Safety Monitoring, Capital Structure, and “Financial Responsibility”*

Eberhard Feess
University of Frankfurt†

Ulrich Hege
ESSEC Business School
and CEPR‡

Current version: February 2001

*An earlier version was circulated as “Lender Liability, Capital Structure, and Mandatory Coverage”. We wish to thank Marcel Boyer, Dominique Demougin and Georg Noldeke for valuable comments and Lori Boughton and Laura Balatao from the EPA for helpful insights. Seminar audiences at the University of Basle and the EALE/Geneva Association Joint Conference in Rotterdam provided useful discussions. Financial support from Thyssen Stiftung is gratefully acknowledged.

†Address: Department of Economics, University of Frankfurt, Schumannstrasse 34a, D-60325 Frankfurt, Germany. E-mail feess@wiwi.uni-frankfurt.de.

‡Address: ESSEC Business School, Dept. of Finance, Avenue Bernard Hirsch, F-95021 Cergy-Pontoise Cedex, France. E-mail hege@essec.fr.
Abstract

We compare different modes of extending liability to third parties motivated by the problem of firms exerting too little preventive care if damages are likely to exceed their equity. In our model, firms can be financed by equity, bank debt or publicly traded debt. There is a moral hazard problem about the choice of care that can be mitigated through stochastic monitoring. We show that the optimal allocation can always be implemented by a liability regime of full “financial responsibility”, that is mandatory liability coverage for total harm that can be fulfilled either by an insurer or by a lender. This result is in contrast to related models which find liability below the level of harm optimal. The difference is due to the inclusion of safety monitoring. The superiority of full financial responsibility is reinforced by the fact that other liability regimes lead to distortions in the capital structure. The robustness of the mechanism vis-a-vis risk aversion is confirmed.

Key words: lender liability, compulsory insurance, choice between private and public debt, limited liability effect.

JEL classification: G32, K13, K32.
1. Introduction

A firm strictly liable for any harm done will generally not choose the efficient care level if there is a possibility that the firm goes bankrupt. Damages large enough to put the firm into bankruptcy - like environmental and product liability or health risks - will be undervalued because some of the losses of the victims will go unclaimed under conventional strict liability. The care level will typically be too low in this case (limited liability effect).\textsuperscript{1} For firms facing considerable liability risks, leveraging up the capital structure may become an effective evasion strategy.\textsuperscript{2}

Extending liability is a natural regulatory response. Attempts to extend liability are most advanced in the United States where managers, shareholders, holding companies and notably secured creditors are among the groups which have been held liable under the Comprehensive Environmental Response, Compensation and Liability Act (CERCLA).\textsuperscript{3}

Several varieties of extended liability have been proposed. The form which is at the center of our analysis is financial responsibility. Financial responsibility means that the operation of hazardous plants and other businesses is only authorized if the operator shows proof that all liability claims are covered either by an insurance company, by a lender, or by another sufficiently solvent party. The principle of financial responsibility is legally mandated in the US (under the Resource Conservation and Recovery Act; RCRA), but the application seems insufficient and distortions are likely.\textsuperscript{4} Financial responsibility is also made possible by law in some European countries, notably in Germany (under the Environmental Liability Act 1990) and in the Netherlands (Environmental Control Act 1979 and Soil Protection Act 1994), but we are not aware of current enforcement efforts.

Two other varieties of extended liability have gained prominence in practice and in the academic discussion. Under lender liability, creditors of an insolvent injurer can be held liable. Under mandatory insurance, hazardous plants need to show proof of insurance coverage in order to be able to operate, just as mandatory liability insurance is prescribed in other areas of torts, like automobile insurance. As far as we were able to gather reliable information on the current situation in the US, financial responsibility appears to be insufficiently enforced as of now. In practice, efforts to enforce extended liability seem to be mostly confined to legal action against identifiable deep pockets like bank lenders, after an accident occurred. That is, the current legal practice in the US can be likened to lender liability. In Europe,

\textsuperscript{1}See e.g. Summers (1983) and Shavell (1986). Beard (1990) and Posey (1993) extend the discussion to care costs that reduce the coverage of the defendant. Kornhauser and Revesz (1990) discuss the problem of multiple tortfeasors.

\textsuperscript{2}Ringleb and Wiggins (1990) indeed find that firms attempt to avoid liability by shielding assets through divestiture. Their analysis suggests that the incentive to avoid liability led to a 20% increase in the number of small corporations between 1967 and 1980.

\textsuperscript{3}See Buente, Crough and Conlan (1996) for an overview.

\textsuperscript{4}Boyd (1996) identifies important loopholes that weaken financial responsibility under RCRA.
extended liability is not yet applied in practice. At the level of the European Union, however, proposals are currently being discussed in the wake of initiatives in favor of strict liability for environmental torts.\(^5\)

If information were perfect, then any of these three forms of extended liability would obviously be fully efficient.\(^6\) In practice, however, there are pervasive problems of asymmetric information between investors and firms, e.g. because investors have difficulties to correctly anticipate environmental risks (hidden information) or cannot monitor the care level perfectly or without suffering large costs (hidden action). Under these circumstances, three questions arise: The first question is whether extending liability to third parties is superior even if the agency costs caused by asymmetric information are taken into account. The second question is whether lender liability, mandatory insurance or financial responsibility should be applied. The third question is whether liability payments should differ from total harm or not.

Our analysis focuses on moral hazard with respect to the firm’s level of care. We assume that the moral hazard problem can be mitigated through monitoring. We demonstrate that financial responsibility for total harm leads to an allocation arbitrarily close to the first best if the contract between the firm and the lender or insurer assuming residual liability is chosen optimally to reduce the agency costs to the minimum. Since it leads to the first best, full financial responsibility is preferable to all competing proposals; in general, those cannot avoid inefficiencies with respect to the care level and the capital structure.

Our finding is in striking contrast to earlier literature on extended liability: It has been repeatedly argued (see references below) that, in the presence of moral hazard, liability should be kept below the level of harm because of the following trade-off: Raising liability towards the level of harm gives better incentives to internalize the social costs of negligent behavior, but comes at the expense of increasing agency costs.

