Liability Design with Information Acquisition

Francisco Poggi and Bruno Strulovici
Mannheim University and Northwestern University*

November 15, 2021

Abstract

How to guarantee that firms perform due diligence before launching potentially dangerous products? We study the design of liability rules when (i) limited liability prevents firms from internalizing the full damage they may cause, (ii) penalties are paid only if damage occurs, and (iii) firms have private information about their products' riskiness before performing due diligence. We show that (i) any liability mechanism can be implemented by a *tariff* that depends only on the evidence acquired by the firm if a damage occurs, not on any initial report by the firm about its private information, (ii) firms that assign a higher prior to product riskiness always perform more due diligence but less than is socially optimal, and (iii) under a simple and intuitive condition, any launch thresholds that are monotonic in the firm's type can be implemented by a monotonic tariff.

1 Introduction

In 2019, a California court sentenced paint maker Sherwin-Williams to pay hundreds of millions of dollars to address the damage caused by lead paint.[1] The sentence was remarkable because even though lead paint became banned in 1978, the suit concerned damage caused

*Emails: poggi@uni-mannheim.de and b-strulovici@northwestern.edu. Strulovici gratefully acknowledges financial support from the National Science Foundation (Grant No.1151410).

[1]See, e.g., "Paint makers reach $305 million settlement in California, ending marathon lead poisoning lawsuit," *Reuters*, July 17, 2019.

during the decades *before* the ban and centered on the accusation that paint makers were aware of the dangers caused by lead paint long before the ban was passed.

In essence, the court's argument was that Sherwin-Williams and other paint makers knew or should have known the dangers caused by lead paint.

While it is difficult for a regulator to guess a firm's private information, it may be easier to assess a firm's due diligence: for example, did paint makers research the risks of lead paint sufficiently well before marketing it? Formally, the problem is not just one of private information, but also one of information acquisition: how can a regulator make sure that firms acquire sufficient evidence before launching potentially harmful products? More generally, how can a regulator induce economic agents to learn sufficiently well the consequences of their actions before taking them?

We model this question as a *delegated Wald problem* (Wald (1945)): the principal is a regulator who relies on an agent (the firm) to acquire information before deciding between launching a product and abandoning it.

If the regulator could arbitrarily penalize the firm in case of damage, she could force the firm to internalize all damage caused by the product and implement the socially-optimal level of information acquisition. Likewise, if the regulator could perfectly observe a product's likelihood of causing damage at the time of launch, it could charge a fee at the time of the launch such that the firm fully internalizes the expected damage that its product may cause. In reality, however, regulators faces two major limitations.

First, firms—and the managers heading them—rarely pay the full damage caused by their actions, which reduces their incentives to take the socially optimal level of precaution. This limitation stems from various reasons, which include: (i) limited liability for managers; (ii) bankruptcy laws for firms; (iii) concentration of vested interest: defending firms typically have more at stake and more financial resources than other parties to engage in expensive and lengthy litigation and reduce the damage actually paid; (iv) burden of proof: firms pay damage only if they are found liable, which may be hard to prove. For example, if each damage caused by the firm can be proved with a fixed probability q and the real damage caused is L, then the firm's expected damage payment is qL; (v) short-termism by managers: managers in charge of corporate decisions do not fully internalize the risk of damage caused by their firm over the longer term, or they may discount future events more heavily than does their firm or a social planner.

Second, a regulator may be able to punish a firm only if some damage occurs, and unable to punish firms that behaved recklessly but were lucky enough that their risky product did not in fact cause harm. Although it would be ideal if regulators could charge firms based on the riskiness of their product regardless of damage occurrence, in reality firms suffer significant costs only if and when they cause damage. Examples abound of firms able to launch risky products, either getting a nominal amount of scrutiny by regulators before the launch, or eschewing any regulation at the time of the launch. In addition to lead paint, another prominent example concerns the use and discharge of perfluorooctanoic acid (PFOA) by DuPont, which went unregulated for years, caused health damage for thousands of individuals in Ohio and West Virginia, and resulted in large fines and cash settlements paid by DuPont.[2] The damage-based penalty structure is hard to avoid because firms have private information that is hard or too costly to evaluate by regulators at the time of product launch, and receives much stronger scrutiny only if and when damage is caused.

Taking these regulatory limitations into account, we analyze the problem of deterring the launch of risky products by firms that choose how much information and are subject to limited liability rules. Our model builds on a Brownian version of the Wald Problem: the firm observes a Brownian process whose drift depends on the riskiness of the product. Information acquisition is costly. The first-best policy is to acquire information until the riskiness of the product becomes sufficiently known and launch the product if this riskiness is low and abandon it if the riskiness is high.

We characterize all incentive-compatible liability rules when (i) the firm has initial private information, (ii) liability is capped, and (iii) the regulator can penalize the firm only when damage occurs.[3]

The set of possible mechanisms is a priori large. For example, one could reward a firm announcing ex ante that its product is risky, by giving it a lower penalty in case of damage if it can demonstrate that it did enough due diligence and a harsher penalty if due diligence was too weak, compared to a firm announcing that its product is less risky, which would face a moderate penalty for a large range of due diligence levels.

A key question in this context concerns whether the regulator should try to extract more pri-

[2]See, e.g., "DuPont, Chemours settle PFOA lawsuit for $670m," *Chemical Watch*, February 16, 2017.

[3]The model can easily accommodate an approval threshold, such that the firm can only launch the product if it can demonstrate that the evidence acquired in favor of safety exceeds a particular threshold. This possibility, briefly discussed at the end of Section 2, does not affect the gist of the analysis.

vate information for the firm at the outset. The regulator may wish to propose at the outset a menu of contracts to the firm in order to extract some of the firm's private information. Indeed, this approach is the one theoretically suggested by the Revelation Principle. This approach would be difficult to implement, because it requires that the firm contracts with the regulator long before launching the product and, in fact, even before knowing whether the firm wishes to launch the product.

