UNIVERSITY OF MANNHEIM

University of Mannheim / Department of Economics

Working Paper Series

Why agents need discretion: The business judgment rule as optimal standard of care

Andreas Engert Susanne Goldlücke

Working Paper 13-04

February 2013

Why agents need discretion:

The business judgment rule

as optimal standard of care*

Andreas Engert† and Susanne Goldlücke‡

February 24, 2013

Should managers be liable for ill-conceived business decisions? One answer is given by U.S. courts, which almost never hold managers liable for their mistakes. In this paper, we address the question in a theoretical model of delegated decision making. We find that courts should indeed be lenient as long as contracts are restricted to be linear. With more general compensation schemes, the answer depends on the precision of the court's signal. If courts make many mistakes in evaluating decisions, they should not impose liability for poor business judgment.

JEL-Codes: K13, K22, M53

Keywords: business judgment rule, manager liability, delegated decision-making

1 Introduction

Agents are legally bound to act in the interest of their principals. If they fail to exercise due care, they must be liable for the losses – or so it seems. For corporate directors and officers ("managers" for short) the law turns out to be quite different. It is governed by the "business judgment rule." The leading corporate law court in the U.S. summarized the rule's effect as follows:

*We thank Matthias Lang and participants of the SFB TR 15 meeting 2012 in Caputh as well as seminar participants in Frankfurt and Bonn for helpful comments and suggestions.

†University of Mannheim, Department of Law, Schloss Westflügel, 68131 Mannheim, Germany. Email: engert@uni-mannheim.de.

‡Corresponding author. University of Mannheim, Department of Economics, L7, 3-5, 68161 Mannheim, Germany. Email: goldluecke@uni-mannheim.de. Phone: +49 621 181 3072.

> "[I]n the absence of facts showing self-dealing or improper motive, a corporate officer or director is not legally responsible to the corporation for losses that may be suffered as a result of a decision that an officer made or that directors authorized in good faith. There is a theoretical exception to this general statement that holds that some decisions may be so 'egregious' that liability for losses they cause may follow even in the absence of proof of conflict of interest or improper motivation. The exception, however, has resulted in no awards of money judgments against corporate officers or directors in this jurisdiction" Chancery Court of Delaware (1996).

This is a staggering statement given the potentially enormous agency costs in public corporations. We therefore attempt to evaluate the business judgment rule in a theoretical model. Assuming a risk-averse agent, we use the basic set-up of Lambert (1986), augmented by a liability rule. In our model, liability is part and parcel of the manager's contract. The compensation terms and the standard of care are each set optimally. We consider two dimensions of manager behavior: the effort in preparing a decision and the subsequent choice between a safe and a risky project. Courts receive a signal of the manager's conduct but the difficulties of adjudicating business decisions (and of predicting the court's judgment) create noise in the court's signal. In our basic setup, the noise affects only the court's assessment of whether the manager rightly chose the risky project.

Our main results are the following: If the compensation scheme under the contract is linear, courts should always be lenient. By contrast, non-linear contracts can work against the risk-deterrent effect of liability. In this setting, courts should use precise signals. As the quality of the signal declines, liability causes increasing costs to the principal. If the signal is very noisy, the court should refrain from using it because liability exposes the agent to too much risk even if he is informed and makes careful decisions. This result conforms to the business judgment rule's prescription that liability should obtain only in "egregious" cases (where signal precision is high) but not in others (where signal precision is low). We thus offer an economic justification for the surprising forbearance that U.S. law affords to corporate managers.

It appears that we are the first to formally analyze the business judgment rule and its leniency towards managers.[1] In a related contribution, Gutiérrez (2003) explains why shareholders may wish

[1] Hakenes and Schnabel (2012) study "manager liability" conditional on readily observable outcomes rather than on a finding of negligence after a court proceeding.

to protect managers by obtaining directors' & officers' (D&O) liability coverage or by eliminating liability in the articles of incorporation. While the question resembles ours, the thrust of Gutiérrez' argument is different: In her model, D&O insurance and liability limitations serve to determine the amount of ex post litigation. Inefficiencies result only from shareholders spending too little or too much on suing managers. Liability does not distort the agent's incentives except when there is too little of it. By contrast, we directly address the chilling effect that liability could have on risk-taking.

The paper begins with explaining the business judgment rule and its justification by courts and legal commentators (section 2). Section 3 presents our model, section 4 contains our main results. In section 5, we consider the possibility that managers can be held liable for a failure to take risk. Section 6 relaxes the assumption that the courts can perfectly observe the manager's effort in preparing a decision, and section 7 concludes.

2 The business judgment rule

The classic statement of the business judgment rule is due to the Supreme Court of Delaware (1984):

> "[There is] a presumption that in making a business decision the directors of a corporation acted on an informed basis, in good faith and in the honest belief that the action taken was in the best interests of the company. [...] Absent an abuse of discretion, that judgment will be respected by the courts."

The passage sets out the legal structure of the business judgment rule in Delaware, the dominant corporate law jurisdiction in the U.S.: To a limited degree, the courts reserve judgment on the *process* of decision-making. To invoke the business judgment rule, managers must have considered "all material information reasonably available to them." Yet even at this stage of "process due care," there is a presumption that managers have lived up to their obligations and courts restrict themselves to a gross negligence standard (Supreme Court of Delaware, 1984 and 2000). Once the business judgment rule applies, the courts completely abstain from reviewing the *substance* of the decision (Supreme Court of Delaware, 1984 and 2000). The effort and choice dimensions in our model thus mirror the stages of "process due care" and "substantive due care" in the legal analysis of manager liability cases.

As a result of the business judgment rule, personal liability is only a remote threat for corporate managers:

The Disney Case. In August 1995 the Walt Disney Company hired Michael Ovitz as President of the corporation, only to terminate his employment fourteen months later with a severance package of $130 million. The large payout was essentially mandated by the terms of the compensation agreement. The board's compensation committee had approved the key employment terms in a one-hour meeting together with other agenda items. Although the courts found this "decision-making process [to fall] far short of corporate governance 'best practices,'" they did not question the directors' decision and dismissed any claim of personal liability (Supreme Court of Delaware, 2006).

The Citigroup Case. Citigroup Inc. sustained very significant losses in 2007 and 2008 from its exposure to the U.S. subprime residential mortgage market. Shareholders sued directors and officers of Citigroup for failing to monitor risk in spite of a multitude of warning signs that market conditions were deteriorating. Yet the Delaware Chancery Court ruled out director liability at the outset. The court went so far as to deny a board duty to oversee business risk because this would "involve courts in conducting hindsight evaluations of decisions at the heart of the business judgment of directors" (Chancery Court of Delaware, 2009)

Judges and legal scholars offer a bunch of justifications for the surprising leniency of the business judgment rule: It is said, first, that business decisions are highly specific to the particular situation in which managers have to act. "The judges are not business experts" (Supreme Court of Michigan, 1919). The very same reason that prevents the corporate contract from being complete also impedes the courts in adjudicating claims of mismanagement (Fischel, 1985). Second, legal commentators point to the perils of judging in hindsight. The business judgment rule could serve to offset an inclination to give too much weight to bad outcomes in evaluating ex ante behavior (Chancery Court of Delaware, 2009; Eisenberg, 1993).

A third line of reasoning relates to the proper amount of risk-taking. If managers face a threat of personal liability, the argument goes, they will shy away from risky but value-enhancing projects. Instead of pursuing innovative opportunities they will stick to business as usual. Such fearful behavior runs against the interest of shareholders, particularly in a public corporation where shareholders are well diversified (American Law Institute, 1992, p. 135; Chancery Court of Delaware, 1996; Allen, Jacobs and Strine, 2002). Fourth, unfettered liability for negligence could reduce the supply of able directors and executives. Compensation would rise as a result or the quality of managers would decline (Allen, Jacobs and Strine, 2002; Black, Cheffins and Klausner, 2006b). A fifth and final argument is that there are other and less costly mechanisms to discipline managers. One element consists of price signals from the capital market that feed into stock-based compensation and the market for corporate control. Reputational concerns in the managerial labor market are believed to provide another powerful incentive (Easterbrook and Fischel, 1991, pp. 96-97; Black, Cheffins and Klausner, 2006b).

In the wake of corporate scandals and the great financial crisis, some commentators question the wisdom of the business judgment rule (Campbell, 2011; Nowicki, 2008; Fairfax, 2005). Jurisdictions outside the U.S. often recognize the business judgment rule only to a limited extent. German courts, for instance, seem far less deferential to the business judgment of corporate directors even under criminal law (as in the "Mannesmann" case, Bundesgerichtshof, 2005; Gevurtz, 2007). German lawyers sometimes argue that the business judgment rule only precludes strict liability for business failure but that managers remain accountable for any negligence (Ulmer, 2004). By contrast, our results counsel to restrict liability to cases where signal precision is high, that is where the court can be certain that the manager violated the duty of care. One way to translate this into law is that the applicable standard should be gross negligence or even knowing neglect of due care.

3 The model

Shareholders delegate decision-making to the managers of the corporation. In the following, we refer to shareholders as the principal and to managers as the agent. The agent has to decide between a safe option (continuation, business as usual) and a risky one (growth, reorganization). The safe project yields a certain cash flow r which is normalized to $r = 0$, while the risky project can either yield a return $R > 0$ or a loss $L < 0$. Ex ante, the probability that the risky project yields the high cash flow R is $p \in [0, 1]$. The agent can acquire additional information about the profitability of the risky project. We identify this information with probabilities: If the agent receives information q, he knows that the posterior probability of the high cash flow R is equal to q. From an uninformed perspective, information is distributed according to a distribution function F with $p = \int q dF$.

We assume that the principal is risk-neutral and the agent is risk-averse. The agent's (money) utility function is given by u, which is assumed to be twice differentiable with $u' > 0$ and $u'' \leq 0$. The agent can choose whether to acquire information or not. Acquiring information ($e = 1$) involves an effort cost of $\kappa \geq 0$ for the agent, while staying uninformed ($e = 0$) has no cost. Effort cost and money utility are additively separable. Whether he learned and what he learned ($e \in \{0, 1\}$ and $q \in [0, 1]$) is the agent's private knowledge.

The agent has limited assets: His wealth cannot sink below $A \geq L$ in any contingency. Moreover, the agent only accepts the contract if his expected utility is not lower than his reservation utility $\bar{u} = u(\bar{w})$.