The optimality of full liability in our analysis is mainly explained by the fact that we introduce an element that has been largely ignored in the previous discussion about extended liability, namely safety monitoring. Safety monitoring seems to be a reasonable extension in the discussion, since big industry insurers and banks possess and actively deploy considerable competencies in screening environmental risks, in safety consulting and in auditing. We demonstrate that increasing agency costs can be avoided if safety monitoring is applied appropriately. We show that an insurance or loan contract containing both penalties and bonus payments in response to the findings of monitoring solves the limited liability problem.

The superiority of financial responsibility over competing forms of full liability, notably lender liability, is a consequence of our treatment of possible distortions in the capital structure which is more encompassing than in the preceding literature. Previously, only the debt-equity choice of firms has been considered. By contrast, we allow for a choice among

\(^6\) See e.g. Shavell (1987).
equity, private debt and publicly traded debt. Distinguishing between private or bank debt and publicly floated debt\textsuperscript{7} seems more realistic in the light of recent trends towards disintermediation. It captures the fact that in practice lender liability can only be implemented for large lenders who are geared up to have a special relationship with the borrower. This is, for example, the current legal practice under RCRA.

As long as firms can escape lender liability by issuing publicly traded debt, and as long as such debt issues can be designed to enjoy priority over liability claims - this is for instance easily possible under current US law by issuing secured debt\textsuperscript{8} -, lender liability alone cannot solve the limited liability problem. Lender liability will then eliminate distortions between equity and private debt, but at the expense of creating new distortions between insider finance and outsider finance, since only large creditors can be held liable. Hence, we propose financial responsibility for total harm not only because safety monitoring can eliminate agency costs, but also because all other liability regimes lead to capital structure distortions.

A number of recent papers have addressed lender liability and compulsory insurance in models with moral hazard. The seminal contribution is Pitchford (1995), who shows that extending liability beyond the manager’s wealth reduces the care level of the manager and hence social welfare. The reason is that with lender liability and competitive capital markets, the manager voluntarily commits her total wealth as collateral if damage occurs. If liability is extended the manager’s wealth, the only effect is an increase in interest rates or insurance premia, which will lower the manager’s incentives to avoid environmental harm. Boyer and Laffont (1997), and Boyd and Ingberman (1997), and Endres and Lüdeke (1998) confirm Pitchford’s conclusion that it might be second best optimal to adjust damages below total harm with somewhat different arguments. In clear contrast to these articles, we argue that full financial responsibility leads to the first best allocation under plausible circumstances.

In the ensuing discussion, it has been pointed out that Pitchford’s result crucially depends on the assumption that the manager has all of the bargaining power (via competitive capital markets). It has been shown that if the lender has the bargaining power, then the interest rate decreases and the manager’s effort increases in the lender’s liability.\textsuperscript{9} Balkenborg (1997)

\textsuperscript{7}This distinction has received a lot of attention recently in the theory of corporate finance, see e.g. Diamond (1991), Rajan (1992), Chemmanur and Fulghieri (1994), and Bolton and Scharfstein (1996).

\textsuperscript{8}Basically, in U.S. bankruptcy proceedings, claims of the Environmental Protection Agency (EPA) for the clean-up of toxic property are ranked as administrative expenses. That is, they are junior to secured claims, but senior to all unsecured claims; see in re. \textit{Chateaugay Corporation} 994 F.2d 997 (2nd Circuit Court 1991); in re. \textit{Hemingway Transport, Inc.}, 993 F.2d 915 (1st Circuit Court) and Pennsylvania vs. Conroy, no. 93-3284 (3rd Circuit Court 1994).

\textsuperscript{9}See e.g. Newman and Wright (1992), Heyes (1996) and Demougin and Fluet (1999). Shavell (1997) demonstrates that the optimal level of extended liability might well be above total harm (punitive damages) if the principal has the bargaining power.
demonstrates that the optimal level of lender liability is increasing in the creditor’s bargaining power. Demougin and Fluet (1999) show that there might be excessive incentives for monitoring, because the principal uses monitoring as a means to reduce the agent’s information rent, and Feess (1999) argues that this effect is aggravated if lenders are subject to the negligence rule rather than strict liability. In contrast to this discussion, full financial responsibility is optimal in our analysis independent of the bargaining power, and no excessive monitoring occurs.

Our model is also related to literature discussing mandatory insurance to avoid environmental harm. Jost (1996) emphasizes the benefits of compulsory insurance if the regulator will approve production only if an efficient insurance contract has been signed. Shavell (1986) argued earlier, on a less optimistic note, that the overall effect of mandatory insurance is ambiguous, because it prevents socially inefficient projects. Polborn (1999) argues that this does not happen if the agent’s wealth is high enough to pay the insurance premium in advance, a result that is not confirmed in our model when the manager is wealth-constrained.

There is empirical evidence in support of our concern that the capital structure and notably the choice between bank debt and bond financing might be distorted if extended liability is insufficiently comprehensive. Kroszner and Strahan (1998) find that lenders are hesitant to be found closely involved in their borrower’s day-to-day operations, because of concerns about future tort liabilities. In a previous theoretical paper (Feess and Hege (2000)), we had already visited this question, but following the conventional assumption that agency costs are increasing in the induced level of care, we found that reduced rather than full liability is optimal. By contrast, the inclusion of a genuine microeconomic model of safety monitoring in the present paper leads to a very different conclusion. Lewis and Sappington (1999) is the only prior article finding that full financial responsibility is optimal, but this is driven by their assumption that the level of harm provides a publicly signal for the unobservable effort level. In our analysis, no such public signal, but only private monitoring efforts exist.

The paper is organized as follows. The model is laid out in Section 2. Section 3 compares the different liability regimes and demonstrates the efficiency of financial responsibility. Section 4 extends the analysis to risk aversion. Section 5 concludes.