Fortunately, our first main result is that it is without loss generality for the regulator to focus on *tariff mechanisms*, which are mechanisms for which the firm does not report its private information and only pays a penalty if damage occurs. This result may be viewed as a Taxation Principle for situations in which transfers take place only after some contingencies (damage occurs), but not others, and builds on our companion paper (Poggi and Strulovici (2020)), which provides a general Taxation Principle with Non-Contractible Events.

With a tariff mechanism, a firm's decision to launch the product depends on its prior information, which affects the probability that the product causes damage. Our second main result is that any incentive-compatible tariff mechanism has the following property: firms whose initial private information assigns a higher probability of damage always acquire more evidence before launching their product. This monotonicity property is not an immediate consequence of incentive compatibility, and would in fact be violated if the regulator could impose evidence-based transfers to the firm regardless of whether a damage occurred.

Our third main result is to show that any launch thresholds that induce the firm to perform more due diligence that it would under a fixed penalty can be implemented by a monotonic tariff, i.e., a tariff whose penalty is decreasing in the strength of evidence acquired by the firm before launching the product.

We also show that for a general specification of the regulator's objective function, setting the tariff at its uniform ceiling induces too little due diligence compared to the social optimum, even when the social benefit from launching the product exceeds the firm's profit from doing so. This result holds under a cost-benefit ratio condition, which stipulates that the social benefit from the product relative to the harm it may cause is smaller that the firm's profit relative to the maximum liability that it may face.

1.1 Literature

This paper contributes to the study of liability in the context of costly information acquisition about risks. Shavell (1992) studies how different liability rules affect the incentives to take precautions and acquire information about risks. In his model, an agent can pay a fixed cost to learn whether a risk exists or not. Shavell compares and analyses strict liability and four other negligence rules. Baumann and Friehe (2015) and Goeschl and Pfrommer (2015) study agents that learn about the risks of an activity or product through experimentation. Goeschl and Pfrommer (2015) find negligence rules to be superior to strict liability rules when the technology is learned-by-doing. Baumann and Friehe (2015) show that both strict liability and negligence rules fail to provide incentives for the agent to take precautions in a socially optimal way. In our paper, we analyse the general question of what can be achieved with any rule that maps what the regulator is able to observe after some damage occurs to the amount that the firm has to pay in compensation.

Some recent papers study settings in which a firm acquires information that is used in the approval process of a product. Friehe and Schulte (2017) study information acquisition when firms require the approval of a regulator to launch a potentially risky product. In their model, the regulator cannot commit to an approval rule based on the evidence presented. Instead, the paper focuses on how the equilibrium outcome changes under different liability rules. Henry et al. (2021) considers a mixture of ex ante and ex post interventions. They study the trade-off between acquiring information about risks of an activity before or when the activity is taking place. In their model, the planner commits to a liability rate per unit of time that the product is on the market. In this paper, instead, we allow the planner to commit to any liability rule that satisfies certain feasibility conditions.

At a broader level, liability rules may be viewed as one of several instruments used to curb risky activities. Shavell (1984) contrasts liability with regulation and observes that liability is potentially cheaper since administrative costs are incurred only if some harm occurs—an event that may have a low ex ante probability. Kolstad et al. (1990) and Schmitz (2000) show that, in settings with heterogeneous agents, a combination of regulation and liability may achieve a more efficient outcome than liability alone. This paper focuses on liability rules, but it would be easy to incorporate other instruments, such as an approval rule requiring that the evidence about a product's safety at the time of its launch be sufficiently strong.

Finally, our results complement recent papers that study the problem of learning before a

irreversible decision. In Wald (1945)'s classical framework, the statistician controls both the research process and the final decision about the product's launch or abandonment. Henry and Ottaviani (2019) and McClellan (2019) study a continuous-time version of the Wald framework in which the research is controlled by a firm but whether the product is launched depends on a decision carried out by a regulator. In our paper, the decision of whether to launch the product or not is carried out exclusively by the firm. The regulator, however, can impose costs to firms that launch the product. The expected costs that the regulator can impose on a firm that launches de product are bounded because the firm can only pay up to a certain maximal amount and only when accident actually occurs.

2 Model

A firm must decide between launching a product and abandoning its development. If launched, the product may cause damage with positive probability. The firm has some private information about the product's riskiness and can acquire additional information ("due diligence"), before making a final decision.

A regulator wishes to encourage the launch of low-risk products and deter the launch of high-risk ones, as well as to encourage the firm to acquire sufficient information before making its decision.

The regulator faces two constraints. First, the firm has limited liability: the social cost caused by product damage is $L > 0$ and the firm's liability is capped at some lower level $l < L$. Second, the regulator can penalize the firm only if damage occurs. In particular, it cannot penalize firms that acquired too little information and took an overly risky decision unless such risk results in actual damage.

The timing of the game is as follows:

1. The firm is endowed with a prior $\theta \in \Theta \subset [0,1]$ about the product's riskiness $y \in \{0,1\}$, with $\theta = \Pr(y = 1)$.

2. The firm can acquire additional information about y according to a dynamic technology to be described shortly.

3. The firm decides between launching and abandoning the product.

4. If the firm launches the product, it causes some damage if the product was faulty ($y = 1$) and doesn't if the product was safe ($y = 0$).

5. In case of damage, the firm pays a penalty $\psi \leq l$ set by the regulator.

The assumption that a faulty product causes damage with probability 1 is without loss of generality: if this probability were less than 1, the same analysis would apply using expected damage and expected penalties.

Information structure: During the information-acquisition stage, the firms observes a process X given by

$$X_t = (-1 + 2y)t + \sigma B_t$$

where B is the standard Brownian motion. The drift of X depend symmetrically on the product's riskiness y: the drift is $+1$ if the product causes damage and -1 if it does not. Therefore, observing X gradually reveals y. This revelation is progressive due to the stochastic component of X.