First best

The model so far – without liability – is almost identical to the one in Lambert (1986), which the reader is referred to for a more detailed derivation of the first best solution. In the absence of any incentive problems, the principal can let the agent stay uninformed and pay him a wage $\bar{w}$, or make the agent acquire information and pay him a fixed wage $\tilde{w}$ with $u(\tilde{w}) = \bar{u} + \kappa$. It is optimal to choose the risky project whenever the probability of success is greater than the cut-off

$$\bar{q}^{FB} = \frac{-L}{R - L}. \tag{1}$$

The principal's payoff if the agent stays uninformed is equal to

$$\max\{pR + (1 - p)L, 0\} - \bar{w}. \tag{2}$$

To write down the principal's payoff if the agent obtains information, we introduce the following notation:

$$\begin{aligned} \rho_R(\bar{q}) &= \int_{\bar{q}}^{1} q dF, && (3) \\ \rho_L(\bar{q}) &= \int_{\bar{q}}^{1} (1 - q) dF, && (4) \\ \rho_0(\bar{q}) &= F(\bar{q}). && (5) \end{aligned}$$

and

$$S(\bar{q}) = \rho_R(\bar{q})R + \rho_L(\bar{q})L. \tag{6}$$

In the first best, the agent becomes informed if the principal's maximum payoff is larger than her payoff in the case without information acquisition:

$$S(\bar{q}^{FB}) - \tilde{w} \geq \max\{pR + (1-p)L, 0\} - \bar{w}. \tag{7}$$

Contracts

Contracts specify transfers conditional on outcomes, $w(x)$ with $x \in \{L, 0, R\}$. We sometimes use the following short-cut notation: $w_x = w(x)$ and $u_x = u(w(x))$ for $x \in \{L, 0, R\}$, and $u_A = u(A)$. By restricting attention to this class of contracts, we adopt an incomplete contracting approach: There is no communication once the agent has acquired information. This means that there is no stage in the game at which the agent is informed and can choose from a menu of contracts. With this assumption we follow the literature on delegated expertise (Lambert 1986, Demski and Sappington 1987, Gromb and Martimort 2007).[2] Another critical assumption is that the contract cannot condition on the outcome of the risky investment if the agent implements the safe project. In fact, it would often be exceedingly difficult to determine the forgone profits or losses if the manager decides not to pursue a risky investment. We make the same assumption for court-imposed liability (see the next subsection).

Following other contributions (e.g., Lambert and Larcker, 2004; Kadan and Swinkels, 2008), we restrict the range of possible contracts further and focus on two subclasses: (Affine-)linear or sharing contracts of the form $w(x) = \beta + \alpha x$ and monotonic, nondecreasing contracts. Linear contracts can only consist of a fixed salary, a variable salary that is proportional to the firm's revenue and/or granting the agent a share in the firm. Monotonic compensation schemes can involve salary, stock, options and/or bonuses for performance above a certain threshold.

The monotonicity constraint is often imposed in models of moral hazard (Innes, 1990; Matthews, 2001), and for good reason: Non-monotonic contracts are rarely if ever used, presumably because

[2] Other papers that allow only a single payment rule and no menus include Matthews (2001) and Malcomson (2009),(2011). Matthews (2001, p.3) argues that "the rarity of menu contracts in reality suggests that they may often have low benefit or high cost". Similarly, Raith (2008) assumes that the agent possesses "specific knowledge" about his productivity that he cannot communicate to the principal.

rewarding an agent for reduced performance would create problems of ex post moral hazard. For instance, when a risky investment starts to pay off, the manager might have to sell it at an undervalue to avoid being penalized for being too successful. By imposing the monotonicity restriction, we show that our results do not turn on such an implausible compensation scheme. They would, however, continue to hold without assuming monotonicity.

Courts

The contract uses readily observable information on outcomes. By contrast, whether the agent has exercised (process or substantive) due care is not obvious and, therefore, has to be determined by a court. In the main part of the paper, we assume that the agent can only be held to account if the risky project fails. There is no liability for taking the safe business-as-usual decision even if the risky project promises a superior expected return. The reason for this assumption is the following: Courts calculate damages by comparing the actual outcome with a hypothetical outcome that would have obtained if the agent had exercised due care. With the risky project, the courts would have to estimate the expected return over various states of the world as the hypothetical outcome. Courts are reluctant to engage in such guesswork, especially if the risky choice consists of a specific business strategy that only the firm's managers can devise. Even if the courts were willing to consider such claims, their estimate of the damages would be very hard to predict. As a result, litigation is significantly more attractive to the principal when the risky project has caused a major loss. Nonetheless, in Section 5 we also consider the case that both the risky and the safe choice can trigger liability.

During trial, the court receives signals regarding the agent's behavior. It can try to answer the following questions: Did the agent obtain enough information? Did he take the correct decision from an ex ante point of view? As to the first question, the court will seek to establish whether the agent was sufficiently informed about his decision, that is, whether he exercised process due care. The court receives a signal e^c about the agent's effort to become informed. Regarding the second question, the court receives a signal q^c about the probability that the risky project has a high return. A liability rule compares these signals to standards of care $\bar{e}^c$ and $\bar{q}^c$. We assume that the court seeks to promote efficient contracting and applies standards that are optimal for the problem at hand, which may well differ from the first best values e^{FB} and q^{FB}.[3] We could say equivalently that the principal herself stipulates the standards in the contract.

Depending on how signals compare with the standards, courts rule that the agent is liable or not. If the agent is held liable, he has to pay damages to compensate the principal. We assume that the wealth constraint always binds in this case.[4]

In the main part of the paper, we focus on liability for substantive due care. We therefore assume that the court's signal about effort is a perfect one, $e^c = e$. As a consequence, it is always optimal to impose process due care liability, with the standard set equal to $\bar{e}^c = e^{SB}$. If $e^c < \bar{e}^c$, the agent is liable; otherwise, the court finds the agent liable if the signal q^c is below the standard of care $\bar{q}^c$.

Further, we assume that $q^c = q + \epsilon$, where ϵ follows a symmetric distribution with mean zero.[5] This distribution is denoted by Φ. We assume that Φ is differentiable with density $\phi > 0$ in the interior of the support. For a given $0 < \bar{q}^c < 1$, we define

$$\lambda(q) = Prob[q^c < \bar{q}^c | q] = \Phi(\bar{q}^c - q) \tag{8}$$

as the probability of being found liable for a violation of substantive due care if the true success probability is q. The function λ is continuous and decreasing in q, with $\lambda(\bar{q}^c) = \frac{1}{2}$ and $\lambda'(q) = -\phi(\bar{q}^c - q)$ for all $0 < q < 1$. There are two interpretations of the function λ: One is that in adjudicating substantive due care, the court errs in perceiving the true success probability q or in applying its standard of care $\bar{q}^c$. The second interpretation is that the agent himself is uncertain about the true success probability, the court's ability or the standard of care.

We furthermore define $\lambda(q)$=0 for $\bar{q}^c = 0$ and $\lambda(q) = 1$ for $\bar{q}^c = 1$. We sometimes single out the case that $\bar{q}^c = 0$ as a rule of "no liability for substantive due care" and denote it by λ^{nl}. A liability rule that is described by standards $\bar{e}^c = 1$ and $\bar{q}^c$ thus translates into the following probabilities of being liable for the agent: If $e = 0$, the agent is liable with probability 1. In case that $e = 1$, his probability of being liable is equal to $\lambda(q)$, where λ depends on the standard $\bar{q}^c$ and therefore stands

[3] Since fiduciary duties are designed to promote the shareholder's interest, it makes sense to assume that the court maximizes the shareholders' payoff, which also results in a Pareto efficient allocation.

[4] This assumption means that if D are the damages, then it must be that $w_L - D \leq A$, which can be shown to hold if, for example, $D = -L$ and $\bar{w} \leq 0$.

[5] Note that we allow $q^c < 0$ and $q^c > 1$. Such signals are treated as 0 and 1, respectively. This yields differentiability of λ_e with respect to q, but also means that there is a discontinuity with regard to the standard of care at $\bar{q}^c = 0$ and $\bar{q}^c = 1$. A standard of zero means no liability, but a small but positive standard could be very different from no liability.

for a whole family of liability rules.

4 Analysis

The ex post stage: choosing projects

We will first look at the agent's choice between projects. With a contract w and a liability rule in place, an informed agent who knows the probability of the good state to be q chooses the risky project if and only if

$$qu_R + (1-q)((1-\lambda(q))u_L + \lambda(q)u_A) \geq u_0. \tag{9}$$

We assume that if for some probabilities the agent is indifferent between the alternatives, then he makes the efficient decision.

Lemma 1. *There exists a cut-off point $\bar{q}(w)$ with*

$$\bar{q}(w)u_R + (1-\bar{q}(w))\big((1-\lambda(\bar{q}(w)))u_L + \lambda(\bar{q}(w))u_A\big) = u_0, \tag{10}$$

such that the agent chooses the safe project for all $q \leq \bar{q}(w)$ and the risky project for $q > \bar{q}(w)$.

If there is more than one solution to this equation, then $\bar{q}(w)$ is the one that is closest to $\bar{q}^{FB}$. Wherever the contract is clear from the context, we drop the reference to w and write only $\bar{q}$. The cut-off in the special case of $\bar{q}^c = 0$ is denoted by $\bar{q}^{nl}$, i.e.,

$$\bar{q}^{nl}u_R + (1-\bar{q}^{nl})u_L = u_0. \tag{11}$$

The analysis of the ex post stage provides a first intuition of how liability can deter efficient risk-taking. Assuming that the agent simply gets a fixed salary equal to his reservation wage $\bar{w}$, he is indifferent between the risky and the safe project and always chooses the one with the higher return. The cut-off point is $\bar{q}^{nl} = \bar{q}^{FB}$. But once there is the slightest probability of liability, the safe project provides a safe harbor against liability and, therefore, strictly dominates the risky project no matter how valuable the latter. Hence, with a fixed wage no liability (with standards $\bar{e}^c = 0$ and $\bar{q}^c = 0$, or any standard leading to $\lambda(\bar{q}^{FB}) = 0$) is optimal. The effort dimension plays no role as it is not possible to induce information acquisition with a fixed wage.

Of course, agents often get paid for performance. In the following, we study how incentive pay affects the agent's choice under the threat of liability. We start with the special case of linear compensation schemes.

Lemma 2. *For every linear contract $w(x) = \beta + \alpha x$ with $\alpha \leq \frac{A-\beta}{L}$, it holds that $\bar{q}(w) \geq \bar{q}^{FB}$. In addition, for any two contracts $w(x) = \beta + \alpha x$ and $\hat{w}(x) = \hat{\beta} + \alpha x$ with $0 \leq \beta - \hat{\beta} \leq \lambda(\bar{q}(w))(1 - \bar{q}(w))(\beta - A + \alpha L)$, it holds that $\bar{q}(w) \geq \bar{q}^{nl}(\hat{w})$.*

This lemma says that with a linear contract for which the limited liability constraint does not bind, there is underinvestment in the risky project. The intuition is simple: With a linear contract and no liability, a risk-neutral agent chooses the first best cut-off, while a risk-averse agent will at the first best cut-off still prefer the safe project. Liability only makes the risky project less attractive. Indeed, as the lemma also shows, the agent will be more inclined to take risk if he is relieved of substantive due care liability.

While risk aversion thus limits the set of decision thresholds that can be implemented under a linear contract, under a more general contract all decision thresholds $\bar{q} \in [0, 1]$ can be implemented. In the following, we consider incentives to acquire information and solve the full optimization problem for both classes of contracts.

The ex ante stage: choosing effort

Every wage scheme w together with the liability rule described by λ induces an effort level $e(w) \in \{0, 1\}$ and a decision threshold $\bar{q}(w) \in [0, 1]$. We can interpret e and $\bar{q}$ as part of the contract. It may be optimal to implement no information acquisition at all, i.e. $e^{SB} = 0$. This can be done with a contract $w(x) = \bar{w}$. Because we assume that the agent makes the efficient choice whenever he is indifferent, there is no need for substantive due care liability, hence $\bar{q}^c = 0$ is optimal.