2. The Model

We consider a risk-neutral manager who needs to raise $I$ to finance a project. There is some risk that the project causes environmental damage with a monetary loss of $D$. Harm $D$ occurs with probability $p(e) > 0$ which depends on the manager’s unobservable care level $e$. As usual, $e$ expresses at the same time the cost of care. We assume that $p' < 0$ and $p'' > 0$, so

---

10See also Gompers and Lerner (1999) for similar results in a somewhat different context.
accidents can be prevented by more care, but at a decreasing rate. It follows that the socially
optimal level of care requires that marginal costs equal the marginal reduction in expected
harm:

\[ p'D = 1 \]  \hspace{1cm} (2.1)

We denote by \( e^* \) the optimal care level satisfying (2.1). The manager decides not only
on the care level \( e \), but also on the capital structure of the firm. All financial variables
are expressed as present values when the financial structure is chosen. The manager has
unobservable initial wealth \( W \) that can be used as equity \( E \). Besides, she can only use debt
instruments to finance \( I \). Capital markets providing debt financing are competitive. Let \( I_D \)
be the part financed by debt and \( E = I-I_D \) be the part financed by equity, where \( E \leq W \). To
analyze the impact of different liability regimes on the capital structure, we assume that the
corporate value (i.e. present value of future cash flows) \( R \) in case there is no environmental
harm depends not only on the project itself, but also on the choice of the capital structure.
Following the standard static-trade-off model usually applied in corporate finance, we assume
that there is a unique optimal debt level \( I_D^* \) that maximizes the corporate value \( R(I_D) \), i.e.
Hence, \( R(I_D^*) \geq R(I_D) \) for all \( I_D \neq I_D^* \). We do not enter into an explicit, and cumbersome,
modelling of the impact of \( I_D \) on \( R \), since we are only interested in potential distortions in
the capital structure caused by liability rules. the only thing requirement is the assmuption
that there is an optimal capital structure, which could for example be explained by the static
trade-off theory.\(^{11}\) We assume \( W \geq E^* = I-I_D^* \) to ensure that the optimal capital structure
is feasible with the available equity.

We also allow for the manager’s choice between bank debt and publicly traded debt. In
case of publicly floated debt, there is a large number of dispersed bond investors contributing
towards the investment \( I_D \), while in the case of bank debt, there is only a single bank lender.
To capture the difference as easily as possible, we assume that bank debt has costs \( c^B \), and
that public debt has costs \( c^P \). Either form of debt can be more cost efficient, i.e. \( c^B \geq c^P \).
To simplify the notation we assume that \( \min(c^B, c^P) = 0 \). Again, we do not model explicitly
the factors influencing the optimal debt structure. Instead, we treat \( c^B \) and \( c^P \) as given to
focus on possible distortions caused by liability rules. We assume that mixing public and
private debt is never better than either pure bank debt or public debt financing.

If harm \( D \) arrives, then the firm is strictly held liable by the court for damages in the
amount of \( L \leq D \).\(^{12}\) We assume that damages have strict priority over equity, but that

\(^{11}\)The basic idea of the static trade-off model ist that the tax burden is decreasing in the level of debt,
whereas expected bankruptcy costs are increasing. As both costs are reflected in the firm’s net present value,
there is an interior optimum for the optimal debt level \( I_D \). See originally Myers (1977) or any standard
textbook on corporate finance.

\(^{12}\)We do not consider negligence rules since \( e \) is unobservable.
debt is secured and is prior to damages. The manager’s wealth not invested into the firm is protected through limited shareholder liability.

In view of the insolvency problem, we consider three liability regimes:

(i) (Un-extended) Strict Liability: liability is neither extended to third parties nor is insurance coverage required to get a permission for the project.

(ii) Lender Liability without financial responsibility: strict liability is extended to bank creditors. Lender liability does not extend to public creditors, since transaction costs would be prohibitively high.

(iii) Financial Responsibility: As explained in the introduction, this means that the investment is permitted only if the manager shows proof that £ is fully covered by either her own wealth, an insurance company, or a lender.

Thus, the problem of the regulator is to choose ex ante among these three liability regimes and to define total liability payments £ ≤ £. Insurance is available on a competitive insurance market.

Next, we turn to the possibility of monitoring. In our model, monitoring means that lenders and insurers may perform costly environmental audits on the firm to make sure that a level of care written into the debt or insurance contract is maintained. For simplicity, we assume that the effort level can perfectly be estimated if monitoring is performed. We denote by £ the costs of monitoring, and by £ the probability that the firm is audited. Hence, expected monitoring costs are ££. To avoid double moral hazard problems, we assume that the monitor can credibly commit to £. Furthermore, we assume that banks and insurers have the same monitoring technology. Finally, duplication of auditing efforts is inefficient, i.e. if both the bank and the insurance company audit, then their total costs are higher than if one of the two were to assume all the auditing.

If the manager and a third party (insurer or lender) agree upon a positive level of monitoring, then total social welfare can be written as

$$W + R(I_D) - I - p(e)D - e - c^i - qm. \quad (2.2)$$

Equation (2) shows that social surplus depends on the capital structure £, the debt structure expressed by £, £ = £, £, the care level £ and the monitoring probability £. Clearly, the investment should only be carried out if expression (2.2) is positive.

13In reality, either banks or insurance companies might have the superior auditing technology. On the one hand, insurance companies have great experience in environmental monitoring, but on the other hand, it has been common bank practice for years to assess the environmental risks of potential borrowers.
3. Analysis of Liability Regimes

3.1. Strict Liability

In this section, we analyze the outcome, and demonstrate in particular distortions to effort choice and capital structure, if there is no extended liability. Note that there will be no distortion between private and public debt, since lenders bear no risk.\(^{14}\) Suppose the regulator chooses \(L = D\). The manager’s expected utility \(U^{S}\) is

\[
U^{S} = W + [1 - p(e^{S})] \left[ R(I_{D}^{S}) - I_{D}^{S} \right] - E^{S} - e^{S}.
\]  

(3.1)

If there is no environmental harm (which happens with probability \(1 - p\)), the manager gets \(R(I_{D}^{S}) - I_{D}^{S}\). Otherwise, the firm goes bankrupt. Note that \(I_{D}^{S}\) matters only if the firm can repay its debt, while \(E^{S}\) and \(e^{S}\) have to be borne in any case. It follows that the manager’s utility maximizing capital structure is given by

\[
\frac{dR}{dI_{D}^{S}} = 1.
\]  

(3.2)

Since the first best debt level \(I_{D}^{*}\) is given by \(\frac{dR}{dI_{D}^{*}} = 0\), it follows immediately that the debt level is too high, because debt is prior to damages and hence the part of harm borne by victims is the higher the higher the level of debt.

