Efficient Incentives from Obligation Law and
the Compensation Principle
PRELIMINARY VERSION

Urs Schweizer, University of Bonn

April 14, 2015

Abstract

The compensation principle provides a link between the legal requirement to compensate for deviations from obligations and the economic desideratum of welfare maximizing incentives. The principle allows for new insights but also for revisiting known results on tort and contract law from a unifying perspective. Quantifying damages in line with the difference hypothesis would ensure the compensating goal being achieved as required for the compensation principle. To economize on transaction costs, parties may establish inefficient obligations in which case reasonable person standards may provide incentives for efficient breach even if obligations fail to be efficient.

JEL classification: K12, K13

Keywords: legal obligation, difference hypothesis, compensation goal, compensation principle, efficient incentives, efficient breach, negligence rule, liability in sequential choice

*address: Urs Schweizer, Department of Economics, University of Bonn, Kaiserstr. 1 53113 Bonn, Germany, schweizer@uni-bonn.de

†Support by the DFG through SFB/TR 15 is gratefully acknowledged.
1 Introduction

Legal rules are called efficient if they generate incentives for strategically acting parties to take decisions leading to a welfare maximizing outcome.

Obligation law provides general rules for contractual and tort relationships, including unjustified enrichment. If a debtor deviates from a (contractual or legal) obligation, the law offers remedies to creditors who suffer from harm caused by the debtor’s deviation. These remedies aim at compensating creditors.

The compensation principle, finally, refers to a link between the legal requirement of compensation and the economic concept of efficient rules. If each party is compensated for unilateral deviations from a reference profile by the other party and if this reference profile maximizes (expected) welfare then all Nash equilibria of the underlying game are efficient and payoff equivalent.

The compensation goal is, in particular, achieved if creditors are awarded damages in line with the difference hypothesis. In Germany, this hypothesis is attributed to Friedrich Mommsen, a legal scholar from the nineteenth century. Accordingly, damages should account for that part of the harm that was caused by the deviation from the obligation and should be calculated as the difference of the hypothetical (hence counterfactual) value of the creditor’s assets if the debtor had met her obligation and their actual value, given that actually she has not.

In tort relationships, courts may face the task of specifying care standards at their efficient level. In contractual relationships, as a matter of principle, the involved parties would be responsible for specifying efficient obligations themselves. Yet, to economize on transaction costs, parties may have stipulated incomplete contracts instead and, hence, may have specified obligations that fail to be efficient (for some contingencies at least).

For maintaining efficient incentives under inefficient obligations, the difference hypothesis must be modified to generate incentives for efficient breach ex post without distorting incentives to invest ex ante.

Moreover, as the difference hypothesis refers to counterfactuals and as counterfactuals may remain uncertain, courts may face difficulties specifying damages in line with the difference hypothesis. Yet, by suitably quantifying such damages on average over the observed event, the first best solution
turns out to remain implementable even if counterfactuals are not observable. These are major findings of the present paper.

In addition, the paper offers a simple method to handle rather disparate cases from tort and contract law in an unifying way. For this purpose, it proves useful expressing strategic interaction in normal form even if parties move sequentially.

In any case, an efficient pair of strategies in the normal form will serve as reference profile. Damages will then be quantified such that the efficient reference profile constitutes a Nash equilibrium in the corresponding normal form game. If more than one exists, all Nash equilibria turn out to be payoff equivalent.

Under sequential choice, the subgame perfect equilibrium will be payoff equivalent as well even if the efficient reference profile itself fails to be subgame perfect such that, off the equilibrium path, parties may wish to renegotiate. Fortunately, as will follow from the compensation principle, renegotiations (if anticipated) will not affect the efficiency of the equilibrium outcome.

The paper also provides a new justification for a damages regime that generates efficient incentives under sequential moves, on and off the equilibrium path. Based on the work of Rea (1987), Grady (1988) and Kornhauser and Revesz (1991), the textbook by Miceli (2008) nicely summarizes earlier findings on liability under sequential moves. Most of the rules he examines, however, fail to generate efficient incentives off the equilibrium path with the exception of marginal cost liability as pioneered by Wittman (1981). In spite of their nice properties, however, (as Miceli argues) courts do not seem to follow marginal cost liability in practice.

The justification proposed by the present paper, in contrast, rests on damages in line with the difference hypothesis combined with reductions for contributory negligence based on a reasonable person standard. As these are all concepts widely recognized by law and courts, the proposed damages regime may be of practical relevance nonetheless.

Furthermore, the paper introduces a damages regime, based on the generalized difference hypothesis, which generates efficient incentives in a hold-up setting of non-contingent contracts and two-sided investments of selfish, cooperative or even hybrid nature. By taking averages over the observed event,
efficient incentives can still be implemented if counterfactuals remain uncertain. In an earlier paper (Schweizer, 2006), I have proposed a similar approach for the accident model which the present model extends to the more general hold-up problem.

On top of such truly new results, the paper also takes a fresh look at well-known findings from the existing literature which has identified a whole bunch of damages regimes generating efficient incentives within the accident model. As all these regimes satisfy the conditions underlying the compensation principle, a single proof turns out sufficient to establish all these efficiency results at once. As a fringe benefit, much weaker assumptions than traditionally imposed are sufficient to establish efficiency by relying on the compensation principle instead.

Göller and Hewer (2014) have examined compensation rules for takings that provide two-sided efficient incentives for purely selfish investments. Their main result turns out to be a special case of the difference hypothesis extended to an inefficient obligation as introduced by the present paper. The same holds true for Schweizer (2006), who has proposed a bilateral damages regime that ensures an efficient outcome in a setting of one-sided investments of the cooperative type.

For sake of completeness, finally, let me also refer to Schweizer (2005) on the economic analysis of obligation law. While, in that paper, I had focussed on the mathematical saddle point property, the present paper takes a more legal perspective by relying on the compensation goal and the difference hypothesis instead.

The paper is organized as follows. Section 2 establishes the compensation principle for strategic interaction expressed in normal form. In particular, the principle applies if damages are quantified in line with the minimum requirement of the difference hypothesis. This result will be referred to as difference principle. If the externality is of essentially unidirectional nature only one party need be held liable to provide efficient incentives for both parties nonetheless.

Section 3 deals with sequential choice. If expressed in normal form, more than one decision profile exists that can serve as an efficient reference profile, including the ex post efficient one. If the externality is of unidirectional nature, even the subgame perfect equilibrium can be implemented by a dam-
ages regime based on the difference hypothesis in combination with reduction of claims for contributory negligence.

Section 4 revisits the traditional accident model. Damages rules as discussed in the literature are shown to fulfil the conditions of the difference principle and, for that reason, they all generate two-sided efficient incentives.

Section 5 extends the difference principle related to an inefficient obligation. It is shown that damages in line with the difference hypothesis but combined with a deduction for contributory negligence, based on a reasonable person standard, still ensure efficient incentives for both parties.

In section 6, these findings are applied to a contractual relationship with two-sided reliance investments under uncertainty and a performance decision to be taken ex post. The inefficiency of the obligation may be due to the fact that parties, to economize on transaction costs, have stipulated non-contingent obligations in their contract. Incentives for efficient breach emerge from the proposed damages regime.

