

**FUNDAMENTAL CHARACTERISTICS OF INCENTIVE STREAMS
CREATED BY LEGAL SYSTEMS**

**ALGEMENE KARAKTERISTIEKEN VAN INCENTIVE STROMEN
GECREËERD DOOR RECHTSSYSTEMEN**

(met een samenvatting in het Nederlands)

Proefschrift

ter verkrijging van de graad van

doctor

aan de Universiteit Utrecht

op gezag van

de **Rector Magnificus Prof. dr. W.H. Gispen**

ingevolge het besluit van het College voor Promoties

in het openbaar te verdedigen

op 14 juni 2002 des namiddags om 2.30 uur

door

Giuseppe Dari Mattiacci

geboren op 29 april 1974

te Latina

Ph.D. Dissertation – Utrecht University

SUPERVISORS:

Prof. Dr. Gerrit De Geest
Utrecht University
Ghent University

Prof. Dr. Jacques J. Siegers
Utrecht University

All right reserved © Giuseppe Dari Mattiacci 2002

ISBN number: 90-393-3036-0

*A papà, mamma e Marco,
che mi hanno sempre
incoraggiato e sostenuto*

ACKNOWLEDGMENTS

This work could not have been accomplished without the indispensable assistance of my supervisors, Gerrit De Geest and Jacques J. Siegers, whom I wish to thank not only for their efforts as supervisors but also for what they have done beyond this function. They have presented me with a perfect research environment and with a pleasant sense of warmth and *gezelligheid*, which has made my work enjoyable. I also thank them for their permission to include our joint papers in this book.

I thank the members of the Ph.D. commission for accepting this task, all my friends and my family, who have always managed to transform an experience into a wonderful experience, and my colleagues at the Institute of Economics for their support, especially Tineke Dupont, Ellie Renes and Carine Toet-Mariette, for their daily assistance, Johan den Hertog, who shared the office with me, and Christine Baaijens, Chantal Remery and Ann-Sophie Vandenberghe, for their continual encouragement and many helpful suggestions.

Many colleagues and friends offered me precious suggestions and comments concerning some of the chapters. I am indebted to Alessandra Arcuri, Mireia Artigot Golobardez, Alexandra Ayala, Kennet McArthur, Marianne Boutellier, Roger Bowles, Eimaer Coffie, Jef De Mot, Robert Dur, Ehrling Eide, Winand Emons, Thomas Eger, Carlo Ferrigno, Gianluigi Galeotti, Nuno Garoupa, Panos Giannopoulos, Mark Grady, Margaret Mary Hammersley, Tina Heubeck, Manfred Holler, Henrik Lando, Jana McLean, Willemijn Lindenhovius, Vicente Mendes, Katja Modric, Antony Oagus, Francesco Parisi, Mitchell Polinsky, Hein Roelfsema, Caspar Rose, Alessandra Ruggeri, Hans-Bernd Schäfer, Pietro Trimarchi, Ann-Sophie Vandenberghe, Ben van Velthoven and Wouter Wils. The usual disclaimer applies.

I wish to thank the organizers of and the participants in the annual conferences of the European Association of Law and Economics, Ghent (2000) and Vienna (2001), the annual meeting of the American Law and Economics Association, Washington (2001), the annual meeting of the Italian Economists Society, Rome (2001), the Marie-Curie workshop at Hamburg University (2001), the seminar at Copenhagen Business School (2001), and the CASLE workshop, Ghent (2002), where most of the chapters were presented.

Chapter 2 is forthcoming in Hatzis, Aritides, *Economic Analysis of Law: A European Perspective*, and included in this book with the permission of the publisher, Edward Elgar.

Assistance for part of this research was provided by Hamburg University and the Max Plank Institute in Hamburg and financially supported by the European Commission, the Marie-Curie Fellowship Programme in Law and Economics.

A note to the reader

The chapters 2 to 9 have been written in the form of separate articles, references are listed at the end of each chapter. As they deal with problems of interest both for lawyers and for economists, I have tried to limit the use of technical terms, unless it was strictly necessary, and have explained them in the text or in the footnotes. In many cases, the context will be a helpful clarifier. Updated versions of the chapters will be available on my web page: www.law.uu.nl/eco/gdm.htm.

Utrecht, March 19th, 2002

CURRICULUM VITAE

Giuseppe Dari Mattiaci (Latina, Italy, 1974) studied law at the University of Rome “La Sapienza” (cum laude, 1997) and took courses on corporate regulation and statistics (1998). He then joined the LL.M. Programme in Law and Economics (summa cum laude, 2000), and the Ph.D. programme at Utrecht University. In 2001/2002 he was a Marie Curie fellow at Hamburg University. The author is currently researcher at the Institute of Economics, Utrecht University.

TABLE OF CONTENTS

1. Introduction

2. Tort law and economics

Forthcoming in Hatzis, Aritides ed., Economic Analysis of Law: A European Perspective, Edward Elgar

3. On the definitions of care and activity level and the choice of liability rules

Presented at the Marie Curie Fellowship Programme workshop, Hamburg University, July 6th, 2001

Presented at the CASLE workshop, Ghent University, February 22nd, 2002

4. Comparative negligence as a filter mechanism when courts cannot observe precaution costs (with Gerrit De Geest)

Presented at the 17th annual conference of the European Association of Law and Economics, Ghent, September 14th-16th, 2000

Presented at the annual meeting of the Italian Economists' Society (SIE), Rome, October 25th-27th, 2001

5. An analysis of the judgment-proof problem under different tort models (with Gerrit De Geest)

Presented at Copenhagen Business School, May 20th, 2001

Presented at the 18th annual conference of the European Association of Law and Economics, Vienna, September 13th-15th, 2001

6. On the combined use of anti-judgment-proof regulation and tort law (with Gerrit De Geest)

7. When will judgment-proof injurer take too much precaution? (with Gerrit De Geest)

To be presented at Hamburg University, May 23rd, 2002

8. Noxae deditio and negligence: Roman and modern solutions to inefficient monitoring under vicarious liability

9. Efficiency wages, conditional bonuses and punitive damages: monitoring levels when sanctions are financed by victims (with Gerrit De Geest and Jacques J. Siegers)

Presented at the annual meeting of the American Law and Economics Association, Washington, May 11th-12th, 2001

10. Conclusions

Summary in Dutch

DETAILED TABLE OF CONTENTS

1. Introduction

- 1.1. Aim and scope
- 1.2. Method
- 1.3. An overview
- 1.4. References

2. Tort law and economics

Forthcoming in Hatzis, Aritides ed., Economic Analysis of Law: A European Perspective, Edward Elgar

- 2.1. Introduction
- 2.2. Goals of the tort law system: incentives towards optimal precaution, the basic model
- 2.3. Why do we have tort law? Between property rights and inalienability
- 2.4. Accident types
- 2.5. Elements of tort liability: harm, cause and fault
 - 2.5.A. Harm
 - 2.5.B. Cause
 - 2.5.C. Fault
- 2.6. Liability rules
 - 2.6.A. Strict liability and no liability
 - 2.6.B. Precaution under negligence rules: care and activity level
 - 2.6.C. The negligence rules: simple negligence
 - 2.6.D. The negligence rules: contributory negligence
 - 2.6.E. The negligence rules: strict liability with defence of contributory negligence and strict liability with defence of dual contributory negligence
 - 2.6.F. Negligence rules: the efficiency-equivalence theorem and the activity-level theorem
 - 2.6.G. Comparative negligence
 - 2.6.H. Causation and the negligence rules
- 2.7. Errors and uncertainty
 - 2.7.A. Errors and uncertainty: harm
 - 2.7.B. Errors and uncertainty: cause
 - 2.7.C. Errors and uncertainty: fault
- 2.8. Goals of a tort law system: incentives to acquire information
- 2.9. Additional incentive streams
- 2.10. Goals of a tort law system: optimal risk allocation. Insurance
- 2.11. Goals of a tort law system: minimisation of administrative cost
- 2.12. Specific issues in tort liability
 - 2.12.A. Product liability and ultra hazardous activities
 - 2.12.B. Vicarious liability
 - 2.12.C. Multiple tortfeasors and the right to contribution
 - 2.12.D. Pure economic losses
 - 2.12.E. Non-pecuniary losses
 - 2.12.F. Low-probability accidents and egg-shell-skull cases
- 2.13. Concluding remarks
- 2.14. References

3. On the definitions of care and activity level and the choice of liability rules

Presented at the Marie Curie Fellowship Programme workshop, Hamburg University, July 6th, 2001

Presented at the CASLE workshop, Ghent University, February 22nd, 2002

- 3.1. Introduction
 - 3.1.A. A criticism of the traditional definitions of care and activity level
- 3.2. The model: assumptions and notation
- 3.3. Torts in a world with no administrative cost
 - 3.3.A. The social cost in a world with no administrative cost
 - 3.3.B. Liability rules in a world with no administrative cost
- 3.4. Torts in a world with positive administrative cost
 - 3.4.A. Administrative costs
 - 3.4.B. Social cost and liability rules in a world with positive administrative cost. Genesis of the distinction between care and activity level
 - 3.4.C. The dangerousness criterion and the verifiability criterion
- 3.5. Characteristics of the new definitions of care and activity level: endogeneity and variability
- 3.6. A diachronic perspective: the evolution of liability rules
 - 3.6.A. A historical case for proposition 5: Roman and English legal history
 - 3.6.B. A historical case for proposition 6 and a criticism of the application of the dangerousness criterion in order to explain the choice of liability regimes in modern societies: liability for industrial accidents and product liability from no-liability-based rules to strict-liability-based rules
- 3.7. A synchronic perspective: “strict liability vs. negligence”, an analysis of liability rules in modern legal systems
 - 3.7.A. The general application of negligence
 - 3.7.B. The implementation of strict liability in specific areas of torts
 - 3.7.C. Room for no liability
- 3.8. Concluding remarks
- 3.9. References

4. Comparative negligence as a filter mechanism when courts cannot observe precaution costs (with Gerrit De Geest)

Presented at the 17th annual conference of the European Association of Law and Economics, Ghent, September 14th-16th, 2000

Presented at the annual meeting of the Italian Economists' Society (SIE), Rome, October 25th-27th, 2001

- 4.1. Introduction
 - 4.2. The model
 - 4.2.A. Standard assumptions
 - 4.2.B. The graph
 - 4.2.C. The social welfare function
 - 4.3. Unilateral care (only one party can prevent the accident at a reasonable cost)
 - 4.3.A. Uniform distributions
 - 4.3.B. Non-uniform distributions
 - 4.3.C. Sharing rules other than 50/50. All-or-nothing rules as limit-cases of comparative negligence.
 - 4.3.D. Under- or overcompensation
 - 4.3.E. Relaxing the assumption of systematic and predictable errors in compensation
 - 4.3.F. Relaxing the assumption of perfect predictability of the judicial decisions
 - 4.4. Alternative care in strict sense (each party can avoid the accident at a reasonable cost)
 - 4.5. Efficient accidents (each party can prevent the accident, but no party at a reasonable cost)
 - 4.6. Concluding remarks
 - 4.7. References
- Appendix
- Appendix 1. Simulation of non-uniform distributions of the costs of care in unilateral care cases
 - Appendix 2. Overcompensation
 - Appendix 3. Undercompensation
 - Appendix 4. Sharing rules other than 50/50

5. An analysis of the judgment-proof problem under different tort models (with Gerrit De Geest)

Presented at Copenhagen Business School, May 20th, 2001

Presented at the 18th annual conference of the European Association of Law and Economics, Vienna, September 13th-15th, 2001

- 5.1. Introduction
 - 5.1.A. Applicability of the analysis
 - 5.1.B. A note on the literature
- 5.2. The models
 - 5.2.A. First model: probability model
 - 5.2.B. Second model: magnitude model
 - 5.2.C. Third model: joint probability-magnitude model
 - 5.2.D. Fourth model: separate (three-dimensional) probability-magnitude model
- 5.3. Stochastic variation of the harm
- 5.4. Correcting inefficient deterrence
 - 5.4.A. Punitive damages (over-compensation of the harm)
 - 5.4.B. Average compensation of the harm
 - 5.4.C. Under-compensation of the harm
 - 5.4.D. Negligence
- 5.5. Concluding remarks
- 5.6. References

6. On the combined use of anti-judgment-proof regulation and tort law (with Gerrit De Geest)

- 6.1. Introduction
 - 6.1.A. A reason why tort law and regulation might fail if employed alone
 - 6.1.B. How to combine regulation and tort law: regulation of minimum levels of precaution
 - 6.1.C. How to combine regulation and tort law: regulation of magnitude reducing precaution
 - 6.1.D. When regulatory compliance should not relieve of liability for negligence
 - 6.1.E. Structure of this chapter
 - 6.1.F. A note on the literature
- 6.2. The one-pocket and the two-pocket magnitude models
 - 6.2.A. Judgment-proofness under strict liability
 - 6.2.B. Judgment-proofness under simple negligence
- 6.3. Anti-judgment-proof regulation
 - 6.3.A. Anti-judgment-proof regulation under strict liability
 - 6.3.B. Anti-judgment-proof regulation under simple negligence
 - 6.3.C. Imperfectly informed regulators
- 6.4. Separate probability-magnitude models
- 6.5. Probability model and joint-probability-magnitude model
 - 6.5.A. The probability model
 - 6.5.B. The joint-probability-magnitude model
- 6.6. Concluding remarks
- 6.7. References

7. When will judgment-proof injurer take too much precaution? (with Gerrit De Geest)

To be presented at Hamburg University, May 23rd, 2002

- 7.1. Introduction. Two-pocket probability model and two-pocket magnitude model.
- 7.2. The one-pocket probability model

- 7.3. The one-pocket magnitude model
- 7.4. The intuition behind the different findings for the one-pocket magnitude and the probability model
- 7.5. Mixed one-pocket probability-magnitude models
- 7.6. Concluding remarks. Policy implications
- 7.7. References

8. Noxae deditio and negligence: Roman and modern solutions to inefficient monitoring under vicarious liability

- 8.1. Introduction
 - 8.1.A. A numerical example
- 8.2. The model
 - 8.2.A. Unilateral accidents
 - 8.2.B. Bilateral accidents
- 8.3. Different combinations of the negligence criterion in employers' liability
- 8.4. Vicarious liability of parents and supervisors: negligence as a solution for over-demanding principals
- 8.5. *Noxae deditio* as a Roman law solution for over-demanding vicariously-liable *patresfamilias*
- 8.6. Concluding remarks
- 8.7. References

9. Efficiency wages, conditional bonuses and punitive damages: monitoring levels when sanctions are financed by victims (with Gerrit De Geest and Jacques J. Siegers)

Presented at the annual meeting of the American Law and Economics Association, Washington, May 11th-12th, 2001

- 9.1. Introduction
- 9.2. The basic framework (one-period model)
 - 9.2.A. No-shirking constraint under efficiency wages
 - 9.2.B. No-shirking constraint under conditional bonuses
 - 9.2.C. No-shirking constraint under punitive damages
 - 9.2.D. Participation Constraint
 - 9.2.E. Maximum sanction constraint (creating a feasibility constraint for x)
 - 9.2.F. Normalized efficiency wages
 - 9.2.G. Normalized conditional bonuses
- 9.3. Chosen monitoring levels (one-period model)
- 9.4. Chosen sanction (one-period model)
- 9.5. Chosen monitoring levels and sanctions under combined sanctions (one-period model)
- 9.6. N -period model
 - 9.6.A. The n -period model
 - 9.6.B. Efficiency wages and retroactivity
 - 9.6.C. Constraints and discount factor in an n -period model
 - 9.6.D. Sanctions and monitoring level in an n -period model
- 9.7. Turnover costs
- 9.8. A note on the appropriation problem
- 9.9. Concluding remarks
 - 9.9.A. Implications for contract law
 - 9.9.B. Implications for labor law
 - 9.9.C. Implications for corruption issues
- 9.10. References