Moreover, the FOC for effort is

\[
-\frac{dp}{de^{S}} \left[ R(I_{D}^{S}) - I_{D}^{S} \right] = 1.
\]  

(3.3)

Hence, \(e^{S} < e^{*}\), and the distortion is increasing in the level of debt. This is the well-known limited liability effect: the safety level is too low whenever harm exceeds the firm’s solvency, and this is aggravated by the incentive to increase the leverage. Finally, the manager has an incentive to overinvest, since the investment should only be undertaken if social welfare as expressed in equation (2) is positive. Under strict liability, however, the manager undertakes the project if \(U^{S}\) is positive. Hence, inefficient investment is undertaken if

\[
[1 - p(e^{S})] \left[ R(I_{D}^{S}) - I_{D}^{S} \right] - E - e > 0 > R(I_{D}^{*}) - I - p(e^{*})D - e^{*}.
\]  

(3.4)

Inequality (3.4) shows that inefficient investments are likely for projects with large accident risk \(D\).\(^{16}\) Note that monitoring plays no role under un-extended strict liability, because

---

\(^{14}\)Recall that the only source of risk is the liability risk, and that claims \(D\) are junior to debt.

\(^{15}\)Subscripts \(S\) denote the case with strict liability only.

\(^{16}\)Nothing could be gained by setting \(L > D\) (punitive damages), because the manager goes bankrupt anyway if an accident happens. The effects of punitive damages would be somewhat more complicated if we allowed for different levels of harm, some of which above and some of which below the insolvency threshold; see Innes (1999), and Lewis and Sappington (1999).
(i) the manager has no incentive to insure the environmental risk and (ii) the lender does not care about the risk because debt is prior to damages.

We summarize:

**Proposition 1.** Under strict liability, the care level is too low and leverage is too high. There is overinvestment into projects with considerable accident risks. Monitoring does not take place.

*Proof:* Immediate from calculations above. ■

### 3.2. Financial Responsibility

Under financial responsibility, the regulator grants permission for the project if and only if the manager shows proof that all damages $L$ are fully covered. Let us assume for a moment that the manager has enough wealth to self-insure all liability risks. It is then straightforward that full liability ($L = D$) is optimal. Since both capital markets and insurance markets are competitive, in the end the manager bears not only the expected costs of environmental harm, but also all costs caused by moral hazard or monitoring. Hence, it is in her self-interest to use $W$ as a liability deposit to avoid the agency costs. It follows that financial responsibility leads to a first best without monitoring, and without distortions in the capital structure, whenever the manager’s initial wealth is sufficiently high.$^{17}$

We can therefore immediately turn to the more interesting case where the manager’s net wealth is insufficient for self-insurance ($W + R(I_D^*) - I < D$). Let us first ignore the possibility of monitoring. In this case, it is again in the manager’s self-interest to pledge her total wealth, so that the insurer or bank (henceforth insurer for short) has to pay only $L - W - R(I_D^*) + I$ if an accident happens. This implies that the insurer requires an insurance premium $r$ if its participation constraint is to be satisfied. For $W$ given, $r$ is strictly increasing in $L$ to fulfill the insurer’s participation constraint. Moreover, the manager’s incentive to avoid the accident is strictly decreasing in $r$, since $r$ reduces her net return only if no accident occurs. It follows immediately that extending the liability beyond the manager’s capacity to self-insure reduces her effort. This is exactly the result derived by Pitchford (1995) who concluded that liability should be restricted to $W$.

Next, we demonstrate that the situation changes fundamentally if safety monitoring is taken into account. The role played by monitoring in our model is in fact that of a bonding device. With extended liability, then the manager will ultimately bear a fraction of damages $L$ that exceeds her net wealth $W + R(I_D^*) - I$: She will have to pay higher risk premia for credit or higher insurance premia. It is again in the manager’s self-interest to minimize

---

$^{17}$To avoid misunderstandings, it should be noted that using $W$ as collateral for liability payments does not mean that $W$ must be fully paid in as equity (only a fraction $E$ is).
total expected costs borne by her and the insurer since insurance markets are competitive. These costs consist of expected liability payments $p^F L$, effort costs $e^F$ and monitoring costs $q^F m^F$.

For any effort level $e^F$ that the manager wishes to implement (i.e., to commit to via monitoring), she wants to minimize $q^F m^F$, but she needs to take her own incentive constraint into account. For this, the difference in her expected profits if she complies to the contractual effort level $e^F$ and if she deviates needs to be maximum. This can be reached by a contract with the following elements:

(i) An (unobservable) effort level $e^F$, a monitoring probability $q^F$, and an insurance premium $r^F$.

(ii) If the manager is monitored and found compliant with the contractual safety standard ($e \geq e^F$), then she receives a bonus payment of $\Pi^F > 0$ regardless of whether an accident occurs or not.

(iii) If the manager is found to have deviated from the contract ($e < e^F$), she loses everything regardless of whether an accident happens or not.

(iv) If the manager is not monitored, she loses everything.

Under these circumstances, the insurance premium $r^F$ and the the bonus payment $\Pi^F$ are interdependent and given by the insurer’s participation constraint. Assuming that the manager chooses the contractual effort level in equilibrium, the insurer’s participation constraint determining $r^F$ is

$$(1 + r^F) I_D^F + q^F m^F = I_D^F + p(e^F) [L - (W + R(I_D^F) - I) - r^F I_D^F] + q^F \Pi^F + q^F m^F. \quad (3.5)$$

The insurer gets $(1 + r^F) I_D^F$ plus a fee covering her expected monitoring costs $q^F m^F$ as up-front payments. On the right hand side, expected costs first consist of the investment $I_D^F$. Second, in case of an accident, the manager pays as much as possible and the insurer the remainder. Furthermore, if the manager is found compliant with the contractual safety standard, the insurance company pays $\Pi^F$. Finally, she has monitoring costs $m^F$ in any case, so that equation (3.5) expresses the zero-profit condition. This can be rewritten as

$$(1 - p(e^F)) r^F I_D^F = p(e^F) [L + I - R(I_D^F) - W] + q^F \Pi^F. \quad (3.6)$$

The manager’s expected utility if she complies with the contractual effort level is

$$U^F = [1 - p(e^F)] [W + R(I_D^F) - I - q^F m^F - r^F I_D^F] + q^F \Pi^F - e^F. \quad (3.7)$$

---

18 It does not make any difference whether a bank or an insurance company assumes this function. We use the term “insurer” to cover both.

