps the firm extract more rents from consumers.

Finally, suppose that $b < 0$. A positive stipulated-damage payment, $w > 0$, subsidizes the infra-marginal consumers with types $x < \overline{x}$, which allows the firm to extract more rent from consumers. Therefore, the firm will offer the fully-compensatory damage payment $w = h$. According to Proposition 2, the firm will invest efficiently and choose $\pi = \pi^{**}$. In this case, with $w = h$, consumers still get positive rents. In fact, if punitive damage payments are feasible, the firm would

[40]Section 4 shows that even if the firm does not sell to all consumers, this result holds when b is sufficiently large.

offer $w = h - b/\pi^{**} > h$ and still choose $\pi = \pi^{**}$, extracting the full surplus from all consumer types. Note that, given $b < 0$, the firm always invests efficiently no matter whether punitive damages are feasible or not.[41]

Note that the firm would never choose a stipulated-damage payment that induces over-investment in safety. If it did, then the consumer's net benefit in (10) would be strictly decreasing in the consumer's type. The stipulated-damage payment is a valuable mechanism for price discrimination here. Increasing the damage payment marginally allows the firm to give a larger discount to consumers with high types while maintaining higher prices for consumers with low types.

Numerical Example 1 (Continued): *Assume that x follows the uniform distribution on $[1/2, 1]$ and $c(\pi) = a\pi + \frac{a}{\pi}$ for $\pi \in (0, 1]$. Normalize $h = 1$. When $b \leq \pi^{**} = (\frac{4a}{4a+3})^{1/2}$, the firm offers $w^* = min\{1 - b/(\frac{4a}{4a+3})^{1/2}, 1\}$ and invests efficiently; when $b > \pi^{**}$, the firm chooses $w^* = 0$ and $\pi^*(0) = \min\{b, (\frac{2a}{2a+1})^{1/2}\} > \pi^{**}$. Figure 3 shows the firm's optimal choices of the stipulated-damage payment and the safety investment.*

The next proposition follows immediately from the analysis above.

Proposition 3. *Legal intervention is necessary to raise the level of product safety if and only if $b > b^{**} = \pi^{**}h$. If the social planner mandates the fully-compensatory damage payment, $w^{**} = h$, then the firm chooses the socially-optimal safety level, $\pi = \pi^{**}$.*

When $b \leq b^{**}$, legal intervention is unnecessary because the firm has a private incentive to stipulate damage compensation in the contract and choose the efficient level of product safety (see Proposition 2). In contrast, when $b > b^{**}$, the firm would disclaim liability for consumer harms ($w^* = 0$) and under-invest in product safety. In this case, legal interventions, which may include products liability and/or limits on contractual waivers and disclaimers, can help to raise the level of product safety. In particular, a legal requirement that forces the firm to make consumers whole (that is, $w^{**} = h$) after suffering accident-related harms aligns the firm's private incentive with the social incentive to invest in product safety and maximizes social welfare.[42]

[41] Punitive damages would affect the distribution of rents between the firm and consumers, which might impact the firm's incentives to invent new products.

[42] From equation (16), the firm's profits are $S(\pi, h) = W(\pi) - (E(x) - x^M)b$. So the firm is the residual claimant of the incremental social benefit associated with product safety choice, π. There may be other mechanisms that implement the social optimum, too.

4 Extensions

□ **Firm Moral Hazard.** Suppose that the quality of the product is not observable to consumers at the time of purchase. We consider the following timing. The firm offers the product for sale with a contract (p, w).[43] Consumers form expectations about the quality of the product, and place their orders. Finally, the firm chooses the quality, π, and delivers the product. We assume that b_0 is sufficiently large so there is full market coverage.

The firm will choose π to minimize its average unit costs. The value chosen $\widetilde{\pi}(w)$ is a function of the damage payment and is the implicit solution to:

$$-wE(x) - c'(\widetilde{\pi}(w)) = 0. \tag{19}$$

Note that $\widetilde{\pi}(w)$ is a strictly decreasing function of w and that $\widetilde{\pi}(h) = \pi^{**}$ so social efficiency is achieved if and only if $w = h$.

Proposition 4. *(Firm Moral Hazard.) If $b \leq 0$, then the firm chooses $w^* = h$ and invests efficiently in product safety, $\widetilde{\pi}(w^*) = \pi^{**}$. If $b > 0$, then the firm chooses $w^* \in [0, h)$ and under-invests in product safety, $\widetilde{\pi}(w^*) > \pi^{**}$. Furthermore, when $b < \widetilde{\pi}(0)h$, then $w^* \in (0, h)$. If the social planner mandates the fully-compensatory damage payment, $w^{**} = h$, then the firm chooses the socially-optimal safety level, $\widetilde{\pi}(h) = \pi^{**}$.*

As in our benchmark model, social welfare is maximized when the firm is required to fully compensate consumers who suffer accident-related harms. The firm would never voluntarily choose the fully-compensatory damage payment $w = h$ when $b > 0$. When $w = h$, the consumer's net benefit in (10) is increasing in the consumer's type x. So, the firm is leaving consumer surplus on the table for all $x > \underline{x}$. The firm will choose a lower payment level, $w < h$, to achieve better price discrimination. The benefit of better price discrimination has a first-order effect on firm profits whereas the loss in profits due to reduced firm incentives is a second-order effect.

□ **Quantity Distortions.** In the benchmark model, we made the simplifying assumption that the firm sells its product to all consumer types. Our main result about the firm's suboptimal incentives to stipulate damage payments and to make safer products remains valid even without full market coverage.[44]

[43] For simplicity, we restrict out attention to the simple contract (p, w), which may not maximize firm profits when moral hazard problems exist. Even if the firm can use a menu of contracts, when $b > 0$, it would not offer $w = h$ and therefore would under-invest in safety.

[44] Similar to the analysis in the baseline model, it can be shown that the optimal contract takes a simple form (π, p, w) even if there is quantity distortion.

Denote x^M as the marginal consumer who buys the product. The firm's optimal price is $p = b_0 + [b - \pi(h-w)]x^M$. For simplicity, assume in this section that the consumer's incremental benefit b is sufficiently large, so that the consumer's net benefit will be increasing in x for all parameter values and consumers with $x \geq x^M$ will buy the product.[45] Then the firm's profit function is:

$$S(w, x^M, \pi) = \int_{x^M} \left\{ b_0 + [b - \pi(h-w)]\, x^M - c(\pi) - \pi w x \right\} dF(x). \tag{20}$$

Note that the firm's profit function is decreasing in w. Therefore, the firm offers $w^* = 0$. Given any $w \geq 0$, the firm's optimal choices of safety and marginal consumer, $\pi^* = \pi^*(w)$ and $x^{M*} = x^{M*}(w)$, satisfy:

$$\left[-hx^{M*} - c'(\pi^*)\right]\left[1 - F(x^{M*})\right] - w\int_{x^{M*}} \left(x - x^{M*}\right) dF(x) = 0, \tag{21}$$

$$-\left[b_0 + (b - \pi^* h)x^{M*} - c(\pi^*)\right] f(x^{M*}) + [b - \pi^*(h-w)]\left[1 - F(x^{M*})\right] = 0. \tag{22}$$

In this section, assume that b_0 is not too large so that the market is not fully covered (i.e., $x^{M*} > \underline{x}$).[46] We also assume that the second-order conditions hold. Then, as we show in the Appendix, conditions (21) and (22) imply that $\frac{d\pi^*(w)}{dw} < 0$ and $\frac{d\widehat{x}^*(w)}{dw} > 0$. An marginal increase in the stipulated-damage payment increase product safety but reduces the total output.

Given $\pi^* = \pi^*(w)$ and $x^{M*} = x^{M*}(w)$, the social welfare function is

$$W(w) = \int_{x^{M*}} \left\{ b_0 + [b - \pi^* h]\, x - c(\pi^*) \right\} dF(x). \tag{23}$$

We show in the Appendix that $\frac{dW(w)}{dw}\mid_{w=h} < 0$. That is, the social planner always sets $w^{**} < h$, whereas the firm does not stipulate any damage payment and may under-invest in safety if $w^{**} > 0$.

The following proposition summarizes these results.

Proposition 5. *(Quantity Distortions.) Suppose that b_0 is not too large and b is sufficiently high.*[47] *The firm chooses $w^* = 0$ and (weakly) under-invests in safety. The first-best outcome cannot be obtained. With the second-best policy, the social planner mandates a less-than-fully-compensatory damage payment, $w^{**} \in [0, h)$, and the firm chooses a (weakly) higher safety level.*

[45] More formally, we assume $b > \pi^H(0, \underline{x})h$, where $-h\underline{x} - c'(\pi^H(0, \underline{x})) = 0$.

[46] This is true when $b_0 < \overline{b_0}$, where $-\left[\overline{b_0} + (b - \pi^H(0, \underline{x})h)\underline{x} - c(\pi^H(0, \underline{x}))\right] f(\underline{x}) + \left[b - \pi^H(0, \underline{x})h\right]\left[1 - F(\underline{x})\right] = 0$.

[47] More formally, $b > \pi^H(0, \underline{x})h$.

Proposition 5 establishes that a less-than-fully-compensatory damage payment ($w < h$) is better than the fully-compensatory payment for social efficiency and the firm prefers to have no damage payments at all. Even without full market coverage, when b is high, the consumer's net benefit will be increasing in the consumer's type x. When $w = 0$, the firm avoids subsidizing the purchases of the consumers with high types. This result is aligned with our findings in the main section. But different from the main section, stipulated-damage payments affect social efficiency by impacting both product safety (quality) and the output level (quantity). Starting at $w = h$, if the social planner reduces w slightly, then the firm's safety investment will fall. However, because the safety level is socially optimal when $w = h$, the reduction in safety has a second-order effect on social welfare. Decreasing w has a first-order effect on the output level chosen by the firm. Therefore, the socially-optimal damage payment is strictly less than h.

Note that, starting from $w = 0$, if the social planner raises the damage payment slightly, product safety will rise and output will fall. In general it is ambiguous whether the socially-optimal damage payment is positive or not. As shown in the following numerical example, the socially optimal damage payment may be strictly positive.