The firm stops acquiring information at some time τ that is adapted to the filtration of X.

The regulator observes nothing about X except if some damage occurs, in which case she observes the last value X_τ taken by the process at the time of the firm's decision. X_τ is a measure of the firm's due diligence to assess the product's riskiness before launching it: in this Brownian model, it is well-known (though not immediate) that for each $t > 0$, the variable X_t is a sufficient statistic for the information about y contained by the entire path $\{X_s\}_{s \leq t}$ of the process X until time t. Mathematically, the likelihood ratio of y associated with a path of X from time 0 to t is only a function of X_t.

Because the stopping time τ is chosen endogenously by the firm, which has private information about y, X_τ is not a sufficient statistic for y once the firm's strategic timing is taken into account. Our assumption that the regulator observes X_τ instead of the entire path $\{X_t\}_{t \leq \tau}$ captures the idea that the regulator does not perfectly observe all the decisions made by the firm during the information acquisition stage. Intuitively, the regulator observes the most informative signal about y contained by the path of X that is independent of the firm's private information.

Payoffs: The firm incurs a running cost c from acquiring information, and a profit π if it launches the product. Let $d = 1$ if the firm launches the product and $d = 0$ if it abandons

it, and τ denote the time spent acquiring information. The firm's realized payoff is

$$d(\pi - y\psi) - c\tau$$

where π is the firm's profit from the launch in the absence of damage. The regulator's objective internalizes the entire damage caused by the product:

$$d(\beta - yL) - c\tau$$

where β is the social benefit from the launch in the absence of damage.

Throughout the paper, we make the following assumption:

Assumption 1 (Ordered Cost-Benefit Ratios) $l/\pi < L/\beta$.

This assumption captures the idea that the risk of damage is more severe for the regulator relative to the benefit of launching the product than it is for the firm. The assumption allows the social benefit from launching the product to exceed the firm's profit (i.e., $\beta > \pi$).

Remark 1 *We could combine the liability rules studied in this paper with the following approval stage: before launching the product, the firm must demonstrates that the evidence in favor of the product's safety exceeds some threshold, i.e., that* $X_\tau \leq x^{approval}$ *for some threshold* $x^{approval}$*. This condition would not affect the thrust of the analysis, which would be applied to penalty functions* ψ *over the restricted domain* $x \in (-\infty, x^{approval}]$ *instead of* $\mathbb{R}$*.*

3 Preliminary Analysis: Symmetric Information

First Best: If the regulator knew the firm's type θ and could dictate the firm's strategy, the optimal strategy would consist in launching the product if the process X drops below some lower threshold x_θ^* and abandoning it if X exceeds some upper threshold $\bar{x}_\theta^* \geq x_\theta^*$.

Tariffs: A tariff is a function $\psi : \mathbb{R} \to \mathbb{R}$ mapping evidence x to a penalty $\psi(x) \leq l$.[4] Given a tariff ψ, a firm with prior θ chooses a stopping time τ and a launch/abandonment decision

[4]We could impose the additional restriction that tariffs be nonnegative. This would not affect the analysis, except for Proposition 4, which considers the implementation of arbitrary ordered launch thresholds. Negative tariffs may be viewed as a subsidy for firms that are revealed to have performed particularly careful inspections before launching their products.

$d \in \{0, 1\}$ to maximize its expected utility

$$E\left[d(\pi - y\psi(X_\tau)) - c\tau \mid \theta\right]. \tag{1}$$

It is straightforward to check that the solution to this problem consists of cutoffs $\underline{x}_\theta^\psi < \bar{x}_\theta^\psi$ such that the firm acquires information until X reaches either of the cutoffs.

Limited liability affects incentives in two ways. First, since the firm does not fully internalize damages, it is willing to take riskier decisions than is socially optimal for a given belief about the product's safety. Second, the value of information is different. For example, if the tariff is $\psi \equiv 0$, the firm has no incentive to acquire any information and always launches its product immediately.

To appreciate the consequences of limited liability, suppose that the regulator sets the tariff uniformly equal to the allowed maximum: $\psi(x) \equiv l$ for all $x \in \mathbb{R}$. In this case, the firm launches the product if X drops below some cutoff $\underline{x}_\theta^l$ and abandons it if X reaches some upper cutoff $\bar{x}_\theta^l$.

This maximum penalty may motivate the firm to perform due diligence before launching the product, but the amount of due diligence is always strictly suboptimal, as the next result shows.

Proposition 1 (recklessness) *$\underline{x}_\theta^* < \underline{x}_\theta^l$ for all $\theta \in \Theta$.*

Proof. We fix some prior $\theta \in \Theta$ throughout the proof and let x^* and x^l denote the socially-optimal and firm-optimal launch thresholds, respectively, when $\psi \equiv l$, given prior θ.

Given a current evidence level x, the firm's expected payoff if it launches the product at x is:

$$u(x) = \pi - p(x)l$$

where $p(x) = \Pr(y = 1|x, \theta)$. The regulator's expected payoff if the firm stops at x is:

$$v(x) = \beta - p(x)L.$$

Assumption 1 implies that

$$v(x) = \frac{L}{l}(u(x) - k) \tag{2}$$

where $k = \pi - \beta l/L > 0$.

Thus, the "launch-payoff functions" faced by the regulator and the firm are related by equation (2), and both parties face a running cost c before launching or abandoning the

product and a payoff normalized to zero if the product is abandoned. Proposition 1 then follows from two observations:

Observation 1: Consider two launch-payoff functions $\hat{u}, u$. If $\hat{u} = \alpha u$ with $\alpha > 1$, then the optimal launch threshold for $\hat{u}$ is lower than the optimal launch threshold for u.