Thus, the interesting case is that the optimal contract requires the agent to collect information. We write an informed agent's utility who faces contract w and decides according to threshold $\bar{q}$ as

$$U(w, \bar{q}) = \rho_R(\bar{q})u_R + \rho_L(\bar{q})u_L - \rho_\lambda(\bar{q})(u_L - u_A) + \rho_0(\bar{q})u_0 - \kappa, \tag{12}$$

where

$$\rho_\lambda(\bar{q}) = \int_{\bar{q}}^{1} (1-q)\lambda(q)dF. \tag{13}$$

In the following, we will derive the constraints that a contract w with $e = 1$ and decision threshold $\bar{q}$ has to satisfy. First, $\bar{q}$ must be the cut-off above which the agent chooses the risky project:

$$\bar{q}u_R + (1-\bar{q})((1-\lambda(\bar{q}))u_L + \lambda(\bar{q})u_A) = u_0. \tag{D}$$

The contract also has to make sure that the agent prefers acquiring information to just choosing the safe project:

$$U(w, \bar{q}) \geq u_0. \tag{SIC}$$

Similarly, the agent should prefer acquiring information to just choosing the risky project:

$$U(w, \bar{q}) \geq pu_R + (1-p)u_A. \tag{RIC}$$

In addition to these two incentive compatibility constraints, we have the participation constraint

$$U(w, \bar{q}) \geq \bar{u}, \tag{PC}$$

and the limited liability constraint

$$w_L \geq A. \tag{LL}$$

Finally, there is either the monotonicity constraint

$$w_R \geq w_0 \geq w_L, \tag{MON}$$

or the stronger linearity constraint

$$w_x = \beta + \alpha x. \tag{LIN}$$

The principal maximizes her payoff

$$\pi(w, \bar{q}) = S(\bar{q}) - \rho_R(\bar{q})w_R - \rho_L(\bar{q})w_L + \rho_\lambda(\bar{q})(w_L - A) - \rho_0 w_0 \tag{14}$$

subject to the constraints $(D), (SIC), (RIC), (PC), (LL)$ and (MON) or (LIN), respectively.

Linear contracts

We start again with the analysis of linear contracts.

Proposition 1. *If contracts are linear, the optimal standard $\bar{q}^c$ is equal to zero.*

If contracts are linear there should be no liability for making wrong judgments, only for careless preparation of the decision. The intuition behind this result is the following. We have shown in Lemma 2 that with a linear contract, the agent chooses the risky project only if the probability of success of the risky project is relatively large, larger than the first best cut-off $\bar{q}^{FB}$. The contract thus has to induce more risk taking and – as we have also shown in Lemma 2 – that is what less liability achieves. An optimal liability rule therefore holds the agent liable with probability zero.

Non-decreasing contracts

Linear contracts are not optimal in this setting. If the principal can freely specify compensation for each contingency, she can offer extra rewards for pursuing the risky project. In fact, with liability the principal could even want to pay more for suffering losses from the risky project than for the safe return. By setting $w_L > w_0$ or even $w_L > w_R$, she may be able to mitigate the risk-deterrent effect of liability. Consequently, the monotonicity constraint may be binding once it becomes optimal to impose substantive due care liability. Remember that we include the monotonicity constraint to rule out compensation schemes that are barely observed in reality. The following results hold even without this additional restriction.

Proposition 2. *Consider the problem of implementing a given $\bar{q}$ at minimum cost. In the optimum, (SIC) is always binding if (LL) is not binding. Let the optimal contract be denoted by w^λ, and the optimal contract without liability by w^{nl}. Then the following comparisons hold:*[6]

$$w_R^{nl} \geq w_R^\lambda \text{ and } w_L^\lambda \geq w_L^{nl}. \tag{15}$$

Proposition 2 states that imposing substantive due care liability compresses the optimal compensation scheme. This result may seem counterintuitive as the contract has to work against the

[6] Note that if the optimal standard is equal to zero, then we have the special case that $w^{nl} = w^\lambda$.

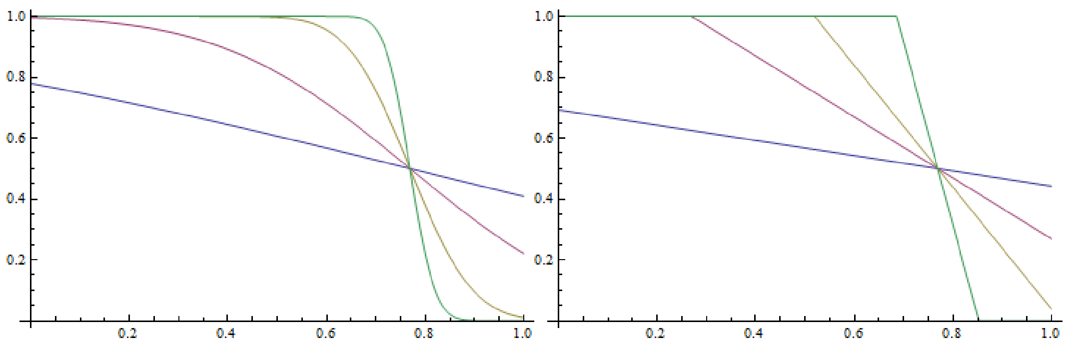

Figure 1: The left panel shows the function $\lambda(q)$ for the case of a normally distributed error term for different levels of precision Δ. A perfect signal would be represented by a step function with a jump at $\bar{q}^c = 0.8$. The right panel shows $\lambda(q)$ for the case of a uniformly distributed error term, again for different levels of precision.

agent's increased incentive to avoid risk. Yet liability also increases the value of making an informed decision and in this regard substitutes for performance-based compensation. Proposition 2 shows that the latter effect always prevails if the optimal standard of care is positive. Intuitively, imposing substantive due care liability can only be optimal if the compressed wage scheme saves the principal more than she has to spend on compensating the agent for the liability risk when he chooses the risky project.

How the standard of substantive due care should be set depends on the precision of the court's signal q^c. To measure the impact of precision, we assume that the signal $q^c = q + \epsilon$ follows a distribution that depends on a parameter Δ, where a larger Δ means greater precision. What is meant by precision is illustrated in Figure 1. There we show the function λ for a normal and a uniform error term, i.e. we consider the case that ϵ follows a normal distribution with mean 0 and variance $1/\Delta$ as well as the case that ϵ follows a uniform distribution with

$$\Phi(\epsilon) = \min\{1, \max\{0, (1 + \epsilon\Delta)/2\}\}. \tag{16}$$

For $\Delta \to \infty$ the case of a perfect signal, $q^c = q$, is approached, while for $\Delta \to 0$ the signal becomes perfectly uninformative, i.e., $\lambda(q) = 1/2$ independent of q. We require that λ depends on Δ in a differentiable way and such that the functions cross only at a single point. We therefore make the following technical assumption.

Assumption 1. *The function Φ is differentiable with respect to Δ for all ϵ in the interior of the* *support, with* $\frac{\partial\lambda(q)}{\partial\Delta} \geq 0$ *for* $q \leq \bar{q}^c$ *and* $\frac{\partial\lambda(q)}{\partial\Delta} \leq 0$ *for* $q \geq \bar{q}^c$*. Moreover we assume that for the inverse function* $\lambda^{-1} : (0,1) \to \mathbb{R}$ *it holds that the derivative of* λ^{-1} *is weakly increasing in* Δ*.*

The last condition can be illustrated as follows using Figure 1: Imagine a horizontal line through any point below 0.5 on the vertical axis. The line is intersected by liability functions for different values of Δ, where functions that correspond to a larger Δ intersect the line more to the left. The assumption says that as one moves from left to right along this line, the slope of the liability functions that intersect the line becomes flatter. Similarly, if one draws a horizontal line through any point above 0.5 on the vertical axis, then if one moves from left to right along this line, the slope of the liability functions that intersect the line becomes steeper.

Remark 1. Assumption 1 holds for the normal and the uniform error terms.

That this remark is true is quite intuitive from Figure 1. Nevertheless, we provide a formal proof in the appendix.

We can now state the main result of this section.

Proposition 3. *There exists a cut-off* $\bar{\Delta}$ *such that the optimal standard is equal to zero if* $\Delta < \bar{\Delta}$*, while it is positive if* $\Delta \geq \bar{\Delta}$*.*

Basically, the trade-off between no liability and liability is between insurance for agents who learn a high q and agents who learn a lower q. If $\rho_\lambda(\bar{q})$ and $\lambda(\bar{q})$ could be set separately, the principal would maximize $\lambda(\bar{q})$ to punish uninformed decisions and minimize $\rho_\lambda(\bar{q})$ because a large $\rho_\lambda(\bar{q})$ exposes the agent to risk that he has to be compensated for. A perfect signal, which essentially sets $\rho_\lambda(\bar{q})$ equal to zero and $\lambda(\bar{q})$ equal to one, is always beneficial. If the signal is not perfect but still very precise, a large $\lambda(\bar{q})$ can punish risk taking with success probabilities closely below $\bar{q}$ without punishing risk taking with larger success probabilities too much. When the signal becomes very imprecise, then exposing the agent to liability risk for values close to $\bar{q}$ implies also exposing him to risk for larger values of q.

Discussion

Proposition 3 enshrines the economic rationale of the business judgment rule. It militates against substantive due care liability if the courts commit too many mistakes in evaluating business decisions or, equivalently, if agents misjudge what courts require. A complementary interpretation is that signal precision varies over cases. While the courts may be able to assess some business decisions, others present "hard cases" with a high probability of error. In this reading, Proposition 3 implies that liability should be confined to straightforward cases of evident mismanagement.

The business judgment rule thus can be a response to the difficulty and error-proneness of evaluating business decisions in the courtroom ("judges are not business experts"). The model also captures other considerations advanced by legal commentators: With a noisy signal, performance pay becomes comparatively more efficient in incentivizing the agent (the alternative-mechanisms argument).[7] If the courts raised the standard of care beyond the optimum, the principal would adjust the compensation scheme to counteract the risk-deterrent effect of liability. Hiring and incentivizing an agent would become more expensive (the higher-compensation argument). Without an adjustment, the agent would be less inclined to choose the risky project, even if it is worthwhile (the risk-deterrence argument).

We can also take up the concern about a potential hindsight bias. It is often said that even experts or professional judges overestimate the ex ante probability of an outcome that they observe ex post. Let $\bar{q}^c$ denote the optimal standard (without hindsight bias). Assume that there is a hindsight bias of the form that if the court observes signal q^c, it actually sees the signal $q^c - h$, for some $\bar{q}^c > h > 0$. With a bias, the probability of being liable from the agent's perspective is equal to $\lambda(q) = Prob[q^c - h \leq \bar{q}^c|q] = Prob[q^c \leq \bar{q}^c + h|q]$. The optimal standard is now $\bar{q}^c - h$, which is weakly decreasing in the size of the hindsight bias h. If courts cannot overcome the hindsight bias, the optimal standard should reflect the bias and be lower. However, the prevalence and extent of the bias are subject to debate.[8] Also, it is not clear how judges would apply a standard that aims at correcting their average hindsight bias.

In our model, the agent is exposed to risk from compensation and liability in order to induce him to become informed. The lower the agent's cost of effort κ, the less risk is needed. In the extreme case that $\kappa \to 0$, a wage of slightly more than $\bar{w}$ can almost achieve the first best. Introducing liability would only introduce unnecessary and costly risk.

[7]Remember, however, that liability can only be desirable if there is a non-linear compensation scheme and in this sense the two act as complementary mechanisms.

[8]In a recent study, Rachlinski, Wistrich and Guthrie (2011) find that court rulings on the legality of searches in criminal proceedings do not vary between foresight and hindsight.