Numerical Example 2: *Suppose that x follows the distribution Gamma (9,0.5) on $(0, \infty)$, $b_0 = 0$, $b = 1$, and $h = 0.5$. (1) If $c(\pi) = \frac{6}{5}(1-\pi)^2$ for $\pi \in (0, 1]$, then a simulation shows that the social welfare has an inverted U-shaped relationship with w, for $w \in [0, h]$. The socially-optimal damage payment is $w^{**} \approx 0.1$. The firm does not stipulate any damage payment, under-invests in safety, and sells more units than the social planner desires.* (2) *In contrast, if $c(\pi) = \frac{3}{2}(1 - \pi)^2$, then a simulation shows that the socially-optimal damage level is $w^{**} = 0$, the same as the firm's choice.*[48]

As a remark, when b is very small, then the consumer's net benefit can be a decreasing function of x and so the consumers with $x \leq x^M$ will buy the product. In this case, with a fixed damage payment, the firm's safety investment may be excessive for those infra-marginal consumers. However, as shown in our benchmark model, given the consumer's net benefit is decreasing in x, the stipulated-damage payment is a valuable mechanism for price discrimination and therefore the firm will increase the damage payment level. Thus, the over-investment problem (for infra-marginal consumers) can be addressed by private contracts with stipulated damages. However, as shown earlier, private contracts cannot solve the under-investment problem, which calls for legal intervention.

[48] In the examples, given the socially-optimal damage payment w^{**}, π^* and x^{M*} are both interior solutions, with $\pi^* x^{M*} \leq 1$.

□ **Versioning.** In the benchmark model, we assume that the firm produces a single standard version of the product rather than a menu of specialized versions. We now allow the firm to offer many different versions of the product with different safety levels, and show that our main insights regarding the firm's suboptimal choices of stipulated-damage payments and safety investments still hold. As in the benchmark model, we assume that b_0 is sufficiently large to ensure full market coverage and we focus on the scenario with $b > b^H = \pi^H(0)h$ so the consumer's net benefit is increasing in the consumer's type, x.[49]

Suppose that the firm can offer a menu $(\pi(x), p(x), w(x))$, where for each consumer type x, $\pi(x)$ is the safety level, $p(x)$ is the price, and $w(x)$ is the stipulated damage payment. A consumer observes the safety levels of all products before choosing which version to purchase. Matthews and Moore (1987) show that risk-neutral consumers only care about their total expected payments and so private warranties cannot generate higher profits for the firm. This logic, however, cannot be applied in our model for the following reason. In Matthews and Moore (1987), the firm offers a "total warranty" $T(x)$ so that a type x consumer's expected payment to the firm is $p(x) - \pi(x)T(x)$. This expected payment only depends on the consumer's choice of contract but not on her real type. In our model, a type x consumer expects to receive a total damage compensation of $\pi(x)w(x)x$, so that her effective payment to the firm is $p(x) - \pi(x)w(x)x$, which depends on the consumer's choice of contract as well as her real type, x.

We can use the mechanism-design approach to derive the firm's optimal menu of contracts $(\pi^*(x), p^*(x), w^*(x))$:

$$\underset{\pi(x),p(x),w(x)}{Max} \int_{\underline{x}}^{\overline{x}} [p(x) - c(\pi(x)) - \pi(x)w(x)x]f(x)dx \tag{24}$$

subject to the (IR) and (IC) constraints:

$$b_0 + [b - \pi(x)(h - w(x))]x - p(x) \geq 0 \quad \forall x \tag{25}$$

$$b_0 + [b - \pi(x)(h - w(x))]x - p(x) \geq b_0 + [b - \pi(\widehat{x})(h - w(\widehat{x}))]x - p(\widehat{x}) \quad \forall x, \widehat{x} \tag{26}$$

$$w(x) \in [0, h] \quad \forall x. \tag{27}$$

Given $b > b^H$, it can be shown that the (IR) constraint is binding for the lowest type $\underline{x}$. As shown in the Appendix, the firm's optimization problem is

[49] The structure of the model is similar to that of Mussa and Rosen (1978). Absent legal intervention or stipulated-damage payments, they showed that the firm would under-invest in product quality.

reduced to the following:

$$\underset{\pi(x),w(x)}{Max} \quad b_0 + b\underline{x} + \int_{\underline{x}}^{\overline{x}} \left[\pi(x)(h - w(x))\frac{1-F(x)}{f(x)} - c(\pi(x)) - \pi(x)hx\right] f(x)dx. \tag{28}$$

Notice that the firm's objective function is strictly decreasing in $w(x)$. That is, the imposition of stipulated-damage payments or the imposition of product liability reduces firm profit. So, absent legal intervention, the firm would choose $w(x) = 0$ for all x.

Given $w(x)$, the firm's optimal choice of safety, $\pi^*(x)$ for any x, satisfies

$$(h - w(x))\frac{1-F(x)}{f(x)} - c'(\pi^*(x)) - hx = 0. \tag{29}$$

Note that, when $w(x) = h$ for all x, the firm's optimal choice of safety for each x satisfies

$$-c'(\pi^{**}(x)) - hx = 0, \tag{30}$$

which is socially optimal. This is not too surprising. When $w(x) = h$, the consumers do not care about product safety (as they are fully compensated for their losses). The price is the same for all versions of the product, $p(x) = b_0 + b\underline{x}$ and consumers, being indifferent, self-select appropriately. The total consumer surplus or rent is $(E(x) - \underline{x})b$. The firm is therefore the residual claimant on the social benefits associated with versioning, and so, when $w(x) = h$, the firm's incentives are aligned with the society's. When $w(x) < h$, comparing conditions (29) and (30), we have $\pi^*(x) > \pi^{**}(x)$ for any $x < \overline{x}$ and $\pi^*(\overline{x}) = \pi^{**}(\overline{x})$.[50]

To summarize, we have the following result:

Proposition 6. *(Versioning.) Suppose that $b > b^H$ and the firm can choose to sell a continuum of versions. The firm chooses $w(x) = 0$ for all x, under-invests for all $x < \overline{x}$, but invests efficiently for $x = \overline{x}$. If the social planner mandates the fully-compensatory damage payment, $w^{**} = h$, then the firm chooses the socially-optimal safety level for each consumer type.*

□ **Non-Linear Net Benefits.** In the benchmark model, the consumer's expected harm and the consumer's gross benefit were both linear functions of the consumer's type, x. This is not crucial for our main results. To see this, consider the following generalization. A type x consumer enjoys a gross benefit $b_0 + bu(x) > 0$, with $u(0) = 0$ and $u'(x) > 0$, and suffers the expected harm $\pi hx > 0$ from using the product. As in the benchmark model, we assume that

[50] It can also be verified that $\pi^*(\underline{x}) > \pi^{**}(\underline{x})$.

b_0 is sufficiently large to ensure full market coverage. With full market coverage, the social welfare is

$$\widetilde{W}(\pi) = b_0 + b\int_{\underline{x}}^{\overline{x}} u(x)dF(x) - \pi h E(x) - c(\pi). \tag{31}$$

Note that the socially-efficient safety level π^{**} is the same as in the benchmark model.

The firm chooses the menu $(\pi, \rho(x), \omega(x))$ to maximize profits subject to incentive compatibility and individual rationality constraints:[51]

$$\int_{\underline{x}}^{\overline{x}} \left[\rho(x) - \pi\omega(x)x - c(\pi)\right] f(x)dx \tag{32}$$

subject to

$$\rho(x) - \pi\omega(x)x \leq \rho(\widehat{x}) - \pi\omega(\widehat{x})x \quad \forall x, \widehat{x} \tag{33}$$

$$\rho(x) - \pi\omega(x)x \leq b_0 + bu(x) - \pi h x \quad \forall x \tag{34}$$

$$\omega(x) \in [0, h] \quad \forall x \tag{35}$$

where (33) is the IC constraint, (34) is the IR constraint, and $\rho(x) - \pi\omega(x)x$ is the effective price paid by a consumer of type x.

As we show in the proof of Proposition 7 in the Appendix, the IC constraint (33) requires $\omega'(x) \geq 0$ and $\rho'(x) = \pi\omega'(x)x$. And we can rewrite the firm's problem as choosing π, $\rho(\underline{x})$, and $\omega(x)$ to maximize

$$\int_{\underline{x}}^{\overline{x}} \left[\rho(\underline{x}) - \pi\omega(\underline{x})\underline{x} - \pi\int_{\underline{x}}^{x} \omega(y)dy - c(\pi)\right] f(x)dx \tag{36}$$

subject to

$$b_0 + bu(x) - \pi h x - \left[\rho(\underline{x}) - \pi\omega(\underline{x})\underline{x} - \pi\int_{\underline{x}}^{x} \omega(y)dy\right] \geq 0 \quad \forall x. \tag{37}$$

The left-hand side of (37) are the rents received by a consumer of type x, with the first derivative $bu'(x) - \pi(h - \omega(x))$. As defined in the benchmark model, the efficient safety level for type $\underline{x}$ is $\pi^H(0)$. Because type $\underline{x}$ is least likely to suffer harm, the firm would never choose any safety level $\pi > \pi^H(0)$. Therefore, when $b > \frac{\pi^H(0)h}{\min\{u'(x)\}}$ and $\omega(x) \geq 0$, the derivative of consumer rent is strictly

[51] In an earlier version, we considered the simple contract form (π, p, w). We showed that, given concave $u(x)$, the firm would stipulate $w^* = \min\{h - \frac{b[u(\overline{x}) - u(\underline{x})]}{\pi^{**}(\overline{x} - \underline{x})}, h\}$ and invest efficiently if $b \leq \frac{\overline{x} - \underline{x}}{u(\overline{x}) - u(\underline{x})}\pi^{**}h$, but choose $w^* = 0$ and under-invest in product safety if otherwise.

positive for all x. That is, the marginal consumer is of type $\underline{x}$. Then as shown in the Appendix, in the optimal mechanism, the firm offers $\omega(x) = 0$ for all x and chooses the safety level $\pi^H(0) > \pi^{**}$.