Observation 2: Consider two launch-payoff functions $\hat{u}, u$. If $\hat{u} = u - \hat{k}$ with $\hat{k} > 0$, the optimal launch threshold for $\hat{u}$ is lower than the optimal launch threshold for u.

Once we justify these observations, Proposition 1 follows from (2) by applying Observation 2 to $u-k$ and u and Observation 1 to $v = L/l(u-k)$ and $u-k$, using the fact that $L/l > 1$.

To prove Observation 1, notice that if $\hat{u} = \alpha u$ with $\alpha > 1$, the dynamic optimization problem with launch payoff $\hat{u}$ and running cost c is equivalent to the problem with launch payoff u and running cost $\hat{c} = c/\alpha < c$, since the problems become identical up to the scaling factor α. With a lower running cost $\hat{c}$, the continuation interval $(\underline{x}(\hat{u}), \bar{x}(\hat{u}))$ contains the continuation interval $(\underline{x}(u), \bar{x}(u))$ with running cost c. In particular, the launch thresholds are ranked: $\underline{x}(\hat{u}) \leq \underline{x}(u)$.

To prove Observation 2, consider the optimal continuation interval $(x^l, \bar{x})$ when the launch-payoff function is u and let $\tau = \inf\{t : X_t \notin (x^l, \bar{x})\}$. Fixing any $x \in (x^l, \bar{x})$, acquiring information is optimal when starting at x, which means that

$$u(x) \leq f(x)u(x^l) - cE_x[\tau] \tag{3}$$

where $f(x)$ is the probability that $X_\tau = x^l$ (as opposed to $\bar{x}$) and $E_x[\tau]$ is the expected value of τ when the process X starts at x. For the launch-payoff function $\hat{u} = u-k$ with $k > 0$, (3) implies that

$$\hat{u}(x) < f(x)\hat{u}(x^l) - cE_x[\tau].$$

This shows that stopping at x to launch the product is strictly dominated by the strategy that consists in launching the product if X reaches x^l and abandoning it X reaches $\bar{x}$. This implies that the optimal launch threshold with $\hat{u}$ is lower than x^l and proves Observation 2. ■

Intuitively, Proposition 1 captures the idea that the regulator values more than the firm having a safer product conditional on launch. Remarkably, however, this result holds even when the social benefit from launching the product exceeds the firm's profit from doing so.

Although the uniform tariff $\psi \equiv l$ brings the firm closest to fully internalizing the damage

that its product might cause, the regulator might choose a different tariff, for example, to reward the firm if it acquired more information. The next section studies the firms' incentives in more details.

4 From Mechanisms to Tariffs

Suppose that the regulator can contract with the firm after the firm has received its initial private information and before it takes any action, and that the regulator has full commitment power.

DEFINITION 1 *A* direct liability mechanism *is a menu* $M = (\{\tau_\theta, d_\theta, \psi_\theta\}_{\theta\in\Theta})$ *such that for all* $\theta \in \Theta$*:*

(i) The stopping time τ_θ *is measurable with respect to the filtration* $\{\mathcal{F}_t^X\}_{t\geq 0}$ *generated by* X*;*

(ii) The decision d_θ *is measurable with respect to the information at time* τ*, i.e., to the* σ*-algebra* $\mathcal{F}_{\tau_\theta}^X$*;*

(iii) The tariff $\psi_\theta : \mathbb{R} \to \mathbb{R}$ *is uniformly bounded above by* l*.*

Since the regulator has full commitment power, the Revelation Principle guarantees that it is without loss of generality to focus on direct liability mechanisms.

Given a direct liability mechanism, the firm chooses an item $f_{\hat\theta} = (\tau_{\hat\theta}, d_{\hat\theta}, \psi_{\hat\theta})$ from the menu. Faced with the tariff $\psi = \psi_{\hat\theta}$, the firm chooses a stopping time and a decision to maximizes its expected utility as given by (1).

DEFINITION 2 *A direct liability mechanism* M *is* incentive compatible *if for each* $\theta \in \Theta$ *it is optimal to chooses the item* f_θ *from* M *and the strategy* (τ_θ, d_θ)*.*

In general, a direct liability mechanism may implement absurd policies: for example, the firm could get a very high reward (i.e., a negative penalty) if it launches the product when X_t is very high (and, hence, the product is very risky). We rule out such a possibility and focus on *admissible* mechanisms:

DEFINITION 3 *An IC direct liability mechanism is* admissible *if each type θ's strategy is characterized by thresholds $\underline{x}_\theta \leq \bar{x}_\theta$ such that θ launches the product if X_t drops below $\underline{x}_\theta$ and abandons it if X_t exceeds $\bar{x}_\theta$.*

In practice, it may be difficult for a regulator to contract with the firm at the outset and agree on penalties that depend finely on a firm's private information before it launches a product and, even earlier, before the firm decides how much due diligence to perform before deciding whether to launch its product. It is therefore valuable to determine when a direct liability mechanism can be implemented by a tariff that is independent of the firm's private information.

DEFINITION 4 *A direct liability mechanism is a* tariff mechanism *if the tariffs $\{\psi_\theta\}_{\theta\in\Theta}$ are independent of θ.*

THEOREM 1 *Any admissible direct liability mechanism is outcome-equivalent to a tariff mechanism.*

Proof. Consider any admisible mechanism M and let $\underline{x}_\theta = \underline{x}_\theta^{\psi_\theta}$ and $\psi_\theta = \psi_\theta(\underline{x}_\theta)$ denote the firm's launch threshold and penalty in case of damage that are implemented under mechanism M when the firm has type θ.