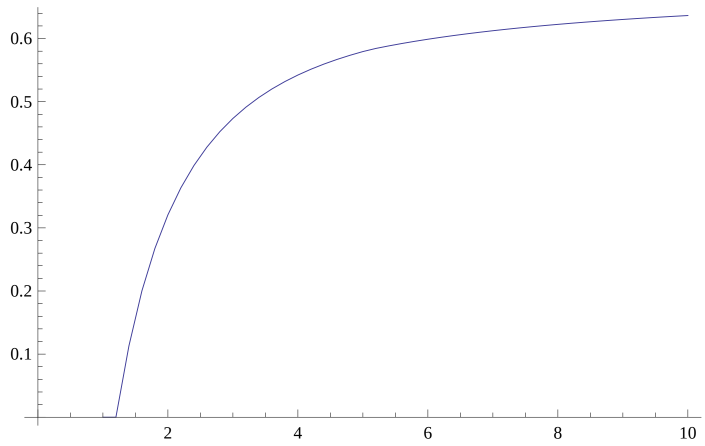

Figure 2: With precision (Δ) measured on the horizontal axis, this graph shows the optimal $\bar{q}^c$ in dependence on Δ for the following example: $u(w) = \sqrt{w}$, $\kappa = 1$, $\bar{u} = 20$, $A = L = 0$, $R = 1000$, $r = 700$, q and ϵ follow uniform distributions. The function is increasing and always below the first best cut-off $\bar{q}^{FB} = 0.7$ (even as $\Delta \to \infty$, the optimal standard does not approach the first best).

Remark 2. If κ is sufficiently low, then the optimal standard is equal to zero.

This finding has immediate policy implications. Certain agents make far-reaching decisions but at relatively little personal cost in terms of time and labor. An example are outside directors on corporate boards. While their responsibility is to monitor the corporate officers and to ratify important decisions, they do so based on information provided by the corporate officers, accountants etc. Their own effort is limited to several board meetings per year.[9] Remark 2 counsels to eliminate substantive due care liability in this setting. In fact, the law appears to be especially lenient towards outside directors (Black, Cheffins and Klausner, 2006a). To the extent that outside directors receive only a fixed wage or stock, our result for linear contracts leads to the same conclusion. In a similar vein, agents in charity and other pro bono activities should not be subject to liability if their decisions involve risky choices.

For agents with significant effort cost and a convex compensation scheme, the implications are less straightforward. Proposition 3 implies that beyond a certain level of noise in the court's signal, the agent should not face substantive due care liability. While we cannot show in general that the optimal standard is decreasing as the signal becomes noisier, Figure 2 shows a typical example, in which this is the case.

It is sometimes argued that a limit on damages borne by the agent can render liability less

[9]In the taxonomy of Fama and Jensen (1983), outside directors exercise "decision control" while "decision management" is vested in the executive directors and officers.

harmful and indeed desirable. In fact, if one can set an appropriate cap for damages (i.e., if A is chosen rather than determined by the agent's wealth), even a very noisy signal can be used. Turning this result into policy advice is difficult though. The court would have to adjust the damage cap to the difficulty of adjudicating the case (or, similarly, the severity of the agent's fault) as well as to the agent's private wealth and risk aversion. While conceivable, this is not what courts usually do in private law litigation. In principle, one could also devise a D&O insurance with a variable deductible reflecting the noise in the court's signal.[10]

5 No safe haven

The basic model presumes that the agent is never liable for taking the safe path of "business as usual." This assumption is plausible when liability for the safe choice would require the court to compare the actual outcome to a hypothetical decision with a broad range of possible consequences. Sometimes, however, the implications of taking the risky decision can be observed quite well. For instance, in a takeover the safe choice on behalf of shareholders consists of selling the firm for cash (as happened in the famous Smith v. Van Gorkom case, Supreme Court of Delaware, 1985). But because the firm often continues to operate after the takeover, there is an indication of how the shareholders would have fared with keeping the firm. In such a setting, courts can use the observed outcome to calculate damages.

To study this possibility, we assume in this section that the outcome of the risky project is observable ex post. What if the principal can (and will) sue, not just in the case of a large loss, but whenever the performance could have been better? If the risky project would have yielded the superior return and the agent is found liable for pursuing the safe project instead, damages amount to $d = R$; but as before, we assume that the limited liability constraint is binding. Again, the agent is always liable if he fails to become informed (and his project performs poorly). Therefore, for an uninformed agent the expected payoff from the safe choice is $(1-p)u_0 + pu_A$.

If the agent has observed process due care, the court evaluates the substantive merit of his decision. In principle, the court could apply the same standard $\bar{q}^c$ to impose liability for taking the

[10]To protect the agent, the insurance would have to cover any damages exceeding the deductible. In such a setting, moral hazard would become a serious concern if liability were imposed in spite of a noisy signal (i.e., for minor mistakes).

risky and the safe decision. When the agent pursues the safe project, he is held liable if $q^c > \bar{q}^c$; his expected payoff would be $u_0 - q(1 - \lambda(q))(u_0 - u_A)$. However, using the same standard would prevent the court from adapting the standard to a noisy signal because relaxing it for the safe choice would imply a tightening for the risky choice (and conversely). Alternatively, two different standards can be used: The court can infer "too much risk-taking" from $q^c < \bar{q}^c$ and "too little risk-taking" from $q^c > \bar{q}^c_0$, with $\bar{q}^c_0 > \bar{q}^c$. We take some function $l(q)$ to measure an informed agent's expected probability of being liable after choosing the safe project. The court can influence this function by setting a standard $\bar{q}^c_0$. The function that we have in mind is $l(q) = q\Phi(q - \bar{q}^c_0)$, but it can be any other function that decreases in the standard $\bar{q}^c_0$ and is differentiable for $0 < q < 1$ and $0 < \bar{q}^c_0 < 1$. It should be noted that the agent's decision is still monotone in q: if he prefers the risky project for $\bar{q}$, then he does so for all $q \geq \bar{q}$.

Proposition 4. *In the setting without a safe haven, the optimal standard $\bar{q}^c_0$ is equal to one. If F is the uniform distribution, then $\bar{q}^c$ is equal to zero.*

An informed agent should never be held liable for taking the safe choice. The reason is that the cheapest way to restrain the agent's inclination towards the safe project is to reduce compensation w_0 in that contingency. True, liability uses more information, namely the court's signal q^c and the hypothetical outcome of the risky project. But this time the information cannot be used to reduce risk: While liability for wrongfully taking the risky choice substitutes for variation in compensation between w_L and w_R, there is no such effect for the safe choice.

Without a safe haven, liability sometimes becomes inefficient for the risky choice as well. Even the first best can be implemented in some cases now. At first blush, eliminating the safe haven offers the best of both worlds: optimal outcomes without imposing liability. One must keep in mind though that Proposition 4 only concerns liability for substantive due care. The result very much depends on the fact that – in contrast to the basic model – liability for process due care now applies to both the safe and the risky choice. Because we still assume that the court observes the agent's effort perfectly, inducing him to become informed requires no exposure to risk. In the next section, we examine how much our results depend on this assumption.

6 Robustness: uncertainty in the effort dimension

So far, we have assumed that the court's signal e^c about the agent's effort to acquire information is perfectly precise. This allowed us to neglect the effort dimension throughout the analysis. A natural question to ask is what happens if we relax this assumption. Does the gist of our analysis hold up if courts commit mistakes in assessing the agent's procedural care? Judging the proper decision-making process can be difficult: Whether more information would have been needed rises similar questions as the validity of the decision. Also, courts may find it hard to distinguish genuine diligence in preparing the decision from a pretense of gathering relevant information that only serves to shield the agent against liability.

To model noise in the court's assessment of procedural care, we introduce a distribution function Ψ_e, $e \in \{0, 1\}$ for the noisy signal e^c and a standard $\bar{e}^c \in [0, 1]$ for process due care. Important are the following two probabilities:

$$\Psi_0(\bar{e}^c) = Prob[e^c < \bar{e}^c | e = 0]) \text{ and } \Psi_1(\bar{e}^c) = Prob[e^c < \bar{e}^c | e = 1]. \tag{17}$$

The court can influence these two probabilities by setting the standard. Again, a standard $\bar{e}^c = 0$ means no liability in the effort dimension.

Proposition 5. *If there exists a standard of care $\bar{e}^c = \tilde{e}$ such that $\Psi_1(\tilde{e}) = 0$ and at the same time*

$$\Psi_0(\tilde{e}) \geq \frac{p\kappa}{(\bar{u} - u_A) \int_p^1 (q - p) dF}, \tag{18}$$

then our results still hold: The standards $\bar{q}^c = 0$ and $\bar{e}^c = \min\{e^{SB}, \tilde{e}\}$ are optimal if contracts are linear (Proposition 1), if contracts are monotone and $\Delta \leq \bar{\Delta}$(Proposition 3) and if q is uniformly distributed in the setting without a safe haven (Proposition 4); imposing an optimal standard $\bar{q}^c > 0$ compresses the compensation scheme as before (Proposition 2).

Proposition 5 essentially states that as long as the signal about effort is sufficiently precise, our earlier results remain unaffected; the process due care standard should be such that an informed agent is never held liable. The basic idea is that process due care only matters because it penalizes uninformed risk-taking and thereby relaxes the risky choice incentive constraint (RIC). With a precise signal, one can ensure that (RIC) does not bind and still relieve the agent of liability risk if he becomes informed. In this case, liability for process due care only confers benefits. With (RIC) out of the way, the analysis becomes the same as with a perfect signal on effort. Therefore, our earlier results are confirmed.

A standard $\tilde{e}$ as in the proposition only exists if the signal e^c is sufficiently precise. Note that the standard can only be set at $\bar{e}^c = 1$ if the signal is perfect. As the signal becomes less precise, the standard has to decrease. Yet $\Psi_0(\tilde{e})$ must remain sufficiently large as required by condition (18).

Condition (18) is easier to satisfy the lower p and κ and the larger $\bar{u} - u_A$. A low p implies that the risky project is less attractive from an uninformed perspective and (RIC) is therefore less likely to bind. With a low κ, the contract can be relatively flat, which again makes (RIC) less likely to bind. Finally, a large $\bar{u} - u_A$ implies that being held liable hurts the agent a lot; the probability $\Psi_0(\tilde{e})$ can then be lower.

It should also be noted that Proposition 5 provides only a sufficient condition. There are large parameter regions for which our results also continue to hold. In fact, liability for neither process nor substantive due care should be imposed in certain cases such as the one set out in the following corollary:

Corollary 1. *If $\bar{q}^{FB} \geq p$, then standards $\bar{q}^c = 0$ and $\bar{e}^c = 0$ are optimal if contracts are linear and if $\Delta \leq \bar{\Delta}$.*

7 Conclusion

We have developed a theoretical argument for why courts in the U.S. routinely abstain from imposing liability for poor business judgment. Shareholders want managers to take risks, but also to be diligent and careful in pursuing risky projects. Following the legal analysis applied by the courts, we distinguish liability for lack of effort in preparing a risk-taking decision (process due care) and liability for the decision itself (substantive due care). Our key insight is the following: As long as the courts administer liability in the effort dimension reasonably well, they should be reluctant to second-guess managerial decisions. This prescription applies if compensation relates linearly to performance, if liability can also be imposed for failure to take risk or, most importantly, if courts (or managers) often err in evaluating business decisions.

Our model has direct policy implications. Outside directors on corporate boards often receive only a flat salary or shares in the corporation. The result for linear contracts suggests that they should not be subject to substantive due care liability. The same is evidently true for pro bono directors in charitable organizations.