10. Conclusions

- 10.1. Some general characteristics of incentive streams
 - 10.1.A. “One bullet for two gangsters”
 - 10.1.B. “Two gangsters in the dark: shoot either one or point in the middle and wait?”
 - 10.1.C. “Shooting bulletproof gangsters”
 - 10.1.D. “Are gangsters easier to capture if police have orders to kill?”
 - 10.1.E. “Are parents too strict with badly behaved children?”
 - 10.1.F. “Should broken promises be sanctioned or should kept promises be rewarded?”
- 10.2. Relation to the literature
 - 10.2.A. Activity level vs. care
 - 10.2.B. Strict liability vs. negligence
 - 10.2.C. Sharing damages among multiple tortfeasors
 - 10.2.D. The judgment-proof problem
 - 10.2.E. The combined use of regulation and tort law
 - 10.2.F. Vicarious liability
 - 10.2.G. The efficiency wage theory
 - 10.2.H. Towards and economic analysis of Roman law
 - 10.2.I. The economic analysis of the evolution of legal systems
- 10.3. Relevance for policy makers
 - 10.3.A. The choice of the residual bearer
 - 10.3.B. The extent of the negligence criterion
 - 10.3.C. Comparative negligence and the rule of contribution
 - 10.3.D. Insolvency
 - 10.3.E. Liability caps
 - 10.3.F. Negligence-based vs. strict vicarious liability
- 10.4. A final remark
- 10.5. References

Summary in Dutch

INDEX

- Accident types 2.4
- Actiones 8.5
- Activity level vs. care 3.1.A, 3.5, 10.2.A
- Administrative costs 3.4.A
- Allocative loss 3.3.A, 3.4.B
- All-or-nothing rules 4.1
- Alternative precaution accidents 4.4 *see also accident types*
- Anti-judgment-proof regulation 6, 10.2.E
- Appropriation problem 9.8
- Average compensation 5.4.B
- Bankruptcy *see maximum upper threshold*
- Bilateral precaution accidents 8.2.B *see also accident types*
- Bonds 9.8
- Burden of proof 3.7
- Care *see activity level vs. care*
- Causation *see elements of tort liability*
- Claim costs *see compensation costs*
- Comparative negligence 4, 2.6.G, 10.3.C
- Compensation costs *see administrative costs*
- Compliance defense 6.1.D
- Conditional bonuses 9
- Contributory negligence 2.6.D
- Corruption 9.9.C
- Criminal law 2.9
- Damages caused by:
 - Things 3.7.B
 - Animals *see damages caused by things*
 - Children *see vicarious liability*
 - Persons under supervision *see vicarious liability*
- Dangerousness criterion 3.4.C

Economic analysis of contracts 9, 8.2

Economics analysis of accident law *see tort law and economics*

Efficiency wages 9, 10.2.G

Efficient accidents *see accident types*

Egg-shell-skull accidents 2.12.F

Elements of tort liability

- Harm 2.5.A
- Cause 2.5.B
- Fault 2.5.C

Errors and uncertainty 2.7

Evolution of English law 3.6.A

Evolution of Roman law 3.6.A, 8.5, 10.2.I

Evolution of tort law systems 3.6, 10.2.I

Ex ante vs. ex post control of externalities 6.1

Extent of the negligence criterion 3.3, 3.4, 10.3.B

Fault *see negligence*

Fault-based vicarious liability 8.4, 10.3.F

Filter argument 4

Goals of the tort law system

- Optimal precaution 2.2
- Acquiring information 2.8
- Risk allocation 2.10
- Minimization of administrative costs 2.11

Harm-reducing precaution *see magnitude-reducing precaution*

Imperfect information 4.1, 6.1.A, 6.3.C

Inalienability 2.3

Incentive stream 1.1, 10.1

Incentives to acquire information about risk *see goals of the tort law system*

Incentives to take precaution *see goals of the tort law system*

Industry subsidy argument 3.6.B

Information costs *see administrative costs*

Insolvency *see maximum upper threshold*

Joint precaution accidents *see accident types*

Judgment-proof problem *see maximum upper threshold*

Labor contracts 9

Liability caps 5.1, 6.1, 10.3.E

Low-probability accidents 2.12.F

Magnitude-reducing precaution 5.2.B

Maximum sanction constraint 9.2.E

Maximum upper threshold

The problem 5.1, 10.2.D, 10.3.D

Tort law solutions 5.4

Regulatory solutions 6

Two-pocket models

Probability models 5.2.A, 6.5.A, 7.1

Magnitude models 5.2.B, 6.2, 7.1

Joint probability-magnitude models 5.2.C, 6.5.B

Separate probability-magnitude models 5.2.D, 6.4

One-pocket models

Probability models 7.2, 6.5.A

Magnitude models 7.3, 6.2

Joint probability-magnitude models 7.5, 6.5.A

Separate probability-magnitude models 7.5, 6.4

Threshold on paterfamilias liability *see nexae deditio*

Minimization of administrative costs *see goals of the tort law system*

Minimum regulation 6.1.B

Monitoring

In vicarious liability 8.1

In labor contracts 9.1, 9.3

Multiple tortfeasors *see rule of contribution*

Negligence 3.7.A, 5.4.D, *see also negligence per se, fault-based vicarious liability, strict liability vs. negligence, strict liability with defense of contributory negligence, strict liability with defense of dual contributory negligence, simple negligence, contributory negligence, comparative negligence*

Negligence per se 6.1

No liability 2.6.A, 3.7.C

No-liability-based rules 3.3.B

Non-pecuniary losses 2.12.E

Non-residual bearer 3.3.B *see also residual bearer*

No-shirking constraint 9.2.A, 9.2.B, 9.2.C, 9.6.C

Normalized efficiency wages 9.2.F

Normalized conditional bonuses 9.2.G

Noxae deditio 8.1, 8.5

One-pocket models *see maximum upper threshold*

Optimal liability rules 3.3, 3.4

Overcompensation 2.7.A, 4.3.D, *see also punitive damages*

Overprecaution 7.2, 7.4, 7.5

Participation constraint 9.2.D, 9.6.C

Probability-reducing precaution 5.2.A

Product liability 2.12.A, 3.6.B

Property rights 2.3

Punitive damages 5.4.A

Pure economic losses 2.12.D

Residual bearer 3.3.B, 10.3.A, *see also non-residual bearer*

Risk allocation *see goals of the tort law system*

Roman law 3.5.A, 8.1, 8.5, 10.2.H

Rule of contribution 2.12.C, 10.3.C

Rules vs. standards 3.8

Sanction

- Positive sanctions 9.1
- Negative sanctions 9.1, 8.2

Selective regulation 6.1.C

Sharing rules 4, 10.2.C

Shirking behavior 9.1

Simple Negligence 2.6.C, 3.3, 3.4, 4.1, 5.4.D, 6.2.B, 6.3.B, 8.3

Stochastic models 5.3

Strict liability 2.6.A, 3.7.B

Strict liability vs. negligence 3.7, 10.2.B

Strict liability with defense of contributory negligence 2.6.E, 3.3

Strict liability with defense of dual contributory negligence 2.6.E

Strict vicarious liability 8.3, 10.3.F

Strict-liability-based rules 3.3.B

Suboptimal regulatory standard *see minimum regulation*

Tort law and economics 2

Turnover costs 9.7

Two-pocket models *see maximum upper threshold*

Ultra-hazardous activities 2.12.A, 3.7.B

Undercompensation 2.7.A, 5.4.C

Unilateral precaution accidents *see accident types*

Verifiability criterion 3.4.C

Vicarious liability 10.2.F

Employers' vicarious liability 8.1, 8.3

Parents' vicarious liability 8.1, 8.4

Paterfamilias vicarious liability 8.1, 8.5

Supervisors' vicarious liability 8.1, 8.4

Writ system 3.6.A

INTRODUCTION

In the last 40 years, law and economics has developed an impressive and innovative method for analyzing legal rules. This has been the result of the application of microeconomic theory to legal analysis, and has led to the appearance of specific law and economics commentaries for all areas of the law.

A traditional handbook on civil or common law contains separate chapters on property law, contract law, accident (tort) law, and so on. A modern handbook on law and economics is subdivided into economic analysis of property, economic analysis of contracts, economic analysis of torts, etc.

What surprised me at the beginning of my study in law and economics was the almost identical structure of the two. I asked myself the question whether law and economic analysis really ought to be parallel to legal analysis or it could take a proper and different path.

That question was the muse for this research.

1.1. Aim and scope

knowing what rhythm holds the humans

(Archilochos)¹

The question does not simply concern the typographical aspect of handbooks of law and economics. Putting new contents into a traditional container helps one understand similarities and differences between the traditional analysis and the innovative one, although the differences might be much weightier.

Legal analysis is traditionally a *de iure condito* analysis: it is an analysis of *the law as it is*. Jurists have to interpret and clarify the law and fill the gaps therein, mainly by analogy. Even when a new problem arises, which requires a new solution, legal analysis is performed upon the fiction that any solution already exists within the law and ought to be found there. Legal analysis translates the (abstract) law into (concrete) rules that fit reality. Inventing rules is a blasphemous attempt.

Law and economics is indeed blasphemous. It deals with rules, regardless of whether or not they are actually implemented in the existing body of the law. It poses questions concerning the social desirability of certain rules. From this simple approach, much more follows.

¹ Archilochos, Fr. 67 a D. See also the Loeb Classical Library series (1931), *Elegy and Iambus*, (trans. John

If the rule being analyzed is actually implemented within the legal system, the economic assessment of the effects of that rule can be used in order to criticize the law as it has been written by the legislature or as it has been interpreted by the judiciary. The conclusion might be either that the existent law is already the most advantageous law or that a change in the law is desirable. Similarly, the economic analysis of a legal rule can be used in order to criticize a certain (*de iure condito*) interpretation of the law. One interpretation might be more desirable than another.

If the rule being analyzed is one which cannot be found in the existing law, then law and economics goes further, and suggests new rules to the rule maker.

Law and economics is indeed a method of *de iure condendo* analysis: it analyzes the *law to be made*. Law and economics always questions the existing law and any hypothetical rule, both in an attempt to test it and in an attempt to shape it. The evidence for such a statement lies in the fact that law and economics also provides an analytical framework for the desirability of perfectly legal contract clauses, for (illegal) criminal behavior, and for international interactions.

Law and economics cannot, therefore, be a substitute for traditional legal analysis, as it is not a new form of legal analysis. It is indeed something different. It follows that the fact that law and economics analysis seem to proceed along the same path as legal analysis, while inquiring into the law, must be due to another reason. Coincidence, path dependence, marketing strategy or a lack of self-consciousness?

With the exception of the first, all these factors have had a role to play. It is often said that we are now in the third generation of law and economics. The first was the one of the founding fathers, and was dominated by lawyers with some understanding of economics, who pioneered the field by applying microeconomic reasoning to specific legal problems. They obviously started from the major legal problems. As lawyers, it was in their interest to analyze the same topics that other lawyers were discussing. They remained as close as possible to their colleagues, in order to enhance communication with them.

The second generation was characterized by economists employed in law faculties, who worked much more on formalization and mathematics but also remained close to the traditional legal classifications for a stronger reason: now the dialogue was not between scholars within the same subject but between economists and lawyers.

The third generation is the actual generation of young professors and researchers, who have studied both law and economics, and of lawyers and economists who are used to discussing with each other in a common language. They benefit from the widespread belief

that law and economics is indeed a useful tool. They are actually free from having to show this at any analysis and can concentrate their efforts in two directions: enlarging the domain of law and economics analysis, by applying it to new issues, and investigating the foundations of law and economics.

As far as the latter is concerned, the boundaries between different areas of law are finally collapsing², allowing the analysis to be made throughout them in search of general concepts. The first generation asked questions such as whether in tort law it is more desirable to have an individualized standard of negligence (such as the *bonus paterfamilias* criterion in Roman and civil law³), determined case by case, or a rule of negligence (such as driving below 50 km per hour), equal for everyone and specified ex ante. In the third generation, scholars inquire into the issue of whether the legal system in general ought to be dominated by standards (a general criterion to be specified case by case) or rules (a criterion fully specified in advance)⁴.

It seems as if we were on the other side of a metaphoric Bacon double ladder. Law and economics started with a bottom-up analysis of specific problems and general concepts just happened to be discovered (externality, incentive, risk allocation,...), which were then used for related areas. General concepts were then *induced* from their specific applications.

Once there are a sufficient number of general concepts, the opposite is also possible and law and economics can start descending from the top of the ladder and *deduce* practical applications thereby.

A number of theoretical papers have been written along these lines. The theory of anti-commons⁵, anti-insurance⁶ and givings vs. takings⁷ are examples of this tendency. Some of these general analyses have found specific and unexpected applications only after their theoretical relevance was recognized⁸.

The process of highlighting the fundamental concepts upon which law and economics analysis builds is still incomplete. My aim is both to denude this incompleteness and to contribute to the research.

² See on this point Cooter (1985).

³ Under the *bonus paterfamilias* criterion, the court controls whether a party took the level of precaution that the “good father of the family” would have taken in the same situation. The level of precaution is, hence, specified with respect to the characteristics of the situation under analysis, on the basis of a rather vague general criterion.

⁴ For a recent general approach to the study of the legal system see Kaplow (2000) and Wittman (2000).

⁵ Michelman (1982).

⁶ Cooter and Porat (2001).

⁷ Bell and Parchomovsky (2001).

⁸ This is the case of anti-commons. See Heller (1998) and Parisi, Depoorter and Schulz (2001).

1.2. Method

The scientist is not a person who gives the right answers, he is one who asks the right questions
(Claude Lévi-Strauss)⁹

My analysis focuses on incentives. I will investigate some fundamental characteristics that have been overlooked in the literature. Legal (and also contract) rules generate incentive streams in order to channel (in a sense, distort) the behavior of individuals, so to attain an efficient outcome.

The potential victims of accidents obviously want potential injurers to take precaution, and employers want employees to work. In order for the objective to be attained, resources ought to be deployed, and addressed to the task assigned. Just as a stream of water running downhill, an incentive stream might have different effects depending on the way it is directed towards the target.

A complete theory of how resources result in compliance is still lacking, also because incentive streams have always been analyzed in specific circumstances only, the effort spent on generalizing the findings being very limited.

My analysis is subdivided into different parts, each of which deals with a specific question. They delve into a field (especially torts) that was largely studied in the 1980s and early 1990s and is currently considered not to be a promising exercise for young scholars, as almost nothing is left to be discovered. I hope that I will be able to prove the opposite.

My first question is indeed a very basic one: what do care and activity level mean? Taken from the economic theory of accidents, the distinction between care and activity level pervades the whole economic analysis of law.

This question happens to be a disruptive one. The efficiency-equivalence theorem¹⁰ and the activity-level theorem¹¹, the two fundamental results in the economic analysis of tort law, are based upon the distinction between care and activity level. The choice between strict liability and negligence is also based thereon. Surprisingly, the distinction has always been taken as given and has never been even slightly challenged.

The question is equivalent to the one of how to use a single incentive stream (the duty to bear the accident loss) in order to induce two individuals to take precaution, and what are the limits thereof.

Once an incentive stream has been set, it ought to be correctly implemented by the judicial system, as ex post decisions provide ex ante effects. There are situations in which the

⁹ Claude Lévi-Strauss (1970), *The Raw and the Cooked*, London, Cape – (1964), *Le cru et le cuit*, Paris, Plon.

¹⁰ “Any negligence rule gives both parties efficient incentives with respect to care”, Landes and Posner (1980).

¹¹ “No negligence rule gives both parties efficient incentives with respect to activity level”, Shavell (1980).

court does not have enough information in order to decide the case appropriately. My second question is the following: how should the court share the loss between the parties when it does not have information concerning their ex-ante behavior, in order to maximize the effect of a single incentive stream in unclear cases?

Should the legal system direct the incentive stream fully toward either party or should it direct a portion of the incentive stream toward each of them?

Even when an appropriate incentive stream has been created and implemented, a problem can arise due to limits on the parties' liability, which tend to reduce the effect of incentives. In order to generalize the analysis as much as possible, I will make use of the concept of maximum upper threshold, referring to any limitation on the parties' exposure to liability, as deriving from the legal system itself or from a shortage of the parties' assets. It has always been believed that a maximum upper threshold progressively reduces the incentives to take precaution.

My third question is whether this is in fact true. As the traditional answer critically depends on the way in which incentives have always been modeled, i.e. on the specific formalization which has been given to the question of whether a maximum upper threshold undermines the incentive stream, I re-ask this question in a different formal way.