Note that, when $\omega(x) = h$ for all x, the firm's profit (36) becomes

$$\int_{\underline{x}}^{\overline{x}} [\rho(\underline{x}) - \pi h x - c(\pi)] f(x) dx, \tag{38}$$

which is maximized by $\pi = \pi^{**}$. When $\omega(x) = h$, the consumers do not care about product safety. The price is the same for all consumers, $\rho(x) = b_0 + bu(\underline{x})$, and the firm's incentives are aligned with the society's.

Proposition 7. *(Non-Linear Net Benefits.) When b is sufficiently large,*[52] *the firm chooses $\omega(x) = 0$ for all x and under-invests in product safety. If the social planner mandates the fully-compensatory damage payment, $w^{**} = h$, then the firm chooses the socially-optimal safety level.*

The results in Proposition 7 resemble the findings in the benchmark model. When the incremental benefit b is large, offering positive damage payments would be counterproductive for the firm's price discrimination quest, as positive damage payments would subsidize the infra-marginal consumers. A legal requirement that forces the firm to make consumers whole (that is, $w^{**} = h$) after suffering accident-related harms increases the firm's private incentive to invest in product safety, and social welfare is maximized.

Recall that in the benchmark model, when b is positive but small, the firm achieves perfect price discrimination by choosing $w^* = h - \frac{b}{\pi^{**}}$ and making the efficient investment. This result holds qualitatively in the non-linear setting as long as $u(x)$ is concave. To see this, suppose that $u''(x) < 0$ and $0 \leq b \leq \frac{\pi^{**} h}{u'(\underline{x})}$. Then for any x, $h - \frac{bu'(x)}{\pi^{**}}$ is positive and increasing in x. If the firm chooses $\omega^*(x) = h - \frac{bu'(x)}{\pi^{**}}$ for all x and the safety level π^{**}, the derivative of consumer rent is

$$bu'(x) - \pi^{**}(h - \omega^*(x)) = 0 \quad \forall x. \tag{39}$$

Because the consumer rent does not vary with type x, all consumer types are marginal. As $u'(x)$ decreases in x, $\omega^*(x) = h - \frac{bu'(x)}{\pi^{**}}$ increases in x and therefore the IC constraint holds. Thus, when b is positive but small enough and $u(x)$ is concave, with a menu of partially-compensatory damage payments $\omega^*(x)$, the firm implements perfect price discrimination and extracts all consumer surplus.[53] The social optimum is achieved.

---

[52]More formally, $b > \frac{\pi^H(0)h}{\min\{u'(x)\}}$.

[53]In this case, the firm offers a menu of prices $\rho(\underline{x}) = b_0 + bu(\underline{x}) - \pi^{**}\underline{x}(h - \omega^*(\underline{x}))$ and $\rho(x) = \rho(\underline{x}) - \pi^{**}\omega^*(\underline{x})\underline{x} + \pi^{**}\omega^*(x)x - \pi^{**}\int_{\underline{x}}^{x} \omega^*(y)dy$ for $x > \underline{x}$.

5 Conclusion

■ Should manufacturers be held liable for injuries suffered by consumers? In the United States, approximately 30,000 new products liability cases are filed in state courts each year.[54] This is just the tip of the iceberg, as many potential lawsuits (some of which are meritorious) are not brought at all. According to the Consumer Product Safety Commission, a government agency that oversees products ranging from coffee makers to toys and table saws, 38 million people sought medical attention in the United States in 2010 for injuries related to consumer products.[55] In addition, the National Safety Council estimates 4.6 million people were required medical attention for injuries sustained while using the roadways in 2016, and that the costs of motor vehicle deaths, injuries, and property damage were $432 Billion.[56]

Our article informs the policy debate about the appropriate role of products liability law by providing a rigorous demonstration of the basic circumstances under which firms will try to limit or disclaim products liability and, as a consequence, under-invest in product safety. We show that firms that possess market power will disclaim responsibility for consumer harms and will under-invest in safety when consumers who face the highest risks of accidents are also those who are willing to pay more for the product. This result persists even when quality is unobservable to consumers at the time of sale, when firms offer multiple versions of their products with a menu of prices and damage payment clauses, and when there is less-than-full market coverage.[57] We illustrate how legal interventions — including products liability and prohibitions on liability waivers and disclaimers in private contracts — can improve social efficiency.

Our insights are also relevant when consumers can take actions to reduce the risks associated with dangerous products. Consumers can and should take reasonable precautions when driving ATVs or operating chain saws, for example.[58] In such settings, a socially-optimal liability rule would hold consumers accountable

[54]There is also a stock of more than 30,000 published case decisions on the subject of product liability. See Owen (2015). Note also that many defective consumer products are repaired or replaced under warranty.

[55]See https://www.cpsc.gov/PageFiles/134720/2010injury.pdf. Not all of these injuries were caused by the products.

[56]See www.nsc.org/Connect/NSCNewsReleases/Lists/Posts/Post.aspx?ID=180. Moreover, the Centers for Disease Control estimate that almost 50 million people suffer from food poisoning each year. See https://www.cdc.gov/foodsafety/foodborne-germs.html

[57]Our formal analysis focuses on monopoly. In a follow-up work, we show that our insights are relevant in settings with imperfect competition.

[58]For other harmful products, like tobacco or certain drugs, consumer precautions are less relevant.

for a failure to take reasonable precautions. In particular, a rule of strict liability with a defense of contributory negligence would align the private interests of firms and consumers with those of the social planner.[59] As in our benchmark model, legal intervention is necessary when the marginal consumer places a relatively low value on product safety. Without legal intervention or regulation, firms would eschew contractual liability and under-invest in product safety.[60]

Although our model was framed in terms of consumer products and public safety, the issues apply to commercial transactions as well. Our buyers could be businesses or corporations, and the harms could include property damage and lost economic profits. In practice, the courts have been less inclined to intervene in business-to-business settings than in consumer-product settings. In a famous case involving a shipbuilder that suffered millions of dollars in harms and losses when its turbine engines malfunctioned, the Supreme Court declined to find tort liability noting that "products liability grew out of a public policy judgment that people need more protection from dangerous products than is afforded by the law of warranty. It is clear, however, that if this development were allowed to progress too far, contract law would drown in a sea of tort."[61]

[59]This rule would hold the firm strictly liable for consumer harms if and only if the consumer took due care while using the product. This presupposes that the court can observe the consumer's actual level of care.

[60]This result also holds without a defense of contributory negligence. A formal analysis of consumer moral hazard is available from the authors upon request.

[61]*E. River S.S. Corp. v. Transamerica Delaval, 476 U.S. 858 (1986).*

References

[1] Arlen, J. and MacLeod, W.B. "Malpractice Liability for Physicians and Managed Care Organizations." *New York University Law Review*, Vol. 78 (2003), pp. 1929–2006.

[2] Baker, S. and Choi, A.H. "Reputation and Litigation: Why Costly Legal Sanctions Can Work Better than Reputational Sanctions." *Journal of Legal Studies*, Vol. 47 (2018), pp. 45-82.

[3] Boom, A. "Product Risk Sharing by Warranties in a Monopoly Market with Risk-Averse Consumers." *Journal of Economic Behavior and Organization*, Vol. 33 (1988), pp. 241-257.

[4] Che, Y.-K. "Customer Return Policies for Experience Goods." *Journal of Industrial Economics*, Vol. 44 (1986), pp. 17-24.

[5] Che, Y.-K. and Kim, J. "Bidding with Securities: Comment." *American Economic Review*, Vol. 100 (2010), pp. 1929-35.

[6] Chen, Y. and Hua, X. "Ex ante Investment, Ex post Remedies, and Product Liability." *International Economic Review,* Vol 53 (2012), pp. 845-866.

[7] Chen, Y. and Hua, X. "Competition, Product Safety, and Product Liability." *Journal of Law, Economics, & Organization,* Vol. 33 (2017), pp. 237-267.

[8] Choi, A.H. and Spier, K.E. "Should Consumers Be Permitted to Waive Products Liability? Product Safety, Private Contracts, and Adverse Selection." *Journal of Law, Economics, & Organization*, Vol. 30 (2014), pp. 734-766.

[9] Cooper, R. and Ross, T.W., "Product Warranties and Double Moral Hazard." *RAND Journal of Economics*, Vol. 2 (1985), pp. 101–113.

[10] Courty, P. and Li, H. "Sequential Screening." *Review of Economic Studies*, Vol. 67 (2000), pp. 697-717.

[11] Cremer, J. and McLean, R. " Full Extraction of the Surplus in Bayesian and Dominant Strategy Auctions." *Econometrica*, Vol. 56 (1988), pp. 1247-1257.

[12] Daughety, A.F. and Reinganum, J.F. "Product Safety: Liability, R&D, and Signaling." *American Economic Review*, Vol. 85 (1995), pp. 1187–1206.

[13] Daughety, A.F. and Reinganum, J.F. "Everybody Out of the Pool: Products Liability, Punitive Damages, and Competition." *Journal of Law, Economics, & Organization*, Vol. 13 (1997), pp. 410–432.

[14] Daughety, A.F. and Reinganum, J.F. "Market, Torts, and Social Inefficiency." *RAND Journal of Economics* Vol. 37 (2006), pp. 300-323.

[15] Daughety, A.F. and Reinganum, J.F. "Communicating Quality: A Unified Model of Disclosure and Signaling." *RAND Journal of Economics*, Vol. 39 (2008b), pp. 973–989.

[16] Daughety, A.F. and Reinganum, J.F. "Imperfect Competition and Quality Signaling." *RAND Journal of Economics*, Vol. 39 (2008a), pp. 163-183.

[17] Daughety, A.F. and Reinganum, J.F. "Cumulative Harm and Resilient Liability Rules for Product Markets." *Journal of Law, Economics, & Organization*, Vol. 30 (2014), pp. 371-400.

[18] DeMarzo, P.M., Kremer, I. and Skrzypacz, A. "Bidding with Securities: Auctions and Security Design." *American Economic Review*, Vol. 95 (2005), pp. 936-959.