For corporate managers with a non-linear, convex compensation scheme, policy advice is less clear-cut: Courts should refrain from imposing liability if they commit too many mistakes in adjudicating the ex ante validity of business decisions, or if it is hard for managers to anticipate the court's ruling. The sweeping business judgment rule in the U.S. suggests that courts have little confidence in their ability to review and guide management decisions. This view is even more appealing if litigation costs are taken into account. However, one could also be more optimistic. Perhaps specialized courts can handle even difficult business cases in a sophisticated and predictable manner. With such expert courts, substantive due care liability could be efficient. Ultimately, the case for or against the business judgment rule must be made on empirical grounds.

We motivated our analysis with the example of corporate directors and officers. However, the model carries over to other settings in which an agent makes risky decisions on behalf of a principal. Asset managers are a case in point. To invest the capital of their clients, they continuously choose among projects with different risk-reward profiles. Our analysis suggests that exposing an asset manager to substantive due care liability can be costly for the client, especially if the asset manager can pay relatively large damage awards. Isolating the agent from liability for his decisions – granting him discretion – can be efficient in this agency relation as well as in many others, such as for business consultants or attorneys. Discretion is, in this sense, a general feature of the law of agency.

In our model, the interests of the agent and the principal diverge with regard to the effort in decision-making, not to the decision itself. The problem only arises because incentivizing the agent can distort his decision: He can avoid having to exert effort (and liability for failure to take effort) by choosing the safe project. Liability in the decision dimension reduces the cost of inducing the agent to take profitable risks – if it is used at all. The analysis changes fundamentally if a conflict of interest arises with regard to the decision. For instance, the agent may obtain a private benefit from choosing the risky project. As the agent strictly prefers the risky choice even in the absence of any incentive contract, there can be a greater role for liability in the decision dimension. In legal terms, this case calls up the agent's "duty of loyalty," to which the business judgment rule does not apply. We leave an inquiry of the duty of loyalty to future work.

Appendix

Proof of Lemma 1. A $\bar{q}$ satisfying (10) exists, since for $q = 1$ the agent's payoff from the risky project is equal to $u_R \geq u_0$ and for $q = 0$ it is equal to $(1 - \lambda(0))u_L + \lambda(0)u_A \leq u_0$. Note that these inequalities would hold in every optimal contract and not just because of the monotonicity constraint. We show first that for all $q < \bar{q}$, the agent prefers the safe project. Because $u_R \geq u_0$, it must be true that

$$(1 - \lambda(\bar{q}))u_L + \lambda(\bar{q})u_A \leq u_0. \tag{19}$$

Because λ is a weakly decreasing function, for $q < \bar{q}$ it holds that

$$(1 - \lambda(q))u_L + \lambda(q)u_A \leq (1 - \lambda(\bar{q}))u_L + \lambda(\bar{q})u_A. \tag{20}$$

It then also follows that

$$qu_R + (1 - q)\big((1 - \lambda(q))u_L + \lambda(q)u_A\big) \leq u_0. \tag{21}$$

Next, we show that for all $q > \bar{q}$ the agent prefers the risky project. It holds that

$$(1 - \lambda(q))u_L + \lambda(q)u_A \geq (1 - \lambda(\bar{q}))u_L + \lambda(\bar{q})u_A, \tag{22}$$

such that if $(1 - \lambda(q))u_L + \lambda(q)u_A \leq u_0$ we again have that

$$qu_R + (1 - q)\big((1 - \lambda(q))u_L + \lambda(q)u_A\big) \geq u_0, \tag{23}$$

because there is more weight on the larger of the two payoffs. If instead $(1 - \lambda(q))u_L + \lambda(q)u_A \geq u_0$, the inequality holds as well. □

Proof of Lemma 2. The first claim is a direct consequence of risk aversion. With a linear contract, the expected wage (without liability) from the risky project at $q < \bar{q}^{FB}$ is $\beta + \alpha(qR + (1 - q)L) < \beta$. Hence, the agent chooses the safe project even if he is risk-neutral and $\lambda(q) = 0$. Liability and risk aversion will only distort the agent's choice towards the safe choice.

For the proof of the second claim, we define $l = \beta + \alpha L - A \geq 0$. We will show that for the probability $\bar{q}$ at which

$$u(\beta) = \bar{q}u(\beta + \alpha R) + (1 - \bar{q})\big((1 - \lambda(\bar{q}))u(\beta + \alpha L) + \lambda(\bar{q})u_A\big), \tag{24}$$

the agent also prefers the risky to the safe project with no liability and the contract $\hat{w}$, i.e.,

$$u(\hat{\beta}) \leq \bar{q}u(\hat{\beta} + \alpha R) + (1 - \bar{q})u(\hat{\beta} + \alpha L). \tag{25}$$

The intuition behind this result is that the agent would be willing to pay at least $(1 - \bar{q})\lambda(\bar{q})l$ to be insured against the additional lottery to lose l with probability $\lambda(\bar{q})$ in the event of failure. Insurance against the additional risk of liability makes the lottery more attractive. The result can be proved by first noting that, due to concavity of u,

$$u(\hat{\beta}) \leq u(\beta) - (\beta - \hat{\beta})u'(\beta). \tag{26}$$

Using (24), we can conclude that

$$u(\hat{\beta}) \leq \bar{q}u(\beta + \alpha R) + (1 - \bar{q})u(\beta + \alpha L - \lambda(\bar{q})l) - (\beta - \hat{\beta})u'(\beta). \tag{27}$$

With reasoning as before, we get

$$u(\hat{\beta}) \leq \bar{q}u(\hat{\beta} + \alpha R) + (1 - \bar{q})u(\beta + \alpha L - \lambda(\bar{q})l) + (\beta - \hat{\beta})(\bar{q}u'(\hat{\beta} + \alpha R) - u'(\beta)) \tag{28}$$

and from this

$$\begin{aligned} u(\hat{\beta}) \quad \leq \quad & \bar{q}u(\hat{\beta} + \alpha R) + (1 - \bar{q})u(\hat{\beta} + \alpha L) - \lambda(\bar{q})(1 - \bar{q})lu'(\hat{\beta} + \alpha L) \\ & +(\beta - \hat{\beta})\big(\bar{q}u'(\hat{\beta} + \alpha R) + (1 - \bar{q})u'(\hat{\beta} + \alpha L) - u'(\beta)\big). \end{aligned} \tag{29}$$

Because $0 \leq \beta - \hat{\beta} \leq (1 - \bar{q})\lambda(\bar{q})l$ the claim holds. □

Proof of Proposition 1. Take any liability rule with probability of being found liable λ and let $w(x) = \beta + \alpha x$ be the principal's optimal contract under this rule. The resulting threshold for risk-taking is denoted by $\bar{q}$. If the agent reacts to this contract by staying uninformed, the principal can get a weakly better outcome under no liability simply by setting a fixed wage $\bar{w}$. We are therefore concerned with the case that contract and liability rule lead to information acquisition. We first discuss the case that $\alpha \geq \frac{A}{L}$, which means that $u_L = u_A$. In this case, the rule λ is not distinguishable from strict liability, as the agent is already at the lower bound with the contractually specified wage. The same outcome can be replicated under no liability, which must therefore be weakly better.

Next, we consider the more interesting case that the wealth constraint only binds if the agent is held liable. We define $l = \beta + \alpha L - A \geq 0$ and a contract $\hat{w} = \hat{\beta} + \hat{\alpha}x$ by

$$\hat{\alpha} = \alpha \text{ and } \hat{\beta} = \beta - l\rho_\lambda. \tag{30}$$

We will show that by offering this contract and excluding liability, the principal gets a higher payoff than under the liability rule. First, because

$$\beta - \hat{\beta} = l \int_{\bar{q}}^{1} (1-q)\lambda(q)dF \leq l(1-\bar{q})\lambda(\bar{q}), \tag{31}$$

we can apply Lemma 2 and get

$$\bar{q}^{FB} \leq \bar{q}^{nl}(\hat{w}) \leq \bar{q}. \tag{32}$$

Next we will show that the agent's payoff is weakly higher under no substantive due care liability and corresponding contract $\hat{w}$. As the first step, we show that

$$U^{nl}(\hat{w}, \bar{q}) \geq U^\lambda(w, \bar{q}), \tag{33}$$

where U^{nl} denotes the agent's expected payoff with no liability and U^λ denotes the agent's expected payoff under liability rule λ. These payoffs are the expected utilities of two lotteries. We denote the distribution function of the lottery induced by $\hat{w}$ and no liability by G^{nl} and the distribution function of the lottery induced by w and λ by G^λ. Figure 3 shows these distribution functions.

We use this figure to show that G^{nl} second order stochastically dominates G^λ. First note that the two lotteries have the same expected value. For second-order stochastic dominance we have to show that for all w

$$\int_A^w G^\lambda(x) - G^{nl}(x)dx \geq 0. \tag{34}$$

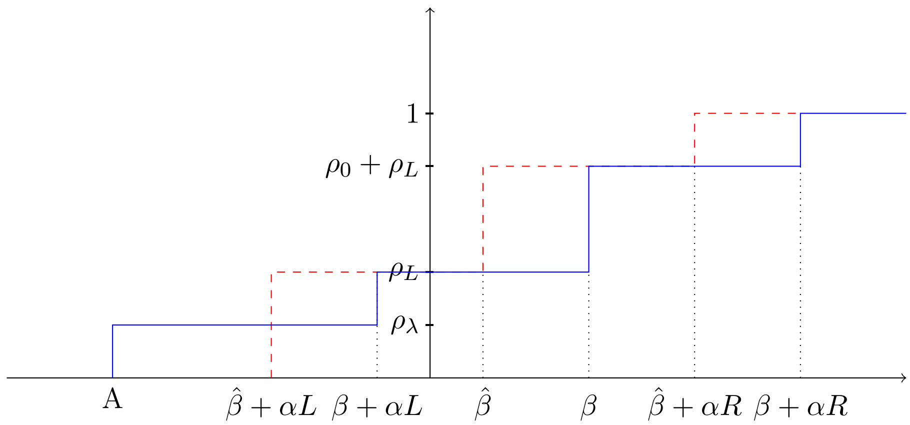

Figure 3: The (red) dashed line is the distribution function G^{nl} and the other line is the distribution function G^{λ}.

Figure 3 shows that this reduces to

$$\int_{A}^{\beta+\alpha R} G^{\lambda}(x) - G^{nl}(x)dx \geq 0, \tag{35}$$

and that this integral is equal to $\rho_\lambda l - (\beta - \hat{\beta})$, which is equal to zero. As the second step, we conclude that the agent's actual utility under no substantive due care liability must be even larger if he can choose the optimal $\bar{q}^{nl}(\hat{w})$ instead of $\bar{q}$:

$$U^{nl}(\hat{w}, \bar{q}^{nl}(\hat{w})) \geq U^{nl}(\hat{w}, \bar{q}) \geq U^{\lambda}(w, \bar{q}). \tag{36}$$

Knowing that the agent's utility is weakly greater under no liability immediately gives us the participation constraint (PC). Because $\hat{\beta} \leq \beta$, we also have the safe-choice incentive constraint (SIC). Similarly, the risky-choice incentive constraint (RIC) holds because $\hat{\beta} + \alpha R \leq \beta + \alpha R$. Finally, the principal's payoff under no liability is larger than the principal's payoff under the rule described by λ :

$$\pi^{nl}(\hat{w}, \bar{q}^{nl}) = (1-\alpha)S(\bar{q}^{nl}) - \hat{\beta} \geq (1-\alpha)S(\bar{q}) + \rho_\lambda l - \beta = \pi^{\lambda}(w, \bar{q}). \tag{37}$$

$\square$

Proof of Proposition 2.