We introduce a ceiling mechanism $\tilde{M}$ as follows: for each θ, $\tilde{\psi}_\theta$ gives the maximal penalty l for all x except at $\underline{x}_\theta$, where it gives ψ_θ. The ceiling mechanism $\tilde{M}$ is IC and implements the same thresholds $\underline{x}_\theta$, because under M the firm faces the penalty only when it launches the product and higher penalties at other levels can only reduce the incentive to deviate.

If M prescribes the same threshold $\underline{x}$ to types $\theta \neq \theta'$, the penalties ψ_θ and ψ'_θ must be identical. Otherwise, one type would want to misreport its type and M would not be incentive compatible.

We define the tariff ψ as follows:

$$\psi(\underline{x}_\theta) = \psi_\theta$$

for all $\theta \in \Theta$ and

$$\psi(x) = l$$

otherwise.

This tariff is independent of the firm's private information. Moreover, it implements the same launch thresholds as M, as is easily checked. ■

Theorem 1 shows that any admissible liability mechanism can be implemented by a tariff. From now on, we invoke Theorem 1 and focus without loss of generality on admissible mechanisms that are implemented by tariffs, hereafter "admissible tariffs".

Given any admissible tariff $\psi : x \mapsto \psi(x)$, each type θ faces a Markovian decision problem in which the state variable at time t is X_t. Therefore, there exist thresholds $\underline{x}_\theta^\psi \leq \bar{x}_\theta^\psi$ such that type θ stops acquiring information when the process X leaves the interval $(\underline{x}_\theta^\psi, \bar{x}_\theta^\psi)$, launches the product at $\underline{x}_\theta^\psi$ and abandons it at $\bar{x}_\theta^\psi$.

Our next result establishes a single-crossing property for the firm.

LEMMA 1 *Consider any admissible tariff ψ, level x, and type $\theta \in \Theta$. If θ prefers acquiring information at x to immediately launching the product at x, then so does any type $\theta' \geq \theta$.*

Proof. We fix a tariff function ψ and a level x, and suppose that $X_t = x$ at some time t that we normalize to 0 for simplicity. Suppose that some type θ prefers the strategy that consists in launching the product at $\underline{x} < x$ and abandoning it at $\bar{x} > x$, and let $p = \Pr(y = 1|\theta)$.

If θ launches the product at x, it gets:

$$\pi - p\psi(x). \tag{4}$$

Let T^g, f^g denote the expected hitting time and the probability of hitting $\underline{x}$ if $y = 0$ (the product is safe or "good"), and T^b and f^b be defined similarly if $y = 1$ (the product is faulty or "bad"). If θ continues until hitting $\underline{x}$ or $\bar{x}$, its expected payoff is

$$p(f^b(\pi - \psi(\underline{x})) - cT^b) + (1 - p)(f^g \times \pi - cT^g). \tag{5}$$

Comparing (4) and (5), continuing is optimal if

$$p(f^b(\pi - \psi(\underline{x})) + \psi(x) - cT^b) + (1 - p)(f^g\pi - cT^g) \geq \pi. \tag{6}$$

The left-hand side is a convex combination of two terms: $a = f^b(\pi - \psi(\underline{x})) + \psi(x) - cT^b$ and $b = f^g\pi - cT^g$. The second term, b, is less than π, because f^g is a probability. Therefore, (6) can hold only if the first term, a, is greater than π.

Rewriting (6), a firm that assigns probability p to $y = 1$ wishes to continue if

$$p(a - b) \geq \pi - b.$$

Since $a > b$, the coefficient of p is strictly positive. This implies that any type that assigns probability $p' > p$ to $y = 1$ also prefers the continuation strategy to launching the product immediately at x. ■

Lemma 1 has the following intuition: If a firm knew that the product were safe, it would optimally launch the product immediately. The return to acquiring more evidence is negative in this case. Given any liability function, if a type wants to acquire more evidence it must be that doing so has a positive return conditional on the product being faulty. The expected return from acquiring more evidence is thus increasing in the probability that the firm assigns to the product being faulty.

Lemma 1 immediately implies the following monotonicity result:

PROPOSITION 2 *For any admissible tariff ψ, the launch thresholds $x^{\psi}(\theta)$ are decreasing in θ.*

This monotonicity result crucially hinges on the fact that the regulator can only charge the firm if it causes some damage. The following example[5] shows that if the regulator can charge the firm even when the product causes no damage, the launch thresholds increase with the type of the firm.

Example: Monotonicity Violation with Damage-Independent Fee

Suppose that all assumptions of the baseline model are maintained with one exception: if the firm launches its product, the regulator charges the firm a "liberating" fee $\eta(x) \in \mathbb{R}_+$ that depends on the evidence x demonstrated when the product is launched, and insulates the firm from any repercussion from damage subsequently caused by the product. This variation corresponds to a form of approval mechanism, where approval comes at a cost that depends on the evidence produced. In fact, the case in which $\eta(x) = 0$ for $x \leq x^*$ and $\eta(x) = +\infty$ for $x > x^*$, where $x^* < 0$ is an evidentiary threshold, corresponds to a standard approval mechanism.

Under this scenario, the principal observes X_τ and charges a fee $\eta(X_\tau)$ to the firm if the firm launches the product. While the firm is not liable for any damage incurred after the launch, the principal can nonetheless discourage reckless behavior by charging a prohibitively large fee to firm that acquire weak evidence.

[5]This example draws some inspiration from the approval mechanisms studied by **?** and Henry and Ottaviani (2019).

We focus on *admissible* fee functions, which are such that $\eta(x) > \pi$ for $x > 0$. This restriction implies that, starting at $x = 0$, the firm never launches the product if acquires bas information ($X_\tau > 0$) about the product. Thus, if a firm decides to acquire information, it will launch the product only if gets positive evidence ($X_\tau < 0$, and will abandon it at some positive threshold.