[19] Dybvig, P.H. and Lutz, N.A. "Warranties, Durability, and Maintenance: Two-Sided Moral Hazard in a Continuous Time Model." *Review of Economic Studies*, Vol. 60 (1993), pp. 575–598.

[20] Emons, W. "Warranties, Moral Hazard, and the Lemons Problem." *Journal of Economic Theory*, Vol. 46 (1988), pp. 16-33.

[21] Epple, D. and Raviv, A. "Product Safety: Liability Rules, Market Structure, and Imperfect Information." *American Economic Review*, Vol. 68 (1978), pp. 80–95.

[22] Grossman, S.J. "Informational Role of Warranties and Private Disclosure about Product Quality." *Journal of Law & Economics*, Vol. 24 (1981), pp. 461–483.

[23] Hamada, K. "Liability Rules and Income Distribution in Product Liability." *American Economic Review*, Vol. 66 (1976), pp. 228–234.

[24] Hansen, R.G. "Empirical Testing of Auction Theory." *American Economic Review, Papers and Proceedings*, Vol. 75 (1985), pp. 156-159.

[25] Hay, B. and Spier, K.E. "Manufacturer Liability for Harms Caused by Consumers to Others." *American Economic Review*, Vol. 95 (2005), pp. 1700–1711.

[26] Heal, G. "Guarantees and Risk Sharing." *Review of Economic Studies*, Vol. 44 (1977), pp. 549–560.

[27] Klein, B. and Leffler, K.B. "The Role of Market Forces in Assuring Contractual Performance." *Journal of Political Economy*, Vol. 89 (1981), pp. 615–641.

[28] Lutz, N.A. "Warranties as Signals Under Consumer Moral Hazard." *RAND Journal of Economics*, Vol. 20 (1989), pp. 239–255.

[29] Lutz, N.A. and Padmanabhan, V. "Warranties, Extended Warranties, and Product Quality." *International Journal of Industrial Organization*, Vol. 16 (1998), pp. 463–493.

[30] Mann, D.P. and Wissink, J.P. "Money-Back Contracts with Double Moral Hazard." *RAND Journal of Economics*, Vol. 19 (1988), pp. 285-292.

[31] Matthews, S. and Moore, J. "Monopoly Provision of Quality and Warranties: An Exploration in the Theory of Multidimensional Screening. " *Econometrica*, Vol. 55 (1987), pp. 441-467.

[32] Moorthy, S. and Srinivasan, K. "Signaling Quality with a Money-Back Guarantee: The Role of Transaction Costs." *Marketing Science*, Vol. 14 (1995), pp. 442-466.

[33] Ordover, J.A. "Products Liability in Markets with Heterogeneous Consumers." *Journal of Legal Studies*, Vol. 8 (1979), pp. 505–525.

[34] Padmanabhan, V. "Usage Heterogeneity and Extended Service Contracts." *Journal of Economics and Management Strategy*, Vol. 4 (1995), pp. 33–53.

[35] Polinsky, A.M. and Rogerson, W.P. "Products Liability, Consumer Misperceptions, and Market Power." *Bell Journal of Economics*, Vol. 14 (1983), pp. 581–589.

[36] Polinsky, A.M. and Shavell, S. "The Uneasy Case for Product Liability." *Harvard Law Review*, Vol. 123 (2010), pp. 1437–1492.

[37] Priest, G.L. "A Theory of the Consumer Product Damages." *Yale Law Journal*, Vol. 90 (1981), pp. 1297–1352.

[38] Riordan, M.H. and Sappington, D.E. "Optimal Contracts with Public Ex Post Information." *Journal of Economic Theory*, Vol. 45 (1988), pp. 189-199.

[39] Rothschild, M. and Stiglitz, J. "Equilibrium in Competitive Insurance Markets: An Essay on the Economics of Imperfect Information." *Quarterly Journal of Economics*, Vol. 90 (1976), pp. 629–649.

[40] Shieh, S. "Price and Money-Back Guarantees as Signals of Product Quality." *Journal of Economics and Management Strategy*, Vol. 5 (1996), pp. 361-377.

[41] Simon, M.J. "Imperfect Information, Costly Litigation, and Product Quality." *Bell Journal of Economics*, Vol. 12 (1981), pp. 171-184.

[42] Spence, A.M. "Monopoly, Quality, and Regulation." *Bell Journal of Economics*, Vol. 6 (1975), pp. 417–429.

[43] Spence, A.M. "Consumer Misperceptions, Product Failure and Producer Liability." *Review of Economic Studies*, Vol. 44 (1977), pp. 561–572.

[44] Wickelgren, A.L. "The Inefficiency of Contractually Based Liability with Rational Consumers." *Journal of Law, Economics, & Organization*, Vol. 22 (2006), pp. 168–183.

Appendix

Proof of Lemma 1. Dividing (7) by $x - \widehat{x}$, and taking the limit as $x - \widehat{x}$ approaches zero gives

$$\rho'(x) = \pi\omega'(x)x. \tag{40}$$

Integrating the right-hand side by parts gives $\rho(x) = \rho(\underline{x}) - \pi\omega(\underline{x})\underline{x} + \pi\omega(x)x - \pi\int_{\underline{x}}^{x}\omega(y)dy$,[62] and rearranging terms gives (8) in the text.

Substituting (8) into (3) and (5), we can rewrite the firm's problem as choosing π, $\rho(\underline{x})$, and $\omega(x)$ to maximize

$$\int_{\underline{x}}^{\overline{x}} \left[\rho(\underline{x}) - \pi\omega(\underline{x})\underline{x} - \pi\int_{\underline{x}}^{x}\omega(y)dy - c(\pi)\right] f(x)dx \tag{41}$$

subject to

$$b_0 + bx - \pi hx - \left[\rho(\underline{x}) - \pi\omega(\underline{x})\underline{x} - \pi\int_{\underline{x}}^{x}\omega(y)dy\right] \geq 0 \quad \forall x. \tag{42}$$

The left-hand side of (42) are the rents received by a consumer of type x. The first derivative is $b - \pi(h - \omega(x))$, and the second derivative is $\pi\omega'(x) \geq 0$.

Let $x^M \in [\underline{x}, \overline{x}]$ be the value (not necessarily unique) where the left-hand side of (42) is minimized.[63] It must be the case that the IR constraint binds at x^M. If not, then the firm could increase profits by increasing the price $\rho(\underline{x})$. Because the IR constraint (42) holds with equality when $x = x^M$ we have

$$\rho(\underline{x}) - \pi\omega(\underline{x})\underline{x} = b_0 + (b - \pi h)x^M + \pi\int_{\underline{x}}^{x^M}\omega(y)dy. \tag{43}$$

Plugging (43) into (42), we rewrite the IR constraint as:

$$(x - x^M)(b - \pi h) - \pi\int_{\underline{x}}^{x^M}\omega(y)dy + \pi\int_{\underline{x}}^{x}\omega(y)dy \geq 0 \quad \forall x. \tag{44}$$

We will now prove that the optimal schedule has the property that $w'(x) = 0$ for all x. We will proceed in two steps. First, we prove that if $\omega(x)$ is the optimal schedule, then a transformed schedule with $\omega(x) \equiv \omega(x^M)$ satisfies IR in (44). Then, we prove that firm profits are higher when the schedule $\omega(x)$ is replaced by one with $\omega(x) \equiv \omega(x^M)$.

[62]Note that one could equivalently show that $\rho(x) - \pi\omega(x)x = \rho(\overline{x}) - \pi\omega(\overline{x})\overline{x} + \pi\int_{x}^{\overline{x}}\omega(y)dy$.

[63]If $b - \pi(h - \omega(\underline{x})) \geq 0$ then the left-hand side of the IR constraint (42) is increasing for all $x \in [\underline{x}, \overline{x}]$. If $b - \pi(h - \omega(\overline{x})) \leq 0$ then the left-hand side of (42) is decreasing for all $x \in [\underline{x}, \overline{x}]$.

We first prove IR is not violated when $\omega(x)$ is flattened out. Plugging $w(x) \equiv \omega(x^M)$ into (44) gives the information rents for type x:

$$\begin{aligned}(x - x^M)(b - \pi h) - \pi\omega(x^M)(x^M - \underline{x}) + \pi\omega(x^M)(x - \underline{x}) \\ = (x - x^M)(b - \pi h) - \pi\omega(x^M)x^M + \pi\omega(x^M)x \\ = (x - x^M)[b - \pi(h - \omega(x^M))]. \end{aligned} \tag{45}$$

We will show that the rents are non-negative for all x. First, if $x^M = \underline{x}$ then $b - \pi(h - \omega(x^M)) \geq 0$. If this was not true, then type $\underline{x} + \varepsilon$ where ε is a small positive number would have lower rents than type $\underline{x}$. So x^M does not have the lowest rents, a contradiction. So if $x^M = \underline{x}$ then $(x - x^M)[b - \pi(h - \omega(x^M)] \geq 0$ for all $x > x^M = \underline{x}$. Second, if $x^M = \overline{x}$ then $b - \pi(h - \omega(x^M) \leq 0$. If this was not true, then type $\overline{x} - \varepsilon$ where ε is a small positive number would have lower rents than type $\overline{x}$. So if $x^M = \overline{x}$ then $(x - x^M)[b - \pi(h - \omega(x^M)] \geq 0$ for all $x < x^M = \overline{x}$. Finally, if $x^M \in (\underline{x}, \overline{x})$ then it must be true that $b - \pi(h - \omega(x^M) = 0$. If this were not true, then type x^M would not have the lowest rents. So for $x^M \in (\underline{x}, \overline{x})$ we must have $(x - x^M)[b - \pi(h - \omega(x^M)] = 0$ for all x. This concludes the proof that changing the warranty schedule to $\omega(x) = \omega(x^M)$ for all x would not violate individual rationality.