19 The superscript “$F$” indicates the case with financial responsibility.

20 Note that there are no reasons for distortions in the capital or debt structure, so that we can safely write $I_D^F$, and neglect $e^i$ since $\min(e^i, e^F) = 0$. 

11
If no accident occurs, the manager keeps her wealth $W$ and the value of the firm $R(I_D^F)$, but she has invested $E^F, q^F m^F$, and $(1 + r^F) I_D^F$ as upfront payments. If she is monitored, she receives additionally $q^F \Pi^F$, and she bears effort costs $e^F$ in any case.

Let $e^v$ be the manger’s utility maximizing effort if she deviates from $e^F$. Her expected utility $U^v$ is then

\[ U^v = (1 - p(e^v)) [1 - q^F] \left[ W + R(I_D^F) - I - m^F - r^F I_D^F \right] - e^v, \]

since the gets the term in brackets only if no accident occurs and if she is not monitored. It follows that the manager complies if $U^F \geq U^v$, or if

\[ e^F - e^v < \left( (1 - p(e^F)) - (1 - p(e^v))(1 - q^F) \right) \left[ W + R(I_D^F) - I - m^F - r^F I_D^F \right] + q^F \Pi^F. \]

The incentive constraint (3.9) shows that a sufficient condition for the manager to comply is that $e^F - e^v < q^F \Pi^F$.\(^{21}\) Clearly, this can be reached by choosing $q^F \Pi^F$ high enough. Furthermore, $q^F$ and $\Pi^F$ are perfect substitutes with respect to the manager’s incentive. Since monitoring is costly whereas the premium is only a redistribution, in the optimal contract $q^F$ is chosen arbitrarily small to save monitoring costs, and $\Pi^F$ is adjusted accordingly to fulfill the manger’s incentive constraint.\(^{22}\) Finally, the manager proposes the contract that maximizes their joint surplus. Since the insurer’s expected profit is zero in equilibrium, this joint surplus is identical to the manager’s utility $U^F$. Substituting (3.5) into (3.7)

\[ U^F = W + R(I_D^F) - I - p(e^F)L - e^F - q^F m^F. \]

It follows that the effort written in the contract is strictly increasing in the liability payment $L$, because the monitoring costs $m^F$ required to guarantee that the manager actually chooses the contractual effort level are arbitrarily small (regardless of the effort level itself). This leads immediately to

**Proposition 2.** Regardless of the manager’s wealth, the optimal liability regime is full financial responsibility ($L = D$).

**Proof:** See Appendix.

The logic of this result is the following. When monitoring is introduced, it is possible to reward the manager not only for the fact that no accident happened, but also if the effort observed by the monitor is indeed compliant with the standard. We introduced the bonus payment $\Pi^*$ to do this trick. By giving very high-powed bonus payments $\Pi^*$ if the firm is

\(^{21}\) $\left( (1 - p(e^F)) - (1 - p(e^v))(1 - q^F) \right) \left[ W + R(I_D^F) - E - m^F - (1 + r^F) I_D^F \right]$ is non-negative.

\(^{22}\) This idea goes back to Becker (1968) who emphasized that the detection probability and the fine are substitutes when deterring potential criminals.
monitored and found compliant, the monitoring costs can be chosen approximately close to zero since only the expected bonus \( q^F \Pi^* \) matters. Thus, monitoring costs can be neglected, and we can conclude that financial responsibility allows to get arbitrarily close to the first best allocation. But then, the choice of \( L \) should be such that \( e^F = e^* \) is chosen. This is precisely the case if \( L = D \), i.e. if the court decides on full liability, since the manager’s objective function under the optimal contract will then just coincide with the maximization of the social surplus. Hence, the bonus payment solves the problems caused by limited liability.

3.3. Lender Liability

Now consider the effect of lender liability that is not backed up by the requirement to prove mandatory coverage. The capital structure decision may now be distorted because of the unequal treatment of private and public debt: Only bank financed firms have to fully internalize the environmental risk. Two distortions can emerge: first, public debt firms will use too much debt and exercise too little care because of the lacking imposition of social costs. Second, some firms which should use bank debt because of \( c^B < c^P \) will now find it privately optimal to switch to public debt to free ride on parts of the social costs of their activities. This is the debt structure bias of lender liability.

Suppose the firm continues to use bank debt in spite of this handicap. Then everything is identical to financial responsibility: the manager uses her initial wealth as deductible, and chooses the optimal contract according to Proposition 2, with \( e^F = e^* \) and \( q^* \to 0 \). Her utility is then again

\[
U^L_B = W + R(I^*_D) - I - p(e^*)L - e^*. \tag{3.11}
\]

On the other hand, if the firm uses public debt, its objective function is the same as under strict liability only, because liability is not extended to third parties:

\[
U^L_P = W + [1 - p(e^S)] \left[ R(I^*_D) - I^*_D \right] - E^S - e^S - e^P. \tag{3.12}
\]

Therefore, as long as the cost advantage of bank debt is small, the possibility to evade liability by choosing dispersed public debt will dominate, and the outcome is comparable to strict liability. It follows

**Proposition 3.** With lender liability, we get the following results:

(i) If \( c^B \geq c^P \), then public debt is preferred and the same distortions as under strict liability arise: underprovision of care, overleverage, and overinvestment.

\[23\] Superscript “L” indicates lender liability. Recall that \( \min(c^B, c^P) = 0 \), and that a necessary condition for using bank debt is that \( c^B < c^P \).
(ii) If $c^B < c^P$ and if the cost advantage is overcompensated by the limited liability effect, then public debt is preferred. A distortion in the debt structure arises in addition to the inefficiencies listed under (i).