OBSERVATION 1 *Let $\{\underline{x}_\theta\}_{\theta\in\Theta}$ denote the optimal adoption thresholds induced by a fee schedule η. Then, $\eta(\underline{x}_\theta) = \min_{x\geq\underline{x}_\theta} \eta(x)$ for all $\theta \in \Theta$.*

This observation is straightforward: it means that a firm never wants to acquire more evidence before launching its product, if the fee corresponding to this additional evidence is higher than when the firm launches with a lower amount of evidence.

The next proposition shows that for admissible fees then the highest types acquire less information before launching the product.

PROPOSITION 3 *For any admissible fee $\eta(\cdot)$, the launch thresholds are increasing in θ.*

When the transfer is independent of whether the product is faulty or not, the only reason to gather information is to reduce the cost of launching the product (the fee). For admissible fees, every type launches the produce when X_t reaches some low threshold $\underline{x}_\theta$. Thus, with admissible fees, the only "good results" (the results that reduce the cost of launching the product) are results about the product safeness. For low types, the probability of getting these type of results is higher, in other words, the expected cost of gathering more 'good information' before launching the product is lower. So, whenever a type $\theta' > \theta$ is willing to gather more 'good information' before launching the product, so is type θ.

Proof.

We use the following notation: given two stopping thresholds $\underline{x}, \bar{x}$ and an initial $x \in (\underline{x}, \bar{x})$ let f^y be the probability of hitting the low threshold first and T^y the expected time before stopping when the product's quality is y.

LEMMA 2 *For a given threshold $\underline{x} < 0$, let v^θ denote the continuation value of a firm of type θ—under its optimal continuation strategy—conditional on reaching $\underline{x}$. If $v^0 \geq v^\theta$, then for any $\bar{x} > 0$ and any $x \leq \frac{1}{2}(\underline{x} + \bar{x})$ the expected value for type 0 of using the launch and abandonment thresholds $\underline{x}$ and $\bar{x}$ starting at x (i.e., conditional on reaching $X_t = x$, and*

viewed from the perspective of time t) is higher than the corresponding expected value for type θ.

Proof. The difference in these expected values is

$$\begin{aligned}&(f^g v^0 - cT^0) - (\theta f^b v^\theta + (1-\theta) f^g v^\theta - cT^\theta)\\ &= f^g(v^0 - v^\theta) + \theta(f^g - f^b)v^\theta + c(T^\theta - T^0)\end{aligned}$$

where by definition $T^\theta = \theta T^1 + (1-\theta)T^0$. Since $v^0 \geq v^\theta$, the last expression must be positive if we prove that $f^g \geq f^b$ and $T^\theta \geq T^0$. The first inequality is a direct consequence of the drift of X.

The second inequality follows from the fact that $T^1 \geq T^0$ and T^θ is a convex combination of T^1 and T^0. To see why $T^1 \geq T^0$, notice that these expected times are equal by symmetry when starting at the midpoint $x = (\underline{x} + \bar{x})/2$, and that T^0 becomes smaller and T^1 larger as x moves closer to $\underline{x}$. ■

To conclude the proof of Proposition 3, consider a type θ, starting at 0 with optimal launch and abandonment thresholds $\underline{x}, \bar{x}$. The continuation value $v_\theta(\underline{x}, \bar{x}, x)$ for type θ of using these thresholds when starting at x must exceed the value of stopping immediately at x—which is equal to $\pi - \eta(x)$ and independent of θ—for all x in $(\underline{x}, \bar{x})$. Notice that the continuation value of type 0 conditional on reaching $\underline{x}$ is weakly higher than the continuation value of type θ at $\underline{x}$, because θ optimally stops at $\underline{x}$ and 0 can obtain the same payoff by stopping. Therefore, for any $x \in (\underline{x}, (\underline{x} + \bar{x})/2)$, Lemma 2 implies that starting at all such x and using the thresholds $\{\underline{x}, \bar{x}\}$ yields a higher continuation value $v_0(\underline{x}, \bar{x}, x)$ for type 0 than for type θ and, in particular, exceeds the stopping value. Moreover, we have

$$v_\theta(\underline{x}, \bar{x}, x) = \theta \cdot v_0(\underline{x}, \bar{x}, x) + (1-\theta) \cdot v_1(\underline{x}, \bar{x}, x).$$

This implies that $v_0(\underline{x}, \bar{x}, x) \geq v_\theta(\underline{x}, \bar{x}, x) \geq v_1(\underline{x}, \bar{x}, x)$. Thus, for any type $\theta' < \theta$, $v_{\theta'}(\underline{x}, \bar{x}, x) \geq v_\theta(\underline{x}, \bar{x}, x) \geq 0$. This shows that type θ' does not find it optimal to stop for any $x \in (\underline{x}, (\underline{x} + \bar{x})/2)$. To conclude the argument, notice that we have showed that for all $x \in (\underline{x}, (\underline{x} + \bar{x})/2)$, type 0 has a higher continuation value upon reaching x than does θ, because type 0 can always the thresholds $\underline{x}, \bar{x}$ that are optimal for θ and we have established that under these thresholds 0 was already getting a higher continuation value, and 0 is getting an even higher value when it uses its own optimal threshold. Therefore, we can apply Lemma 2 with any lower bound $\underline{x}' \in [\underline{x}, (\underline{x} + \bar{x})/2]$, and in particular for $\underline{x}' = (\underline{x} + \bar{x})/2$, to show that 0 has a higher continuation value than θ for all $x \in (\underline{x}, 1/4\underline{x} + 3/4\bar{x})$. Repeating an earlier argument,

this implies that no type θ' wishes to stop when reaching such x. Proceeding iteratively yields the result for all $x \in (\underline{x}, \bar{x})$ and concludes the proof. ■

5 Monotonic Tariffs

Proposition 1, shows that the regulator would like to implement lower thresholds than the firm when the firm faces with a uniform penalty, regardless of the firm's private information. The next proposition shows that under these circumstances, it is without loss of generality to focus on tariffs that are nondecreasing functions of x, i.e., which impose a lower penalty, the more due diligence is demonstrated by the firm.