Next, we prove that firm profits are highest when $\omega(x) \equiv \omega(x^M)$. Using (43) we rewrite (41) as

$$b_0 + (b - \pi h)x^M + \pi \int_{\underline{x}}^{\overline{x}} \int_{\underline{x}}^{x^M} \omega(y)f(x)dydx - \pi \int_{\underline{x}}^{\overline{x}} \int_{\underline{x}}^{x} \omega(y)f(x)dydx - c(\pi) \tag{46}$$

which can be rewritten as

$$b_0 + (b - \pi h)x^M + \pi \int_{\underline{x}}^{x^M} \int_{x}^{x^M} \omega(y)f(x)dydx - \pi \int_{x^M}^{\overline{x}} \int_{x^M}^{x} \omega(y)f(x)dydx - c(\pi). \tag{47}$$

Using Fubini's theorem, the optimand is

$$b_0 + (b - \pi h)x^M + \pi \int_{\underline{x}}^{x^M} \int_{\underline{x}}^{y} \omega(y)f(x)dxdy - \pi \int_{x^M}^{\overline{x}} \int_{y}^{\overline{x}} \omega(y)f(x)dxdy - c(\pi) \tag{48}$$

which becomes

$$b_0 + (b - \pi h)x^M + \pi \int_{\underline{x}}^{x^M} \omega(y)F(y)dy - \pi \int_{x^M}^{\overline{x}} \omega(y)(1 - F(y))dy - c(\pi) \tag{49}$$

or equivalently

$$b_0 + (b - \pi h)x^M + \pi \int_{\underline{x}}^{x^M} \omega(x)F(x)dx - \pi \int_{x^M}^{\overline{x}} \omega(x)(1 - F(x))dx - c(\pi). \quad (50)$$

Recall that IC implies that $\omega'(x) \geq 0$ for all x. If $x < x^M$ then $\omega(x) \leq \omega(x^M)$ and profits would be raised by raising $\omega(x)$ so $\omega(x) = \omega(x^M)$ for all $x < x^M$. If $x > x^M$ then $\omega(x) \geq \omega(x^M)$ and profits would be raised by lowering $\omega(x)$ so $w(x) = \omega(x^M)$ for all $x < x^M$. Therefore in the optimal mechanism, $\omega(x) \equiv \omega(x^M)$.

This completes the general proof that in the optimal mechanism the warranty is constant, so $\omega'(x) \equiv 0$ and $\rho'(x) \equiv 0$. So, we may restrict attention to simple contracts of the form (π, p, w). □

Proof of Lemma 2. Conditions (13) and (15) can be re-written as

$$-h\underline{x} - w(E(x) - \underline{x}) - c'(\pi^H(w)) = 0 \quad (51)$$

$$-h\overline{x} - w(E(x) - \overline{x}) - c'(\pi^L(w)) = 0. \quad (52)$$

Accordingly, $d\pi^L(w)/dw = \frac{-(E(x)-\overline{x})}{c''(\pi^L(w))} > 0$ and $d\pi^H(w)/dw = \frac{-(E(x)-\underline{x})}{c''(\pi^H(w))} < 0$. □

Proof of Proposition 1. We begin by proving a claim.

Claim 1. *Taking $w \in [0, h]$ as fixed, the firm's profit-maximizing safety level $\pi^*(w)$ has the following properties:*

1. *If $\widehat{\pi}(w) < \pi^L(w) < \pi^H(w)$ then $\pi^*(w) = \pi^L(w)$.*
2. *If $\pi^L(w) \leq \widehat{\pi}(w) \leq \pi^H(w)$ then $\pi^*(w) = \widehat{\pi}(w)$.*
3. *If $\pi^L(w) < \pi^H(w) < \widehat{\pi}(w)$ then $\pi^*(w) = \pi^H(w)$.*

This claim tells us that given the stipulated-damage payment or liability level w, the firm will choose one of three safety levels: $\widehat{\pi}(w)$, $\pi^H(w)$, or $\pi^L(w)$ defined in (11), (13), and (15), respectively. In part 1 of the claim, the consumer with type $\overline{x}$ is the marginal consumer and the firm chooses safety level $\pi^*(w) = \pi^L(w) < \pi^{**}$ from Lemma 2. Product safety is socially excessive in this case. In part 2, the firm chooses $\pi^*(w) = \widehat{\pi}(w)$ and all consumer types are marginal (the net benefit of the consumers is independent of x). Product safety can be excessive or insufficient. In part 3, the consumer with type $\underline{x}$ is the marginal consumer and the firm's safety level reflects that, $\pi^*(w) = \pi^H(w) > \pi^{**}$. In this case, product safety is insufficient.

Proof of Claim 1. We know from Lemma 2 that $\pi^L(w) < \pi^H(w)$. $\widehat{\pi}(w)$ may be below these two values, above them, or in between. So, we consider these three cases.

Suppose $\widehat{\pi}(w) < \pi^L(w) < \pi^H(w)$. If $\pi < \widehat{\pi}(w)$ then the profit function is $\underline{S}(\pi, w)$ which is concave in π and reaches its maximum at $\pi^H(w) > \widehat{\pi}(w)$. Because $\underline{S}(\pi, w)$ is increasing in π for all $\pi < \widehat{\pi}(w)$ the firm will not choose $\pi < \widehat{\pi}(w)$. If $\pi > \widehat{\pi}(w)$ then the profit function is $\overline{S}(\pi, w)$ which is concave in π and reaches its maximum at $\pi^L(w)$. So the firm chooses $\pi^*(w) = \pi^L(w)$.

Suppose $\pi^L(w) < \pi^H(w) < \widehat{\pi}(w)$. If $\pi > \widehat{\pi}(\mathrm{w})$ then the profit function is $\overline{S}(\pi, w)$ which is concave in π and reaches its maximum at $\pi^L(w) < \widehat{\pi}(w)$. Because $\overline{S}(\pi, w)$ is decreasing in π when $\pi > \widehat{\pi}(w)$, the firm will not choose $\pi > \widehat{\pi}(w)$.If $\pi < \widehat{\pi}(w)$ then the profit function is $\underline{S}(\pi, w)$ which reaches its maximum at $\pi^H(w) < \widehat{\pi}(w)$.So the firm chooses $\pi^*(w) = \pi^H(w)$.

Finally, suppose $\pi^L(w) \le \widehat{\pi}(w) \le \pi^H(w)$. Following the logic above, if $\pi < \widehat{\pi}(w)$ then the profit function is $\underline{S}(\pi, w)$ which reaches its maximum at $\pi^H(w) \ge \widehat{\pi}(w)$. If $\pi > \widehat{\pi}(w)$ then the profit function is $\overline{S}(\pi, w)$ which reaches its maximum at $\pi^L(w) \le \widehat{\pi}(w)$. The firm will therefore choose $\pi^*(w) = \widehat{\pi}(w)$. □

When $w = 0$ then from (11) we have $\widehat{\pi}(0) = b/h$. By Lemma 2, $\pi^L(w) < \pi^{**} < \pi^H(w)$ for all $w \in [0, h)$. Recall that functions $\pi^H(w)$ and $\pi^L(w)$ in (13) and (15) do not depend on the parameters b_0 or b. When $\widehat{\pi}(0) < \pi^L(0)$, or equivalently $b < b^L = \pi^L(0)h$, then from Claim 1 we have $\pi^*(0) = \pi^L(0)$. By Lemma 2, $\pi^L(0) < \pi^{**}$ and so there is over-investment in product safety. When $\widehat{\pi}(0) > \pi^H(0)$, or equivalently $b > b^H = \pi^H(0)h$, then by Lemma 2 and Claim 1 we have $\pi^*(0) = \pi^H(0) > \pi^{**}$ and so there is under-investment in product safety. When $\widehat{\pi}(0) \in [\pi^L(0), \pi^H(0)]$ or equivalently $b \in [b^L, b^H]$, then we have $\pi^*(0) = \widehat{\pi}(0) = b/h$. If $b < b^{**}$ the firm is over-investing in product safety and if $b > b^{**}$ the firm is under-investing. □

Proof of Proposition 2. We begin by proving a claim.

Claim 2. *For any continuous and non-empty set of w, if $\pi^*(w) = \pi^H(w)$, then the firm's optimal stipulated-damage payment is the lower bound of the set. If $\pi^*(w) = \pi^L(w)$, then the firm's optimal stipulated-damage payment is the upper bound of the set. If $\pi^*(w) = \widehat{\pi}(w)$, then the firm's optimal stipulated-damage payment is the lower bound of the set when $\widehat{\pi}(w) > \pi^{**}$ and the upper bound when $\widehat{\pi}(w) < \pi^{**}$.*

Proof of Claim 2. When $\pi^*(w) = \pi^H(w)$ then it must also be the case that $\pi^H(w) < \widehat{\pi}(w)$ (Claim 1). The firm's associated profit function $\underline{S}(\pi, w)$ is a strictly decreasing function of w so the firm will choose the smallest value of w from the set.

When $\pi^*(w) = \pi^L(w)$, then it must also be the case that $\pi^L(w) > \widehat{\pi}(w)$ (Claim 1). The firm's associated profit function $\overline{S}(\pi, w)$ is a strictly increasing function of w so the firm will choose the largest value of w from the set.

When $\pi^*(w) = \widehat{\pi}(w)$, each consumer's net benefit is $b_0 - p$ which is of course independent of type x. The firm charges price $p = b_0$ and its average unit cost is $\widehat{\pi}(w)wE(x) + c(\widehat{\pi}(w))$. Taken together, the firm's associated profit function is:

$$\underline{S}(\widehat{\pi}(w), w) = \overline{S}(\widehat{\pi}(w), w) = b_0 - \widehat{\pi}(w)wE(x) - c\left(\widehat{\pi}(w)\right). \tag{53}$$

The profit function in (53) may not be monotonic in w, as a higher w increases $\widehat{\pi}(w)$ and then decreases $c(\widehat{\pi}(w))$. If $\pi^*(w) = \widehat{\pi}(w)$, the profit function is the same as social welfare, Recall that social welfare is concave in π and reaches its maximum at π^{**}. Therefore, given $\widehat{\pi}(w)$ strictly increasing in w, the profit function in (53) decreases in w if $\widehat{\pi}(w) > \pi^{**}$, and increases in w if $\widehat{\pi}(w) < \pi^{**}$. □

Now to prove Proposition 2, we consider four cases in turn.