Section 7 examines the difference hypothesis adapted to an informational setting where counterfactuals fail being observable by courts. It is shown how courts should quantify damages on average over observable events to still implement a given efficient reference profiles as a Nash equilibrium of the corresponding normal form game. Section 8 concludes.

2 Compensation principle and difference hypothesis

I consider the following contractual or legal relationship between two parties A and B. Party A (you may think of a buyer in a contractual relationship or a victim in a tort relationship) takes a decision $a$ from the set $A$ of her alternatives. Party B (think of a seller/producer or an injurer/tortfeasor) takes a decision $b$ from the set $B$ at his disposal. The expected value (net of costs for $a$) of party A’s assets is assumed to be a function $v(a, b)$ of the decision profile $(a, b)$ as chosen from the Cartesian product $A \times B$ by parties A and B. In general, expected costs $k(a, b)$ to be borne by party B may also be a function of the entire profile. Expected welfare amounts to $w(a, b) = u(a, b) - k(a, b)$. 

5
As no formal structure is imposed on the strategy sets \(A\) and \(B\) (they may combine elements of discrete and continuous choice and they may be one- or multi-dimensional, including complete contingent plans under sequential choice) this model is quite flexible and allows for disparate applications in tort and contract law. Moreover, while I am talking about values and costs, no definite sign is imposed on any of these functions which allows to extend the range of interpretations even further.

Throughout the paper, I think of a fixed reference profile \((a^o, b^o)\), which maximizes welfare, i.e.

\[
(a^o, b^o) \in \text{arg} \max_{(a, b) \in A \times B} w(a, b). \tag{1}
\]

The payoff functions \(\phi(a, b)\) and \(\psi(a, b)\) of parties \(A\) and \(B\), respectively, reflecting the legal regime in place (including possibly damages claims and other compensation transfers), may depend on this efficient reference profile even if, for better readability, the notation suppresses such dependence.

Obviously, the sum of payoffs cannot exceed the maximum welfare, i.e.

\[
\phi(a, b) + \psi(a, b) \leq w(a^o, b^o) \tag{2}
\]

must hold for any profile \((a, b) \in A \times B\). If, however, parties stick to the reference profile, this sum is assumed equal to the maximum welfare, i.e.

\[
\phi(a^o, b^o) + \psi(a^o, b^o) = w(a^o, b^o). \tag{3}
\]

To justify this assumption, as both parties stick to the reference profile, neither costly litigation takes place nor costly transfer of values is required and, in the absence of such institutional frictions, the maximum welfare may be attained indeed.

By definition, the compensation goal for unilateral deviations from the reference profile is achieved if no party suffers from a unilateral deviation from the efficient reference profile by the other party or, more precisely, if

\[
\phi(a^o, b^o) \leq \phi(a^o, b) \tag{4}
\]

holds for any deviation \(b\) from \(b^o\) by party \(B\) and, similarly, if

\[
\psi(a^o, b^o) \leq \psi(a, b^o) \tag{5}
\]

holds for any deviation \(a\) from \(a^o\) by party \(A\).
The following proposition deals with Nash equilibria of the game in normal form with strategy sets $A$ and $B$ and payoff functions $\phi(a, b)$ and $\psi(a, b)$.

**Proposition 1 (compensation principle)** Suppose conditions (1) – (5) are met. Then the reference profile $(a^o, b^o)$ forms a Nash equilibrium and all Nash equilibria (if more than one exists) are payoff equivalent.

**Proof.** By (2), (1), (5) and (3) it follows that

$$\phi(a, b^o) \leq w(a^o, b^o) - \psi(a, b^o) \leq w(a^o, b^o) - \psi(a^o, b^o) = \phi(a^o, b^o)$$

must hold and, hence, $a^o$ is a best response by party A to $b^o$. For similar reasons, $b^o$ is also a best response by party B to $a^o$ and, hence, the reference profile must be a Nash equilibrium indeed.

Suppose $(a^N, b^N)$ is any other Nash equilibrium consisting of mutually best responses. It then follows from (4) and (5) that

$$\phi(a^N, b^N) \geq \phi(a^o, b^o) \geq \phi(a^o, b^o)$$

as well as

$$\psi(a^N, b^N) \geq \psi(a^N, b^N) \geq \psi(a^N, b^N)$$

and hence, from (2) and (3), that

$$w(a^o, b^o) \geq \phi(a^N, b^N) + \phi(a^N, b^N) \geq \phi(a^o, b^o) + \psi(a^N, b^N) = w(a^o, b^o)$$

must hold. Yet, as the reference profile maximizes welfare (see (1)), none of the above inequalities can hold in the strict sense. In particular,

$$\phi(a^N, b^N) = \phi(a^o, b^o) \text{ and } \psi(a^N, b^N) = \psi(a^o, b^o)$$

must be true such that payoff equivalence is established as well. ■

Notice, as the Nash equilibrium will be efficient, in equilibrium, there is no scope for voluntary renegotiations. Off the equilibrium, however, a range for voluntary renegotiations may possibly present itself. Yet, even if parties anticipate to renegotiate off the equilibrium, the conditions needed for the compensation principle continue to hold.

In fact, if inefficient outcomes are possibly renegotiated, post renegotiation payoffs are denoted by $\phi^+(a, b)$ and $\psi^+(a, b)$. As renegotiation remains voluntary, the participation constraints

$$\phi(a, b) \leq \phi^+(a, b) \text{ and } \psi(a, b) \leq \psi^+(a, b)$$

hold.
have to be met and, hence, the compensation goal for unilateral deviations from the reference profile is achieved all the more. Moreover, maximum welfare remains of course an upper bound for the sum of payoffs even post renegotiations. Since the reference profile maximizes welfare, it follows that, at the reference profile,

$$w(a^o, b^o) = \phi(a^o, b^o) + \psi(a^o, b^o) \leq \phi^+(a^o, b^o) + \psi^+(a^o, b^o) \leq w(a^o, b^o)$$

must hold and, hence, all the above inequalities must be binding. As a consequence, the compensation principle also applies to the game with payoff functions $\phi^+$ and $\psi^+$ such that equilibrium payoffs remain unaffected by anticipated renegotiations (in the above sense) off the equilibrium.

The assumptions needed for the compensation principle will, in particular, be fulfilled if damages are granted in line with the difference hypothesis. Depending on the chosen strategy profile $(a, b)$, party A receives damages $d(a, b)$ from party B (if $d(a, b) < 0$, party B receives damages $-d(a, b)$ from party A), leading to payoff functions

$$\phi_d(a, b) = u(a, b) + d(a, b) \text{ and } \psi_d(a, b) = -k(a, b) - d(a, b)$$

of party A and B, respectively (subscript $d$ refers to difference hypothesis).

Damages are in line with the difference hypothesis if, as a minimum requirement, the following conditions are met for unilateral deviations from the efficient reference profile $(a^o, b^o)$ at least:

$$d(a^o, b) \geq u(a^o, b^o) - u(a^o, b) \text{ and } d(a, b^o) \leq k(a^o, b^o) - k(a, b^o) \quad (6)$$

which allows for the following interpretation.

If $u(a^o, b^o) - u(a^o, b)$ is positive, this difference corresponds to the loss suffered by party A from party B’s deviation. Damages in line with the difference hypothesis must cover this loss. Yet, not even overcompensation (punitive damages) would be ruled out by the minimum requirement (6).