PROPOSITION 4 *Suppose that Θ is finite and consider any thresholds $\{x_\theta\}_{\theta\in\Theta}$ that are (i) decreasing in θ and (ii) such that $x_\theta < \underline{x}_\theta^l$[6] for all $\theta \in \Theta$. Then, there exists a non-decreasing, piecewise-constant tariff ψ such that $\underline{x}_\theta^\psi = x_\theta$ for all $\theta \in \Theta$.*

Proof. We index the elements of Θ from the smallest θ_1 to the largest $\theta_{|\Theta|}$ and construct the tariff ψ by moving from large values of x to lower ones. We start by setting $\psi(x) \equiv l$ for all $x \geq x_{\theta_1}$. At x_{θ_1}, we lower the tariff to a level ψ_1 that makes θ_1 indifferent between launching the product at x_{θ_1} and at $\underline{x}_{\theta_1}^l$. We keep ψ constant at the level ψ_1 for $x \in (x_{\theta_2}, x_{\theta_1}]$. Since a firm's launch threshold when it faces a constant tariff $\hat{l}$ is decreasing in $\hat{l}$, and since $\psi_1 < l$, we have

$$x_{\theta_1} < \underline{x}_{\theta_1}^l \leq \underline{x}_{\theta_1}^{\psi_1}$$

where $\underline{x}_{\theta_1}^{\psi_1}$ is the launch threshold used by type θ_1 when the tariff is constant and equal to ψ_1. This implies that type θ_1 prefers threshold x_{θ_1} to any level $x \in (x_{\theta_2}, x_{\theta_1})$. The reason is that when facing a fixed penalty level (here, ψ_1), a type's preference over launch thresholds is monotonic increasing up to this type's ideal threshold. This result is proved in the appendix (Lemma 3). Applied to the present setting, type θ_1's optimal threshold when facing a constant penalty of ψ^1 is by definition $\underline{x}_{\theta_1}^{\psi_1}$. Lemma 3 then implies that type θ_1 prefers, among stopping thresholds associated with penalty ψ_1, any stopping threshold x to any stopping threshold x', whenever $x' < x < \underline{x}_{\theta_1}^{\psi_1}$. This shows that θ_1 prefers x_{θ_1} to any $x \in (x_{\theta_2}, x_{\theta_1})$.

[6]Recall from Section 3 that $\underline{x}_\theta^l$ is the launch threshold for type θ when $\psi \equiv l$.

At x_{θ_2}, we lower the tariff ψ to a level ψ_2 that makes type θ_2 exactly indifferent between launching the product at x_{θ_2} and at its preferred level $\hat{x}_2$ among all $x > x_{\theta_2}$, given the tariff ψ constructed so far. By the single-crossing property established in Lemma 1, this implies that θ_1 prefers $\hat{x}_2$ to any x_{θ_2} and, combined with the previous paragraph, that θ_1 prefers x_{θ_1} to any $x \geq x_{\theta_2}$.

We set ψ equal to ψ_2 for all $x \in (x_{\theta_3}, x_{\theta_2}]$. Since $x_{\theta_2} \leq \underline{x}^l_{\theta_2} \leq \underline{x}^{\psi_2}_{\theta_2}$, type θ_2 prefers x_{θ_2} to any $x \in (x_{\theta_3}, x_{\theta_2})$. Another application of Lemma 1 guarantees that type θ_1 also prefers x_{θ_2} to any $x \in (x_{\theta_3}, x_{\theta_2})$.

Proceeding iteratively, we then lower ψ at x_{θ_3} to a level ψ_3 that makes type θ_3 exactly indifferent between launching the product at x_{θ_3} and at its preferred level $\hat{x}_3 > x_{\theta_3}$ given the tariff ψ constructed so far. Repeated applications of Lemma 1 guarantee that types θ_1, θ_2 prefer their respective thresholds $x_{\theta_1}, x_{\theta_2}$ to x_{θ_3}. We extend ψ by setting it constant, equal to ψ_3 for all $x \in (x_{\theta_4}, x_{\theta_3}]$. The proof is completed by induction. ■

6 Taxation Principle with Identifiable Information Acquisition

Theorem 1 is a corollary of the Taxation Principle with Non-Contractible Events of our companion paper (Poggi and Strulovici (2020)).

In that paper, we introduce the concept of an *observably injective* mechanism.[7] This concept requires that two conditions be satisfied by the mechanism. We explain these conditions in the present setting.

Let A denote the set of all possible strategies by the firm. Each element of A consists of a pair (τ, d), where τ is a stopping time adapted to the filtration of X and d is measurable with respect to $\mathcal{F}^X_\tau$.

For any strategy $a \in A$, let μ_a denote the distribution of observable outcomes by the regulator *if the firm chooses that action and causes some damage.*

DEFINITION 5 *An IC mechanism M is* observably injective *if there exists a partition $\mathcal{A} =$*

[7] In that paper, the concept is used for social choice functions rather than mechanisms. For simplicity we omit this distinction here.

$\{A_k\}_{k=1}^K$ *of* A *such that*

(i) $a \in A_j, a' \in A_k$ *with* $j \neq k \Rightarrow$ $supp(\mu_a) \cap supp(\mu_{a'}) = \emptyset$.

(ii) If for two types θ, θ' *the mechanism prescribes an actions* $a, a' \in A_k$ *then* $\mu_a = \mu_{a'}$.

Recall from Section 4 that a mechanism is *admissible* if for each type θ, the firm's optimal strategy consists in launching the product at some threshold $\underline{x} \leq 0$ and abandoning it at some threshold $\bar{x} > \underline{x}$.[8]

PROPOSITION 5 *If* M *is admissible, then it is observably injective.*

Proof. If the firm launches the product and causes damage the regulator observes the evidence X_τ, which is the firm's threshold used by the firm to launch the product. Therefore, an observable outcome in the present setting simply consists in the adoption threshold used by the firm. For any strategy a, the distribution μ_a thus reduces to a single point, the firm's adoption threshold.