(iii) If $c^B < c^P$ and if bank debt is optimal, then the first best is achieved as under financial responsibility.

Proof: Part (i): Follows from the analysis of strict liability. Parts (ii) and (iii): Follow from the comparisons of equations (3.11) and (3.12).

3.4. Comparison of the Liability Rules

The main result of our analysis is that financial responsibility either leads to the first best allocation (if the manager’s wealth is sufficiently high) or comes arbitrarily close to that (if her wealth is too low for self-insurance) if the regulator requires full liability, $L = D$. This striking difference to the literature is explained by the fact that our contract proposes the optimal combination of contract penalties and bonus payments in response to the findings of the monitoring: The proponents of restricted liability argue that in the presence of agency costs, the trade-off between improving the risk internalization (by increasing $L$ toward $D$) and the rising agency costs leads to an internal optimum. In our model, safety monitoring is introduced, and we show that with the optimal contract, almost no agency costs $m$ (for monitoring the manager) are required. It follows that full liability ($L = D$) is optimal.

By contrast, both strict liability alone and lender liability lead to distortions: The limited liability effect is present under strict liability, and also under lender liability if banks are not substantially more cost efficient. In fact, the relative ranking between strict liability and lender liability is ambiguous. On the one hand, for firms which prefer bank debt under both regimes, efficiency is improved through lender liability because they are now internalizing the environmental risk in their investment and care decisions. On the other hand, for firms which prefer bank debt under strict liability but are induced to switch to public debt, efficiency will deteriorate because the environmental risk will still only be partially internalized, and there will be an additional loss from using a less preferable capital structure. Finally, firms preferring public debt under both regimes will see no change in their incentives.

4. Risk Aversion

The optimal liability rule proposed in Proposition 2 relies on the assumption of risk neutrality. The optimal contract was based on the highest possible remuneration $\Pi^F$ if the manager was monitored and found compliant. As the monitoring probability $q^F$ was reduced to zero, the reward $\Pi^F$ paid to a compliant management (if monitored) was raised towards infinity.

In this Section, we discuss to what extent our findings are robust if we allow for risk aversion of the manager. Risk aversion means that it is not possible any more, in a sequence where
$q^F$ converges to zero, to keep the manager’s utility constant without raising the expected cost of the bonus, $q^F \Pi^F$.

To derive the optimal amount of damages $L$, we compare the manager’s optimization problem expressed above to the social planner’s objective function. To deduce the second-best optimal contract, we need to assume that the planner could not only choose $L$, but also a monitoring level $m$ and incentive payments to the manager for all possible outcomes of monitoring and harm. The planner would choose these instruments such that the manager behaves second-best efficient. The question is then whether a liability level $L$ exists that sets incentives for the manager to implement the same allocation via an insurance contract.

As before and as in the analysis of Pitchford (1995), we first maintain the assumption that changes in the liability rule have two simultaneous effects: they alter incentives for effort prevention, and they alter the income distribution or the net wealth of the manager. That is, we assume that it is not possible to engineer a fixed transfer of funds, say via a tax from the potential victims of an accident to the manager, in order to neutralize the income distribution effect of the liability rule.

We will argue informally that in this case, the optimal liability rule would be partial financial responsibility, $L < D$. To understand the reason, we need to look the effects as one considers to gradually increase the liability $L$ towards $D$. This will induce two countervailing effects: the positive effect is an amelioration in the incentives for the insurer to invest into monitoring. The negative effect is the increase in the expected liability and the insurance premium, which leaves less to the manager if no accident occurs or if no effort deviation is detected. The optimum $L^*$ is reached when the cost from second effect becomes as important as the benefit from the first effect, and this optimum must be strictly below $D$ because the first effect vanishes as $L \to D$. Thus, the reason why partial liability is optimal is that changes in the liability rule entail wealth effects which, by assumption, cannot be neutralized by a redistributive policy. The optimality of partial liability in the model of Pitchford (1995) requires similarly that income redistribution is not possible.\footnote{A simple way of seeing this is that if direct bargaining between victims and manager were possible in Pitchford’s model, then welfare-increasing, Coasian bargaining would always be possible. The victims would offer a direct transfer to the manager in exchange for accepting full liability (via a private contract recognizing complementary liability).}

To emphasize the critical role of the assumption on the possibility of income transfers we will turn to a formal analysis of the case where this assumption is relaxed. That is, we now explicitly allow for the possibility that the liability rule is determined jointly with a scheme of transferring money between victims and manager. Note that this transfer scheme could either be the outcome of state-mandated fiscal policy, or be the result of efficient private bargaining between victims and manager.

To keep the notation simple, we assume without loss of generality that the manager’s expected utility is separable in money and effort and exhibits risk aversion only in its money
component, while it is risk neutral in effort. Consider then a risk-averse manager with a (concave) utility function \( U(.) - e \) who receives all the surplus of running the firm. We normalize the manager’s lowest monetary utility level to \( U(0) = 0 \). Two well-known effects will emerge with this transition. First, since the manager is risk averse, introducing liability insurance will always allow for better risk sharing whenever the insurer is less risk averse. Second, introducing insurance reduces the manager’s effort incentives. The trade-off between insurance and incentives explains that the first best can no longer be achieved. We are therefore looking for a second best, which is a pure wealth effect, is negative and explains why the optimum.

As with risk-neutrality, the manager will suggest an insurance contract that maximizes her own and the insurer’s joint surplus, subject to her incentive constraint and the insurer’s participation constraint.\(^{26}\) The optimal contract will again consist of three different state-contingent payments from the insurer to the manager: A payment \( \Pi_1 \) if the manager is monitored and found compliant, a payment \( \Pi_2 \) if there is no monitoring and no accident, and a payment \( \Pi_3 \) if there is an accident, but there was no monitoring. Clearly, \( \Pi_1 > \Pi_2 > \Pi_3 \).\(^{27}\) The income transfer that is now possible by assumption takes the form of a lump-sum tax \( T' > 0 \) that the victims need to pay to the manager, independently of the outcome concerning accident and monitoring (respectively a lump-sum subsidy that victims receive if \( T' < 0 \)). These payments must guarantee that the insurer breaks even in expectation.