If $u(a^o, b^o) - u(a^o, b)$ is negative then party A enjoys an enrichment from B’s deviation. While party B may keep such an enrichment for free, the minimum requirement (6) would neither rule out that party A must reverse the enrichment to party B.

The interpretation of condition (6) with respect to the difference $k(a^o, b^o) - k(a, b^o)$ is symmetric. Moreover, let me point out that the exact specification of damages for strategy profiles where both deviate from the efficient
reference profile does not matter for the following proposition to hold. This proposition deals with Nash equilibria of the game in normal form with A and B as the strategy sets and $\phi = \phi_d$ and $\psi = \psi_d$ as payoff functions.

**Proposition 2** (difference principle)

(i) If the minimum requirement (6) of the difference hypothesis is met then $d(a^o, b^o) = 0$ as well as (4) and (5) must hold.

(ii) Conversely, if $d(a^o, b^o) = 0$ as well as (4) and (5) are met then the minimum requirement (6) of the difference hypothesis must hold.

(iii) If the minimum requirement (6) of the difference hypothesis is met and the reference profile $(a^o, b^o)$ maximizes welfare then the reference profile $(a^o, b^o)$ forms a Nash equilibrium and all Nash equilibria (if more than one exists) are payoff equivalent.

**Proof.** The proof of claims (i) and (ii) follows simply from rearranging terms and claim (iii) is then an immediate consequence of the compensation principle (i.e. proposition 1).

The decisive properties of the games above have been derived from the difference hypothesis, a concept of obligation law. Efficient incentives, in contrast, are an economic desideratum. The above propositions provide a simple link between the legal and the economic perspectives that proves useful for establishing the efficiency of provisions from obligation law quite generally.

The existing literature on the economic analysis of tort law has concentrated on externalities of unidirectional nature in the sense that party B’s cost function $k = k(b) \text{ remains independent of party A’s decision } a$. Precaution expenditures of party B are typically assumed to affect the expected wealth of party A but A’s own precautions do not affect B’s costs. More generally, I call the externality to be of essentially unidirectional nature if

$$k(a, b^o) = k(a^o, b^o) \quad (7)$$

holds at least for unilateral deviations by party A from the efficient reference profile. If condition (7) is met, damages

$$d(a, b) = \max [u(a, b^o) - u(a, b), 0]$$

would, in particular, be in line with the difference hypothesis (6). As this scheme does not take account of potentially contributory negligence of party
B, reduction of damages for contributory negligence as familiar from legal practice would not be necessary to guarantee the efficiency of the Nash equilibrium.

3 Sequential choice

As a first application of the difference principle (i.e. proposition 2), I consider sequential choice. Suppose it is party B who, at stage 1, takes a decision \( y \in Y \) before party A, after having observed the decision \( y \), chooses \( x \) from \( X \) at stage 2. The payoffs of A and B amount to \( U(x, y) \) and \( -K(x, y) \), respectively.

The first best solution can be calculated by backwards induction. For any decision \( y \) of party B at stage 1, let

\[
x^+(y) \in \arg \max_{x \in X} U(x, y) - K(x, y)
\]

denote a socially best response at stage 2 and

\[
y^o \in \arg \max_{y \in Y} U(x^+(y), y) - K(x^+(y), y)
\]
an efficient decision at stage 1. For later reference, let \( x^o = x^+(y^o) \) denote the socially best response at stage 2 to the efficient decision \( y^o \) at stage 1.

Before making use of the difference principle (see proposition 2 above), the sequential interaction must be expressed in normal form. While the set of B’s strategies is \( B = Y \), that of party A now consists of \( A = X^Y \), i.e. the set of all reaction functions \( \xi : Y \rightarrow X \) with the following interpretation: if B decides \( y \) then, under the complete contingent plan \( \xi \), party A will react by choosing \( x = \xi(y) \). Therefore, if expressed in the normal form, the payoffs of party A and B amount to

\[
u(\xi, y) = U(\xi(y), y) \quad \text{and} \quad -k(\xi, y) = -K(\xi(y), y),
\]

respectively.

In the following, I fix an efficient reference profile \((\xi^o, y^o)\) which maximizes welfare \( w(\xi, y) = u(\xi, y) - k(\xi, y) \). Notice, the constant reaction function \( \xi^o(y) = x^o \) but also the ex post efficient reaction function \( \xi^o(y) = x^+(y) \) would both qualify. In any case, \( \xi^o(y^o) = x^o \) is assumed to hold such that the reference profile is efficient indeed.
Damages will have to be quantified (if at all) ex post when, by assumption, courts can observe the actual decisions \( x \) and \( y \) but not party A’s complete contingent plan \( \xi \). Let \( D(x, y) \) denote the damages claims of party A (if \( D(x, y) > 0 \)) and \(-D(x, y)\) the damages claims of party B (if \(-D(x, y) > 0\)).

To be in line with the difference hypothesis in the sense of (6), I specify damages as follows.

If party B invests efficiently, \( y = y^o \), then damages are restricted by the condition

\[
D(x, y^o) \leq K(x^o, y^o) - K(x, y^o) \quad \text{for all} \quad x \text{ but } D(x^o, y^o) = 0 \quad (8)
\]

If party B deviates, by deciding \( y \neq y^o \), damages depend on party A’s response. If she decides in line with the reference profile, \( x = \xi^o(y) \), damages are restricted by the condition

\[
D(x, y) = D(\xi^o(y), y) \geq U(x^o, y^o) - U(\xi^o(y), y) \quad (9)
\]

whereas if she does not (i.e. \( x \neq \xi^o(y) \)), we would be free to restrict them as we wish. For later reference, let me restrict them by the condition

\[
D(x, y) \leq D(\xi^o(y), y) \quad (10)
\]

which allows for the following interpretation.

Since party B has deviated from his obligation (i.e. \( y \neq y^o \)), party A’s obligation would be to respond with \( x = \xi^o(y) \). If she deviates, by deciding \( x \neq \xi^o(y) \), from her obligation, the difference \( U(x, y^o) - U(x, y) \) evaluated at \( x \) may be higher than \( D(\xi^o(y), y) \). In this case, (10) would require to reduce damages \( U(x, y^o) - U(x, y) \) in excess of \( D(\xi^o(y), y) \).

Such reductions are in line with legal provisions. The German civil code (§ 254 BGB), e.g., requires that where fault (including the fault of failing to reduce the damage) on the part of the injured person contributes to the occurrence of the damage, the extent of compensation to be paid depend on the circumstances, in particular to what extent the damage is caused mainly by one or the other party.

In any case, the following proposition refers to the game with strategy sets \( A = X^Y \) and \( B \) and payoff functions

\[
\phi_d(\xi, y) = U(\xi(y), y) + D(\xi(y), y) \quad \text{and} \quad \psi_d(\xi, y) = -K(\xi(y), y) - D(\xi(y), y).
\]
Proposition 3 Let \((\xi^o, y^o)\) denote an efficient reference profile.

(i) If damages \(D(x, y)\) satisfy conditions (8) and (9), then the efficient reference profile forms a Nash equilibrium of the corresponding normal form game and all Nash equilibria (if more than one exists) are payoff equivalent.