My fourth question is strictly related to the third. Once I have clarified the effect of a maximum upper threshold on incentive streams, I inquiry into which regulatory solutions can be adopted, in order to correct possible inefficiencies and on the relationship between ex post tort law incentives and ex ante incentives created by regulation.

My fifth question is what are the situations in which a maximum upper threshold can lead injurers to be more careful rather than less careful. As for the third question, I re-ask this question once again by using a different formal model from the one adopted in the literature, in order to unveil the logic and the conditions for such a counter-intuitive result to hold.

My sixth question applies the analysis of maximum upper thresholds to a totally different field: vicarious liability in Roman law. Roman *paterfamilias* were responsible for damage caused by members of the family, but not without limit; they had the possibility to release the tortfeasor to the victim's family instead of paying damages, a rule called *noxae deditio*. This mechanism is rather atypical and actually results in a limitation of liability. Nothing similar exists today in the case of employers' or parents' vicarious liability. I investigate the economic justification of *noxae deditio* and ask the related question of why today employers' vicarious liability is generally strict, while parents' vicarious liability is normally fault-based.

My seventh question explores those situations in which incentive streams are directly financed by victims, and asks what kind of device rational victims will use in order to create the most effective incentive stream at the minimum cost for them.

The answers to all the former questions should provide sufficient evidence that an inquiry into the general characteristics of incentive streams is far from complete and this is indeed a fruitful field of research.

All these questions are dealt with in a law and economics fashion: they are first translated into a formal model, then analyzed and discussed. The results of the formal analysis are finally exploited in two directions: their theoretical contribution to the subject and their policy implications.

1.3. An overview

This thesis is divided into ten chapters; the first (this chapter) and the last are an introduction and conclusions respectively, the remaining chapters are written in the form of separate articles and deal with specific problems.

The second chapter ushers the reader into the field of tort law and economics, and is meant to clarify the main results obtained during forty years of research by leading law and economics scholars.

*Chapter 2 – Tort law and economics*¹²

The accident loss cannot be removed once it has occurred; it can only be transferred from one party to another for the reason of creating incentives toward optimal precaution

(Pietro Trimarchi, *Rischio e responsabilità oggettiva*)¹³

Tort law has been one of the first fields of law to be analyzed from an economic point of view. After the early works by Calabresi and Trimarchi most scholars have devoted their effort to studying the economics of tort law. Noteworthy are the books by Calabresi, Shavell, Landes and Posner, Miceli and some well-known books on law and economics by Polinsky, Posner, Kaplow and Shavell, Cooter and Ulen.

In this chapter, I give a survey of the main results obtained over time and systematize them into a coherent framework that will be the basis for the following chapters. This is particularly suited a ground to begin with, for it has received the attention of the most prominent scholars for more than forty years. It is tacitly believed that most has been already discovered, and that some refinements only are possible.

This makes my analysis particularly meaningful, as if it is possible to go deeper into the economic analysis of accident law, then it might be rather profitable an attempt also in other areas of the law.

¹² Forthcoming in Hatzis, Aritides, *Economic Analysis of Law: A European Perspective*, Edward Elgar.

¹³ Trimarchi (1961), at 15-17, translation is mine.

The remaining seven chapters pose the seven questions described above. The problems are both analyzed by employing a formal model and discussed in the text. The formal models take inspiration from the standard theory, developed in the 1980s¹⁴. Nevertheless, the inability of the literature to answer the questions asked here mainly depends on the way in which the problems have always been modeled. Therefore, I try to provide a new formalization, staying as close as possible to tradition.

*Chapter 3 – On the definitions of care and activity level and the choice of liability rules*¹⁵

Stat rosa pristina nomine, nomina nuda tenemus

(Bernard of Morlay, *De contemptu mundi*)¹⁶

This chapter presents a model in which parties to an accident can take many different precautionary measures in order to reduce the expected accident loss. As controlling parties' precaution through the negligence inquiry triggers an administrative cost, it is efficient to limit the extent of the negligence criterion to some precautionary measures only.

This approach permits us to redefine the concepts of care and activity level in a non-tautological way: care is simply the set of precautionary measures included in the negligence criterion and activity level is the residual set of precautionary measures not included in the negligence criterion. The same framework explains the evolution of liability systems and the choice of liability rules in different areas of modern tort law and helps us to reinterpret the two basic theorems of tort law and economics: the efficiency-equivalence theorem and the activity-level theorem.

¹⁴ See Shavell (1987) and Landes and Posner (1987).

¹⁵ Presented at the final workshop of the Marie Curie Fellowship Program in Law and Economics at Hamburg University, July 2001, and at the CASLE workshop at Ghent University, February 2002.

¹⁶ As quoted in Eco, Umberto (1980), *Il nome della rosa*, Milano, Bompiani – (1983), *The Name of the Rose*, San Diego, Harcourt.

*Chapter 4 – Comparative negligence as a filter mechanism when courts cannot observe precaution costs*¹⁷

*Utque ferant aequos et caelum et terram calores,
nec preme nec summum molire per aethera currum.
Altius egressus caelestia tecta cremabis,
inferius terras: medio tutissimus ibis.*

(Ovidii *Metamorphoseon* II, 134-137)¹⁸

This chapter presents the filtering properties of a 50/50 sharing rule. In unilateral precaution accidents, if the court cannot determine which party could have avoided the accident, all-or-nothing rules completely direct the incentive stream towards one party, and randomly prevent half the accidents only. On the contrary, a 50/50 sharing rule directs a weaker incentive stream on both parties but works as a filter that prevents the most harmful accidents and allows the least harmful ones to occur.

A 50/50 sharing rule is likely to generate a smaller social loss than all-or-nothing rules (or other sharing rules); I discuss the conditions under which this result holds. The analysis also considers alternative and joint precaution cases and the possibility that the court over- or under-compensate the victim's harm.

When it comes to discussing limits of the parties' exposure to liability, the standard model assumes that they can reduce the probability of accidents occurring, but not the magnitude of the harm, which remains exogenous. I consider those frequent cases in which the magnitude of the harm can also be reduced, as for example in car accidents. My results are discussed in chapters 5, 6 and 7.

*Chapter 5 – An analysis of the judgment-proof problem under different tort models*¹⁹

Discovery consists of looking at the same thing as anyone else and thinking something different

(Albert Szent-Gyorgyi)²⁰

The judgment-proof problem arises from a threshold imposed by nature (limited assets) or by a rule maker (liability caps) on the maximum amount of damages that the injurer has to pay in the case of an accident. This chapter analyzes the effects of the judgment-proof problem on

¹⁷ Presented at the 17th annual conference of the European Association of Law and Economics, Ghent, September 14th-16th, 2000 and at the annual meeting of the Italian Economists' Society (SIE), Rome, October 25th-27th, 2001.

¹⁸ "And, that the sky and earth may have equal heat, go not too low, nor yet direct thy course along the top of heaven; for if thou goest too high thou wilt burn up the skies, if too low the earth. In the middle is the safest path", see the Loeb Classical Library series (1956), *Ovid, Metamorphoses* (trans. Frank Justus Miller), London, Heinemann.

¹⁹ Presented at Copenhagen Business School, May 20th, 2001 and at the 18th annual conference of the European Association of Law and Economics, Vienna, September, 13th-15th 2001.

²⁰ (1893-1986) 1937 Nobel laureate in medicine.

the incentive for the injurer to take precaution. It attempts to draw a more generalized picture than previous contributions in the literature, which tended to focus on specific solutions under specific models.

It is demonstrated that the effects of the judgment-proof problem depend on the type of accident or externality. Four tort models are analyzed: the probability model (where more precaution on the part of the injurer only reduces the probability of an accident), the magnitude model (where more precaution only reduces the magnitude of the harm), the joint probability-magnitude model (where more precaution at the same time reduces the probability of the accident and the magnitude of the harm) and the separate probability-magnitude model (where one precautionary measure reduces the probability of an accident and another precautionary measure the magnitude of the harm). In addition, a stochastic variant for each of the four models is discussed.

Different legal solutions to the problem are analyzed: punitive damages, average compensation of harm, under-compensatory damages, and negligence. It is shown that in general subsidizing potentially judgment-proof injurers yields a more desirable outcome than increasing the damages that they have to pay.

Chapter 6 – On the combined use of anti-judgment-proof regulation and tort law

He who does not punish evil commands it to be done

(Leonardo da Vinci, *Scritti letterari*)²¹

Although tort law and regulations are two alternative means of controlling externalities, tort liability is often combined with minimum regulatory standards; this chapter analyzes these instances and builds on the four-model framework of the previous chapter. The reason for their joint use is that the incentives created by tort law might be diluted by the judgment-proof problem, while regulators might lack appropriate information on risk. Thus, they may be used to complement each other. We demonstrate that minimum regulatory standards, which are insufficient to give injurers an incentive to take optimal precaution, might be sufficient to completely remove the effect of judgment-proofness. Once the judgment-proof problem has been corrected, tort law can provide incentives toward optimal precaution, and hence remedy the information deficit of regulatory bodies.

²¹ See Pedretti, Carlo (1977), *The Literary Works of Leonardo da Vinci*, London, Phaidon.

*Chapter 7 – When will judgment-proof injurers take too much precaution?*²²

Noli esse nimis iustus neque sapiens supra modum! Cur te perdere vis?

*(Liber Ecclesiastes 7, 16)*²³

An injurer is said to be judgment proof when his total assets are less than the harm he may cause. In principle, judgment-proof injurers can be expected to take less than optimal precaution, as they only bear a part of the accident losses (as formally proven by Shavell, 1986, using a probability model, i.e. a model where more precaution results in a lower accident probability). Beard (1990), however, demonstrated that under certain conditions the judgment proof problem can lead to overprecaution. Beard used a (stochastic) one-pocket probability model (where it was assumed that the more the injurer spends on precaution, the less assets will remain available to pay damages).

While probability models can be appropriate for analyzing some accident types (such as aircraft accidents), magnitude models (where more precaution reduces the magnitude of the loss and not the probability of an accident) are more appropriate for analyzing other externality problems (such as nuisance, many types of environmental pollution or safety measures). In this chapter it is argued that overprecaution will never occur in a magnitude model. In addition to that, I consider a non-stochastic probability model and show that overprecaution is possible. Mixed models are also analyzed and some policy implications are discussed.

Chapter 8 - Noxae deditio and negligence: Roman and modern solutions to inefficient monitoring under vicarious liability

Patres, nolite ad indignationem provocare filios vestros, ut non pusillo animo fiant

*(Paulus, ad Colossenses 3, 21)*²⁴

Why is vicarious liability usually strict for employers and duty-based for parents and supervisors? Why did Roman law limit *paterfamilias*' vicarious liability to the monetary value of the tortfeasor by means of *noxae deditio*? The answer to these two questions is the same: parents, supervisors and *paterfamilias* do not fully bear the cost of their agent's precaution, while employers do.

Since employers pay employees for their efforts, they internalize the employees' cost of precaution and, while deciding on the level of precaution to enforce, they take an optimal decision.

²² To be presented at Hamburg University, May 23rd, 2002.

²³ Ecclesiastes 7, 16 :“Do not be overrighteous, neither be overwise! Why destroy yourself?”

²⁴ Paul, *Colossians* 3, 21: “Fathers, don't provoke your children, so that they won't be discouraged”.

On the contrary, parents, supervisors and *patresfamilias* do not pay their agents (children, persons under supervision, slaves and members of the family respectively) for the precaution they take. Therefore, they might have an incentive to require too much precaution.

I show that modern law often solves this problem by using fault-based instead of strict vicarious liability for parents and supervisors, while Roman law retained an old rule, *noxae deditio*, which resulted in a maximum upper threshold on the *paterfamilias*' liability and had the effect of correcting the incentives to take precaution.

*Chapter 9 – Efficiency wages, conditional bonuses, and punitive damages: monitoring levels when sanctions are financed by victims*²⁵

Si quis non vult operari, nec manducet

(Paulus, *II ad Thessalonicenses* 3, 10)²⁶

This chapter compares three different incentive devices in a contractual setting: efficiency wages, conditional bonuses and punitive damages. Efficiency wages are a form of positive sanction consisting of a constant overpayment of performances. According to the efficiency wage theory by Shapiro and Stiglitz, it may be rational for an employer to pay his employees above market-clearing wages in order to render the sanction of dismissal more severe and save on monitoring costs.

On the contrary, conditional bonuses (another form of positive sanction) are only paid if the employee is monitored and found non-shirking. Punitive damages (a form of negative sanction) are only paid by shirking employees.

In this chapter, it is argued that only under very implausible assumptions do efficiency wages make sense. Rational parties will nearly always prefer negative sanctions to positive sanctions. In the few cases where positive sanctions will be preferred, the parties will choose conditional bonuses rather than efficiency wages.

It will be shown that efficiency wages lead to higher than optimal monitoring levels while monitoring levels are optimal under punitive damages and conditional bonuses. The reason is that efficiency wages are a wasteful positive sanction that is also paid to shirking employees: in order to limit such a waste the employer over-invests in monitoring.

1.4. References

Bell, Abraham and Parchomovsky, Gideon (2001), 'Givings', *John M. Olin Program in Law, Economics and Business, Discussion Paper No. 320*, Harvard Law School.

²⁵ Presented at the annual meeting of the American Law and Economics Association, Washington, May, 11th-12th 2001.

²⁶ Paul, *2 Thessalonians* 3, 10: "If anyone will not work, neither let him eat".

- Cooter, Robert D. (1985), 'Unity in Torts, Contracts and Property: The Model of Precaution', **73** *California Law Review*, 1-51.
- Cooter, Robert and Porat, Ariel (2001), 'Anti-Insurance', *UC Berkeley Public Law Research Paper No. 71*.
- Heller, Michael A. (1998), 'The Tragedy of Anticommons: Property in the Transition from Marx to Market', **111** *Harvard Law Review*, 621.
- Kaplow, Louis (2000), 'General Characteristics of Legal Rules', in Boudewijn Bouckaert and Gerrit De Geest (eds.), *Encyclopedia of Law and Economics*, vol. V, Cheltenham, Edward Elgar.
- Landes, William M. and Posner, Richard A. (1980), 'Multiple Tortfeasors: An Economic Analysis', **9** *Journal of Legal Studies*, 517-555.
- Michelman, Frank I. (1982), 'Ethics, Economics, and the Law of Property', in Pennock, J. Roland and Chapman, John W. (eds.), *Nomos XXIV: Ethics, Economics and the Law*, New York, New York University Press, 3-40.
- Parisi, Francesco, Depoorter, Ben and Schulz, Norbert (2001), 'Duality in Property: Commons and Anticommons', *George Mason University School of Law, Law and Economics Research Paper No 00-32*.
- Shavell, Steven (1980), 'Strict Liability versus Negligence', **9** *Journal of Legal Studies*, 1-25.
- Shavell, Steven (1987), *Economic Analysis of Accident Law*, Cambridge (MA), Harvard University Press.
- Trimarchi, Pietro (1961), *Rischio e responsabilità oggettiva* (Risk and Strict Liability), Milano, Giuffrè.
- Wittman, Donald (2000), 'General Structure of the Law', in Boudewijn Bouckaert and Gerrit De Geest (eds.), *Encyclopedia of Law and Economics*, vol. I, Cheltenham, Edward Elgar.

CHAPTER TWO
TORT LAW AND ECONOMICS*

*Giuseppe Dari Mattiacci***
Utrecht University

ABSTRACT

Tort law has been one of the first fields of law to be analysed from an economic point of view. After the early works by Calabresi (1961) and Trimarchi (1959a, 1959b, 1961)¹ most scholars have devoted their effort to studying the economics of tort law. Noteworthy are the books by Calabresi (1970), Shavell (1987), Landes and Posner (1987), Miceli (1997) and some well-known books on Law and Economics by Polinsky (1989), Posner (1998), Kaplow and Shavell (2002), Cooter and Ulen (2000)². In this chapter I will give a survey of the main results obtained over time.

JEL Classification: K13.

Keywords: tort, damage, compensation, liability, fault, negligence, precaution, care, activity level, risk.

* Forthcoming in Hatzis Aritides ed., *Economic Analysis of Law: A European Perspective*, Edward Elgar.