Suppose first $b \geq b^H = \pi^H(0)h$. In this case, $\pi^L(0) < \pi^H(0) \leq \widehat{\pi}(0)$. Because $d\widehat{\pi}(w)/dw > 0$, $d\pi^H(w)/dw < 0$ and $\pi^L(w) < \pi^H(w)$ for all $w \in [0, h)$ from Lemma 2, we have $\pi^L(w) < \pi^H(w) < \widehat{\pi}(w)$ for all $w \in [0, h)$. By Claim 1, $\pi^*(w) = \pi^H(w)$. Given Claim 2, the firm's optimal stipulated-damage payment is the lower bound of the set, $w^* = 0$. Accordingly, the firm is under-investing in safety, $\pi^*(0) = \pi^H(0) > \pi^{**}$.

Second, suppose $b \in (b^{**}, b^H)$. In this case, $\pi^L(0) < \widehat{\pi}(0) < \pi^H(0)$. Therefore, $\pi^*(0) = \widehat{\pi}(0) > \pi^{**}$. As $\widehat{\pi}(w)$ increases in w, we have $\widehat{\pi}(w) > \pi^{**} \geq \pi^L(w)$ for any $w \in [0, h]$. Then Claim 1 implies that, for any $w \in [0, h]$, the firm's optimal safety choice is either $\widehat{\pi}(w) > \pi^{**}$ or $\pi^H(w) > \pi^{**}$. But for both choices, Cliam 2 suggests that the firm's optimal stipulated-damage payment is the lower bound of the set, $w^* = 0$. Accordingly, the firm is under-investing in safety.

Third, suppose $b \in (0, b^{**}]$. If the firm chooses $\pi = \pi^{**}$ and $w^* = h - b/\pi^{**} \in [0, h)$ then $S(\pi^{**}, w^*) = W(\pi^{**})$, that is, firm profit equals the maximal social surplus.

Finally, suppose that $b \leq 0$. Then for all $w < h$, $\widehat{\pi}(w) \leq 0 \leq \pi^L(w) < \pi^H(w)$. By Claim 1, $\pi^*(w) = \pi^L(w)$. Given Claim 2, the firm's optimal stipulated-damage payment is the upper bound of the set, $w^* = h$. Accordingly, the firm invests efficiently, $\pi^*(h) = \pi^{**}$. □

Proof of Proposition 4. Following the logic of the last section, the price charged by the firm is equal to $p = b_0 + bx^M - \widetilde{\pi}(w)(h - w)x^M$ where $x^M \in \{\underline{x}, \overline{x}\}$ is the marginal consumer. The firm's profit is the price minus the firm's unit cost $w\widetilde{\pi}(w)E(x) + c(\widetilde{\pi}(w))$:

$$b_0 + bx^M - \widetilde{\pi}(w)(h - w)x^M - [w\widetilde{\pi}(w)E(x) + c(\widetilde{\pi}(w))]. \tag{54}$$

Taking the derivative with respect to w gives

$$-\widetilde{\pi}'(w)(h - w)x^M + \widetilde{\pi}(w)x^M - \widetilde{\pi}(w)E(x) - \widetilde{\pi}'(w)[wE(x) + c'(\widetilde{\pi}(w))]. \tag{55}$$

The last term in the square bracket is zero (because the firm chose $\widetilde{\pi}(w)$ to minimize the unit costs, so the slope of the firm's profit function simplifes to:

$$-\widetilde{\pi}'(w)(h - w)x^M + \widetilde{\pi}(w)(x^M - E(x)). \tag{56}$$

Suppose first that $b > 0$. Recall from Lemma 2 that $\widehat{\pi}'(w) > 0$ and that it approaches positive infinity as w approaches h. Because $\widetilde{\pi}'(w) < 0$, the two curves cross at most once.

When $\widetilde{\pi}(0) < \widehat{\pi}(0) = b/h$ the curves do not cross at all: $\widetilde{\pi}(w) < \widehat{\pi}(w)$ for all $w \in [0, h)$. In this case, $x^M = \underline{x}$ is the marginal consumer. The slope of the firm's profit function is

$$-\widetilde{\pi}'(w)(h - w)\underline{x} + \widetilde{\pi}(w)(\underline{x} - E(x)). \tag{57}$$

Given $-wE(x) - c'(\widetilde{\pi}(w)) = 0$, we have $\widetilde{\pi}'(w) = \frac{-E(x)}{c''(\widetilde{\pi}(w))}$. Then the slope of the profit function can be rewritten as

$$\frac{E(x)}{c''(\widetilde{\pi}(w))}(h - w)\underline{x} + \widetilde{\pi}(w)(\underline{x} - E(x)). \tag{58}$$

When $w = 0$, the slope is $\frac{E(x)}{c''(\widetilde{\pi}(0))}h\underline{x} + \widetilde{\pi}(0)(\underline{x} - E(x))$, which can be positive or negative, depending on the distribution of x. When $w = h$, the slope is strictly negative. So there is a value $w^* \in [0, h)$ that maximizes firm profits.

When $\widetilde{\pi}(0) \geq \widehat{\pi}(0) = b/h$ then $\widetilde{\pi}(w)$ and $\widehat{\pi}(w)$ cross exactly once. Let ω be the value where $\widetilde{\pi}(\omega) = \widehat{\pi}(\omega)$. If $w < \omega$, then $\widetilde{\pi}(w) > \widehat{\pi}(w)$ and so $x^M = \overline{x}$ is the marginal consumer. Then, the slope of the profit function is:

$$-\widetilde{\pi}'(w)(h - w)\overline{x} + \widetilde{\pi}(w)(\overline{x} - E(x)). \tag{59}$$

Because $\widetilde{\pi}'(w) < 0$ and $\overline{x} - E(x) > 0$ this slope is strictly positive for all $w < \omega$. So the firm would want to raise $w = \omega$.

Now suppose $w \geq \omega$. As above, $\widetilde{\pi}(w) < \widehat{\pi}(w)$ and so $x^M = \underline{x}$ is the marginal consumer. The slope of the firm's profit function is

$$-\widetilde{\pi}'(w)(h - w)\underline{x} + \widetilde{\pi}(w)(\underline{x} - E(x)). \tag{60}$$

When $w = h$ the first term of (60) is equal to zero and the second term is strictly negative. Therefore the firm will set $w = \omega < h$. In particular, if $\widetilde{\pi}(0) > b/h$, $w = \omega > 0$.

Finally, suppose that $b \leq 0$. Then for all $w < h$, $\widetilde{\pi}(w) \geq 0 \geq \widehat{\pi}(w)$ and so $x^M = \overline{x}$ is the marginal consumer. Then, the slope of the profit function is:

$$-\widetilde{\pi}'(w)(h - w)\overline{x} + \widetilde{\pi}(w)(\overline{x} - E(x)). \tag{61}$$

Because $\widetilde{\pi}'(w) < 0$ and $\overline{x} - E(x) > 0$ this slope is strictly positive for all $w < h$. So the firm would want to offer $w = h$. □

Proof of Proposition 5. Following the analysis in this section, the left-hand side of condition (21) decreases in w and the left-hand side of condition (22) increases in w. That is, $S_{\pi w}(w, x^{M*}, \pi^*) = \frac{d^2 S(w,x^M,\pi)}{d\pi dw} < 0$ and $S_{xw}(w, x^{M*}, \pi^*) = \frac{d^2 S(w,x^M,\pi)}{dx^M dw} > 0$. It can also be verified that $S_{\pi\pi}(w, x^{M*}, \pi^*) < 0$, $S_{xx}(w, x^{M^*}, \pi^*) < 0$ and $S_{\pi x}(w, x^{M*}, \pi^*) < 0$. The second order condition also implies that the Hessian matrix is definite-negative, that is,

$$S_{\pi\pi}(w, x^{M*}, \pi^*)S_{xx}(w, x^{M^*}, \pi^*) - [S_{\pi x}(w, x^{M*}, \pi^*)]^2 > 0. \tag{62}$$

Condition (21) implies that, at the optimal solutions,

$$\begin{aligned} 0 &= S_{\pi w} + S_{\pi\pi}\frac{d\pi^*}{dw} + S_{\pi x}\frac{dx^{M*}}{dw} \\ &= S_{xx}S_{\pi w} + S_{xx}S_{\pi\pi}\frac{d\pi^*}{dw} + S_{xx}S_{\pi x}\frac{dx^{M*}}{dw}. \end{aligned} \tag{63}$$

Similarly, condition (22) implies that

$$\begin{aligned} 0 &= S_{xw} + S_{xx}\frac{dx^{M*}}{dw} + S_{\pi x}\frac{d\pi^*}{dw} \\ &= S_{\pi x}S_{xw} + S_{\pi x}S_{xx}\frac{dx^{M*}}{dw} + S_{\pi x}S_{\pi x}\frac{d\pi^*}{dw}. \end{aligned} \tag{64}$$

Therefore, we have

$$S_{\pi x}S_{xw} - S_{xx}S_{\pi w} = [S_{xx}S_{\pi\pi} - (S_{\pi x})^2]\frac{d\pi^*}{dw}. \tag{65}$$

Note that the left-hand side of condition (65) is negative and $S_{xx}S_{\pi\pi} - (S_{\pi x})^2 > 0$. Therefore, we have $\frac{d\pi^*(w)}{dw} < 0$. Then, given $0 = S_{xw} + S_{xx}\frac{dx^{M*}}{dw} + S_{\pi x}\frac{d\pi^*}{dw}$, we have $\frac{dx^{M*}(w)}{dw} > 0$.