(ii) If damages \(D(x, y)\) satisfy conditions (8) – (10), the reference profile includes the ex post efficient response \(\xi^o(y) = x^+(y)\) and the externality is of unidirectional nature in the sense that party B’s cost function \(K = K(y)\) does not depend on party A’s decision, then this reference profile forms a subgame perfect equilibrium of the game in extensive form.

Proof. To establish claim (i), let \((a, b) = (\xi, y)\) be any strategy profile in the normal form of the game. It then follows from (8) and (9) that

\[
d(\xi^o, y) = D(\xi^o(y), y) \geq U(x^o, y^o) - U(\xi^o(y), y) = u(\xi^o, y^o) - u(\xi^o, y)
\]

must hold such that the first part of condition (6) is met. Moreover, it follows that

\[
d(\xi, y^o) = D(\xi(y^o), y^o) \leq K(x^o, y^o) - K(\xi(y^o), y^o) = k(\xi^o, y^o) - k(\xi(y^o), y^o)
\]

must hold as well and, hence, the second part of condition (6) is also satisfied. Claim (i) then immediately follows from the difference principle (i.e. proposition 2).

To establish claim (ii), it follows from claim (i) that the efficient reference profile forms a Nash equilibrium of the normal form game. In particular, \(y^o\) must be a best response of party B to \(\xi^o\) such that the inequality

\[
\psi_d(x^+(y), y) \leq \psi_d(x^+(y^o), y^o)
\]

will hold for all \(y\). For the (ex post efficient) reference profile to be a subgame perfect equilibrium, \(\xi^o(y) = x^+(y)\) must be a best response, on and off the equilibrium path, by party A i.e.

\[
x^+(y) \in \arg\max_{x \in X} U(x, y) + D(x, y)
\]

must hold. Claim (i) has established that \(\xi^o(y^o)\) is a best response to \(y^o\) (on the equilibrium path). Off the equilibrium path, due to the unidirectional
externality,

\[ x^+(y) \in \arg \max_{x \in X} U(x, y) - K(y) = \arg \max_{x \in X} U(x, y) \]

and, hence, \( U(x, y) \leq U(x^+(y), y) \) must hold. Due to (10), \( D(x, y) \leq D(x^+(y), y) \) holds as well and, hence, \( \xi^o(y) = x^+(y) \) is a best response by party B off the equilibrium path as well. Claim (ii) is established. ■

Claim (i) holds, in particular, for the efficient reference profile with constant reaction \( \xi^o(y) = x^o \) and, hence, to ensure an efficient outcome, reduction of damages for contributory negligence would not be needed. Moreover, any subgame perfect equilibrium gives rise to a Nash equilibrium in the corresponding normal form game and, according to claim (i), would have to be payoff equivalent to \( \phi_d(x^o, y^o) \) and \( \psi_d(x^o, y^o) \).

Claim (i), if applied to the ex post efficient reaction function \( \xi^o(y) = x^+(y) \), would ensure that

\[ \psi_d(x^+, y^o) \geq \psi_d(x^+, y) \]

holds for any \( y \) (Nash property). But notice, it would not necessarily mean that the ex post efficient reference profile constitutes a subgame perfect equilibrium of the extensive form game. In fact, to qualify as a subgame perfect equilibrium, the additional condition

\[ \phi_d(x^+, y) \geq \phi_d(\xi, y) \]

would have to be met for all \( \xi \) and all \( y \). Claim (i), however, only ensures that this condition is met for \( y = y^o \) but not necessarily off the equilibrium path.

Claim (ii), in contrast, provides conditions sufficient for the ex post efficient reference profile to form a subgame perfect equilibrium of the game in extensive form. The underlying damages regime is related to marginal cost liability LMC as introduced by Wittman (1981). Surprisingly enough, Wittman himself did not find it worthwhile to examine what, in his notation, would correspond to the negligence version of LMC\(^*\) explicitly. Yet, it is exactly this version of all that can be justified in terms of the difference hypothesis combined with the appropriate reduction of damages for contributory negligence.

13
4 The accident model

In the traditional accident model, party A (victim) and party B (injurer) decide on their precaution expenditures $x \in X = \mathbb{R}_+$ and $y \in Y = \mathbb{R}_+$, respectively. The probability of an accident is assumed to be a monotonically decreasing function $\varepsilon(x, y)$ of both expenditures. If an accident occurs, party A’s wealth diminishes by a loss $L$ of fixed size. The payoff functions of party A and B are

$$U(x, y) = -\varepsilon(x, y) \cdot L - x \text{ and } K(y) = y$$

and, hence, the externality is of unidirectional nature.

The first best solution is derived by backwards induction. For any precaution $y$ of the injurer A, let

$$x^+(y) \in \arg \max_{x \in X} -\varepsilon(x, y) \cdot L - x = \arg \max_{x \in X} -\varepsilon(x, y) \cdot L - x - y$$

denote a socially best response by victim A. Efficient precaution then solves

$$y^o \in \arg \max_{y \in Y} -\varepsilon(x^+(y), y) \cdot L - x^+(y) - y$$

and $x^o = x^+(y^o)$ denotes the socially best response to it.

I first deal with the case where parties decide simultaneously and the decision profile $(x^o, y^o)$ serves as the efficient reference profile. In the accident model, damages $\Delta(x, y)$ (if at all) are usually granted ex post and only if an accident has actually happened. Expected damages then amount to $d(x, y) = \varepsilon(x, y) \cdot \Delta(x, y)$. The condition to be in line with the difference hypothesis (see (6)!) refers to expected damages.

As the externality is of unidirectional nature, damages will satisfy (6) if the conditions

$$\Delta(x, y^o) = 0 \text{ and } \Delta(x^o, y) \geq \frac{\varepsilon(x^o, y) - \varepsilon(x^o, y^o)}{\varepsilon(x^o, y)} \cdot L \quad (11)$$

are met for any profile $(x, y)$. It then follows from the difference principle (see proposition 2) that condition (11) is sufficient for the efficient reference profile $(x^o, y^o)$ to form a Nash equilibrium under simultaneous choice of $x$ and $y$.

In the literature, the following damages rules have been discussed, which all satisfy condition (11) sufficient for efficiency: the negligence rule with and
without a defense of contributory negligence; contributory negligence (damages are reduced below $L$ but only if both parties have deviated), proportional liability quantified as
\[
\Delta(x^o, y) = \frac{\varepsilon(x^o, y) - \varepsilon(x^o, y^o)}{\varepsilon(x^o, y)} \cdot L
\]
as well as damages for the loss of a chance specified as
\[
\Delta(x^o, y) = (1 - \varepsilon(x^o, y^o)) \cdot L
\]
all satisfy condition (11) and, hence, the efficient reference profile $(x^o, y^o)$ forms a Nash equilibrium of all corresponding games and all these Nash equilibria must be payoff equivalent.

Under sequential choice, first $y$ then $x$, and under any damages rule satisfying (11), the equilibrium outcome will also be the same. This follows from proposition 3 claim (i) applied to the efficient reference profile $(\xi^o, y^o)$ in the normal form with constant response $\xi^o(y) = x^o$.

To implement the ex post efficient reference profile $(\xi^o, y^o) = (x^+, y^o)$ as a subgame perfect equilibrium in the sequential move game, however, a damages regime in line with claim (ii) of proposition 3 is needed.