** Giuseppe Dari Mattiacci, Institute of Economics/CIAV, Utrecht University, Kromme Nieuwegracht 22, 3512 HH Utrecht, the Netherlands. Email: g.darimattiacci@econ.uu.nl. I wish to thank for their precious comments Alexandra Ayala, Kennet McArthur, Gianluigi Galeotti, Gerrit de Geest, Mark Grady, Margaret Mary Hammersley, Katja Modric, Hans Bernd Schäfer, Pietro Trimarchi, Ann-Sophie vandenBerghe, Wouter Wils. Updated versions of this chapter will possibly be available on my web page: www.law.uu.nl/eco/gdm.htm.

¹ On Trimarchi see Frezza and Parisi (1998).

² For a complete bibliography also on non-English books see Bouckaert and De Geest (2000), vol. I.

2.1. Introduction.

Ruling on torts is one of the central issues of any legal system, starting from general statements about the conditions under which a victim of a harmful event is entitled to obtain compensation.

Arising from the Roman *lex Aquilia*³, different legal systems have been developed in Europe with respect to torts. English law has followed the Roman approach in retaining specific heads of tortious liability for circumscribed torts, while French law encompasses one general clause applicable to all torts. German tradition steers somewhat in between the English conception of torts and the French unified vision of tort; the German civil code does not include a general clause for torts but lists only three major heads of tort liability.

The judge-made-law principle has spurred a certain degree of homogeneity amongst countries within the common-law tradition, whilst codification has created a richer variety of solutions. Nevertheless, a comparison of the different rules enforced in civil-law countries unveils a predominant conformation to either French or German tradition or a mix of the two. This can be verified also for countries of socialist tradition.

More than differences in legal techniques, we want to underline similarities, in order to provide a common ground for the application of an economic analysis of tort law. The results obtained can be used at a further stage to test the efficiency of specific national solutions and even of the original Roman formants.

The civil codes consider tort(s) as a source of obligation, a broad category of *iuris vincula quibus necessitate adstringimur alicuius solvendae rei*, which encompasses also contracts. If the legal requirements are fulfilled, the injurer (or even a different subject singled out by the law) ought to compensate the victim. The measure of the compensation depends on the magnitude of the harm, which is the first element of tort liability. A second common element is fault. Together with an increasing number of cases of no-fault liability, the no-liability-without-fault principle is present in any modern legal tradition, often based on general and even vague standards (*bonus pater familias*' behaviour, *bonos mores*), but sometimes laying on more precisely defined requirements. Moreover, an analysis of causation is generally required, only the harm caused by the injurer leads to compensation. The legal definition of causation is one of the most controversial issues in tort-law theory; we will suggest an economic approach to that.

³ Probably datable around 287 BC, the *lex Aquilia* provided the cornerstone for the modern tort liability systems. For a view on the continuity of legal ideas on civil liability over centuries see Parisi (1994).

Before being ushered into a three-headed economic model for tort law, some preliminary questions should be addressed. First we have to clarify the ends tort law serves, which means giving an economic foundation for the legal principle of compensation. We will show that the main goal of tort law is to give parties appropriate incentives to internalise externalities, insofar as tort law should allocate the risk to the one who can translate it into cost (Trimarchi, 1959a and 1967), both with respect to precaution to be taken (section 3) and information to be acquired (section 9). When tort law alone is proved to be insufficient, efficient incentives toward optimal precaution are to be provided by other means (section 10). We will discuss also the goals of optimal risk sharing and minimising administrative cost in sections 11 and 12.

A second important issue is the comparison of tort law with other means for internalising externalities, namely property rights and inalienability (section 4). The last part of the chapter is devoted to the analysis of specific issues (section 13).

2.2. Goals of a tort law system: incentives towards optimal precaution, the basic model.

A first intuitive end of tort law is to compensate the victims for losses due to accidents. This is indeed an important issue but is not the central one. It has been shown that tort law is a very expensive means of compensating harms, because it involves high cost due to the functioning of the judicial system; insurance, to the contrary, is much cheaper and much quicker a system (Shavell, 1987, at 263), so that, if the only goal is to compensate victims, a victims' first party insurance is to be preferred over tort liability⁴. Moreover, the cost of insurance can be paid by the potential injurers, shared among potential victims or financed by taxpayers, in order to redistribute the costs (McEwin, 2000).

On the contrary, the main reason of the introduction of tort law is the need to give the injurer appropriate incentives to avoid the accident, i.e. internalise the externalities. To make this point clearer consider two individuals involved in an accident. Iris is the injurer, the one who does not suffer harm and is carrying out an activity. Victoria is the victim. She suffers harm because of Iris' activity and she cannot do anything to avoid it, while Iris can act carefully in order to reduce the probability of an accident happening.

This is a very general situation. A bus driver is a potential injurer and his passengers are potential victims; the latter can drive at a lower speed in order to reduce the probability of an accident. A physician is a potential injurer and his patient a potential victim; the physician can take more precaution in choosing the therapy for the patient in order to limit the likelihood of errors. The owner of a chemical plant is a potential injurer and a farmer living nearby is a

potential victim; the chemical plant can be checked more frequently in order to lower the probability of accidental pollution.

The first aspect to be noticed is that in all those situations the injurer decides on the level of precaution to take (speed, accuracy, controlling) and bears the costs of such a decision. The victim bears the benefits (reduction in the probability of an accident⁵). The benefits (of precaution) are external with respect to the decision (on how much precaution). If the same person bears cost and benefit of his own decision, then he will decide optimally. On the contrary, when either costs or benefits are external, the decision will not take into account all the costs and benefits involved and, hence, will not be optimal.

The main goal of tort law is to internalise the externalities in order to enhance optimal decisions on the level of precaution. This statement is to be formalised. Our model includes two utility-maximising, risk-neutral and perfectly informed⁶ agents: the injurer and the victim. The injurer acts and decides how much precaution to take. The victim is passive⁷ and, depending on the injurer's decision, faces the probability of having an accident and, hence, suffering harm denoted as H . Figure 1 displays this situation. On the horizontal axis is measured the level of precaution, x^8 , and on the vertical axis the costs associated with each level of precaution.

The line denoted as $c(x)^9$ shows the cost for the injurer. The more precaution he takes, the more he spends. Lowering the speed costs time to a car driver, checking up a patient costs time to a physician, increasing the number of controls in a chemical plant is costly for the owner.

The curve denoted as $p(x)H$ shows the cost for the victim of an accident. Of course, after an accident has happened, the victim suffers the harm H , but we look at the accident ex ante, and denote the ex ante measure of the cost of the accident as 'expected accident cost'. The expected accident cost is given by the probability of an accident occurring, $p(x)$, times the magnitude of the harm, H^{10} . If the injurer does not take enough precaution a harm H will take place with probability $p(x)$. The more precaution the injurer takes, the less probable is the

⁴ This solution has been adopted in New Zealand in 1974.

⁵ See note 10 infra.

⁶ The assumptions of risk-neutrality and perfect information will be implicitly relaxed in the following sections.

⁷ This does not need to be always the case; in section 5, we will account for different accident types.

⁸ $x \in [0, \infty)$.

⁹ We will make several assumptions on the functional form, which are standard in the literature on tort law and economics, see Shavell (1987) and Landes and Posner (1987). $c(x) \in [0, \infty)$, $c'(x) > 0$, $c''(x) \geq 0$.

¹⁰ $p(x) \in [0, 1]$, $p'(x) < 0$, $p''(x) > 0$. For simplicity, the harm, H , is supposed to be exogenous. The analysis does not change if we relax this assumption, but gets more complex. So we simplistically assume that the injurer can reduce the probability of an accident by taking more precaution, but cannot do anything about the magnitude of the harm.

accident, as shown in figure 1. So, the benefit of the injurer's decision about the level of precaution consists of a reduction in the expected accident cost, and is gained by the victim¹¹.

What is the best outcome for the injurer? His cost increases as the level of precaution increases. He will spend the smallest amount of money if he takes no precaution at all. This will result in a highly probable accident for the victim¹².

What is the best outcome for the victim? The victim would prefer the injurer to take as much precaution as possible in order to lower the probability of having an accident nearly to zero. This will result in an extremely high cost of precaution for the injurer¹³.

Which is the most efficient solution: no precaution, full precaution or something in between?

To answer this question, imagine that a physician is choosing a therapy for herself or that the owner of a chemical plant is also the owner of the farm nearby. Which level of precaution will he choose? In these cases the same individual bears both the cost of taking precautions and the benefit of taking precaution (a reduction in the expected accident cost). To each level of precaution, x , corresponds a certain cost of precaution and a certain expected accident cost, by adding them together we obtain the total cost that the physician or the owner of the chemical plant have to bear. The best decision is to choose the level of precaution x^* , which corresponds to the lowest total cost¹⁴. We can conclude that if an individual has to bear the total cost, he will choose the level of precaution x^* , which is indeed something in between the two extremes, no precaution at all and full precaution.

Since there is not any other cost involved in this simple model, the sum of the cost of precaution and of the expected accident cost is also equal to the social cost of an accident, that is what society spends in total for an accident; therefore, x^* is the socially optimal level of precaution, because it minimises the sum of the cost of precaution and the expected accident cost and, hence, the social cost. This is the most efficient solution to the problem of how much precaution the injurer should take.

Therefore, the main goal of the tort-law system giving the injurer appropriate incentives to avoid the accident, i.e. internalise the externalities, can be specified as follows: the tort-law system should give the injurer incentives to take the x^* level of precaution.

In addition the tort law systems gives parties incentives to acquire information about the accident. With respect to risk, the tort law system should enhance an optimal allocation of the risk between victim and injurer, but this goal can be reached via insurance. With respect to

¹¹ Parties maximise their own utility.

¹² In figure 1, no precaution is the point where $x=0$, and corresponds to $c(x)=0$ and $p(x)H \leq H$, since $p(0) \leq 1$.

¹³ The maximal precaution is given by $x=\infty$, and corresponds to $c(x)=\infty$ and $p(x)H=0$.

¹⁴ The optimal level of precaution x^* is the level of precaution which solves $\min_x [p(x)H + c(x)]$; the minimum is

transaction costs, the goal of the tort law system is to minimize the administrative cost associated with the functioning of the system itself (mainly the costs of courts and lawyers and the indirect costs borne by litigants). The first formulation of the ends of liability in those terms is due to Calabresi (1970), while the formalisation of an economic model of accidents goes back to Brown (1973). We will focus on incentives toward optimal precaution and discuss the other aspects aside.

2.3. Why do we have tort law? Between property rights and inalienability.

The Coase theorem (Coase, 1960) shows that if parties are allowed to negotiate, no matter what is the rule in force, they will reach an efficient allocation of resources. Applied to our example the Coasian result would be the injurer's choosing the optimal level of precaution, x^* , independently from the liability rule in force.

However, the Coase theorem applies only if there are no transaction costs. In the case of accidents, transaction costs are intuitively very high. First of all, the parties potentially involved are more than two, not all of them identifiable exactly *ex ante*. A car driver should negotiate his speed with all the pedestrians he might spot during his trips; a physician should negotiate his accuracy with all his patients. Moreover, the cost of acquiring the relevant information for the bargaining can be high. Therefore, giving a sort of 'right to create risk' to potential injurers or a 'right to be free from risk' to the potential victims and then leaving room to bargain does not guarantee an efficient outcome. Because the parties cannot negotiate, the outcome would be no precaution in the former case and too high a level of precaution¹⁵ in the latter.

While property rights are protected against unwanted takings *ex ante* (Dworkin, 1975), tort law allows potential injurer to create risk, but obliges them to internalise that cost by imposing compensation for victims *ex post* (Krauss, 2000, see also Kaplow and Shavell, 1996b).

If negotiation fails, a completely different solution is to allocate the rights and forbid negotiation. This is the case for many situations in which bargaining would be likely to produce undesirable effects (Rose-Ackerman, 1985, Radin, 1987). Many transactions on constitutional rights are forbidden by law (inalienability), because a rational party would never do it. Sometimes the law allows donation but not selling. This has, however, a cost. If the (dangerous) activity is likely to produce a benefit bigger than the risk of accidents, the cost of banning is given by the loss of such a benefit.

given by $p'(x)H+c'(x)=0$, since that is a strictly concave function, because $p''(x)>0$ and $c''(x)\geq 0$.

¹⁵ In section 7 it will be shown that stopping the activity is an extreme form of precaution.

There are indeed many situations in which both inalienability and property rights are unsuitable solutions. In those cases, the ideal would be to give the injurer appropriate incentives to act ‘as if’ he had to bear the total social cost. This is the goal of tort law, an intermediate solution between the two extremes. Tort law is therefore justified when bargaining is not possible because of too high transaction costs and banning is too severe a sanction to be imposed on profitable activity (Calabresi and Melamed, 1972).

Inalienability mainly consists of unenforceability of transactions regarding certain rights, and is indeed a form of regulation (Ogus, 1994, at 258). Other forms of regulation are often paired with a tort liability system or are a substitute for it, when it does not produce efficient results (Shavell, 1984 and 1987). The topic is interesting but exceeds the scope of this chapter.

2.4. Accident types.

Before going into the analysis of legal rules, we have to give a precise classification of accident types. In the simple model we have described, we defined the injurer as that individual who does not suffer harm in an accident and the victim as that individual who does suffer harm; we will maintain this simple assumption of only two-party accidents for simplicity. In the same model we considered the situation in which only the injurer can take precautions in order to avoid the accident. This is just one of several different cases.

In a two-party accident there are two fundamental possibilities. The first is when both parties have to take precaution in order to avoid the accident (bilateral precaution). The second is given by situations in which either party has to take precaution (alternative precaution). In this case if both parties take precaution there is a waste of precaution cost, since one party’s precaution would have been already enough. A particular and common case of alternative precaution is unilateral precaution. As in alternative precaution one party’s precaution is enough to prevent the accident¹⁶; however only one party has the actual possibility of avoiding the accident¹⁷. The unilateral precaution case is the simplest; therefore, we will start from this case while describing specific legal rules.

It is important to underline that we defined as injurer the party who does not suffer harm, and as victim the other party. This has nothing to do with the question about which party has the possibility of avoiding the accident; we will denote that party as avoider. In unilateral-

¹⁶ Since we are using continuous variables and since the optimum is not a complete reduction of the probability of accidents, this statement has to be intended in the sense that the marginal contribution of the other party’s precaution to the reduction of the probability of accidents is zero.

¹⁷ Impossibility to take precaution has to be intended in economic terms: the precaution cost is always higher than the expected accident cost.

precaution cases there is only one avoider; in alternative precaution cases both parties are avoiders, but one of them will be the least-cost avoider: the one who could avoid the accident at the least cost; in bilateral-precaution cases there are two avoiders as well, both necessary in order to prevent the accident from occurring. The case analysed in the basic model above is a unilateral-precaution accident where the injurer is the avoider.

2.5. Elements of tort liability.

We have seen that efficiency means the minimisation of the social cost and that this is given by a level of precaution we denoted as x^* : The tort law system should provide parties with the right incentives to choose such a level of precaution. Different legal rules will be analysed and checked whether they are efficient or not.

The present section will be organised as follows. We will discuss the three elements of tort liability in this order: harm, cause and fault. In each subsection we will analyse whether the legal rules give incentives to parties toward an efficient precaution. In the subsection dealing with fault the two fundamental theorems of tort law and economics will be presented: the efficiency-equivalence theorem and the activity-level theorem.

2.5.A. Harm

The discussion about harm encompasses two different aspects. On the one hand, the harm is a condition for the duty of paying damages; on the other hand, the harm is a measure for the magnitude of the damages to be paid. In other words we can set two basic questions: Does the injurer have to pay damages to the victim? And how much?

With respect to the first question the harm is a necessary condition for the victim to be entitled to compensation. If there is not any harm there is not any liability. In economic terms harm is whatever decreases the wealth of the victim: someone damages his car, wounds his nose, kills his most loved pet. However, not every harm leads to tort liability. Firms in a competitive market can daily harm each other by attracting consumers but they are not brought to court; parents can punish their children for not studying enough. Legal doctrines sometimes define the harm to be compensated as unjust, or contrary to the law; economics looks at the harm from a social point of view, harm has to be compensated only if it consists of a social loss.

Competition law provides a good example. Competitors individually harm each other, one's better product causes a decrease in the others' profits, but competition leads to a social gain due (synthetically) to better quality and lower prices, so that this reciprocal harm is not also a social harm: the total wealth increases because of competition. On the contrary, certain

ways of competing are forbidden by the law; for instance, a big firm might be sanctioned for fixing the price at such a low level that the other small firms are driven out of the market. From an economic point of view such behaviour reduces competition by limiting the number of firms on the market and, hence, it is to be considered a social harm.