Also note that, if $w = h$, condition (21) becomes $\int_{x^{M*}} (-hx - c'(\pi^*))\, dF(x) = 0$. We then have

$$\left.\frac{dW(w)}{dw}\right|_{w=h} = \frac{d\pi^*}{dw}\frac{dW}{d\pi^*} + \frac{dx^{M*}}{dw}\frac{dW}{dx^{M*}}$$

$$= \frac{d\pi^*}{dw}\int_{x^{M*}}[-hx - c'(\pi^*)]dF(x) + \frac{dx^{M*}}{dw}\left\{-b_0 - (b - \pi^* h)x^{M*} + c(\pi^*)]\right\} f(x^{M*})$$

$$= \frac{dx^{M*}}{dw}\left\{-b\left[1 - F(x^{M*})\right]\right\} < 0. \tag{66}$$

So the socially-optimal warranty is strictly less than h. As shown in the main text, the firm offers $w^* = 0$. Because $\frac{d\pi^*}{dw} < 0$ and $\frac{dx^{M*}}{dw} > 0$, if $w^{**} > 0$, we have $\pi^*(w^{**}) < \pi^*(0)$ and $x^{M**}(w^{**}) > x^{M*}(0)$. □

Proof of Proposition 6. To simplify notations, denote $t(x) = \pi(x)(h - w(x))$. The firm's optimization problem can be rewritten as

$$\underset{\pi(x),p(x),t(x)}{Max} \int_{\underline{x}}^{\overline{x}} [p(x) - c(\pi(x)) + t(x)x - \pi(x)hx] f(x)dx \tag{67}$$

subject to the IR and IC constraints:

$$p(x) + t(x)x \le b_0 + bx \quad \forall x \tag{68}$$

$$p(x) + t(x)x \le p(\widehat{x}) + t(\widehat{x})x \quad \forall x, \widehat{x} \tag{69}$$

$$t(x) \in [0, \pi(x)h] \quad \forall x. \tag{70}$$

Given $b > b^H = \pi^H(0)h$, we have $[b - \pi(x)(h - w(x))] > 0$ for any x. Therefore, the IR constraint (68) must be binding for the lowest type $\underline{x}$ but not binding for any $x > \underline{x}$. Also, the IC constraint (69) implies

$$-[t(x) - t(\widehat{x})]\widehat{x} \le p(x) - p(\widehat{x}) \le -[t(x) - t(\widehat{x})]x \quad \forall x, \widehat{x}. \tag{71}$$

Thus the constraints (68) and (69) can be re-written as

$$p(\underline{x}) = b_0 + [b - \pi(\underline{x})(h - w(\underline{x}))]\underline{x} \tag{72}$$

$$p'(x) = -t'(x)x \quad \forall x. \tag{73}$$

These two expressions imply that

$$p(x) = p(\underline{x}) + \int_{\underline{x}}^{x} p'(y)dy = p(\underline{x}) - \int_{\underline{x}}^{x} yt'(y)dy. \tag{74}$$

Substituting $p(x)$ into the objective function, and using Fubini's theorem to change the order of integration, the firm's profits are

$$\int_{\underline{x}}^{\overline{x}} \left[p(\underline{x}) - \int_{\underline{x}}^{x} yt'(y)dy - c(\pi(x)) + t(x)x - \pi(x)hx\right] f(x)dx$$

$$= p(\underline{x}) - \int_{\underline{x}}^{\overline{x}} xt'(x)(1-F(x))dx - \int_{\underline{x}}^{\overline{x}} [c(\pi(x)) - t(x)x + \pi(x)hx]\, f(x)dx. \quad (75)$$

Integrating this expression by parts, we let $u = x(1-F(x))$ and $dv = t'(x)dx$, so $du = ((1-F(x) - xf(x))dx$ and $v = t(x)$. And using the binding IR constraint for the lowest type $\underline{x}$, we can re-write the firm's profits as

$$p(\underline{x}) + \underline{x}t(\underline{x}) + \int_{\underline{x}}^{\overline{x}} t(x)[1-F(x)-xf(x)]dx - \int_{\underline{x}}^{\overline{x}} [c(\pi(x)) - t(x)x + \pi(x)hx]\, f(x)dx$$

$$= b_0 + b\underline{x} + \int_{\underline{x}}^{\overline{x}} \left[\pi(x)(h - w(x))\frac{1-F(x)}{f(x)} - c(\pi(x)) - \pi(x)hx\right] f(x)dx. \quad (76)$$

□

Proof of Proposition 7. We can rewrite the IC constraint (33) as

$$\pi[\omega(x) - \omega(\widehat{x})]\widehat{x} \leq \rho(x) - \rho(\widehat{x}) \leq \pi[\omega(x) - \omega(\widehat{x})]x \quad \forall x, \widehat{x}. \quad (77)$$

This implies that $\omega'(x) \geq 0$ and $\rho'(x) \geq 0$. Dividing by $x - \widehat{x}$, and taking the limit as $x - \widehat{x}$ approaches zero gives

$$\rho'(x) = \pi\omega'(x)x. \quad (78)$$

Integrating the right-hand side by parts gives $\rho(x) = \rho(\underline{x}) - \pi\omega(\underline{x})\underline{x} + \pi\omega(x)x - \pi\int_{\underline{x}}^{x} \omega(y)dy$, and rearranging terms, the effective price paid by type x is:

$$\rho(x) - \pi\omega(x)x = \rho(\underline{x}) - \pi\omega(\underline{x})\underline{x} - \pi\int_{\underline{x}}^{x} \omega(y)dy. \quad (79)$$

The derivative of the right-hand side of (79) is $-\pi\omega(x) \leq 0$. So the effective price, $\rho(x) - \pi\omega(x)x$, is weakly decreasing in the consumer's type, x.

Substituting (79) into (32) and (34), we can rewrite the firm's problem as choosing π, $\rho(\underline{x})$, and $\omega(x)$ to maximize

$$\int_{\underline{x}}^{\overline{x}} \left[\rho(\underline{x}) - \pi\omega(\underline{x})\underline{x} - \pi \int_{\underline{x}}^{x} \omega(y)dy - c(\pi) \right] f(x)dx \tag{80}$$

subject to

$$b_0 + bu(x) - \pi hx - \left[\rho(\underline{x}) - \pi\omega(\underline{x})\underline{x} - \pi \int_{\underline{x}}^{x} \omega(y)dy \right] \geq 0 \quad \forall x. \tag{81}$$

The left-hand side of (81) are the rents received by a consumer of type x. The first derivative is $bu'(x) - \pi(h - \omega(x))$.

Let $x^M \in [\underline{x}, \overline{x}]$ be the value (not necessarily unique) where the left-hand side of (81) is minimized. It must be the case that the IR constraint binds at x^M. If not, then the firm could increase profits by lowering the price $\rho(\underline{x})$. Because the IR constraint (81) holds with equality when $x = x^M$ we have

$$\rho(\underline{x}) - \pi\omega(\underline{x})\underline{x} = b_0 + bu(x^M) - \pi hx^M + \pi \int_{\underline{x}}^{x^M} \omega(y)dy. \tag{82}$$

Similar to the proof of Lemma 1, using (82) and Fubini's theorem, we can rewrite (80) as

$$b_0 + bu(x^M) - \pi hx^M + \pi \int_{\underline{x}}^{x^M} \omega(x)F(x)dx - \pi \int_{x^M}^{\overline{x}} \omega(x)(1 - F(x))dx - c(\pi). \tag{83}$$

Suppose that $b > \frac{h\pi^H(0)}{\min\{u'(x)\}}$. Then for any $\pi \leq \pi^H(0)$ and $\omega(x) \geq 0$, the derivative of consumer rent is strictly positive, that is, $bu'(x) - \pi(h - \omega(x)) > 0$ for all x. Therefore, the marginal consumer must be $x^M = \underline{x}$ and the IR constraint is not binding for any $x > \underline{x}$. In the optimal mechanism, we must have $\omega(x) = 0$ for all x. If otherwise $\omega(x) > 0$ for some x, then by replacing $\omega(x)$ with a new scheme $(1-\varepsilon)\omega(x)$, where ε is arbitrarily close to 0, the firm can increase its profit (83). Note that the IR constraint under the original scheme $\omega(x)$ is not binding for all $x > \underline{x}$. Therefore, under the new scheme $(1 - \varepsilon)\omega(x)$ the IR constraint (81) still holds. Because $-(1 - \varepsilon)\omega'(x) \leq 0$, the IC constraint also holds under the new scheme.

Given $x^M = \underline{x}$ and $\omega(x) = 0$ for all x, we can rewrite (83) as

$$b_0 + bu(\underline{x}) - \pi h\underline{x} - c(\pi), \tag{84}$$

which is maximized by $\pi^H(0) > \pi^{**}$. □

Online Appendix – Not for Publication

In our benchmark model, only the firm can mitigate the accident risks. This appendix presents two extensions where consumers also take precautions to mitigate the accident risks. We prove that the private and social incentives to compensate consumers for their injuries diverge. As in our benchmark model, the social planner mandates a damage payment that is strictly higher than the privately-stipulated damage payment.

Consumer Moral Hazard with Variable Costs

In this extension, we assume that consumer precautions involve variable costs. Consumers put forth (unobservable and non-contractible) effort, e, each time they use the product. As in our benchmark model, quality π is observable to consumers at the time of purchase (so there is no moral hazard problem for the firm), and b_0 is sufficiently large for full market coverage.

Suppose that the probability of an accident per use is $\pi - e$ and the consumer's incremental benefit is $b - k(e)$ where $k(e)$ is the consumer's cost of effort per use. This is motivated by the idea that the consumer is choosing the precautions e each and every time they use the product. Assume that $k'(e) > 0, k''(e) > 0$, and $Lim_{e \to 0} k'(e) = 0$. We will also restrict attention to situations where $\pi - e > 0$ to avoid corner solutions. The social welfare function is

$$W(\pi, e) = b_0 + (b - k(e))E(x) - (\pi - e)hE(x) - c(\pi). \tag{85}$$

Social welfare is maximized at π^{**} and e^{**} where $-hE(x) - c'(\pi^{**}) = 0$, the same value defined in our benchmark model, and e^{**} satisfies $-k'(e^{**}) + h = 0$.

The consumer's net benefit from buying the product is

$$b_0 + (b - k(e))x - (\pi - e)(h - w)x - p. \tag{86}$$

The consumer chooses effort e to maximize this expression, so the consumer's effort $e^*(w)$ is implicitly defined by

$$-k'(e^*(w)) + (h - w) = 0. \tag{87}$$

Note that $e^*(w)$ does not depend on the consumer's type, x, and is a decreasing function of w with $e^*(0) = e^{**} > 0$ and $e^*(h) = 0$. In other words, the consumer chooses the socially-optimal effort if $w = 0$, and puts in no effort if the damage payment is fully compensatory, $w = h$.