In the following, we study the problem of implementing a given $\bar{q}$ at minimum cost under the assumption that the limited liability constraint is not binding (if it is binding then $w^{\lambda} = w^{nl}$ and the proposition holds). We find this lowest cost $C(\bar{q})$ by minimizing the expected wage payment subject to the constraints $(D), (SIC), (RIC), (PC)$ and (MON). First, we hold $\bar{q}^c$ fixed and minimize only with respect to the wages. The Lagrangian for this problem is

$$\begin{aligned}
\min_{w_L, w_0, w_R} \quad & \rho_R w_R + \rho_L w_L - \rho_\lambda (w_L - A) + \rho_0 w_0 \\
& + \mu_1 (\bar{q} u_R + (1-\bar{q})((1-\lambda(\bar{q}))u_L + \lambda(\bar{q}) u_A) - u_0) \\
& + \mu_2 (\rho_R u_R + \rho_L u_L - \rho_\lambda (u_L - u_A) + \rho_0 u_0 - \kappa - u_0) \\
& + \mu_3 (\rho_R u_R + \rho_L u_L - \rho_\lambda (u_L - u_A) + \rho_0 u_0 - \kappa - p u_R - (1-p) u_A) \\
& + \mu_4 (\rho_R u_R + \rho_L u_L - \rho_\lambda (u_L - u_A) + \rho_0 u_0 - \kappa - \bar{u}) \\
& + \mu_5 (u_0 - u_L)
\end{aligned} \tag{38}$$

It holds that $\mu_i \geq 0$ for $i = 2, 3, 4, 5$, but we cannot yet conclude the sign of μ_1. This optimization problem is well-behaved with concave constraint functions and a linear objective function. In any optimum, the following first order conditions with respect to w_0, w_R, and w_L have to hold:

$$\frac{1}{u'(w_0)} = -\frac{1}{\rho_0}\mu_1 + \frac{(\rho_0 - 1)}{\rho_0}\mu_2 + \mu_3 + \mu_4 + \frac{\mu_5}{\rho_0} \tag{39}$$

$$\frac{1}{u'(w_R)} = \frac{\bar{q}}{\rho_R}\mu_1 + \mu_2 + \frac{(\rho_R - p)}{\rho_R}\mu_3 + \mu_4 \tag{40}$$

$$\frac{1}{u'(w_L)} = \frac{(1-\bar{q})(1-\lambda(\bar{q}))}{(\rho_L - \rho_\lambda)}\mu_1 + \mu_2 + \mu_3 + \mu_4 - \frac{\mu_5}{\rho_L - \rho_0} \tag{41}$$

with the usual complementary slackness conditions. Using these necessary conditions we can prove that in any optimum it holds that

$$\mu_4 + \mu_3 + \mu_2 \geq \frac{1}{u'(w_L)}. \tag{42}$$

This follows from (41) if $\mu_1 \leq 0$. If $\mu_1 > 0$ and $\mu_5 = 0$ it follows from the first condition (39), and if $\mu_1 > 0$ and $\mu_5 > 0$ (which implies $w_0 = w_L$), it follows from (39) and (41), which together yield

$$\frac{1}{u'(w_L)} = \frac{(1-\bar{q})(1-\lambda(\bar{q})) - 1}{(\rho_L - \rho_\lambda + \rho_0)}\mu_1 + \mu_2 (1 - \frac{1}{(\rho_L - \rho_\lambda + \rho_0)}) + \mu_3 + \mu_4. \tag{43}$$

We can immediately determine the sign of μ_1 if (MON) is not binding ($\mu_5 = 0$). In this case it must be true that $\mu_1 < 0$, because else w_L would be larger than w_0. This in turn implies that $\mu_2 > 0$, because else w_0 would be larger than w_R. The optimal contract w is then defined as the solution to the equations $(D),(SIC),(PC)$ or $(D),(SIC),(RIC)$.

If (MON) is binding ($w_0 = w_L$), then the first order conditions are

$$\frac{1}{u'(w_R)} = \frac{\bar{q}}{\rho_R}\mu_1 + \mu_2 + \frac{(\rho_R - E[q])}{\rho_R}\mu_3 + \mu_4 \tag{44}$$

$$\frac{1}{u'(w_L)} = \frac{(1-\bar{q})(1-\lambda(\bar{q})) - 1}{\rho_L - \rho_\lambda + \rho_0}\mu_1 + \mu_2(1 - \frac{1}{(\rho_L - \rho_\lambda + \rho_0)}) + \mu_3 + \mu_4 \tag{45}$$

It can be seen that if $\mu_1 < 0$, it must hold that $\mu_2 > 0$ (because else w_R would be smaller than w_L). Therefore, $\mu_1 < 0$ implies that (MON),(D),(SIC) are binding.

In the following, we solve the problem of choosing $\bar{q}^c \in (0,1)$ optimally. The objective function and the constraints are differentiable in $\bar{q}^c$ for all $\bar{q}^c \in (0,1)$.[11] If an interior optimum $\bar{q}^c \in (0,1)$ exists, then it must hold there that

$$-\frac{\partial \rho_\lambda}{\partial \bar{q}^c}(w_L - A) + (\mu_4 + \mu_3 + \mu_2)(u_L - u_A)\frac{\partial \rho_\lambda}{\partial \bar{q}^c} + \mu_1(1-\bar{q})\frac{\partial \lambda(\bar{q})}{\partial \bar{q}^c}(u_L - u_A) = 0. \tag{46}$$

We know that $\frac{\partial \rho_\lambda}{\partial \bar{q}^c} \geq 0$ and $\frac{\partial \lambda(\bar{q})}{\partial \bar{q}^c} \geq 0$. Because

$$(\mu_4 + \mu_3 + \mu_2) \geq \frac{1}{u'(w_L)} \geq \frac{w_L - A}{u_L - u_A} \tag{47}$$

the first two terms in (46) add up to something positive. We can therefore conclude that $\mu_1 < 0$ (which we knew for the case that (MON) is not binding, and now follows also for the case that (MON) is binding, which can only be true for a positive standard). This also implies that (SIC) is always binding.

We can exploit that (SIC) and (D) are binding to compare the optimal contract w to the optimal contract w^{nl} for $\bar{q}^c = 0$. From (SIC) and (D) we can derive that

$$u_L^{nl} = u_0^{nl} - u_0 + (1-x)u_L + xu_A, \tag{48}$$

[11] It is also straightforward to show that a regularity condition holds such that we indeed identify necessary conditions for an optimum. One can also show that for $\bar{q}^c \to 0$ and $\bar{q}^c \to 1$, the cost converges against something that is strictly larger than the cost attained with $\bar{q}^c = 0$ and w^{nl}.

where

$$x = \frac{\int_{\bar{q}}(1-\bar{q})\lambda(\bar{q})q - (1-q)\lambda(q)\bar{q}dF}{\int_{\bar{q}}(q-\bar{q})dF}. \tag{49}$$

Assume first that $u_0^{nl} \leq u_0$, which implies $u_L^{nl} \leq (1-x)u_L + xu_A$. Since it holds that $\frac{\rho_\lambda}{\rho_L} \leq \lambda(\bar{q}) \leq x$, this in turn implies that $u_L^{nl} \leq (1-\frac{\rho_\lambda}{\rho_L})u_L + \frac{\rho_\lambda}{\rho_L}u_A$, i.e $w_L^{nl} \leq (1-\frac{\rho_\lambda}{\rho_L})w_L + \frac{\rho_\lambda}{\rho_L}A$. We also know that $\bar{w} \leq w_0$, so that if it were true that $w_R^{nl} \leq w_R$ we would have

$$\rho_R w_R + \rho_L\big((1-\frac{\rho_\lambda}{\rho_L})w_0 + \frac{\rho_\lambda}{\rho_L}A\big) + \rho_0 w_0 \geq \rho_R w_R^{nl} + \rho_L w_L^{nl} + \rho_0 w_0^{nl}, \tag{50}$$

which means that a standard of zero is actually optimal. Hence, it must hold that $w_R^{nl} \geq w_R$ if the contract w is optimal.

Assume now $u_0^{nl} > u_0$. This can only be the case if for no liability (RIC) is binding. Hence we know that

$$pu_R^{nl} + (1-p)u_A = u_0^{nl} > u_0 \geq pu_R + (1-p)u_A, \tag{51}$$

and therefore $u_R^{nl} > u_R$. (D) and (RIC) together yield

$$u_R^{nl}\int_0^{\bar{q}}(\bar{q}-q)dF + (\rho_L + \rho_0(1-\bar{q}))u_L^{nl} \quad = \quad \kappa + (1-p)u_A \tag{52}$$

and

$$u_R\int_0^{\bar{q}}(\bar{q}-q)dF + (\rho_L + \rho_0(1-\bar{q}))u_L - (\rho_\lambda + \rho_0(1-\bar{q})\lambda(\bar{q}))(u_L - u_A) \geq \kappa + (1-p)u_A. \tag{53}$$

We can conclude that

$$(u_R - u_R^{nl})\int_0^{\bar{q}}(\bar{q}-q)dF + (\rho_L + \rho_0(1-\bar{q}))(u_L - u_L^{nl}) - (\rho_\lambda + \rho_0(1-\bar{q})\lambda(\bar{q}))(u_L - u_A) \geq 0, \tag{54}$$

hence $u_L \geq u_L^{nl}$.

□

Proof of Remark 1. We have to show that in these two cases $\frac{\partial}{\partial\Delta\partial\lambda}\lambda^{-1}(\lambda) \geq 0$. With $\phi = \Phi'$, the derivative of the function $\lambda^{-1}(\lambda) = \bar{q}^c - \Phi^{-1}(\lambda)$ is

$$\lambda^{-1\prime}(\lambda) = -\frac{1}{\phi(\Phi^{-1}(\lambda))}. \tag{55}$$

We have to show that this expression is increasing in Δ. To this end, we take the derivative with respect to Δ of the function $\phi(\Phi^{-1}(\lambda))$. If we can show that this is positive in the two cases, we are done. In general, this derivative is equal to

$$\frac{d\phi(\Phi^{-1}(\lambda))}{d\Delta} = \frac{\partial\phi}{\partial\Delta} - \frac{\phi' \frac{\partial\Phi}{\partial\Delta}}{\phi}\bigg|_{\Phi^{-1}(\lambda)}. \tag{56}$$

For the linear error term with distribution $\Phi(\epsilon) = \frac{1+\epsilon\Delta}{2}$ on the interval $[-\frac{1}{\Delta}, \frac{1}{\Delta}]$ the second term is zero so that this derivative is

$$\frac{\partial\phi(\Phi^{-1}(\lambda))}{\partial\Delta} = \frac{1}{2} > 0. \tag{57}$$

For the normal error term with distribution $\Phi(\epsilon) = \sqrt{\frac{\Delta}{2\pi}} \int_{-\infty}^{\epsilon} e^{-\frac{1}{2}x^2\Delta} dx$ we can compute

$$\frac{\partial\phi(\epsilon)}{\partial\Delta} = (\frac{1}{\Delta} - \epsilon^2)\frac{1}{2}\phi(\epsilon), \tag{58}$$

$$\frac{\phi'(\epsilon)}{\phi(\epsilon)} = -\epsilon\Delta, \tag{59}$$

and, using integration by parts,

$$\frac{\partial\Phi(\epsilon)}{\partial\Delta} = \frac{1}{2\Delta}\epsilon\phi(\epsilon). \tag{60}$$

Putting everything together, we get

$$\frac{\partial\phi(\Phi^{-1}(\lambda)))}{\partial\Delta} = \frac{1}{2\Delta}\phi(\Phi^{-1}(\lambda)) > 0. \tag{61}$$

□

Proof of Proposition 3.