We now turn towards the second question: how much does the injurer have to pay? The measure of the damages paid by the injurer is to be considered in two respects: it influences the incentives for the injurer to take optimal care and it determines who bears the risk of an accident (Arlen, 2000). For both concerns damages exactly equal to the harm are defined as perfectly compensatory.

In the simple model described perfect compensation makes the injurer internalise completely the expected accidents cost and, therefore, act efficiently. In the course of the discussion we will analyse cases in which damages are over compensatory (punitive damages).

From the victim's side, since perfect compensation means him being entitled to a compensation exactly equal to the harm, he is indifferent between not having an accident and having an accident but being compensated. The risk of the accident is borne by the injurer.

Perfect compensation is easily defined for market goods. However, not every accident causes easily measurable harm. If the goods do not have a market value, it is not possible to apply a direct correspondence, since the compensation is always monetary. The criterion here is that the victim should receive an amount of money, which compensates him as if the accident had never occurred, so that the victim's wealth before the accident is exactly equal to his wealth after it. This is an ex post criterion, and it is indeed difficult to apply in the case of death and serious physical injury.

A different solution is to evaluate the harm on the base of what the victim would have been willing to pay ex ante to be free of the risk of accident.

Both criteria give an evaluation of what the victim should receive to be fully compensated. On the other hand, the magnitude of damages awarded influences the size of the incentives given to the injurer. In this respect, perfectly compensatory damages make the injurer perfectly internalise the cost of the accident. However, we will see while discussing the effect of errors that also under compensatory damages can give injurers efficient incentives under a negligence rule.

2.5.B. Cause

The second element of tort liability is the cause: the harm has to be connected by the means of a causal chain to an act of the injurer¹⁸.

Early law and economic analysis rejected the requirement of causation as an autonomous element of liability following Coase's view, describing the mutual nature of harm, in which both the victim and the injurer are necessary (Ben-Shahar, 2000). In causal terms, the presence of the victim is equally necessary as the injurer's wrongful conduct for the accident to take place, so that both are to be considered as the cause of the accident. In this view the only determinant of civil liability is inefficient behaviour (Coase, 1960, Brown, 1973).

Under these premises, an economic description of causation has been successfully attempted (Landes and Posner, 1983). An analysis on causation restricts the cases in which liability is applied, defined as scope of liability. Liability is not applicable when the injurer's behaviour is found not to have raised the probability of an accident, in the sense that, all other things being equal, had the injurer acted differently, the accident would have happened with the same probability. It follows that the application of the tort-law incentive to the injurer does not produce a reduction in the expected accident cost.

The concept of causation here applied is the *prospective (ex ante) cause*: before an accident occurs we address the question whether the injurer's conduct raises the probability of the accident. This concept is different and narrower than the *retrospective (ex post) cause*: after the accident has occurred, we address the question whether the accident would have occurred had the injurer behaved differently. The latter is a simple *but-for* test: the injurer is considered to have caused the accidents, which would not have occurred without his action. In the former we consider not only the fact that an accident has happened, but also the probability that a similar accident would have happened under different circumstances (Shavell, 1980b).

The restriction of the scope of liability to the accident actually caused by the injurer serves two goals. First, it reduces the administrative cost of the tort law system, by limiting the number of cases in which liability is applicable. Second, it excludes from the application of liability all the cases in which the injurers' precaution would not have affected the likelihood of an accident, and, hence, would have been a waste of resources. By avoiding injurers' over precaution, the restriction of the scope of liability enhances efficiency in the tort law system. An over-restriction, however, would have the opposite effect of spurring injurers' under precaution (Shavell, 1980b).

¹⁸ For a comparative approach see Honoré (1971).

In the following we will examine the effect of the requirement of causation in the case of negligence rules, with specific reference to the interaction of the requirements of cause and fault.

2.5.C. Fault

The third element of tort liability is fault. Here I will give just a definition of such a requirement, leaving the description of fault-based rules to the next section. In the basic model we have built up, one party chooses the level of precaution to be taken, so that the decision about how much precaution to take is completely left to him. By introducing the requirement of fault, the legal system intervenes in such a decision, by setting a due level of precaution, that we denote as x_d ; for instance, a speed limit. If the party takes less than the due level of precaution he is considered to be negligent and tort liability applies, if the party takes precautions equal to or more than the due level, he is considered to be careful and tort liability does not apply.

While the requirements of harm and cause are always present in the case of tort liability, the analysis of fault is not required for every rule. We will see why and in what cases we need to introduce such a requirement and how it reacts with differences in accident types.

2.6. Liability rules.

2.6.A. Strict liability and no liability

The Roman *lex Aquilia* was a no-fault liability rule: in the presence of harm (and, implicitly, cause) the injurer has to pay damages to the victim (strict liability). Modern legal systems are more fault oriented and strict liability is seen as an exception to the norm; moreover no legal tradition defines another specific liability rule, which is a mirror image of strict liability: no liability. No liability is not a rule in the legal sense, because it is not a source of obligation for the injurer to pay damages, but it is indeed a tort law rule in an economic sense, because, as strict liability does, it allocates the loss due to an accident (Landes and Posner, 1987, at 62). While strict liability allocates the losses to the injurer by entitling the victim to compensation, no liability allocates the losses to the victim, by denying the right to compensation.

No liability and strict liability are the simplest rules, they do not encompass the requirement of fault and on them it is possible to build up a consistent description of the existent rules about tort. This procedure may look far from the legal reality, but enhances a

better systematisation of the different legal solutions into an economic framework. Those simple rules will be tested against different accident types, already described.

In the case of unilateral precaution, the avoider can be either the injurer or the victim. If the avoider is the injurer, strict liability, by allocating to the injurer the loss due to the accident, achieves an efficient result. Indeed, the injurer has to bear both the cost of precaution and the expected accident cost and, hence, he will minimise the sum of those costs; we have already seen that the level of precaution which minimises the social cost (the sum of precaution cost and expected accident cost) is the efficient level of precaution (see figure 1). On the contrary, no liability does not achieve an efficient result because it makes the victim bear the expected accident cost, while the injurer (the avoider, in this case) bears the cost of taking precaution. If the injurer does not have to pay damages, his choice will be no precaution at all, which is an inefficient result.

On the other hand, if the avoider is the victim the rule of strict liability is inefficient while the rule of no liability is efficient. In this case the victim has the possibility of avoiding the accident and bears the cost of taking precaution, while the injurer cannot do anything at all. Strict liability results in the victim being perfectly compensated for the accident by the injurer, whatever level of precaution he takes, so that he has an incentive not to take precaution in order to minimise the precaution cost. On the contrary, no liability makes the victim bear also the expected accident cost, so that he will choose the level of precaution, which minimises the sum of expected accident cost and precaution cost, and that is efficient.

It is clear that strict liability and no liability can give incentives to take efficient precaution only to one party, respectively either the injurer or the victim. Strict liability will fail to produce an efficient outcome when the avoider is the victim, and no liability will fail when the avoider is the injurer.

With respect to alternative precaution, the result is slightly different. While in the case of unilateral precaution, if the tort law system fails to target the avoider, he will take no precaution at all, in the case of alternative care either party can take precaution, therefore, hitting the party which is not the least cost avoider will result in a less than efficient precaution level and a too high cost of precaution, as shown in figure 2.

In the case of bilateral precaution both strict liability and no liability always fail, because neither rule can simultaneously give both parties incentives to take precaution.

Strict liability and no liability can give efficient results only in the case of unilateral or alternative precaution, provided that the incentive stream be directed toward the (least cost) avoider; in bilateral-precaution situations a different rule is needed, capable of creating an incentive stream for both parties at the same time.

2.6.B. Precaution under negligence rules: care and activity level

In order to enhance efficiency also in the case of bilateral precaution, fault is introduced as an element of tort liability. Introducing fault means setting a due level of precaution, defined by the legislator or by the judge. In principle, the due level of precaution, x_d , should be equal to the efficient level of precaution, x^* , we will assume this equality and later on we will relax such an assumption.

Under strict liability and no liability the judge decides the case on the base of harm and cause, and does not control whether the level of precaution actually taken by the parties was efficient or not. On the contrary, under any negligence rule the judge has to perform such a test, by confronting the level of precaution actually taken by the parties with the due level of precaution. This increases the administrative cost of the trial and generates some major complexities.

In the model we described precaution as a continuous variable, but in reality it can assume a variety of forms. In order to avoid being hit by a car, a pedestrian can take different types of precaution; he can walk only on the pavement, can look several times to the left and to the right before crossing the street, can refrain from going out during the night or when the traffic is heavy, can even decide to take a bus or not to go out at all. Some forms of precaution are easily observable ex post, when the accident has occurred and the case has come to court, but some others are very difficult, or even impossible to assess and to compare with the legal standard of precaution. The judge can easily declare that the pedestrian was at fault because he should have crossed the street at the zebra crossing, but it would be difficult to declare that the pedestrian should have taken a bus or stayed home.

A similar example can be performed with respect to a chemical plant owner who has to decide on several variables influencing the likelihood of having an accident, training of the employees, safety equipment, technology, working hours, and so forth. Clearly, not all these variables are observable for the judge. Intuitively, individuals will prefer to invest in observable precaution in order to avoid being declared at fault and refrain from unobservable-precaution investments. We will soon clarify this statement.

We denote as care every observable precaution and as activity level any unobservable precaution. The concept of activity level includes not only the number of times we repeat a certain action (how many times a week one drives one's car or how many bicycles a factory produces) but also any other kind of precaution which cannot be observed ex post. With the general term of precaution we denote both care and activity level. It is clear that the level of precaution required for the judgement on fault is by definition a due level of care, the activity level not being possible to check.

Therefore, the introduction of the concept of fault creates a distinction between care and activity level, which does not exist under strict liability and no liability, and generates a problem: some kinds of precautions (the activity level) are left out.

2.6.C. The negligence rules: simple negligence

The requirement of fault can be introduced both into a no liability rule and into a strict liability rule. In both cases the analysis on fault can be performed either on only one or on both parties. The simplest case is simple negligence, a no-liability-based negligence rule where the requirement of fault is only on the injurer.

Under simple negligence if the injurer is negligent, that is if he does not take at least as much care as the due level, he has to pay damages to the victim. On the contrary, if the injurer is careful the victim bears the harm, no matter whether he was negligent or not. The victim is the residual bearer, in the sense that he has to bear the consequences of the accident if nobody else can be blamed negligent. In this sense, simple negligence is a no-liability-based rule, because as no liability it targets the victim last.

In the case of unilateral-precaution accidents when the victim is the avoider, the injurer cannot be declared negligent since he does not have any possibility to take care (his due level of care is equal to zero); therefore, the victim bears the cost of the accident and he will take the optimal level of precaution (care and activity level) as under no liability (see figure 1).

If the avoider is the injurer, he will have to pay only in the event of his not taking at least the due level of care. Figure 3 describes the incentive stream provided to the injurer (continuous line). We still assume that the due level of care coincides to the optimal level of care. If the injurer takes less than due care (negligent), he has to bear the cost of care and the expected accident cost (pay damages to the victim), $p(x)H+c(x)$. On the contrary, if he takes at least due care (careful), he bears only the cost of care, $c(x)$. Therefore the injurer has an incentive to take at least due care. Moreover, he has an incentive to take just due care and not more, because, by taking more care he will increase the cost of care without further benefits (the reduction of expected accident cost will be gained by the victim). We can conclude that in unilateral precaution cases both the injurer and the victim have the right incentive to take optimal care.

However, with respect to activity level, only the victim has an incentive to take the optimal level of precaution, since he bears the full cost of the accident (residual bearer, figure 1). The injurer will not pay any damages if he is careful, so that he will invest only in care (observable by the judge), in order to avoid liability and will not reduce the activity level (that is invest in unobservable precautions). We can conclude that simple negligence gives efficient incentives with respect to activity level only to the victim.

In the case of alternative precaution, we can apply the same kind of reasoning to conclude that both will take optimal care. However, we know that efficiency requires only the least-cost avoider to take care. If the other party or both take care there is an inefficient result. A way to obtain the desired outcome is to implement such a requirement into the negligence criterion, so that the injurer is declared negligent only if he was the least-cost avoider. This way we obtain the same result as in unilateral-precaution cases both for care and activity level. This is just a transposition of the well-known Hand formula¹⁹.

In the case of bilateral precaution we have seen that both strict liability and no liability fail to give the appropriate incentives to both parties, they direct the incentive stream only towards one party. The introduction of the requirement of fault improves the performance of the rule. In deeds, on the one hand, under the simple negligence rule the injurer can escape liability by being careful, so that he has an incentive to be careful (figure 3). On the other hand, the victim is the residual bearer and, if the injurer is careful, has to bear the expected accident cost, so also the victim has an incentive to take optimal care. It is clear that the introduction of fault in a no liability rule makes both parties take the optimal level of care. However, as in unilateral-precaution cases, simple negligence gives efficient incentives with respect to activity level only to the victim.

We can conclude that the advantage given by the introduction of fault is an improvement in the performance of the rule in bilateral-precaution situations. The negligence criterion makes both parties take the optimal level of care in any case (unilateral, alternative and bilateral precaution), but always gives incentives to choose an optimal activity level only to the residual bearer, the victim.

2.6.D. The negligence rules: contributory negligence

A second negligence rule is contributory negligence, another no-liability-based negligence rule. The residual bearer is again the victim but in this case the negligence criterion is applied to both the victim and the injurer. The injurer has to pay damages only if he was negligent and the victim was careful, in all the remaining cases the victim bears the loss. Therefore, the victim does not have the right to compensation in the case of both being negligent, in the case of his being negligent and the injurer being careful and in the case of both being careful.

By taking the due level of care, the injurer can surely escape liability, so that he receives the appropriate incentive to comply with the legal standard. Given that the injurer will behave carefully, the victim will not be entitled to compensation and he will internalise the expected accident cost (figure 1) and minimise it by taking optimal care. Contributory negligence

¹⁹ *United States v. Carroll Towing Co.*, 159 F.2d 169 (2d Cir. 1947). The principle expressed in this case by Judge Learned Hand is that in alternative-precaution situation only the least-cost avoider is to be held responsible.

enhances efficient behaviour by both parties with respect to care; however, only the residual bearer, the victim, has the right incentive to take an optimal activity level. The injurer will be free of liability only if he takes due care and will take a too high level of activity (that is under-invest in unobservable precautions).

Contributory negligence produces in theory the same effects with respect to care and activity level as simple negligence, but in practice it reduces the number of cases in which the victim is entitled to compensation because it denies compensation also in the case of the victim being negligent. Therefore, contributory negligence should be preferred with respect to the goal of minimising administrative cost.

It is noteworthy that in the literature the term ‘negligence’ (without further specifications) is widely used to refer both to simple negligence and contributory negligence in the comparison to strict liability.

2.6.E. The negligence rules: strict liability with defence of contributory negligence and strict liability with defence of dual contributory negligence

A negligence rule can be based also on the framework of strict liability, in which case the residual bearer is the injurer. As we have done before, we will consider two cases: the case of the negligence criterion being applied only to the victim’s behaviour, and the case of the negligence criterion being applied to both parties.

Under strict liability with defence of contributory negligence, there is a test on the victim’s fault. If the victim was at fault, he does not have the right to compensation. On the contrary, were he not at fault, the injurer pays damages. The victim is spurred to take the due level of care, because did he not do so, he would not be entitled to compensation in the case of an accident. Provided that the victim is careful, the injurer bears the expected accident cost and will take the level of precaution (care and activity level) which makes him minimise the total cost, that is the optimal level of precaution, x^* . Therefore, also strict liability with defence of contributory negligence enhances efficient behaviour with respect to care.

Moreover, for the victim it is sufficient to take due care in order to be compensated for the accident loss, so that he does not have any incentive to take (additional) unobservable precaution. On the contrary, since the injurer is the residual bearer, any additional precaution reduces his expected accident cost and increases his precaution cost; the injurer will invest both in observable and unobservable precaution as long as the sum of precaution cost and expected accident cost is minimised, that is the efficient level of precaution, x^* .