With consumer precautions e and marginal consumer type $x^M \in \{\underline{x}, \overline{x}\}$, the firm's price is

$$p = b_0 + (b - k(e))x^M - (\pi - e)(h - w)x^M. \tag{88}$$

The firm's profit function is

$$S(\pi, e, w) = b_0 + (b - k(e))x^M - (\pi - e)(h - w)x^M - (\pi - e)wE(x) - c(\pi). \quad (89)$$

Note that firm profits are maximized at $\pi^*(w)$ where $-hx^M - w(E(x) - x^M) - c'(\pi^*(w)) = 0$. The firm's choice of safety, $\pi^*(w)$, is the same as in the benchmark model without consumer moral hazard. The firm's profit function can be written as:

$$S(\pi, e, w) = W(\pi, e) - (E(x) - x^M)[b - k(e) - (\pi - e)(h - w)]. \quad (90)$$

Firm profits equal social welfare minus the rents paid to the infra-marginal consumers. Thus, as before, the private and social incentives to invest in product safety diverge.

As in the benchmark model, when the consumer's incremental benefit b is sufficiently large, then the consumer's net benefit will be increasing in x for all parameter values and so $x^M = \underline{x}$. In this case, we have the following result.

Proposition. (Consumer Moral Hazard with Variable Costs.) *Suppose that the incremental benefit b is sufficiently high.*[64] *The firm chooses $w^* = 0$ and under-invests in product safety, $\pi^*(0) > \pi^{**}$. The consumers invest efficiently, $e^*(0) = e^{**}$. In contrast, the social planner would choose $w^{**} \in (0, h)$. The firm invests more and the consumers invest less (compared to the unregulated outcome).*

Proof of Proposition. Following the analysis in this section, we have

$$\begin{aligned} \frac{dS}{dw} &= -k'(e^*(w))\underline{x}\frac{de^*(w)}{dw} + (h - w)\underline{x}\frac{de^*(w)}{dw} \\ &\quad +(\pi^*(w) - e^*(w))\underline{x} - (\pi^*(w) - e^*(w))E(x) + wE(x)\frac{de^*(w)}{dw} \\ &= (\pi^*(w) - e^*(w))[\underline{x} - E(x)] + wE(x)\frac{de^*(w)}{dw} \\ &< 0, \end{aligned} \quad (91)$$

where the second equality follows from the definition of $e^*(w)$ and the inequality holds given $\frac{de^*(w)}{dw} < 0$. Therefore, the firm will set $w^* = 0$. Accordingly, the firm will choose $\pi^*(0) = \pi^H(0) > \pi^{**}$ and consumers will choose $e^*(0) = e^{**}$. That is, the firm under-invests in safety and consumers put in the socially-optimal precaution.

[64] A sufficient condition is that $b > k(1) + b^H$, where $b^H = \pi^H(0)h$. It can be verified that $x^M = \underline{x}$.

Now consider a social planner who chooses w but cannot regulate the firm's investment or the consumers' effort. The social welfare is

$$W(w) = b_0 + (b - k(e^*(w))E(x) - (\pi^*(w) - e^*(w))hE(x) - c(\pi^*(w)). \qquad (92)$$

Note that $-k'(e^*(w)) = -(h-w)$ and $c'(\pi^*(w)) = -h\underline{x} - w(E(x) - \underline{x})$. Thus,

$$\begin{aligned}\frac{dW}{dw} &= -k'(e^*(w))E(x)\frac{de^*(w)}{dw} - hE(x)\frac{d\pi^*(w)}{dw} + hE(x)\frac{de^*(w)}{dw} \\ &\qquad - c'(\pi^*(w))\frac{d\pi^*(w)}{dw} \\ &= -(h-w)E(x)\frac{de^*(w)}{dw} - hE(x)\frac{d\pi^*(w)}{dw} + hE(x)\frac{de^*(w)}{dw} \\ &\qquad -[-h\underline{x} - w(E(x) - \underline{x})]\frac{d\pi^*(w)}{dw} \\ &= wE(x)\frac{de^*(w)}{dw} - (h-w)[E(x) - \underline{x}]\frac{d\pi^*(w)}{dw}. \qquad (93)\end{aligned}$$

Recall that $\frac{de^*(w)}{dw} < 0$ and $\frac{d\pi^*(w)}{dw} < 0$. If $w = 0$, $\frac{dW}{dw} > 0$; if $w = h$, $\frac{dW}{dw} < 0$. Therefore, the social planner would set $w^{**} \in (0, h)$.

□

When the consumers' incremental benefit b is sufficiently high, then an unregulated firm will not offer damage payments at all. With $w = 0$, the consumers will invest efficiently each and every time they use the risky product. This benefits the firm's bottom line, because as a monopolist the firm extracts the surplus created from the consumer's higher effort. Raising the damage payment would be counterproductive for the firm. As in the main section of the article, when b is high then the consumer's net benefit will be increasing in the consumer's type x. When $w = 0$, the firm avoids subsidizing the purchases of the infra-marginal consumers with high types. Note that because type $\underline{x}$ is marginal, the firm is under-investing in safety relative to the social optimum.

If the social planner could regulate the damage payment, it would choose a payment or liability that is greater than zero. It is not hard to see why this is socially beneficial. Starting at $w = 0$, if the social planner raises w slightly, then the consumer's effort will fall. However, because the consumer's effort is socially optimal when $w = 0$, the reduction in effort has a second-order effect on social welfare. But raising w has a first-order effect on the safety level chosen by the firm.[65]

[65] The social planner will choose $w < h$ in order to get some effort from the consumers.

Consumer Moral Hazard with Fixed Costs

We now assume that consumer precautions involve fixed costs. As in our benchmark model, quality π is observable to consumers at the time of purchase (so there is no moral hazard problem for the firm), and b_0 is sufficiently large for full market coverage.

Suppose that the probability of an accident per use is $\pi - e$ and the consumer's lump-sum precaution cost is $K(e)$. Assume that $K'(e) > 0, K''(e) > 0$, and $Lim_{e \to 0} K'(e) = 0$. We will also restrict attention to situations where $\pi - e > 0$ to avoid corner solutions.

The consumer's net benefit from buying the product is

$$b_0 + bx - K(e) - (\pi - e)(h - w)x - p. \tag{94}$$

The consumer chooses effort e to maximize this expression, so the consumer's effort $e^*(w; x)$ is implicitly defined by

$$-K'(e^*(w; x)) + (h - w)x = 0. \tag{95}$$

Note that $e^*(w; x)$ increases in x and decreases in w with $e^*(0; x) = e^{**}(x) > 0$ and $e^*(h; x) = 0$. In other words, the consumer (of type x) chooses the socially-optimal effort if $w = 0$, and puts in no effort if the damage payment is fully compensatory, $w = h$.

When b is sufficiently large, consumer net benefit increases in x. Therefore, the marginal consumer is of type $\underline{x}$. The firm's price is

$$p = b_0 + b\underline{x} - K(e^*(w; \underline{x})) - (\pi - e^*(w; \underline{x}))(h - w)\underline{x}. \tag{96}$$

The firm's profit function is

$$S(\pi, e, w) = b_0 + b\underline{x} - K(e^*(w; \underline{x})) - (\pi - e^*(w; \underline{x}))(h - w)\underline{x} - \int (\pi - e^*(w; x))wxdF(x) - c(\pi). \tag{97}$$

Proposition. (Consumer Moral Hazard with Fixed Costs.) *Suppose that the incremental benefit b is sufficiently high and $\pi^*(w) > \max_x [e^*(w; x) + \frac{K'(e^*(w;x))}{K''(e^*(w;x))}]$ for any $w \in [0, h]$. The firm chooses $w^* = 0$ and under-invests in product safety, $\pi^*(0) > \pi^{**}$. The consumers invest efficiently, $e^*(0; x) = e^{**}(x)$. In contrast, the social planner would choose $w^{**} \in (0, h)$. The firm invests more and the consumers invest less (compared to the unregulated outcome).*

Proof of Proposition. The profit-maximizing $\pi^*(w)$ satisfies $\frac{dS(\pi,e,w)}{d\pi} = 0$. Using the envelope theorem, we have

$$\begin{aligned}
\frac{dS}{dw} &= K'(e^*(w;\underline{x}))\frac{de^*(w;\underline{x})}{dw} - (h-w)\underline{x}\frac{de^*(w;\underline{x})}{dw} \\
&\quad +(\pi^*(w) - e^*(w;\underline{x}))\underline{x} - \int [(\pi^*(w) - e^*(w;x))x - w\frac{de^*(w;x)}{dw}]dF(x) \\
&= -\int [(\pi^*(w) - e^*(w;x))x - (\pi^*(w) - e^*(w;\underline{x}))\underline{x}]dF(x) \\
&\qquad + \int w\frac{de^*(w;x)}{dw}dF(x),
\end{aligned}$$

where the second term is negative but the first term can be positive or negative.

If $\pi^*(w) > \max_x [e^*(w;x) + \frac{K'(e^*(w;x))}{K''(e^*(w;x))}]$ given any $w \in [0,h]$, then we have $(\pi^*(w) - e^*(w;x))x$ increasing in x. In this case, $\frac{dS}{dw} < 0$ and therefore the firm sets $w^* = 0$. Moreover, the impact of w on social welfare is

$$\frac{dW}{dw} = \int w\frac{de^*(w;x)}{dw}dF(x) - \int (h-w)(x-\underline{x})\frac{d\pi^*(w)}{dw}dF(x).$$

Obviously, the socially efficient damage payment satisfies $w^{**} \in (0,h)$. To summarize, if $\pi^*(w) > \max_x [e^*(w;x) + \frac{K'(e^*(w;x))}{K''(e^*(w;x))}]$ given any $w \in [0,h]$, we have $w^{**} > w^* = 0$.

□