Now suppose that the liability rule is fixed at \( L \). The problem for a manager of designing the optimal insurance contract can then be written as:

\[
\max_{\Pi_1, \Pi_2, \Pi_3} q U(\Pi_1) + (1 - q)(1 - p(e))U(\Pi_2) + (1 - q)p(e)U(\Pi_3) - e - qm, \tag{4.1}
\]

subject to the insurer’s participation constraint

\[
q \Pi_1 + (1 - q)(1 - p(e))\Pi_2 + (1 - q)p(e)\Pi_3 = W + R(I_D) - I - p(e)L + T', \tag{4.2}
\]

and the firm’s incentive constraint

\[
q(\Pi_1) + (1 - q)(1 - p(e))U(\Pi_2) + (1 - q)p(e)U(\Pi_3) - e \geq (1 - q)(1 - p(e^v))U(\Pi_2) + (1 - q)p(e^v)U(\Pi_3) - e^v. \tag{4.3}
\]

\(^{25}\) Since \( D \) can be interpreted as the monetary certainty equivalent of the utility losses of the victims, so we do not need to specify their risk attitude.

\(^{26}\) The reason is again that capital and insurance markets are competitive, so that the manager bears the agency costs.

\(^{27}\) The only difference to our earlier exposition is that we do not state the insurance premium \( (1 + r)I_D \) explicitly. Instead, \( (1 + r)I_D \) is part of the expected payments \( \Pi_1, \Pi_2 \) and \( \Pi_3 \).
Also, since transfers are now possible, we stipulate that in addition the social planner decides to impose a lump-sum tax $\Pi^t$ that the insurer needs to pay (respectively a lump-sum subsidy if $\Pi^t < 0$), independently of the outcome concerning accident and monitoring. These payments must guarantee that the insurer breaks even in expectation.

A closer look at the insurer’s problem (4.1) - (4.3) shows that this problem is congruent to the planner’s problem (5.3) - (5.4): Both will have the same solution if the same parameters are chosen, that is $L = D$, as well as $\Pi^t_i = \Pi^p_i$, for $i = 1, 2, 3$ and finally $\Pi^t = \Pi^p$ for the lump-sum transfer. We find then and prove in the Appendix that:

**Proposition 4 (Risk Aversion).** *If the manager is risk averse, then a liability regime of financial responsibility coupled with full liability $L = D$, and possibly a lump-sum transfer $\Pi^t$, is optimal.*

This analysis confirms and generalizes our main finding, that the optimal liability regime is financial responsibility and a level of damages $L = D$. To repeat the intuition, imposing full liability $L = D$ is necessary to make sure that the insurer proposes indeed the steepness of the manager’s incentives that guarantees Pareto efficient effort levels. The introduction of monitoring allows to use bonus payments in case of compliance as the main incentive instrument instead of penalties for the manager if an accident occurs. This additional dimension in the contracting space makes the argument in favor of partial liability largely redundant.

To put it differently, the optimal liability regime should still be full financial responsibility, because, as long as compensating wealth transfers are permitted, monitoring allows to efficiently shift the burden of incentive schemes from the “stick” in the case of an accident to the “carrot” in the case of compliance.

### 5. Conclusion

We compared financial responsibility, strict liability without extension, and lender liability in a model where the manager can finance the project through equity (out of her wealth), publicly floated debt and bank debt. Lenders and insurers can perform audits to mitigate the moral hazard problem with respect to the manager’s effort.

We demonstrate that financial responsibility is superior to each other regime. Under financial responsibility, the efficient care level is chosen, and the agency costs of monitoring can be reduced to a negligible quantity, if contracts are optimized. The social welfare can then attain the first best.

There are three driving forces behind our findings. First, the manager voluntarily invests her total wealth as deductible and insurance premium to reduce agency costs. Second, financial responsibility avoids distortions in the capital structure. Third, the agent chooses the optimal monitor independently of the capital structure. If partial liability is applied instead, distortions in the capital or debt structure are inevitable: the lower the level of extended
liability, the higher the bias towards debt instead of equity. However, the higher the level of lender liability, the higher the incentive to substitute bank credits by publicly floated debt. These distortions together with well developed monitoring technologies lead to the conclusion that financial responsibility is referable.

Certainly, one may remain skeptical whether the reduction of monitoring costs to an arbitrarily small amount is possible in practice. In our view, this feature of the optimal contract nonetheless captures an important aspect of the reality of environmental audits: Since firms so often voluntarily engage in contracts with their insurers where they are routinely monitored, they demonstrate by revealed preferences that in practice, the benefit of implementing the efficient care level typically more than dominates the additional agency costs of continuous safety monitoring.
Appendix

Proof of Proposition 2.

Since the insurance market is competitive, for each choice of bonus $\Pi^F$ the competitive (or fair) premium $r^F$ will be

$$r^F = \frac{p(e^F)}{1 - p(e^F)} \left[ L - W - R(I_D^F) + I \right].$$

(5.1)

For any given level $e^F$, consider all pairs $(q^F, \Pi^F)$ such that the ensuing fair premium $r^F$ satisfies (5.1), and such that incentive constraint (3.9) holds with equality. Inspection of (3.9) reveals that all pairs $(q^F, \Pi^F)$ where the product $q^F \cdot \Pi^F$ is identical will ensure that the manager prefers $e^F$ over any other care level. Thus, among all pairs $(q^F, \Pi^F)$ with identical product $q^F \cdot \Pi^F$, the one maximizing welfare will be optimal. But then consider the sequence of all pairs all pairs $(q^F, \Pi^F)$ with identical product $q^F \cdot \Pi^F$ where $\Pi^F \to 0$: obviously, the limit pointalong this sequence must have $q^F \to 0$. It follows that $q^F$ can be chosen arbitrarily close to zero. But then, monitoring costs $q^F m^F$ can also be chosen arbitrarily close to zero. Thus, in the limit the optimal care level $e^F$ and capital strucutre $I_D^F$ will be the maximizer of utility expression (3.10), which becomes

$$U^F = W + R(I_D^F) - I - p(e)D - e$$

and hence $e^F \to e^*$ and $I_D^F \to I_D^*$ as $\Pi^* \to \infty$, demonstrating that an arbitrary close approximation of the first best allocation can be achieved. ■