Due to the unidirectional nature of the externality, the specification
\[
D(x, y^o) = 0 \text{ for all } x
\]
would be consistent with (8). Moreover, if party B deviates by precautions $y \neq y^o$, then damages $D(x, y) \leq D(x^+(y), y)$ and
\[
D(x^+(y), y) \geq [\varepsilon(x^+(y), y) - \varepsilon(x^o, y^o)] \cdot L + [x^+(y) - x^o]
\]
would ensure that conditions (9) and (10) are also met. Hence, by claim (ii) of proposition 3, the ex post efficient reference profile $(x^+, y)$ constitutes even a subgame perfect equilibrium of the sequential move game if the above conditions are met.

To illustrate ex post efficient regimes, suppose (due to party B’s deviation $y \neq y^o$) party A was obliged to spend $x^+(y) > x^o = x^+(y^o)$ on defensive measures. Then party A can recover extra expenses $[x^+(y) - x^o]$ from party A even if no accident has occurred. On top of it, if an accident has occurred, party A can claim additional damages of size
\[
\frac{\varepsilon(x^+(y), y) - \varepsilon(x^o, y^o)}{\varepsilon(x^+(y), y)} \cdot L
\]
(at least) from party B provided that A has responded in a socially best way (i.e. \( x = x^+(y) \)).

Tort law may be reluctant to award any damages as long as no accident has occurred. While the efficient outcome could be implemented as a sub-game perfect equilibrium nonetheless, such a damages regime is unlikely to generate ex post efficient incentives off the equilibrium path.

5 Efficient breach of inefficient obligations

So far, damages in line with the difference hypothesis have been based on efficient obligations. In the present section, a difference principle related to inefficient obligations is spelled out such that efficient incentives still prevail.

The basic setting is the same as in section 2 and \((a^o, b^o)\) still denotes an efficient reference profile satisfying (1). In addition, however, I consider a (contractual) obligation \(b^{oo} \in B\) faced by party B which need not be efficient. Efficient breach then means that party B chooses \(b^o\) even though his obligation would have been to choose \(b^{oo}\).

To be in line with the minimum requirement of the difference hypothesis, two conditions have to be met. First, if party A, by deciding \(a = a^o\), sticks to the efficient reference profile then damages are restricted by the conditions

\[
d_r(a^o, b) \geq u(a^o, b^{oo}) - u(a^o, b) \quad \text{but} \quad d_r(a^o, b^o) = u(a^o, b^{oo}) - u(a^o, b^o) \quad (12)
\]

and, second, if party A deviates (i.e. \(a \neq a^o\)) they are restricted by the condition

\[
d_r(a, b^o) \leq -[-k(a^o, b^o) + k(a, b^o)] + [u(a^o, b^{oo}) - u(a^o, b^o)] \quad (13)
\]

whenever party B, by deciding \(b = b^o\), has breached efficiently (subscript \(r\) refers to reasonable person standard as explained below).

If, finally, both deviate from the efficient reference profile, the specification of damages will not affect the equilibrium outcome and, on this account, does not matter.

Condition (12) reflects the difference hypothesis but with respect to the inefficient obligation \(b^{oo}\). Notice, however, overcompensation must now be ruled out if party B breaches efficiently.
Condition (13) allows for the following interpretation. As party B’s damages claims are equal to $-d_r(a, b^o)$, condition (13) requires that party B is also entitled to damages $[-k(a^o, b^o) + k(a, b^o)]$ in line with the difference hypothesis (with respect to party A’s efficient obligation $a^o$) from which, however, the difference $[u(a^o, b^{oo}) - u(a^o, b^o)]$ must be deducted. This deduction can be interpreted as reduction for contributory negligence but based on a reasonable person standard, namely $a^o$.

In fact, a reasonable person is a hypothetical person who exercises average care, skill, and judgment in conduct and who serves as a comparative standard for determining liability. If the efficient decision $a^o$ serves as reasonable person standard then the deduction $[u(a^o, b^{oo}) - u(a^o, b^o)]$ corresponds to damages in line with the difference hypothesis as owed to this reasonable person (even if party A has actually deviated from the reasonable person standard). In this sense, damages satisfying (12) and (13) are in line with the difference hypothesis based on a reasonable person standard.

Notice, in principle, the inequality in condition (13) allows to punish party A for deviations $a \neq a^o$ without limit. Yet, as such punitive damages flow into the pockets of party B, party B would end up being overcompensated. Therefore, under legal regimes rejecting overcompensation, condition (13) would have to hold as an equality.

Notice further, if party B’s obligation were efficient, i.e. $b^{oo} = b^o$, then conditions (12) and (13) coincide with the minimum requirement of the difference hypothesis relative to an efficient obligation, i.e. condition (6). In this sense, conditions (12) and (13) are generalizing the difference hypothesis to inefficient obligations indeed.

The following proposition concerns the game in normal form with strategy sets $A$ and $B$ and payoff functions

$$\phi_r(a, b) = u(a, b) + d_r(a, b) \quad \text{and} \quad \psi_r(a, b) = -k(a, b) - d_r(a, b)$$

of party A and B, respectively.

**Proposition 4** (difference principle extended to an inefficient obligation)

(i) If the difference hypothesis is based on a reasonable person standard in the sense that conditions (12) and (13) are met then the compensation goal for unilateral deviations from the efficient reference profile is achieved (i.e. (4) and (5) will hold).
(ii) Conversely, if conditions (4) and (5) as well as the second requirement of (12) are met then (12) and (13) must hold.

(iii) If the requirements (12) and (13) are met then the efficient reference profile \((a^o, b^o)\) forms a Nash equilibrium and all Nash equilibria (if more than one exists) are payoff equivalent.

Proof. Claims (i) and (ii) are established by simple algebra and claim (iii) then follows from the compensation principle (i.e. proposition 1).

In the next section, the above proposition will be applied to a setting of non-contingent contracts and (reliance) investments in an uncertain environment. To economize on transaction costs, parties have specified non-contingent obligations nonetheless such that party B’’s performance obligation may become inefficient ex post (under some contingencies at least). To induce efficient breach, parties possibly rely instead on remedies as offered by contract law.

6 Non-contingent contracts

At stage 0, parties A and B agree on a simple (i.e. non-contingent) contract involving a binary performance decision to be taken by party B at stage 3. After having signed the contract, at stage 1, parties A and B decide on investments \(x \in X\) and \(y \in Y\), respectively. Investment costs \(g(x)\) are borne by party A, investment costs \(h(y)\) by party B. At the investment stage, benefits from investments are still uncertain as, not until stage 2, nature randomly selects the move \(\omega\) of nature from outcome space \(\Omega\) with an exogenously given probability distribution. At stage 3, knowing the actual move of nature, party B takes the (binary) performance decision \(q \in \{0, 1\}\).