Strict liability with defence of dual contributory negligence encompasses a double test on fault; the negligence criterion is applied to both the victim and the injurer. In this case the victim has to bear the loss due to the accident only if he was negligent and the injurer careful,

in all the remaining cases he is entitled to compensation. The injurer has to pay damages if both were negligent, if only the victim was careful and if both were careful. This rule gives the same result as strict liability with defence of contributory negligence but it probably leads to fewer cases of compensation and, hence, saves some administrative cost.

2.6.F. Negligence rules: the efficiency-equivalence theorem and the activity-level theorem

We have examined four possible combinations of a negligence criterion with the two basic rules of strict liability and no liability; from that analysis we can derive two general statements. First, any negligence rule we have examined gives efficient incentives to both parties with respect to care (observable precaution). This statement is known as efficiency-equivalence theorem and can be generalised as follows.

The difference between simple and contributory negligence consists of the fact that the individual who pays for the loss when both are negligent is different under the two rules. Under the former rule the injurer pays, under the latter the victim does. Intuitively, it is clear that this result applies also for mixed solutions, that is when the injurer pays a portion of the accident cost and the victim bears the rest, whatever the shares are. The reason for this is that at least one party (usually the party who bears the larger share when both are negligent, but depends on the cost structure) will always have an incentive to be careful. Once one party is careful the other bears the full cost and will be careful too. This result applies also in the case of strict-liability-based negligence rules (for a mathematical approach see Landes and Posner, 1980). It follows that a negligence rule which makes the injurer pay only 1% of the accident cost in the case of both parties being negligent and lets the victim bears 99% of the loss is as efficient with respect to care as a 50-50 rule, as a 90-10, as a 75-35 and so forth. Therefore, any negligence rule, whatever the division of losses is, gives efficient incentives to both parties with respect to optimal care.

The second general characteristic we have noticed is that the no-liability-based negligence rules (simple negligence and contributory negligence) give efficient incentives with respect to activity level only to the victim, the residual bearer in those cases, while the strict-liability-based negligence rules (strict liability with defence of contributory negligence and strict liability with defence of dual contributory negligence) give efficient incentives with respect to activity level only to the injurer, the residual bearer in those remaining cases.

Therefore, any negligence rule directs efficient incentives with respect to activity level only towards the residual bearer, and fails in enhancing the other party's efficient behaviour. This is the Shavell's theorem on activity level (Shavell, 1980). More precisely, he shows that no negligence rule exists, which can give both parties efficient incentives with respect to activity level. This follows from the fact that the distinction between care (precaution the

judge can observe ex post) and activity level (precaution the judge cannot observe ex post) is due to the introduction of the negligence criterion. The party who can escape liability by simply taking the due level of care will not invest in other unobservable precautions, while the other, the residual bearer, will.

In theory, a rule of decoupled liability gives both parties efficient incentives with respect to care and to activity level. Decoupling liability (Polinsky and Che, 1991) means making both the injurer and the victim be the residual bearer, by denying to the victim any compensation (as under no liability) and having the injurer pay a fine equal to compensatory damages (as under strict liability), regardless to their level of precaution.

2.6.G. Comparative negligence

Comparative negligence is a negligence rule based on no liability; the victim is the residual bearer. A rule of comparative negligence could be constructed on strict liability as well. The sole difference is that in this case the residual bearer would be the injurer.

As under simple negligence and contributory negligence, also under comparative negligence, if the injurer is negligent and the victim is careful, damages have to be paid, and if the victim is negligent and the injurer is careful, there is no right to compensation.

The difference between the three rules arises in the case of both the injurer and the victim being negligent. Simple negligence makes the injurer pay, contributory negligence makes the victim pay and comparative negligence shares the loss in proportion to the degree of negligence of each party, by entitling the victim to less than compensatory damages. Roman law adopted a system of contributory negligence, by denying negligent victims compensation, while most of the contemporary legal systems are in favour to a varying degree of apportionment (Honoré, 1971, Curran, 1992).

It follows directly from the efficiency-equivalence theorem that comparative negligence is as efficient as any other negligence rule but it has one major disadvantage: it is likely to increase the administrative cost of the trial. Not only does the judge have to decide whether or not the injurer is to be held liable and evaluate the magnitude of the harm, but he also has to share the loss between the parties. This is likely to increase the cost of acquiring information during the trial and in general to produce more litigation.

In theory, contributory negligence is the rule to be preferred, since it leads to the lowest administrative cost, but in fact comparative negligence is a most common rule. This is still one of the most interesting puzzles of tort law and economics. Several explanations have been attempted by assuming an imperfect information environment.

In the case of judges making random errors in comparing the due level of care to the level of care actually taken by the parties (evidentiary uncertainty), parties are spurred to take

an inefficient level precaution (Calfee and Crasweel, 1984) in order to minimise their expected liability. Contributory and simple negligence concentrate such an effect on one party only. Comparative negligence, by distributing the loss between the parties, reduces the effect of errors and leads to less distortions (Cooter and Ulen, 1986). Similarly, comparative negligence has been said to be more efficient than contributory negligence when the standard of care is uniform for all parties, but the individual costs of care differ (Rubinfeld, 1987), and when judges err regarding the level of care cost actually borne by parties (Haddock and Curran, 1985).

Under a different perspective, comparative negligence can enhance more efficient outcome when individuals are not identical (Emons, 1990a, 1990b) and when parties make precaution decisions sequentially, so that one party can observe the other party's level of precaution before choosing his (Grady, 1990)²⁰. A completely different approach leads eventually to fairness and equity arguments in favour of comparative negligence, insofar as sharing losses when parties are similarly at fault is socially regarded as just²¹.

2.6.H. Causation and the negligence rules

We have already briefly discussed the meaning and the characteristics of the requirement of cause. Here we will describe the interaction between the requirements of cause and fault under negligence rules. We will implicitly refer to no-liability-based negligence rules, but the analysis can be extended to strict-liability-based negligence rules as well.

The incentive given by the introduction of fault in tort liability has been already described. The peculiar characteristic of such an incentive stream is a discontinuity in the cost borne by the party (figure 4). If the party complies with the legal standard of care, he pays only the cost of care, but if the level of care is lower than the due level he pays also the full expected accident cost. This does not affect the efficiency of the rule but has particular effects in the case of errors, as we will see.

However, this is not always the case. One of the implicit principles of tort law is that a reduction to zero of the probability of having an accident is not to be desired, because it would involve too high a precaution cost (unless in extreme situations). The optimal level of precaution always corresponds to a certain probability of accidents occurring.

Does the injurer pay also for those too-expensive-to-avoid accidents? In the model of negligence we described above he does, and that is the reason for the observed discontinuity in the continuous line of figure 3. In reality, the injurer can escape liability by demonstrating that the accident would have occurred anyway even had he been careful. When this is possible

²⁰ It is worth to recall that throughout this excursus we have assumed identical parties, and simultaneous decisions.

²¹ On comparative negligence see also Orr (1991), Rea (1987), Chung (1993), Edlin (1994).

the abovementioned model is to be refined. The point has been clarified in the literature by a simple example (Kahan, 1989). A cricket field owner builds up a fence surrounding the playing field in order to avoid passers-by being harmed by balls accidentally flying over. Law requires a three-meter-high fence but the owner decides for a two-and-a-half-meter-high fence. If a ball bounces out at 2,8 meters and cause harm to someone, the owner has to pay; however, if the ball flies at 3,5 meters, even a three-meter fence would not have stopped it. If we apply the requirement of causation the owner does not have to pay damages.

The example shows that a negligent injurer does not pay for all accidents but only for the ones caused by his fault (Grady, 1983, Kahan 1989, Marks, 1994). Figure 4 describes this situation. Point *C* on the expected-accident-cost curve corresponds to the accidents, which occur anyway if the party takes due care, and segment *CD* indicates the corresponding cost. Segment *AD* is the sum of precaution cost and expected-accident cost when the party takes due care. According to the model in figure 3, the injurer pays the social cost from the moment he takes a little less than due care, starting from point *A*.

According to Grady's model in figure 4, the injurer pays less, because he does not pay a part of the expected-accident cost. Since the social cost is the sum of precaution cost and expected accident cost, if we subtract an expected accident cost from the social cost we obtain the precaution cost. Therefore, if the injurer takes just a little less than due care, he does not pay *AD* but *AD* (social cost) minus *CD* (expected-accident cost occurring anyway) which is equal to *BD* (precaution cost). If he takes even less care, he will pay more, because the likelihood of an accident will increase. The new curve starting from *B* increases to the left and is always equal to the social-cost curve minus *CD*. On the right it follows the care-cost curve.

Grady's model has the same characteristics as the standard one with respect to incentive to optimal care; since the minimum point of the curve is still corresponding to the due level of care, parties are spurred to take due care. Nevertheless, the conclusions about errors will be different.

2.7. Errors and uncertainty.

So far we have described what incentives tort law gives towards efficient precaution if the judicial system produces unbiased decisions. We will now consider the hypothesis that the judge (or the legislator) errs with respect to the three elements of tort liability. We have already mentioned some arguments pro comparative negligence in the case of errors and uncertainty.

2.7.A. Errors and uncertainty: harm

The judge can make errors²² in determining the precise magnitude of the harm and, hence, entitling the victim to over or under-compensation. This causes various effects depending on the rule in force. First we consider systematic errors.

Under strict liability systematic over or under-compensation affects the injurer's decision in the same direction: the injurer will take more or less care than optimal since he has to pay more or less damages. The same effect can be found under the strict-liability-based negligence rules, since the injurer is the residual bearer and will pay the damages to the victim in the case of both being careful. On the contrary, random errors concerning the correct determination of the harm do not cause distortions in the parties' behaviour.

Of course under no liability there is no room for compensation and errors do not produce any effect, and also if we add negligence²³ to a no liability rule (simple or contributory negligence) a careful injurer never pays damages, so that there is no distortion on the level of care he takes. However, if we consider Grady's model of negligence (figure 4), the result will be different. Over-compensatory damages do not change the injurer's behaviour, since he can still escape liability by simply taking due care. On the contrary, under-compensatory damages will spur the injurer to take less than optimal care, since they reduce the expected accident cost, and, hence, create a new minimum on the left-hand side of the due level of care. Random errors around the correct determination of the harm will have the same effect.

Under a different perspective, the judicial determination of the magnitude of the harm can be correct on average but not accurate in individual cases. Damages are accurate if they correspond exactly to the harm borne by the victim; they are not accurate if they do correspond to the mean of the possible harm but not to the harm actually suffered by the victim in a specific case. The question is actually whether the judge has to invest resources in determining the harm case by case or whether he can rely on average measures, which will necessarily be correct only on average while over or under-compensatory in individual cases (Kaplow, 1994, Kaplow and Shavell, 1996a). The subject is relevant in practice since it encompasses the question whether the judge has to determine the value of an injured leg case by case or whether he can entitle a football player and an housekeeper to the same average compensation (which will result under-compensatory for the former and over-compensatory for the latter).

An accurate decision is more expensive than a non-accurate one (time-consuming, experts bill) and, hence, the investment in accuracy is justified only if it spurs a more efficient

²² By errors we mean a different decision from the one economically desirable, which can be due also to the judge pursuing some goals other than efficiency.

²³ The following conclusions are drawn under the assumption that the due level of care be equal to the optimal level of care when damages are perfectly compensatory.

behaviour of the parties involved. On the one hand, if parties cannot foresee *ex ante* the exact magnitude of the harm, they will face an average expectation, so that an accurate decision *ex post* cannot be anticipated and it will not improve efficiency. On the contrary, if parties can exactly foresee the magnitude of the harm resulting from an accident, they will take over or under-precaution if the measurement of the harm by the judge is based on an average. If for instance parties know *ex ante* that the average compensation the judge will determine is less than the actual harm they are going to cause, they will take less precaution. Therefore, accuracy is needed if parties can foresee the magnitude of the harm; if parties face average expectations, average compensation is cheaper than, and as effective as, accurate compensation.

With respect to risk, an accurate compensation always makes the victim indifferent, while an average compensation makes him bear the risk of over or under-compensation. On the contrary, average compensation removes the injurer's risk of very high or very low damages due to accurate decisions.

In addition, accuracy spurs parties to invest *ex ante* in acquiring information about the magnitude of the harm. If that investment is socially desirable then so is accuracy.

2.7.B. Errors and uncertainty: cause

Systematic errors about causation produce effects of the same direction on the level of precaution taken by the parties (Shavell, 1980a). Random errors are likely not to produce effects if the mean is correct. With respect to accuracy the argument is slightly different from the case of harm. Inaccuracy refers to the possibility that a liable injurer be found not liable and that a non-liable injurer be found liable (Kaplow and Shavell, 1994). It is noteworthy that both kinds of errors reduce the incentives to take optimal precaution (Png, 1986). The first (false negative) decreases the expected cost of causing accident, by overly restricting the scope of liability, and, therefore, leads to under precaution. The second (false positive) decreases the benefit of precaution; as this kind of (false) liability cannot be avoided by taking precaution, there is an incentive towards sub-optimal precaution.

It is clear that a higher level of accuracy increases the incentives toward optimal precaution (particularly, activity level). Administrative cost also increases insofar as accuracy increases; therefore there is a trade-off between optimal incentives and minimisation of administrative cost to be taken into account while setting the level of accuracy.

Somewhat related is the topic of probabilistic liability. In certain situations multiple indistinguishable causes produce a harmful effect. A lung cancer can be due both to the use of a toxic substance and to other natural causes. It is not possible to determine the cause for each

individual but the percentage of lung cancers caused by a particular substance on the total is statistically determinable.

Since the judge has to decide individual cases, he has actually to decide whether a specific case of cancer was caused by the toxic substance or by other carcinogenic factors. A possible criterion is to define a threshold on the probability that a certain event is caused by a certain cause. If that probability in the specific case is higher than the threshold, then compensation will be awarded, otherwise the injurer will be declared non-liable.

This criterion has a major incentive problem: it over-targets injuries above the threshold by always awarding compensation and under-targets injuries below the threshold by never awarding compensation. If the cancer incidence for substance A is 30% and for substance B is 60%, the producer of A will over-produce the substance (since he will always escape liability) and the producer of B will under-produce it (since he internalises 100% of the expected accident cost).

A different criterion is to make an injurer pay in proportion to the probability of having caused the harm in every case. Therefore, producer A will always compensate 30% of the harm to any victim and producer B 60%. This way the result is statistically correct and, since causes are not separable, will induce efficient behaviour of the parties (Shavell, 1985).

With respect to *ex ante* incentives to take optimal precaution, proportional liability guarantees a more efficient result than threshold liability. However, the use of the latter rule minimises the cost of errors in *ex post* adjudications (Kaye, 1982), reduces the degree of uncertainty (Levmore, 1990) and leads to lower administrative cost (Shavell, 1985).

2.7.C. Errors and uncertainty: fault

The introduction of the requirement of fault expands the role of the judge. Under strict liability and no liability the decision about the level of precaution is left completely to the parties while under any negligence rule, the judge, the legislator, or an independent agency sets a due level of care. So far we have assumed that the due level of care corresponds to the optimal level of care, but that does not need to be always the case. Courts can have poor information about the cost of care and the expected accident cost, can underestimate the probability of a certain accident, or can also pursue other goals than the minimisation of social cost. The result could be a biased due level of care, higher or lower than the optimal level of care. We will hereby analyse the reaction of the parties involved in an accident under a negligence rule, when the due level of care is biased (Calfée and Craswell, 1984).

Under the standard model of negligence (figure 3), if the due level of care is lower than optimal, parties will take the due level of care: if they do so they escape liability, and by

taking more care they would just increase the cost of care they bear. The outcome is therefore inefficient.

If the due level of care is higher than optimal, parties will take the due level of care: if they do so they pay just the cost of care, on the contrary, if they take less care they have to pay also the expected accident cost. The outcome is inefficient²⁴.

It is also possible that the due level of care is not clearly established, but based on vague criteria the judge has to apply case by case. On average due care corresponds to optimal care, but there are random variations around that mean; therefore, parties are not sure whether the optimal level of care will be considered by the judge as sufficient for complying with the legal standard. If so, the party pays just the cost of care, if not, the party has to pay also the expected accident cost. In some cases, the party can reduce that risk by investing in more care, so that the result will be his taking more than optimal care. In some other cases, the party will take less than optimal care, as this could be enough to escape liability (Calfée and Craswell, 1984, and Craswell and Calfée, 1986). The outcome is again inefficient.