Proof of Proposition 4: Let $\Pi_1^P$, $\Pi_2^P$ and $\Pi_3^P$ denote the monetary payments to the manager that the social planner would dictate in analogy to the payments in the insurance contracts, i.e. $\Pi_1^P$ is received if monitored and found compliant, $\Pi_2^P$ if there is no monitoring and no accident, and $\Pi_3^P$ if there is an accident, but there was no monitoring. Let $\Pi^P$ denote the expected difference between the (state-contingent) insurance premia and the expected harm,

$$\Pi^P \equiv q\Pi_1^P + (1 - q)(1 - p(e))\Pi_2^P + (1 - q)p(e)\Pi_3^P - (R(I_D) - I + W) - p(e)D.$$  (5.2)

The planner uses the Pareto criterion. Using the utility function of the victims as the numeraire in the planner’s objective function and denoting the (arbitrary) weight that the planner gives to the utility of the manager by $\mu$, the planner’s objective function can be written as:

$$\mu \left[ qU(\Pi_1^P) + (1 - q)(1 - p(e))U(\Pi_2^P) + (1 - q)p(e)U(\Pi_3^P) - e - qm \right] - p(e)D - \Pi^P,$$  (5.3)
subject to the firm’s incentive constraint

\[ qU(\Pi_1^P) + (1 - q)(1 - p(e))U(\Pi_2^P) + (1 - q)p(e)U(\Pi_3^P) - e \geq (1 - q)(1 - p(e^v))U(\Pi_2^P) + (1 - q)p(e^v)U(\Pi_3^P) - e^v. \] (5.4)

In this solution, the planner can dictate the manager’s final payoff in each of the relevant states. This program will give rise to the second best effort level \( e^a \). The incentive constraint (5.4) illustrates that the manager will only take her private cost into account, which for a marginal increase in effort is equal to \((1 - q)(1 - p(e))U(\Pi_2^P) + (1 - q)p(e)U(\Pi_3^P)\). Thus, less effort can be implemented at the margin than would be in the first best, i.e. \( e^* < e^a \).

The planner fixes instruments \((\Pi_1^P, \Pi_2^P, \Pi_3^P)\) and \(\Pi^P\) so as to maximize the entire social objective function, \(\mu [.] - p(e)D - \Pi^P - e - m\). For any level of \(\Pi^P\) chosen and any target level of \(e\) there is a triple \(\Pi_1^P, \Pi_2^P, \Pi_3^P\) which is optimal, i.e. offers best possible risk sharing while giving incentives that the manager chooses \(e\). Call the ensuing level of expected utility of the manager \(EU(\Pi^P, e)\).

Thus, the planners problem can be more simply rewritten as:

\[ \max_{\Pi^P, e} \mu (EU(\Pi^P, e) - e - qm) - p(e)D - \Pi^P. \] (5.5)

The first order condition with respect to \(\Pi^P\) is:

\[ \frac{\partial EU}{\partial \Pi^P} = \frac{1}{\mu} \] (5.6)

and the first order condition with respect to \(e\) is:

\[ \frac{\partial EU}{\partial e} = 1 + \frac{p'D}{\mu} \] (5.7)

There are then optimal instruments \((\Pi_1^P, \Pi_2^P, \Pi_3^P)\) and \(\Pi^P\) solving these conditions and inducing the second-best effort \(e^a\).

Now consider the insurer’s problem of proposing the optimal insurance contract. Suppose that the planner has chosen to impose the optimal instruments, and \(\Pi^I = \Pi^P\), as well as \(L = D\). The insurer’s problem can then be written as a Langrangian of the form:

\[ \max_{(\Pi_1^I, \Pi_2^I, \Pi_3^I)} L = qU(\Pi_1^I) + (1 - q)(1 - p(e))U(\Pi_2^I) + (1 - q)p(e)U(\Pi_3^I) - e - qm \]

\[ -\lambda_1 \left( q\Pi_1^I + (1 - q)(1 - p(e))\Pi_2^I + (1 - q)p(e)\Pi_3^I - W - R(I_P) + I + p(e)D - \Pi^P \right) \]

\[ -\lambda_2 \left( qU(\Pi_1^I) + (1 - q)(1 - p(e))U(\Pi_2^I) + (1 - q)p(e)U(\Pi_3^I) - e - (1 - q)(1 - p(e^v))U(\Pi_2^I) - (1 - \mu) \right). \]

We can use the same trick as for the social planner and represent directly the induced expected utility \(EU(\Pi^I, e)\) that the insurer can optimally propose to the manager, where
\[ \bar{\Pi}^I = q\Pi_1^I + (1 - q)(1 - p(e))\Pi_2^I + (1 - q)p(e)\Pi_3^I - W - R(I_D) + I + p(e)D \]

is a given level of expected net payments to the manager and \( e \) is the target care level. In this representation, the incentive constraint is implicitly contained in the expression \( EU(\bar{\Pi}^P, e) \) and the second constraint can be omitted. Thus, the insurer’s problem becomes:

\[
\max_{\bar{\Pi}^I, e} L = EU(\bar{\Pi}^I, e) - e - qm
- \lambda_1 \left( \bar{\Pi}^I + p(e)D - \bar{\Pi}^P \right)
\]

(5.8)

The first order conditions are:

\[
\frac{\partial EU}{\partial \bar{\Pi}^I} = \lambda_1
\]

(5.9)

and the first order condition with respect to \( \bar{\Pi}^P \) is:

\[
\frac{\partial EU}{\partial e} = 1 + \lambda_1 p'D
\]

(5.10)

Moreover, since the insurance market is competitive, the equilibrium insurance contract satisfies \( \bar{\Pi}^I + p(e)D = \bar{\Pi}^P \). Clearly then, choosing \( \lambda_1 = \frac{1}{p} \) shows that the system of equations (5.5)-(5.7) gives exactly the same solution as the system (5.8)-(5.10), since they are collinear.

References


[4] Bergkamp, ADD