Note that if the standard is set to zero, the optimal contract w^{nl} and the principal's payoff $\pi^{nl}(w^{nl}, \bar{q}^{nl})$ do not depend on Δ.

In a first step, we show that as the signal becomes more precise, eventually a positive standard must be better. We show that with a perfect signal, the same threshold as under no liability, $\bar{q}^{nl}$, can be implemented at lower cost. To this end, we set the legal standard $\bar{q}^c = \bar{q}^{nl}$ and consider a contract with $w_0 = w_L = w_0^{nl}$ and w_R defined by

$$\rho_R(\bar{q}^{nl})(u_R - u_0^{nl}) = \kappa. \tag{62}$$

For the so defined contract it holds that $w_R^{nl} \geq w_R$ and $w_L \geq w_L^{nl}$. In the limit $\Delta \to \infty$, with a perfect signal, this contract implements $\bar{q}^{nl}$ because

$$\bar{q}^{nl} u_R + (1 - \bar{q}^{nl}) u_A \leq u_0 \text{ and } \bar{q}^{nl} u_R + (1 - \bar{q}^{nl}) u_0 \geq u_0. \tag{63}$$

It exposes the agent to a lottery between w_0^{nl} (with probability $\rho_L + \rho_0$) and w_R (with probability ρ_R). The agent's expected utility of this lottery is equal to $U^{nl}(w^{nl}, \bar{q}^{nl})$. Because the lottery exposes the agent to less risk in the sense of second-order stochastic dominance, it must have a lower expected value than the lottery induced by the no liability contract, which means lower cost for the principal.

In a second step, we consider the limit $\Delta \to 0$ and show that a completely uninformative signal is worthless. For an uninformative signal it holds that $\lambda(q) = \frac{1}{2}$ for all q and $\bar{q}^c \in (0, 1)$. If w is the optimal wage, then we define a new contract $\tilde{w}$ by $\tilde{u}_R = u_R$, $\tilde{u}_0 = u_0$ and $\tilde{u}_L = \frac{1}{2} u_L + \frac{1}{2} u_A$. This contract, with the standard set at zero, implements the same cut-off and the same agent's utility at a lower cost for the principal. It must therefore be that the optimal standard is already equal to zero and $w = \tilde{w}$.

It remains to show that the principal's payoff $\pi^\lambda(w, \bar{q})$, once it is equal to $\pi^{nl}(w^{nl}, \bar{q}^{nl})$ for some $\bar{\Delta}$, stays larger than $\pi^{nl}(w^{nl}, \bar{q}^{nl})$ for all $\Delta > \bar{\Delta}$. Let $\bar{q}$ be the optimal threshold for $\bar{\Delta}$, and $\bar{q}^c$ the optimal standard. First note that if the limited liability constraint (LL) is binding at this point, then the same payoff can be achieved for all $\Delta > \bar{\Delta}$ as well. We assume in the following that (LL) is not binding and show that as Δ increases, this particular $\bar{q}$ becomes easier to implement. To do this, we select standards $\bar{q}^c(\Delta)$ such that $\lambda(\bar{q})$ is the same for all $\Delta \geq \bar{\Delta}$. That is, we define a function $\bar{q}^c(\Delta)$ by

$$\Phi(\bar{q}^c(\Delta) - q, \Delta) = \Phi(\bar{q}^c - q, \bar{\Delta}), \tag{64}$$

where we have modified the earlier notation to make the dependence on precision explicit. This function is differentiable in Δ as long as $\bar{q}^c(\Delta) \in (0,1)$. We now look at the principal's cost of implementing $\bar{q}$ as the decision threshold, taking the standard $\bar{q}^c(\Delta)$ as given. This cost $C(\bar{q})$ is derived in the proof of Proposition 2. We take the derivative of the cost with respect to Δ, which varies ρ_λ. Note that by definition of the standard $\bar{q}^c(\Delta)$, $\lambda(\bar{q})$ does not vary with Δ.

$$\frac{\partial C(\bar{q})}{\partial \Delta} = -\frac{\partial \rho_\lambda}{\partial \Delta}(w_L - A) + (\mu_4 + \mu_3 + \mu_2)(u_L - u_A)\frac{\partial \rho_\lambda}{\partial \Delta}. \tag{65}$$

It follows from Assumption 1 that $\frac{\partial \rho_\lambda(\bar{q})}{\partial \Delta} \leq 0$. To see this, note that when we choose $\bar{q}^c(\Delta)$ such that $\lambda(q, \Delta)$ and $\lambda(q, \bar{\Delta})$ intersect at $q = \bar{q}$, then it must hold that

$$\frac{\partial}{\partial q}\lambda(\bar{q},\Delta) < \frac{\partial}{\partial q}\lambda(\bar{q},\bar{\Delta}) \tag{66}$$

and consequently $\lambda(q, \Delta) \leq \lambda(q, \bar{\Delta})$ for all $q \geq \bar{q}$, which means that ρ_λ must be decreasing in Δ. Furthermore, as in the proof of Proposition 2, here it holds that

$$\mu_4 + \mu_3 + \mu_2 \geq \frac{1}{u'(w_L)}. \tag{67}$$

Because $(w_L - A)u'(w_L) \leq u_L - u_A$, the derivative in (65) is negative. Hence, we have shown that cost decreases in precision if we take $\bar{q}^c(\Delta)$ as the standard for Δ. If we take the optimal standard, cost can only decrease further. Hence, the principal's payoff is increasing in the precision of the signal. $\square$

Proof of Proposition 4. This time, we cannot a priori exclude the case that the monotonicity constraint is binding in the other direction ($w_0 = w_R$). However, as we will show in the following, if $w_0 = w_R$ is optimal then also $\bar{q}^c = 0$ and $\bar{q}^c_0 = 1$. To this end, let λ the probability of being liable after the risky choice and l the probability of being liable after the safe choice, defined by the optimal standards $\bar{q}^c$ and $\bar{q}^c_0$, and let w be the optimal wage. This contract induces a threshold $\bar{q}$, given by

$$\bar{q}u_R + (1 - \bar{q})((1 - \lambda(\bar{q}))u_L + \lambda(\bar{q})u_A) = u_0 - l(\bar{q})(u_0 - u_A) \tag{68}$$

Let $U^\lambda(w, \bar{q})$ denote the agent's payoff under the liability rule and contract w. The other constraints

$(PC), (RIC), (SIC)$ are

$$U^\lambda(w, \bar{q}) \geq \bar{u} \tag{69}$$

$$U^\lambda(w, \bar{q}) \geq pu_R + (1-p)u_A \tag{70}$$

$$U^\lambda(w, \bar{q}) \geq (1-p)u_0 + pu_A \tag{71}$$

For the regime of no liability for a violation of substantive due care, we set $\tilde{w}_R = \tilde{w}_L = \tilde{w}_0 = U^\lambda(w, \bar{q})$ such that the agent's payoff is the same, i.e. $\tilde{w}$ is defined to be the wage at which $U^{nl}(\tilde{w}, \bar{q}) = U^\lambda(w, \bar{q})$. It then holds that $\tilde{w}_R \leq w_R$. The agent is indifferent between all decisions and will provide the efficient one. The constraints are still satisfied:

$$U^{nl}(\tilde{w}, \bar{q}^{FB}) \geq \bar{u} \tag{72}$$

$$U^{nl}(\tilde{w}, \bar{q}^{FB}) \geq p\tilde{u}_R + (1-p)u_A \tag{73}$$

$$U^{nl}(\tilde{w}, \bar{q}^{FB}) \geq (1-p)\tilde{u}_0 + pu_A \tag{74}$$

In case that $u_0 = u_R$, the last inequality (i.e. the safe constraint) follows from the original safe constraint, and for the case that $p = \frac{1}{2}$, it follows from the risky constraint. Since the agent is completely insured, the principal's payoff is higher.

In general, it may be the case that the so defined $\tilde{w}_0$ is larger than the original w_0. In these cases, we can only show that $\bar{q}_0^c = 1$ optimally, but not necessarily $\bar{q}^c = 0$. Compared to the main part of the paper, the problem of minimizing cost has changed in the following way:

$$\begin{aligned}
\min_{w_L, w_0, w_R} \quad & \rho_R w_R + \rho_L w_L - \rho_\lambda (w_L - A) + \rho_0 w_0 - \rho_l (w_0 - A) \\
& + \mu_1 (\bar{q} u_R + (1-\bar{q})((1-\lambda(\bar{q}))u_L + \lambda(\bar{q})u_A) - u_0 + l(\bar{q})(u_0 - u_A)) \\
& + \mu_2 (\rho_R u_R + \rho_L u_L - \rho_\lambda (u_L - u_A) + \rho_0 u_0 - \rho_l (u_0 - u_A) - \kappa - u_0(1-p) - u_A p) \\
& + \mu_3 (\rho_R u_R + \rho_L u_L - \rho_\lambda (u_L - u_A) + \rho_0 u_0 - \rho_l (u_0 - u_A) - \kappa - pu_R - (1-p)u_A) \\
& + \mu_4 (\rho_R u_R + \rho_L u_L - \rho_\lambda (u_L - u_A) + \rho_0 u_0 - \rho_l (u_0 - u_A) - \kappa - \bar{u}) \\
& + \mu_5 (u_0 - u_L)
\end{aligned} \tag{75}$$

with

$$\rho_l(\bar{q}) = \int_0^{\bar{q}} l(q)dF. \tag{76}$$

The first order conditions are now

$$\frac{1}{u'(w_0)} = -\frac{1-l(\bar{q})}{\rho_0-\rho_l}\mu_1 - \frac{1-p}{\rho_0-\rho_l}\mu_2 + \mu_2 + \mu_3 + \mu_4 + \frac{\mu_5}{\rho_0-\rho_l} \tag{77}$$

$$\frac{1}{u'(w_R)} = \frac{\bar{q}}{\rho_R}\mu_1 + \mu_2 + \frac{(\rho_R - p)}{\rho_R}\mu_3 + \mu_4 \tag{78}$$

$$\frac{1}{u'(w_L)} = \frac{(1-\bar{q})(1-\lambda(\bar{q}))}{(\rho_L-\rho_\lambda)}\mu_1 + \mu_2 + \mu_3 + \mu_4 - \frac{\mu_5}{\rho_L-\rho_\lambda} \tag{79}$$

Only the first constraint has changed, and it follows immediately that

$$\mu_4 + \mu_3 + \mu_2 \geq \frac{1}{u'(w_0)} \tag{80}$$

if either $\mu_1 < 0$ (from (78)) or if $\mu_1 > 0$ and $\mu_5 = 0$ (from (77)). For the case that $\mu_1 > 0$ and (MON) is binding, the first and the third condition together yield

$$\frac{1}{u'(w_0)} = \frac{(1-\bar{q})(1-\lambda(\bar{q})) - 1 + l(\bar{q})}{\rho_0 - \rho_l + \rho_L - \rho_\lambda}\mu_1 - \frac{1-p}{\rho_0-\rho_l+\rho_L-\rho_\lambda}\mu_2 + \mu_2 + \mu_3 + \mu_4.$$