Those results change if we consider Grady's model of negligence. This is due to the fact that under this model the discontinuity in the cost curve between being careful and being negligent is filled; instead a continuous curve describes the cost borne for any level of care.

If the due care is lower than optimal, parties take due care and the outcome is inefficient as above²⁵, but, if the due level of care is higher than optimal, parties will take optimal care and the result is efficient. This argument can be easily shown by figure 5. By definition the social cost reaches its minimum when parties take optimal care, if we move the curve downwards by subtracting vertically the cost of accidents which would occur anyway even if parties take optimal care, the shape of the curve stays the same and the minimum still corresponds to optimal care (point *E*). Point *B*, which corresponds to (a biased level of) due care is on the right of point *E* but still on the same curve, and since *E* is the minimum, then *B* is higher than *E*, that is the party is bearing a higher cost by taking due care than by taking optimal care. Therefore, if due care is higher than optimal, parties will take optimal care anyway and the result will be efficient notwithstanding the bias in the rule. That has been said to be a justification for increasing the standard of due care, since a too low standard produces inefficiency, while a too high standard does not have any effect (Kahan, 1989, at 443).

If there are random errors, parties are likely to take less than optimal care, since there is no gain in taking more than optimal care, but just an additional cost, while by taking less than optimal care parties can save care cost.

²⁴ There could be an efficient result if the bias is so large that the required level of care costs more than optimal care plus expected accident cost; in this case parties take optimal care.

²⁵ As under the standard model, by taking due care parties escape liability, therefore there is no incentive to take

It is clear that the most evident difference between the two models of negligence we have examined concerns how parties react to errors. The choice between the two depends on the assumption made about the courts' evaluation of causation.

2.8. Goals of a tort law system: incentives to acquire information.

So far the analysis has been led on the track of the main goal of the tort law system: giving parties incentives towards optimal precaution. Upon the same model we can base a related analysis on the incentives the tort law system gives towards the acquisition of information by parties. Two kinds of information will be considered: information on the degree of risk and information on the risk-abatement technologies.

In order to take optimal precaution parties need to have information about the activity they are undertaking. Although full information is desirable for optimal decision on the level of precaution, information is most of the time expensive to acquire. Therefore, it can be efficient to acquire information only up to a certain limit or even not to acquire information at all. The optimal level of information depends on the cost and benefit (reduction in expected accident cost) of the acquisition. Information on the degree of risk is needed in order to make accurate valuation of the expected accident cost and, therefore, optimal decision regarding the level of precaution. Gross estimations about risk lead to inaccurate decision on the level of precaution, which might turn out to be over or under-sized (Shavell, 1992). Information on the risk abatement technologies is needed in order to make optimal decision on the type of precaution to take and reduce the cost of precaution.

Under strict liability the expected accident cost is borne by the injurer, and so is the cost of precaution. Therefore the injurer benefits from any investment in information; the injurer also bears the cost of acquiring information. The conclusion can be easily drawn that the injurer will choose the optimal level of information. That means that the decision on the level of precaution will be optimally accurate and that the chosen type of precaution will be the most efficient. A consequence is that under strict liability the injurer faces optimal incentives to invest in research and development of new risk abatement technologies (Shavell, 1980a, Schäfer and Schönemberger, 2000). On the contrary the victim has no incentives at all.

The same kind of reasoning can be reversed for the case of no liability: the victim will behave optimally with respect to information acquisition, while the injurer will receive no incentives at all.

On the other hand, the negligence rules produce slightly different and more variable results, also depending on the extension of the definition of due care. In fact, under

more than due care.

negligence rules only the residual bearer has the appropriate incentive to acquire optimal information, since he faces fully the expected accident cost. On the contrary, the other party can escape liability by taking due care and, therefore, has incentives to conform his investment in information to what is required as due care. Shavell (1992) clarifies that complete definitions of due care (definitions that include the requirements of party's optimally acquiring information) lead to an efficient result, while incomplete definitions of due care tend towards inefficiency. Therefore, negligence rules require judicial control on whether or not parties invest optimally in information on the degree of risk. With respect to information on the risk-abatement technology, the negligence criterion still spurs to optimal incentives towards research and development of new technologies only if the standard of due care is flexible towards new solutions; parties will not invest otherwise (Shavell, 1980a, at 23, Posner, 1998).

This is an argument for strict liability insofar as there is asymmetric information (and asymmetric cost of acquiring information) between injurer and victim (or potential victims). This is likely to happen in the case of ultra-hazardous activities but also, as Trimarchi (1959a and 1961) notes, in every case of enterprise that probably produces injuries. In addition, it is worthwhile to consider that under strict liability and no liability the cost of acquiring information is completely borne by the parties, while under negligence rule is partially borne also by the due-care setter (judge, legislator or public agency), which raises the administrative cost of the trial.

An analysis of incentives to acquire information can also provide a justification for a shift of liability from the actual injurer to his principal. We will discuss vicarious liability in section 13.

2.9. Additional incentive streams.

A wider view on torts permits us to notice that some behaviours are targeted not only by tort law but also through the means of other sanctions, even by criminal punishment. Since tort law already provides an incentive stream towards optimal precaution, a second additional incentive stream can be justified only if the first one proves to be insufficient²⁶.

First, intentional torts (Landes and Posner, 1981) are criminally sanctioned under any legal system. An accidental harm follows an action with a certain probability lower or equal to 1. An intentional harm follows an action always with probability 1²⁷, which increases the

²⁶ From an historical point of view the question should be posed differently, since tort law mainly evolved from criminal liability. We will discuss two cases.

²⁷ Sometimes this simplified description could not fit perfectly the real situation; however the substance of the argument does not change.

expected accident cost but does not change the conclusion of the model: if we let the injurer internalise the expected accident cost he will behave efficiently.

However, intentional torts are mainly directed towards specific goods or individuals and, therefore, the transaction costs in those cases are lower than the transaction costs usually paired with unintentional torts. Once bargaining over rights is possible, then property rights are a better solution than tort liability, because the price is set by the parties on the base of their own preferences, while under tort law the compensation is awarded by the judge and is much more likely to be biased. Hence, the solution is to punish intentional torts in order to spur contractual outcomes. In addition, unpunished intentional torts will induce potential victims to invest in private protection and potential injurers to search for torts to commit (Polinsky and Shavell, 2000), which are both socially wasteful expenses.

Second, apart from intentional torts (and seldom in combination with them), additional sanctions are desirable when the first incentive stream is insufficient because the injurer does not fully internalise the cost he causes. That happens when damages are not equal to the social cost of the accident, but cover just a part of it, for theoretical or practical lack of perfect compensation (Cooter and Ulen, 2000). Following Polinsky and Shavell (2000) we account for two cases: injurer's illicit gain and difficult-to-estimate loss.

In economic terms the injurer's gain is to be taken into account while analysing torts. Indeed a harmful behaviour can provide the injurer with a gain, which exceeds the loss to the victim. If this is the case, the loss will occur even if the legal system entitles the victim to compensation, and this is indeed an efficient outcome; as in the case of someone breaking a window in a hotel room to save his child from fire. In those cases we want the harm to occur, and an additional sanction would prevent this solution²⁸. In other cases, however, the injurer's gain is considered to be illicit: the pleasure one receives from punching another individual will not be taken into account, and one would probably be declared liable both under tort law and under criminal law. A very general explanation for the difference between licit and illicit gains is that in the latter case there exists a social loss other than the victim's loss, to be considered. In the example, the rise of violence in society can be taken into account as a social cost and, hence, it will rebalance the cost-benefit analysis, by giving more weight to costs. Once we consider not only victim's harm but also social harm, then we see that the injurer internalises under tort law only the former (damages to be paid to the victim), therefore we need an additional sanction to make him internalise also the latter.

A similar argumentation can be used for difficult to estimate losses, usually left out from the assessment of damages. Also in this case the internalisation of costs is imperfect and an

²⁸ In those cases legal systems give different solutions with respect to whether tort damages have to be paid, see Limpens, Kruithof, Meinertzhagen-Limpens (1979).

additional sanction is required. As it can be easily noticed, the two problems described share the same nature and also the same limit: social harm and victim's difficult-to-estimate losses are not easily expressible in monetary terms.

Another case of imperfect internalisation can be due to too low an apprehension rate. We denote as apprehension rate the probability of an injurer being caught and found liable. Normally the apprehension rate is equal to 100%, so that the injurer always pays the due compensation to the victim. However, there are cases in which the apprehension rate is lower than 100%, as for instance in the case of sexual harassment, where the injury is difficult to prove, or in the case of small but repeated injuries with different victims not all willing to bring unimportant cases to court, or in the case of not easily identifiable injurers.

If the apprehension rate is lower than 100%, the injurer faces just a portion of the expected accident cost. For instance, if the apprehension rate is only 25%, the injurer knows that he has to pay damages only in 25% of the accidents he will cause and, hence, he will not internalise the full expected accident cost but just 25% of it. That will result in a sub-optimal level of precaution, which calls for an additional incentive stream²⁹ (Cooter, 1989, Polinsky and Shavell, 1998 and 2000).

The foregoing analysis shows that additional incentive stream is needed when the injurer intentionally causes harm or when he does not fully internalise the expected accident cost. A related subject is how to create an additional incentive stream. Different legal systems adopt different solutions; amongst the feasible are criminal sanctions, administrative sanctions, and over-compensatory tort damages. Civil law countries opt for a combination of the first two, while the third (punitive damages) is a peculiar American solution. The difference between fines and punitive damages (besides differences in the applicability) is that under the latter regime the victim gets what exceeds compensation, while the fine always goes to the state. For that reason, if a lower-than-one apprehension rate is due to victim's unwillingness to sue, punitive damages can give a counter-incentive, while fines have no effect on that.

2.10. Goals of a tort law system: optimal risk allocation. Insurance.

As we have already pointed out, were the tort law system solely aimed to optimal risk allocation, then a generalised system of insurance (Shavell, 1982) would be the only rational

²⁹ The solution is to make the injurer pay more when he is caught, so that the incentive stream is brought back to the efficient level. In our example, since the injurer pays damages only in $\frac{1}{4}$ of the cases, he should pay damages equal to 4 times the expected harm. The expectation is to pay $\frac{1}{4}$ of the time 4 times the expected harm, which is equal to $\frac{1}{4} * 4 * \text{expected harm} = \text{expected harm}$. In general if the apprehension rate is e the damages have to be multiplied for $1/e$. This way the incentive stream is again at the efficient level.

answer (Calabresi, 1970, at 43-44), because it creates less administrative cost and guarantees a wider risk spreading than the tort liability system.

The need to shield parties from risk rises from the assumption that parties are generally risk-averse. This can be true for both parties: the victim faces the risk of harm, the injurer faces the risk of having to pay damages. That gives room for first-party insurance in the former case and third-party insurance in the latter. Risk-aversion makes it preferable for parties to pay an insurance premium *ex ante* than to accept (the probability) to pay accident cost *ex post*.

Liability rules allocate the risk of accident losses to the residual bearer, since the other party will tend to avoid liability by behaving carefully. The injurer bears the risk under strict liability (and strict-liability-based negligence rules) and the victim does under no liability (and no-liability-based negligence rules). A system of social insurance has the result of spreading the risk amongst a large number of individuals, so that it minimizes the effect of risk. Moreover, it can be paired with a deep-pocket financing rule, justified by a decreasing-marginal-utility-of-money assumption, insofar as rich people pay more than poor, since they value a payment of one euro less than poor people do. A system of private insurance will have the same risk-spreading effect.

However, insurance, by shielding parties against the risk of bearing accident cost, actually removes the incentive toward optimal precaution and optimal information and frustrates the main end of the tort liability system. Moral hazard heavily affects this issue (Shavell, 1979); parties will have an incentive to buy insurance and then, since they are free from risk, take a sub-optimal level of precaution. Insofar as the insurer can control parties' behaviour, that incentive stream can be replaced by other means, as *bonus-malus* clauses or partial coverage.

In addition, mandatory insurance can be desirable because of the 'judgement-proof problem' (Shavell, 1986): injurers are considered to be judgement-proof with respect to the portion of harm exceeding their wealth that they are not able to compensate³⁰. If this is the case, injurers do not bear the full expected accident loss, since for some extremely large harm they will pay just a portion of the damages. We have noted in section 8 that over-compensatory damages do not affect efficiency under the standard model of negligence, but they do reduce the incentives to take precaution under Grady's model of negligence and under strict liability. On the assumption that the problem of moral hazard can be solved by the insurer controlling the injurer's behaviour, compulsory insurance can be a possible solution for the judgement-proof problem, since it transforms the threat of paying *ex post* a huge amount of money (which exceeds the injurer's wealth) in compensation in an obligation to

pay ex ante a (relatively) small premium. Within the European Union mandatory insurance is for example required against civil liability in respect of the use of motor vehicles³¹, and should cover compulsorily both damage to property and personal injuries³². Vicarious liability can also be required in order to deal with the same problem (section 13).

2.11. Goals of a tort law system: minimisation of administrative cost.

The administrative cost due to the tort liability systems mainly consists of the cost of administering the treatment of accidents. The goal of minimising administrative cost is tertiary in the sense that its aim is to reduce the cost of achieving the other two goals of optimal incentives and optimal risk sharing (Calabresi, 1970, at 26). We consider the direct cost of courts and lawyers and the cost borne by parties also in the case of out of court settlements. Those costs depend both on the cost per single case and on the volume of cases brought to court.

In the foregoing sections we mentioned some aspects of rules related to administrative cost. Briefly, on the one hand, the wider the role of the judge, the greater the cost per case. No liability is indeed inexpensive under that respect. Strict liability involves an analysis of harm and cause, and therefore costs per case less than any negligence rule, which involves an additional control on fault. Among the negligence rules, the cost depends positively on the complexity of the standard of due care and on whether or not losses are to be shared among the parties (comparative negligence).

On the other hand, under no-liability-based negligence rules the number of cases in which compensation is awarded is smaller than under strict liability, and this reduces the volume of cases brought to court. Which of those two effects prevails is difficult to assess. Moreover, the administrative cost of litigation matters for the incentives to sue; chapter 6 of this book will be devoted to that topic.

2.12. Specific issues in tort liability.

In this section we will provide an analysis of specific issues in tort law and economics, on the basis of the models described above.

³⁰ This argument applies not only to individuals but also to corporations.

³¹ Directive 72/166 of 24 April 1972, Second Directive 84/5 of 30 December 1983 and Third Directive 90/232 of 14 May 1990.

³² Article 1, Directive 84/5.

2.12.A. *Product liability and ultra hazardous activities*

In product liability the situation is slightly more complex when the relation between injurer (producer) and victims (customers) is contractual. Therefore, tort liability cannot be justified on the ground of prohibitive transaction costs. Injurer and victim actually bargain over the price and the quality of the products; therefore, risk will also be taken into account while fixing the price. The Coase theorem endorses the outcome's being efficient whatever the liability rule is (Coase, 1960). However, the Coase theorem applies if parties are perfectly and symmetrically informed.

In the case of asymmetric information the allocation of rights does matter for the sake of an efficient outcome. Only the informed party will perfectly internalise expected accident cost, while the uninformed party will base his cost-benefit analysis on the (imperfect) information he possesses, leading to an inefficient result (Spence, 1977). Therefore, the legal system should target the liability incentive on the informed party, which means making the informed party the residual bearer. If we assume that the producers are better informed than consumers, then we have an argument for the use of strict liability (or strict-liability-based negligence rules) in product liability (Shavell, 1980a, Schäfer and Schönenberger, 2000). The Directive 85/374³³ enhances this solution among the members of the European Union, who were, in most cases, already developing their own laws in the direction of a strict product liability.

A similar approach can be applied to ultra-hazardous activity. The use of strict liability creates appropriate incentives towards optimal precaution (both for care and activity level) and optimal acquisition of information (including research and development of better technologies) for the injurer. Moreover, insofar as the victim is likely to play an almost completely passive role the situation can be regarded as a unilateral-precaution case with the injurer in the position of the avoider; therefore, strict liability is likely to lead to an efficient outcome.