The payoff functions net of investment costs of party A and B are denoted as \(V(x, y, \omega) \cdot q - g(x)\) and \(-C(x, y, \omega) \cdot q - h(y)\), respectively. For illustration, think of \(v = V(x, y, \omega)\) as the buyer’s utility, \(x\) as the buyer’s reliances and \(c = C(x, y, \omega)\) as the seller’s cost of performance. Since the investment profile is allowed to enter both payoff functions, investments may be of hybrid type. In contrast, investments are called purely selfish, if payoffs \(v = V(x, \omega)\) and \(c = C(y, \omega)\) were independent of the other party’s investments whereas they are called purely cooperative, if payoffs \(v = V(y, \omega)\) and \(c = C(x, \omega)\) depend exclusively on the other party’s investments.
The first best solution can be derived by backwards induction. If parties have invested \( x \) and \( y \) and if nature contributes move \( \omega \) then the ex post efficient performance decision solves

\[
q^+(x, y, \omega) \in \arg \max_{q \in \{0, 1\}} [V(x, y, \omega) - C(x, y, \omega)] \cdot q,
\]

whereas the efficient investment profile solves

\[
(x^o, y^o) \in \arg \max_{(x, y) \in X \times Y} E \left[ (V(x, y, \omega) - C(x, y, \omega)) \cdot q^+(x, y, \omega) \right] - g(x) - h(y).
\]

Notice, the efficient investment profile would also maximize expected welfare if the ex post efficient performance decision \( q^+(x, y, \omega) \) were replaced by \( q^+(x^o, y, \omega) \) or \( q^+(x^o, y^o, \omega) \). These alternative performance decisions may fail to be ex post efficient (off the equilibrium path) as they do not depend on actual investments \( x \) by party A. But they turn out to be implementable as a Nash equilibrium of the normal form game in a simpler way than the ex post efficient performance decision. In the present section, the two cases will be treated separately.

Before making use of the findings of the previous section, the sequential interaction between the two parties must be expressed in normal form. Since party A decides exclusively on her investments \( x \) from \( X \), her strategy set in the normal form game remains to be \( A = X \). Party B decides, at stage 1, on his investments \( y \) from \( Y \) and, at stage 3, on the performance decision \( q \in \{0, 1\} \). At stage 3, party B may condition his performance decision \( q = \eta(x, y, \omega) \) on actual investments and the move of nature. Let \( \{0, 1\}^{X \times Y \times \Omega} \) denote the set of all complete contingent performance plans \( \eta : X \times Y \times \Omega \to \{0, 1\} \) such that the Cartesian product \( B = Y \times \{0, 1\}^{X \times Y \times \Omega} \) constitutes party B’s strategy set in the normal form of the game.

In spite of uncertainty, to economize on transaction costs, parties have stipulated a simple contract which requires party B to perform non-contingently, i.e. to take the performance decision \( \eta^\infty(x, y, \omega) = 1 \) independent of actual investments and move of nature in exchange of which party A pays the sum \( T \) to party B. Let \( b^\infty = (y^\infty, \eta^\infty) \) denote the strategy of party B to invest efficiently and to perform under all circumstances. For obvious reasons, the obligation \( b^\infty \) may fail to be efficient (for some moves of nature at least).

Moreover, as incentives for efficient breach are at stake, let me also fix an efficient reference profile \((a^o, b^o) = (x^o, y^o, \eta^o(x, y, \omega))\) in the normal form
of the game which, under the appropriate damages regime, will form a Nash equilibrium.

With this notation at hand, the expected payoffs (not including damages) can be calculated explicitly. For any given strategy profile \((a, b) = (x, y, \eta) \in X \times Y \times \{0, 1\}^{X \times Y \times \Omega}\) of the normal form game, the expected payoffs and costs of party A and B amount to

\[
u(a, b) = E\left[ V(x, y, \omega) \cdot \eta(x, y, \omega) \right] - g(x) - T
\]

and

\[
k(a, b) = E\left[ C(x, y, \omega) \cdot \eta(x, y, \omega) \right] + h(y) - T,
\]

respectively.

Damages (if at all) will be quantified ex post when, by assumption, courts can observe the actual investment profile \((x, y)\), the move \(\omega\) of nature and party B’s performance decision \(q\) but not his complete contingent performance plan \(\eta \in \{0, 1\}^{X \times Y \times \Omega}\).

To begin with, suppose the efficient reference profile includes a performance decision \(\eta^o(y, \omega)\) that does not depend on actual investments \(x\) of party A. In this case, the following restrictions on damages will ensure that expected damages are in line with the difference hypothesis based on a reasonable person standard (see proposition 4).

If party A invests efficiently, \(x = x^o\), then damages are restricted by

\[
D_r(x^o, y^o, \omega, q) \geq V(x^o, y^o, \omega) \cdot (1 - q)
\]

but

\[
D_r(x^o, y^o, \omega, q) = V(x^o, y^o, \omega) \cdot (1 - q)
\]

(for all \(y, \omega\) and \(q\)). If party A deviates unilaterally, \(x \neq x^o\), damages are restricted by

\[
D_r(x, y^o, \omega, q) \leq V(x^o, y^o, \omega) \cdot (1 - q) + [C(x^o, y^o, \omega) - C(x, y^o, \omega)] \cdot q
\]

(for all \(x, \omega\) and \(q\) again).

If both deviate from the efficient investment profile, the specification of damages does not matter for the following proposition to hold. The proposition concerns the game in normal form with strategy sets \(A = X\) and
$B = Y \times \{0, 1\}^{X \times Y \times \Omega}$ and payoff functions

$$\phi_r(a, b) = u(a, b) + d_r(a, b)$$
$$\psi_r(a, b) = -k(a, b) - d_r(a, b)$$

where $d_r(a, b) = E[D_r(x, y, \omega, \eta(x, y, \omega))]$ denotes expected damages at strategy profile $(a, b) = (x, y, \eta)$.

**Proposition 5** Suppose the efficient reference profile $(a^o, b^o) = (x^o, y^o, \eta^o)$ is such that party B’s performance decision $\eta^o = \eta^o(y, \omega)$ remains independent of party A’s actual reliance investments $x$. Moreover, damages $D_r(x, y, \omega, \eta)$ are awarded satisfying (14) – (16). Then the efficient reference profile forms a Nash equilibrium of the corresponding normal form game and all Nash equilibria (if more than one exists) are payoff equivalent (in expected terms).

**Proof.** It will first be shown that damages $d_r(a, b) = E[D_r(x, y, \omega, \eta(x, y, \omega))]$ satisfy the conditions (12) and (13) of the extended version of the difference principle. In fact,

$$k(a^o, b^o) - k(a, b^o) + u(a^o, b^{oo}) - u(a^o, b^o)$$
$$= E [ (C(x^o, y^o, \omega) - C(x, y^o, \omega)) \cdot \eta^o(y^o, \omega) + V(x^o, y^o, \omega) \cdot (1 - \eta^o(y^o, \omega)) ] .$$

It then follows from (14) that

$$d_r(a^o, b^o) \geq E [ V(x^o, y^o, \omega) \cdot (1 - \eta(x^o, y, \omega))] = u(a^o, b^{oo}) - u(a^o, b^o)$$

and from (15) that $d_r(a^o, b^o) = u(a^o, b^{oo}) - u(a^o, b^o)$ and, hence, (12) must hold indeed.

Similarly, it follows from (16) that

$$d_r(a, b^o) = E[D_r(x, y^o, \omega, \eta^o)]$$
$$\leq E \left[ (C(x^o, y^o, \omega) - C(x, y^o, \omega)) \cdot \eta^o(y^o, \omega) + V(x^o, y^o, \omega) \cdot (1 - \eta^o(y^o, \omega)) \right] =$$
$$= k(a^o, b^o) - k(a, b^o) + u(a^o, b^{oo}) - u(a^o, b^o)$$

and, hence, (13) must hold as well.