As before, μ_1 is multiplied with a negative term. This follows directly if $l(\bar{q}) = \bar{q}\Phi(\bar{q} - \bar{q}_0^c)$, and more generally it follows form (D), which takes the form

$$\bar{q}(u_R - u_0) = ((1-\bar{q})\lambda(\bar{q}) - l(\bar{q}))(u_0 - u_A), \tag{81}$$

and hence implies $(1-\bar{q})\lambda(\bar{q}) - l(\bar{q}) \geq 0$. We can then deduce as in the proof of Proposition 2 that $\mu_1 < 0$. Taking the derivative with respect to $\bar{q}_0^c$ yields

$$-\frac{\partial \rho_l}{\partial \bar{q}_0^c}(w_0 - A) + (\mu_4 + \mu_3 + \mu_2)(u_0 - u_A)\frac{\partial \rho_l}{\partial \bar{q}_0^c} - \mu_1 \frac{\partial l(\bar{q})}{\partial \bar{q}_0^c}(u_0 - u_A) \leq 0. \tag{82}$$

Hence, the cost of implementing any $\bar{q}$ is decreasing in $\bar{q}_0^c$. □

Proof of Proposition 5. We will show that if the signal is as precise as stated and the standard is set at $\bar{e}^c = \tilde{e}$, then we can achieve the same outcome as with a perfect signal. First, we will show that given the assumption about $\Psi_0(\tilde{e})$, the safe constraint always binds. If (SIC) did not exist, the best way to implement $\bar{q}$ and $e = 1$ would be by setting $u_R = u_0 = u_L = \bar{u} + \kappa$ and $\bar{q}^c = 0$. The constraints $(D), (PC), (MON)$ would be naturally satisfied, and (RIC) would take the form

$$\bar{u} \geq \bar{u} + \kappa - \Psi_0(\tilde{e})(1-p)(\bar{u} + \kappa - u_A). \tag{83}$$

To conclude from the assumption on $\Psi_0(\tilde{e})$ that this condition is satisfied, we have to show that

$$(\bar{u} + \kappa - u_A)(1-p)p \geq (\bar{u} - u_A) \int_p^1 (q-p)dF, \tag{84}$$

which holds because $(1-p)p \geq \int_p^1 (q-p)dF$ is equivalent to

$$\int_0^p qdF + \int_p^1 pdF \geq \int_0^1 qpdF. \tag{85}$$

The result that (SIC) binds holds true both for linear contracts and monotonic ones. It implies the result of Proposition 2, because in the proof only (SIC) and (D) were used to show how wages compare. Next, we will show that (RIC) is not binding if $\bar{q}^c = 0$. This follows immediately for $\bar{q} \geq p$ because in that case (RIC) follows directly from (SIC) :

$$u_0 = \bar{q}u_R + (1-\bar{q})u_L \geq pu_R + (1-p)u_L - (1-p)\Psi_0(\tilde{e})(u_L - u_A). \tag{86}$$

Next we consider the case $\bar{q} < p$ and assume to the contrary that (RIC) is binding. Together with (D) it yields

$$(u_R - u_L) \int_0^{\bar{q}} (\bar{q} - q)dF + (u_L - u_A)\Psi_0(\tilde{e})(1-p) = \kappa, \tag{87}$$

while (SIC) and (D) together yield

$$(u_R - u_L) \int_{\bar{q}}^1 (q - \bar{q}) = \kappa, \tag{88}$$

and (PC) and (D) together yield

$$u_L - u_A + (u_R - u_L)(\rho_R + \rho_0 \bar{q}) \geq \kappa + \bar{u} - u_A. \tag{89}$$

From (87) and (89) we take that

$$(u_R - u_L)\int_0^{\bar{q}}(\bar{q}-q)dF + \Psi_0(\tilde{e})(1-p)\big(\kappa + \bar{u} - u_A - (\rho_R + \rho_0\bar{q})(u_R - u_L)\big) \leq \kappa, \tag{90}$$

which implies

$$\Psi_0(\tilde{e})(1-p)(\bar{u} - u_A) \leq \kappa + (1-p)\big((\rho_R + \rho_0\bar{q})(u_R - u_L) - \kappa\big) - (u_R - u_L)\int_0^{\bar{q}}(\bar{q}-q)dF. \tag{91}$$

Exploiting (88) this can be rearranged to yield

$$\Psi_0(\tilde{e}) \leq \frac{(1-\bar{q})p\kappa}{(1-p)(\bar{u}-u_A)\int_{\bar{q}}^1 (q-\bar{q})dF} \tag{92}$$

Since the right-hand side is increasing in $\bar{q}$, this contradicts our assumption on $\Psi_0(\tilde{e})$. Hence, we can deduce also for $\bar{q} < p$ that (RIC) is not binding. The outcome of the optimal contract for $\bar{q}^c = 0$ must hence be the same as with a perfect signal. This shows the result of Proposition 1, and it also implies that for monotonic contracts and all $\Delta \leq \bar{\Delta}$, the outcome of the optimal contract is still the same. In the case that there is also liability following the safe project and F is the uniform distribution, the contract $u_R = u_L = u_r = \bar{u} + \kappa$ is implementable because of (83). □

Proof of Corollary 1. For linear contracts we know that $\bar{q} \geq \bar{q}^{FB} \geq p$ always holds. It follows from (86) in the proof of Proposition 5 with $\Psi_0(\tilde{e}) = 0$ that for $\bar{q} \geq p$, (SIC) is stricter than (RIC). Hence, the outcome with $\bar{e}^c = 0$ is the same as with a perfect signal and a standard of $\bar{e}^c = 1$.

For the case of more general contracts, Lambert (1986) treats the case of $\bar{q}^c = 0$ and $\bar{e}^c = 0$ in detail and shows that if $\bar{q}^{FB} > p$ holds(RIC) is not binding, which gives us the result. □

References

Allen, William T., Jack B. Jacobs, and Leo E. Strine Jr., "Realigning the Standard of Review of Director Due Care with Delaware Public Policy: A Critique of Van Gorkom and Its Progeny as a Standard of Review Problem," *Northwestern University Law Review*, 2002, *96* (2), 449–466.

Black, Bernard S., Brian R. Cheffins, and Michael Klausner, "Outside Director Liability," *Stanford Law Review*, 2006a, *58*, 1055–1159.

____ , ____ , and ____ , "Outside Director Liability: A Policy Analysis," *Journal of Institutional and Theoretical Economics*, 2006b, *162* (1), 5–20.

Bundesgerichtshof, "Judgment 3 StR 470/04 of December 21, 2005," *Entscheidungen des Bundesgerichtshofs in Strafsachen*, 2005, *50* (1), 331–346.

Campbell, Rutherford B., "Normative Justifications for Lax (or No) Corporate Fiduciary Duties: A Tale of Problematic Principles, Imagined Facts and Inefficient Outcomes," *Kentucky Law Journal*, 2011, *99* (2), 231–257.

Chancery Court of Delaware, "Gagliardi v. Trifoods International, Inc., Atlantic Reporter," *Atlantic Reporter, 2d Series*, 1996, *683*, 1049–1055.

____ , "In re Citigroup Inc. Shareholder Derivative Litigation," *Atlantic Reporter, 2d Series*, 2009, *964*, 106–140.

Demski, Joel S. and David E. M. Sappington, "Delegated Expertise," *Journal of Accounting Research*, April 1987, *25* (1), 68–89.

Eisenberg, Melvin Aron, "The Divergence of Standards of Conduct and Standards of Review in Corporate Law," *Fordham Law Review*, *62* (3).

Fairfax, Lisa, "Spare the Rod, Spoil the Director? Revitalizing Directors' Fiduciary Duty Through Legal Liability," *Houston Law Review*, 2005, *42*, 393–456.

Fama, Eugene F. and Michael C. Jensen, "Separation of Ownership and Control," *Journal of Law and Economics*, June 1983, *26* (2), 301–325.

Fischel, Daniel R., "The Business Judgment Rule and the Trans Union Case," *The Business Lawyer*, 1985, *40*, 1437–1455.

Gevurtz, Franklin A., "Disney in a Comparative Light," *American Journal of Comparative Law*, 2007, *55* (3), 453–492.

Gromb, Denis and David Martimort, "Collusion and the organization of delegated expertise," *Journal of Economic Theory*, November 2007, *137* (1), 271–299.

Gutierrez, Maria, "An Economic Analysis of Corporate Directors' Fiduciary Duties," *The RAND Journal of Economics*, October 2003, *34* (3), 516–535.

Hakenes, Hendrik and Isabel Schnabel, "Bank Bonuses and Bail-Outs," *CEPR Discussion Paper No. 8852*, 2012.

Innes, Robert D, "Limited liability and incentive contracting with ex-ante action choices," *Journal of Economic Theory*, October 1990, *52* (1), 45–67.

Institute, American Law, *Principles of Corporate Governance: Analysis and Recommendations*, Vol. 1, St. Paul, Minnesota: American Law Institute Publishers, 1992.

Kadan, Ohad and Jeroen M. Swinkels, "Stocks or Options? Moral Hazard, Firm Viability, and the Design of Compensation Contracts," *Review of Financial Studies*, January 2008, *21* (1), 451–482.

Lambert, Richard A., "Executive Effort and Selection of Risky Projects," *The RAND Journal of Economics*, April 1986, *17* (1), 77–88.

—— **and David F. Larcker**, "Stock Options, Restricted Stock, and Incentives," *SSRN eLibrary*, April 2004.

Malcomson, James M, "Principal and Expert Agent," *The B.E. Journal of Theoretical Economics*, January 2009, *9* (1).

Malcomson, James M., "Do Managers with Limited Liability Take More Risky Decisions? An Information Acquisition Model," *Journal of Economics and Management Strategy*, 2011, *20* (1), 83–120.

Matthews, Steven A., "Renegotiating Moral Hazard Contracts under Limited Liability and Monotonicity," *Journal of Economic Theory*, March 2001, *97* (1), 1–29.

Nowicki, Elizabeth, "Director Inattention and Director Protection under Delaware General Corporation Law Section 102(B)(7): A Proposal for Legislative Reform," *Delaware Journal of Corporate Law*, 2008, *33* (3), 695–718.

Rachlinski, Jeffrey John, Andrew J. Wistrich, and Chris Guthrie, "Probable Cause, Probability, and Hindsight," *Journal of Empirical Legal Studies*, 2011, *8* (S1), 72–98.

Raith, Michael, "Specific knowledge and performance measurement," *RAND Journal of Economics*, 2008, *39* (4), 1059–1079.

Supreme Court of Delaware, "Aronson v. Lewis," *Atlantic Reporter, 2d Series*, 1984, *473*, 805–818.

——, "Smith v. Van Gorkom," *Atlantic Reporter, 2d Series*, 1985, *488*, 858–899.

——, "Brehm v. Eisner," *Atlantic Reporter, 2d Series*, 2000, *746*, 244–268.

——, "In re the Walt Disney Company Derivative Litigation," *Atlantic Reporter, 2d Series*, 2006, *906*, 27–75.

Supreme Court of Michigan, "Dodge v. Ford Motor Co," *Northwestern Reporter*, 1919, *170*, 668–685.

Ulmer, Peter, "Haftungsfreistellung bis zur Grenze grober Fahrlässigkeit bei unternehmerischen Fehlentscheidungen von Vorstand und Aufsichtsrat?," *Der Betrieb*, 2004, *57* (16), 859–863.