We partition the set of admissible strategies according to their launch thresholds. If two strategies a, a' are in different elements of the partition, then μ_a and $\mu_{a'}$ have disjoint support conditional on damage occurring, since the strategies have different launch thresholds. (There is no accident and nothing observed when the product is abandoned.) If the mechanism prescribes to two types θ, θ' strategies that are in the same element of the partition, this means that their launch thresholds are identical: $x_\theta = x_{\theta'}$. In this case, the distributions μ_{a_θ} and $\mu_{a_{\theta'}}$ of observable outcomes are identical since, conditional on damage occurring, the observed outcome is $X_\tau = x_\theta = x_{\theta'}$. Therefore, in all cases, the conditions required for the mechanism to be observably injective are satisfied. ■

COROLLARY 1 *If an IC mechanism* M *is admissible, then it can be implemented by a tariff mechanism.*

Proof. Proposition 5 implies that M is observably injective. The result then follows from Theorem 1 in Poggi and Strulovici (2020). ■

[8]Since the firm faces a Markovian decision problem, it is always optimal for the firm to use a threshold policy. Admissibility imposes the further restriction that the optimal policy consists in launching the product at the lower threshold and abandon it at the upper threshold.

A Appendix

This appendix establishes a lemma used in the proof of Proposition 4.

Fix a type θ, a constant penalty ℓ in case of damage, and the starting value $X_0 = 0$.

For each $x < 0$, consider the policy that consists in launching the product for all $z \leq x$, keep acquiring information for all $z \in (x, x')$, and abandoning the product for all $z \geq x'$, where x' is the optimal abandonment threshold given x. Let $V(x)$ denote the expected payoff of the corresponding policy, and let x^* denote the optimal launch threshold.

LEMMA 3 *$V(x)$ is increasing in x for $x < x^*$.*

Proof. Since θ is fixed, we can work in the space of posteriors p instead of x. The firm faces a standard Wald problem with Brownian learning. The value of stopping at posterior p is $\pi - p\ell$, since p is the probability of damage. The firm's optimal value function satisfies the smooth-pasting property (i.e., it is continuous differentiable) at the optimal launch and abandon thresholds, p^* and $a(p^*)$. For any launch threshold p, let $v_p(q)$ denote the expected payoff when using threshold p and the optimal abandonment threshold $a(p)$ when the initial value is q. By optimality of p^*, we have $v_{p^*}(q) \geq v_p(q)$ for all p, q. Moreover, this implies that $a(p) \leq a(p^*)$: it is optimal to abandon earlier when using a suboptimal launch threshold.

Now consider any threshold $p < p^*$. By construction, we have $v_p(q) = v_{p^*}(q) = \pi - q\ell$ for all $q \leq p$, and $v_p(q) < v_{p^*}(q)$ for all $q \in (p, a(p^*))$. Finally, consider a third threshold $p' < p$. By construction we have $v_{p'}(q) = v_p(q) = v_{p^*}(q)$ for all $q \leq p'$. However, for q in a right neighborhood of p', we have $v_{p'}(q) < v_{p^*}(q)$, because $v_{p'}$ leaves the line corresponding the adoption function $A : q \mapsto \pi - \ell q$, and it must leave it from below because function v_{p^*} still coincides with A. But we also still have $v_p(q) = v_{p^*}(q)$ for $q \in (p', p)$.

This shows that the value functions v_p and $v_{p'}$ are pointwise ranked with the first one strictly above the second one for $q \in (p', p)$. Moreover, these functions cannot cross because for $q \in (p, a(p'))$ they both solve the same differential equation and they must satisfy distinct smooth pasting conditions at their respective abandonment thresholds.

This shows that v_p is everywhere above $v_{p'}$, strictly so for $q \in (p', a(p))$, and hence that launching at p is strictly better than launching at $p' < p$. ■

References

Baumann, F. and Friehe, T. (2015). Learning-by-doing in torts: Liability and information about accident technology. *Economics Letters*, 138.

Friehe, T. and Schulte, E. (2017). Uncertain product risk, information acquisition, and product liability. *Economics Letters*, 159:92–95.

Goeschl, T. and Pfrommer, T. (2015). Learning by negligence - torts, experimentation, and the value of information. Discussion Paper Series 598, Heidelberg. urn:nbn:de:bsz:16-heidok-191977.

Henry, E., Loseto, M., and Ottaviani, M. (2021). Regulation with Experimentation: Ex Ante Approval, Ex Post Withdrawal, and Liability.

Henry, E. and Ottaviani, M. (2019). Research and the Approval Process. *American Economic Review*, 109(3):911–955.

Kolstad, C. D., Ulen, T. S., and Johnson, G. V. (1990). Ex Post Liability for Harm vs Ex Ante Safety Regulation: Substitutes or Complements? *American Economic Review*, 80(4):888–901.

McClellan, A. (2019). Experimentation and Approval Mechanisms.

Poggi, F. and Strulovici, B. (2020). A Taxation Principle with Non-Contractible Events.

Schmitz, P. W. (2000). On the joint use of liability and safety regulation. *International Review of Law and Economics*, 20(3):371–382.

Shavell, S. (1984). Liability for Harm versus Regulation of Safety. *Journal of Legal Studies*, XIII:357–374.

Shavell, S. (1992). Liability and the Incentive to Obtain Information about Risk. *Journal of Legal Studies*, XXI(June):259–270.

Wald, A. (1945). Sequential Tests of Statistical Hypotheses. *The Annals of mathematical Statistics*, 16(2):117–186.