The claim of the proposition then immediately follows from the extended version of the difference principle (see proposition 4). ■

The above proposition generalizes several findings from the existing literature. To begin with, the result is in the spirit of what Cooter (1985) has originally introduced as efficient expectation damages.
Second, the above result also captures and generalizes the findings of Göller and Hewer (2014) who have proposed an efficient compensation regime for takings. Their case corresponds to one of purely selfish investments.

Third, Schweizer (2006) has introduced bilateral damages regimes in a setting of one-sided cooperative investments. The above result generalizes my earlier findings to two-sided investments and to fully hybrid investments. Notice further, any subgame perfect equilibrium in the extensive form of the game constitutes a Nash equilibrium in the corresponding normal form game and, hence, must be payoff equivalent to the Nash equilibria of the above proposition as well.

To conclude the present section, I now consider the ex post efficient reference profile \((x^o, y^o, q^o = q^+)\). For this reference profile to constitute a Nash equilibrium of the corresponding normal form game, the following restrictions on damages \(D^+(x, y, \omega, q)\) turn out to be sufficient. First, if party A has invested efficiently, \(x = x^o\), then damages

\[
D^+(x^o, y, \omega, q) = D_r(x^o, y, \omega, q)
\]  

are as before (see (14) and (15)). Second, if party A has deviated, \(x \neq x^o\), but party B has invested efficiently, \(y = y^o\), then

\[
D^+(x, y^o, \omega, q) \leq V(x^o, y^o, \omega) - S^o(\omega) - C(x^o, y^o, \omega) \cdot q
\]  

is assumed to hold where

\[
S^o(\omega) = \max \{V(x^o, y^o, \omega) - C(x^o, y^o, \omega), 0\}
\]

denotes the maximum surplus as a function of the move \(\omega\) of nature at the efficient investment profile \((x^o, y^o)\).

Finally, if both parties have deviated from the efficient investment profile, the exact specification of damages does not matter for the following proposition to remain valid.

The proposition concerns the game in normal form with \(A = X\) and \(B = Y \times \{0, 1\}^X \times Y \times \Omega\) as strategy sets and

\[
\phi^+ (a, b) = u(a, b) + d^+ (a, b)
\]

and

\[
\psi^+ (a, b) = -k(a, b) - d^+ (a, b)
\]
as payoff functions where $d^+(a, b) = E[D^+(x, y, \omega, \eta(x, y, \omega))]$ denotes expected damages under the above regime.

**Proposition 6** Suppose the efficient reference profile $(a^o, b^o) = (x^o, y^o, \eta^o)$ is such that party B’s performance decision $\eta^o(x, y, \omega) = q^+(x, y, \omega)$ is ex post efficient. Moreover, damages $D^+(x, y, \omega, q)$ are awarded satisfying (17) and (18). Then the ex post efficient reference profile forms a Nash equilibrium of the corresponding normal form game and all Nash equilibria are payoff equivalent (in expected terms).

**Proof.** At strategy profile $(a^o, b) = (x^o, y, \eta)$, expected damages are the same as in proposition 5 and, hence, conditions (12) must be satisfied for damages $d^+(a^o, b)$ as well.

and (13)

If, however, $a \neq a^o$, then expected damages satisfy

$$d^+(a, b^o) = E[D^+(x, y^o, \omega, \eta^o(x, y^o, \omega))] \leq E[V(x^o, y^o, \omega) - S^o(\omega) - C(x, y^o, \omega) \cdot \eta^o(x, y^o, \omega)] =$$

$$= E \left[ C(x^o, y^o, \omega) \cdot q^+(x^o, y^o, \omega) - C(x, y^o, \omega) \cdot \eta^o(x, y^o, \omega) + V(x^o, y^o, \omega) \cdot (1 - q^+(x^o, y^o, \omega)) \right]$$

$$= k(a^o, b^o) - k(a, b^o) + u(a^o, b^oo) - u(a^o, b^o)$$

(recall that $\eta^o = q^+$) and, hence, (13) holds for $a \neq a^o$ as well.

The claim of the proposition then immediately follows from the extension of the difference principle (see proposition 4).

Notice, as in section 3 also dealing with sequential choice, the ex post efficient reference profile in the above proposition forms a Nash equilibrium in the normal form of the game though not necessarily a subgame perfect equilibrium in its extensive form. While efficient investments are a best response of party A to the anticipated ex post efficient performance decision of party B, the ex post efficient performance decision of party B need not be a best response to inefficient investments at stage 1 (i.e. off the equilibrium path).

Yet, the ex post efficient reference profile is of particular interest in a setting where counterfactuals fail to be observable. In fact, the knowledge of the actual values and costs of performance, $v = V(x, y, \omega)$ and $c = C(x, y, \omega)$, is sufficient to take the ex post efficient performance decision. In such an
informational setting, the ex post efficient performance decision remains to be an available strategy, in contrast to a performance decision that would be ex post efficient under efficient but possibly counterfactual investments $x^o$ only (i.e. the strategy underlying the legal regime of proposition 5). As an application of proposition 6, the next section deals with such an informational setting.

7 Unobservable counterfactuals

The setting is still that of the previous one except that counterfactuals are no longer observable. More precisely, the following informational setting is considered. Before party B decides on performance, he learns the actual utility $v$ and costs $c$ of performance and, hence, he would still be able to reach the ex post efficient performance decision, namely $q^*(v, c) = 1$ if $c \leq v$ and $q^*(v, c) = 0$ otherwise. Courts when called in (ex post) can also observe $v$ and $c$ and, on top of it, actual investments $x$ and $y$ by both parties. They cannot observe, however, counterfactuals as would be needed to implement the damages regime (17) – (18) of proposition 6 directly. Yet, by implementing essentially the same regime but conditional on the observed event, the equilibrium payoffs in the normal form of the game and, hence, the incentives will remain the same.

To spell out details of the approach, some more notation is needed. Let $I$ denote an index set such that, for all $i \in I$, there exists a pair $(v_i, c_i)$ of realizations of utility and costs of performance. This index set is assumed large enough to capture all possible realizations.

Suppose parties have actually invested $x$ and $y$ and the pair $(v_i, c_i)$ has been realized. While counterfactuals may remain uncertain, parties and courts can still infer that the event

$$\Omega_i(x, y) = \{\omega \in \Omega : V(x, y, \omega) = v_i \text{ and } C(x, y, \omega) = c_i\}$$

must have occurred. Notice, for any given investment profile $(x, y)$, the observable events $\Omega_i(x, y)$ form a partition of the outcome space

$$\Omega = \bigcup_{i \in I} \Omega_i(x, y).$$

In this setting of unobservable counterfactuals, the set $H_{uc}$ of available performance decisions of party B forms a subset of the set $\{0, 1\}^{X \times Y \times \Omega}$ of all
complete contingent plans. In fact, as party B cannot distinguish different moves of nature from the event $\Omega_i$, his performance decision $\eta(x, y, \omega)$ must be constant on the event $\Omega_i = \Omega_i(x, y)$ and, hence, functions $q_i(x, y)$ must exist such that

$$\eta(x, y, \omega) = q_i(x, y)$$

holds for all $\omega \in \Omega_i(x, y)$ and all $(x, y) \in X \times Y$. Notice, the ex post efficient performance decision characterized by $q_i = q^*(v_i, c_i)$ is even independent of actual investments and, in particular, remains to be a feasible plan for party B also in the setting of unobservable counterfactuals. In the following, the ex post efficient performance decision is denoted by $\eta^\ast \in H_{uc}$.

Suppose courts have observed the event $\Omega_i(x, y)$ and party B has reached performance decision $q_i$. To preserve efficient incentives, courts should award damages according to the legal regime (17) – (18) but on average over the observed event, i.e.

$$D_i^+(x, y, q_i) = E \left[ D^+(x, y, \omega, q_i) \mid \Omega_i(x, y) \right].$$

The use of conditional probabilities ensures that, from the ex ante perspective, expected damages claims and, hence, incentives in the normal form of the game remain the same as in proposition 6.

In fact, for any available strategy profile $(x, y, \eta) \in X \times Y \times H_{uc}$ in the normal form of the game, the expected damages claims under the legal regime as proposed above amount to

$$\sum_{i \in I} \text{prob} \{ \Omega_i(x, y) \} \cdot D_i^+(x, y, q_i(x, y)) = E \left[ D^+(x, y, \omega, q_i(x, y)) \right]$$

and are equal to the corresponding damages $d^+(x, y, \eta)$ underlying proposition 6 (the observable sets form a partition of the outcome space $\Omega$!). As a consequence, the ex post efficient reference profile $(x^o, y^o, \eta^o)$ remains to be a Nash equilibrium in the setting of unobservable counterfactuals.

Moreover, if more than one Nash equilibrium exist, they are all payoff equivalent. But keep in mind, in the setting of unobservable counterfactuals, only those performance decisions can be part of a complete contingent plan that are constant on observable events. On this account, performance decisions as reached in the Nash equilibria according to proposition 5 may have to be ruled out.
To further illustrate the approach, let me explicitly quantify damages along the above lines for the following example of purely selfish investments (i.e. utility $v = V(x, \omega)$ and costs $c = C(y, \omega)$ of performance do not depend on the other party’s investment decision).

If both parties deviate from the efficient investment profile $(x^o, y^o)$, the specification of damages does not matter for the efficiency of incentives. If only party B deviates (i.e. $y \neq y^o$) while party A invests efficiently (i.e. $x = x^o$) and the event $\Omega_i(x^o, y)$ has been observed then damages not lower than

$$D_i^+ \geq V(x^o, \omega) \cdot (1 - q_i) = v_i \cdot (1 - q_i)$$

should be awarded. In fact, in this event, the observed utility $v_i$ coincides with the hypothetical utility $V(x^o, \omega)$ because the utility is assumed to be independent of party B’s investment decision $y$.

If both parties have invested efficiently (i.e. $(x, y) = (x^o, y^o)$) and the event $\Omega_i(x^o, y^o)$ has occurred then the above inequality must even be binding, i.e.

$$D_i^+ = V(x^o, \omega) \cdot q_i = v_i \cdot (1 - q_i)$$

must hold. If, finally, only party A has deviated (i.e. $x \neq x^o$ but $y = y^o$) and the event $\Omega_i(x, y^o)$ has occurred then damages not higher than

$$D_i^+ \leq E[V(x^o, \omega) - S^o(\omega) \mid \Omega_i(x, y^o)] - c_i \cdot q_i$$

should be awarded as, in this event, the observed costs $c_i$ of performance coincide with the hypothetical costs $C(y^o, \omega)$ (investments $y$ are also of purely selfish nature). Notice, in this event, the identity

$$V(x^o, \omega) - S^o(\omega) = V(x^o, \omega) - \max [V(x^o, \omega) - c_i, 0] = \min [c_i, V(x^o, \omega)]$$

must hold. Moreover, if (by assumption) utility $V(x^o, \omega)$ is distributed independently of party B’s investments $y$, then the upper bound in (19) simplifies to

$$D_i^+ \leq E [\min [c_i, V(x^o, \omega)] \mid V(x, \omega) = v_i] - c \cdot q_i.$$  

For legal systems rejecting overcompensation, the above condition must even hold as an equality. In any case, to calculate the upper bound explicitly, the conditional distribution of the counterfactual utility $V(x^o, \omega)$ will be needed.
To proceed further, let me impose two simplifying assumptions. First, the utility $V(x, \omega)$ from performance is assumed to attain two values $v_L < v_H$ only and party A’s investment decision concerns the probability $x = \text{prob} (V(x, \omega) = v_H)$ with which the utility attains its high value. This assumption is reminiscent of the accident model. Second, contingencies under which the utility from performance would be lower in spite of higher investments are ruled out such that the utility is a monotonously increasing function of $x$ (i.e. if $x < x'$ then $V(x, \omega) \leq V(x', \omega)$ for all moves $\omega$ of nature).

Under these simplifying assumptions, let me quantify damages also for the case where party A has invested excessively (i.e. $x^o < x$). If, nonetheless, the low utility of performance $V(x, \omega) = v_L$ has been realized then the counterfactual utility $V(x^o, \omega) = v_L$ must a fortiori be low (as follows from the imposed monotonicity) and, hence, the upper bound in (19) amounts to

$$D^+_i \leq \min [c_i, v_L] - c_i \cdot q_i$$

for sure.

If, however, the high utility $V(x, \omega) = v_H$ has been realized the counterfactual utility $V(x^o, \omega)$ remains uncertain. In fact, with probability $x^o/x$, it attains the high value $V(x^o, \omega) = v_H$ whereas, with probability $(x - x^o)/x$, it attains the low value $V(x^o, \omega) = v_L$. On average over the observed event, the upper bound in (19) is then given by

$$D^+_i \leq \frac{x - x^o}{x} \cdot \min [c_i, v_L] + \frac{x^o}{x} \cdot \min [c_i, v_H] - c_i \cdot q_i.$$

To sum up, courts can implement the above damages regime even if they are unable to observe the counterfactuals. To maintain efficient incentives, courts should award damages in line with proposition 6 though, for lack of information, on average over the observed event.

8 Conclusion

From the legal perspective, compensation of the creditor is considered as the primary goal of obligation law whereas the economic analysis sees the efficiency of incentives as the ultimate goal. Some texts suggest a conflict between the two perspectives. Cooter (1985) even talks of a compensation paradox.
The present paper, in contrast, propagates the compensation principle as a tool to establish the efficiency of provisions from obligation law. While exact compensation in a bilateral setting may be beyond reach indeed, efficient incentives require compensation for unilateral deviations only. In particular, any regime that awards damages in line with the difference hypothesis achieves the compensation goal and, according to the compensation principle, generates efficient incentives.

Achieving the compensation goal is a sufficient but not a necessary condition for efficient incentives. The present paper presents a series of efficiency results where this sufficient condition is met and where the compensation principle establishes efficiency from an unifying perspective. No doubt, many more applications are around where the argument would also work even if the existing literature, instead of the compensation principle, has relied on the methodologically more demanding first order approach.

There exist, however, efficiency results where the first order approach works but the compensation principle does not (see Edlin and Reichelstein, Ohlendorf (2009) or Stremitzer (2012), e.g.). But such efficiency results are of a less robust nature as they require more restrictive assumptions.

9 References


Ohlendorf, S.- (2009), "Expectation Damages, Divisible Contracts and