2.12.B. *Vicarious liability*

Sometimes the duty to compensate the victim moves from the injurer to another subject. Rules laid down in German and French law state the liability of the supervisor for the harmful action of a child or of adult under supervision³⁴. Similarly employers are liable for the actions of employees, within the scope of their employment³⁵. In the internal relationship employer-

³³ Directive 85/374 of 25 July 1985, which states in article 1 that 'the producer shall be liable for damage caused by a defect in his product' if (article 4) the injured person proves 'the damage, the defect and the causal relationship between defect and damage'. Such a system of strict liability is, however, subject to many restrictions listed in the Directive itself.

³⁴ In those cases there is a presumption of fault of the supervisor (Le Gall, 1976).

³⁵ For employer vicarious liability French and German systems differ slightly. Under the French system the

employee, the employee's liability is usually removed by means of employment or insurance contract provision, or directly by tort law. Apparently, such rules, while giving to the victim the appropriate compensation in any case, (to the victim it is indifferent who pays), remove the incentive stream from the injurer and the result may be an increase in the number of accidents. Nevertheless, there is an economic justification for it (Sykes, 1981 and 1984, Kornhauser, 1982, Kraakman, 2000).

Vicarious liability recurs between parties of a principal-agent relation. The principal controls and gives instructions to the agents, as employers do with employees, and parents with children. The power of controlling and instructing substitutes the incentive stream given by tort law and tends to avoid inefficient result, similar to what happens in the case of insurance. The limitation of vicarious liability to harm, which occurred within the scope of the employment, reflects the need to give vicarious liability the same limits of the principal's power of controlling the agent.

Moreover, the principal can have better information on the risk involved in a certain activity, and, hence, is in the best position to decide what precautions have to be taken and even whether or not the activity should be continued. A second advantage given by vicarious liability is that the principal is likely to be more solvable than the agents and, therefore, victims are better shielded against the risk of not being compensated because of judgement-proof problems. Third, vicarious liability solves the problem of individualising which agent, among the principal's agents, is responsible for the loss.

2.12.C. Multiple tortfeasors and the right to contribution

In several cases the accident is due to the action of more than one injurer. We have seen that any negligence rule gives to both the victim and the injurer the appropriate incentives to take optimal care. This is due to the fact that if one party is careful the other bears the cost of the accident and, therefore, both have an incentive to take optimal care. This conclusion can be extended to the case of multiple injurers (Landes and Posner, 1980), so that each injurer will be spurred to optimal care.

A further question is how to divide the loss amongst the liable injurers after the accident. Joint liability allows the victim to sue only one of them and to obtain the full compensation, instead of suing each of them for a portion. This saves transaction costs and is the economically desirable solution. The second problem is whether the injurer who paid damages should be entitled to contribution towards the other injurers. In economic terms we can notice that if the liable injurers are three a rule which makes each of them pay 1/3 of the

employer is strict liability 'for the other person', while under the German system the employer is liable for his own fault. (Eörsi, 1975).

harm gives exactly the same expectation as a rule which makes one pay the whole amount with probability 1/3. In terms of incentives the outcome is exactly the same. Therefore, in principle, contribution does not matter concerning the incentives to take optimal precaution and increase the administrative cost of the trial (Landes and Posner, 1980). A different solution can be reached after introducing some complexities into the model (see Kornhauser and Revesz, 2000), for example a judgement-proof injurer.

However, as in the case of pure economic losses, a different distribution can be desirable and justify the transaction costs involved in the rule of contribution.

2.12.D. Pure economic losses

In legal terms a loss is purely economic if there is lack of physical damage. For example, losses of profits due to false information or defective services are considered to be purely economic. Common law traditionally denies awards but for exceptional cases, while under civil law pure economic losses are in principle accepted (German law is closer to the common law solution than to the civil law one).

In economic terms, the definition of pure economic loss happens to be slightly different. As pure economic losses we denote a pure transfer from an individual to another: one individual gains exactly what someone else loses. Those losses are usually referred as losses of earning or profits. Since the total wealth of society (given by the sum of individual wealth) is unchanged there is not any social loss, so there is no room for tort law (Bishop, 1982, Rizzo, 1982, Rabin, 1985, Arlen, 2000).

That will be clarified by a simple example. During the New Year celebration for the year 2000 fireworks burn the cable connection of a small travel agency selling flight tickets and other services only on the Internet. The owner suffers two kinds of losses: first he has to repair the cable connection, second he cannot work for one month (the time needed to repair everything) and loses quite high earnings. The question is whether the firework organiser has to pay for both losses or just for the repairing. The question is relevant because the levels of care and activity taken while performing fireworks depends on the expected accident cost and, hence, on the amount to be paid as compensation.

The repair costs have to be paid; however, (in our simplified example) the loss of earnings is purely economic, since the other firms selling the same products on line will gain from it to the same extent³⁶. Therefore, the social loss consists only of the repairing expenses and only these costs have to be compensated. By entitling the victim to a larger compensation,

³⁶ In this example the marginal cost of serving one more customer is practically zero, therefore the remaining firms will be able to provide services for all customers. Goods are in this example perfectly identical (the firms sell tickets of the same flight companies, hotel reservations, and similar services) and the customers are indifferent with respect to where to buy. In the case of the consumer suffering a loss, compensating pure economic losses may

the legal system would create an over-incentive to take precaution and, hence, produce an inefficient result. In reality, however, it could be difficult to ascertain whether a loss is truly purely economic. A different question is whether there should be a transfer between the firms, which matters for distribution of resources.

English and German law adopt a limiting approach towards pure economic losses (particularly when there are no property losses or physical injury), while French law traditionally shows more generous entitlements.

2.12.E. Non-pecuniary losses

We have seen that perfect compensation serves two goals: it makes the victim indifferent with respect to the accident and makes the injurer internalise the expected accident cost. However, not every kind of loss can be perfectly compensated. Pain and suffering, loss of affection, emotional and moral harm are difficult to translate in monetary terms. Not simply does the actual impossibility of finding a monetary equivalent matter, but it is crucial that such losses generally relate to interests, which are never the object of commercial exchange (Stoll, 1972).

These characteristics lead to the impossibility of any compensation rather than to the practical difficulty of assessing the magnitude of an appropriate compensation. Any amount of money will be insufficient for compensating a parent for the death of his child, or an honest man for a public defamation. Since compensation is actually impossible (Friedman, 1982), then the loss should not be compensated.

On this point there is a variety of different solutions in different legal systems. French law displays a total acceptance of the principle that non-pecuniary loss should be compensated both with respect to injury and death claims. German law starts with a general rejection of non-pecuniary losses, but provides several specific exceptions. The common law system seems oriented toward compensation of injury claims but not of death claims (McGregor, 1972). A similar variety can be found in the case of violation of the rights of the personality, the compensation of which finds wide acceptance in French law, while common and German law adopt a more restrictive approach (Ollier and Le Gall, 1981). Court decisions show, nevertheless, a constant evolution of the jurisprudence in this field.

From the incentives point of view, however, the fact that the injurer does not pay any compensation makes his incentive to take precaution poor (Shavell, 1987, at 133-135). A correct incentive stream can be restored by other means, as noted in section 10.

lead to a better internalization of the harm by the injurer. See Shavell (1987) at 135.

2.12.F. *Low-probability accidents and egg-shell-skull cases*

Is the fact that an accident is extremely unlikely to happen a good reason for excluding liability? The question is more relevant if we consider the case of an extremely unlikely but very harmful accident. A car driver is responsible for an accident and has to pay for the damage he has caused to another car. The accident occurred at a very low speed so that normally the harm would have been about 500 euro. However, the other car was a unique model of a very old Ferrari and it is nearly impossible to find the original components to repair it: the loss amounts to 10000 euro. Similar cases are known as egg-shell-skull cases. When the (physical or even psychological) condition of the victim determines the magnitude of the harm, the principle ‘tortfeasor takes the victim as he finds him’ applies in most of the legal systems, which consequently entitle the victim to full compensation (Honoré, 1971, at 182).

This rule is fully justified also from an economic perspective. Since, even if improbable, the accident is possible, it should be considered among the possible consequences of the injurer’s action, and if the victim is not entitled to full compensation, the injurer will internalise only a portion of the expected accident loss, which we know would lead to sub-optimal precaution. An improbable accident will be discounted by a low probability, but has to be internalised in order to spur to efficient precaution. Therefore, even very-low-probability large-harm accidents should lead to compensation of the victim’s loss (Shavell, 1980b).

2.13. **Concluding remarks.**

This chapter consists of a general survey of the main results obtained by the law and economics approach in the field of tort law. It is meant to give a basic overview of the main framework of analysis and suggestions for further insights. Since so far the analysis has centred mainly on American law, much is still to be said in the economic analysis of tort liability in civil law countries.

2.14. **References.**

- Arlen, Jennifer (2000), ‘Tort Damages’, in Boudewijn Bouckaert and Gerrit De Geest (eds.), *Encyclopedia of Law and Economics*, vol. II, Cheltenham, Edward Elgar.
- Ben-Shahar, Omri (2000), ‘Causation and Foreseeability’, in Boudewijn Bouckaert and Gerrit De Geest (eds.), *Encyclopedia of Law and Economics*, vol. II, Cheltenham, Edward Elgar.

- Bishop, William (1982), "Economic Loss in Tort", **2** *Oxford Journal of Legal Studies*, 1-29.
- Bouckaert, Boudewijn and De Geest, Gerrit (eds.), *Encyclopedia of Law and Economics*, Cheltenham, Edward Elgar.
- Brown, John Prather (1973), 'Toward an Economic Theory of Liability', **2** *Journal of Legal Studies*, 323-350.
- Calabresi, Guido (1961), 'Some Thoughts on Risk Distribution and the Law of Torts', **70** *Yale Law Journal*, 35-46.
- Calabresi, Guido (1970), *The Costs of Accidents: A Legal and Economic Analysis*, New Haven, Yale University Press.
- Calabresi, Guido and Melamed, A. Douglas (1972), 'Property Rules, Liability Rules and Inalienability: One View of the Cathedral', **85** *Harvard Law Review*, 1089-1128.
- Calfee, John E. and Craswell, Richard (1984), 'Some Effects of Uncertainty on Compliance with Legal Standards', **70** *Virginia Law Review*, 965-1003.
- Chung, Tai-Yeong (1993), 'Efficiency of Comparative Negligence: A Game Theoretic Analysis', **22** *Journal of Legal Studies*, 395-404.
- Coase, Ronald H. (1960), 'The Problem of Social Cost', **3** *Journal of Law and Economics*, 1-44.
- Cooter, Robert D. (1989), 'Punitive Damages for Deterrence: When and How Much?', **40** *Alabama Law Review*, 1143-1196.
- Cooter, Robert D. and Ulen, Thomas S. (1986), 'An Economic Case for Comparative Negligence', **81** *New York University Law Review*, 1067-1110.
- Cooter, Robert D. and Ulen, Thomas S. (2000), *Law and economics*, 3rd ed., Reading, Massachusetts, [etc.]: Addison-Wesley.
- Craswell, Richard and Calfee, John E. (1986), 'Deterrence and Uncertain Legal Standards', **2** *Journal of Law, Economics, and Organization*, 279-303.
- Curran, Christopher (1992), 'The Spread of the Comparative Negligence Rule in the United States', **12** *International Review of Law and Economics*, 317-332.
- Dworkin, Ronald M. (1975), *Taking Rights Seriously*, Cambridge (MA), Harvard University Press.
- Edlin, Aaron S. (1994), 'Efficient Standards of Due Care: Should Courts Find More Parties Negligent under Comparative Negligence?', **14** *International Review of Law and Economics*, 21-34.
- Emons, Winand (1990a), 'Efficient Liability Rules for an Economy with Non identical Individuals', **42** *Journal of Public Economics*, 89-104.
- Emons, Winand (1990b), 'Some Recent Developments in the Economic Analysis of Liability Law: An Introduction', **146** *Journal of Institutional and Theoretical Economics*, 237-248.

- Eörsi, Gyula (1975), 'Private and Governmental Liability for the Torts of Employees and Organs', in André Tunc (ed.), *Torts, International Encyclopedia of Comparative Law*, Tübingen, Mohr.
- Faure, Michael (1995), 'The Limits to Insurability from a Law and Economics Perspective', *Geneva Papers on Risk and Insurance*, 454-462.
- Faure, Michael G. and Van den Bergh, Roger (eds.) (1989), *Essays in Law and Economics: Corporations, Accident Prevention and Compensation for Losses*, Antwerpen, Maklu.
- Frezza, Giampaolo and Parisi, Francesco (1998), 'Pietro Trimarchi', in Backhaus, Jürgen G. (ed.), *Elgar Companion to Law and Economics*, Elgar- Ashgate.
- Friedman, David (1982), 'What Is Fair Compensation for Death or Injury?', *2 International Review of Law and Economics*, 81-93.
- Grady, Mark (1990), 'Multiple Tortfeasors and the Economy of Prevention', *19 Journal of Legal Studies*, 653-678.
- Grady, Mark F. (1983), 'A New Positive Economic Theory of Negligence', *92 Yale Law Journal*, 799-829.
- Haddock, David D. and Curran, Christopher (1985), 'An Economic Theory of Comparative Negligence', *14 Journal of Legal Studies*, 49-72.
- Honoré, A.M. (1971), 'Causation and Remoteness of Damage', in André Tunc (ed.), *Torts, International Encyclopedia of Comparative Law*, Tübingen, Mohr.
- Kahan, Marcel (1989), 'Causation and Incentives to Take Care under the Negligence Rule', *18 Journal of Legal Studies*, 427-447.
- Kaplow, Louis (1994), 'The Value of Accuracy in Adjudication: An Economic Analysis', *23 Journal of Legal Studies*, 307-401.
- Kaplow, Louis and Shavell, Steven (1994), 'Accuracy in the Assessment of Liability', *37 Journal of Law and Economics*, 1-15.
- Kaplow, Louis and Shavell, Steven (1996a), 'Accuracy in the Assessment of Damages', *39 Journal of Law and Economics*, 191-209.
- Kaplow, Louis and Shavell, Steven (1996b), 'Property Rules versus Liability Rules: An Economic Analysis', *109 Harvard Law Review*, 713-790.
- Kaplow, Louis and Shavell Steven (2002), 'Economic Analysis of Law', in Haverback, Alan J. and Feldstein, Martin S. eds., *Handbook of Public Economics*, Amsterdam, Elsevier Science.
- Kornhauser, Lewis A. (1982), 'An Economic Analysis of the Choice Between Enterprise and Personal Liability for Accidents', *70 California Law Review*, 1345-1392.
- Kornhauser, Lewis A. and Revesz, Richard L. (2000), 'Joint Tortfeasors', in Boudewijn Bouckaert and Gerrit De Geest (eds.), *Encyclopedia of Law and Economics*, vol. II, Cheltenham, Edward Elgar.

- Kraakman, Reinier H. (2000), 'Vicarious and Corporate Civil Liability', in Boudewijn Bouckaert and Gerrit De Geest (eds.), *Encyclopedia of Law and Economics*, vol. II, Cheltenham, Edward Elgar.
- Krauss, Michael I. (2000), 'Property Rules versus Liability', in Boudewijn Bouckaert and Gerrit De Geest (eds.), *Encyclopedia of Law and Economics*, vol. II, Cheltenham, Edward Elgar.
- Landes, William M. and Posner, Richard A. (1980), 'Multiple Tortfeasors: An Economic Analysis', **9** *Journal of Legal Studies*, 517-555.
- Landes, William M. and Posner, Richard A. (1981), 'An Economic Theory of Intentional Torts', **1** *International Review of Law and Economics*, 127-154.
- Landes, William M. and Posner, Richard A. (1983), 'Causation in Tort Law: an Economic Approach', **12** *Journal of Legal Studies*, 109-134.
- Landes, William M. and Posner, Richard A. (1984), 'Tort Law as a Regulatory Regime for Catastrophic Personal Injuries', **13** *Journal of Legal Studies*, 417-434.
- Landes, William M. and Posner, Richard A. (1985), 'A Positive Economic Analysis of Products Liability', **14** *Journal of Legal Studies*, 535-567.
- Landes, William M. and Posner, Richard A. (1986), 'New Light on Punitive Damages', **10** *Regulation*, 33-36.
- Landes, William M. and Posner, Richard A. (1987), *The Economic Structure of Tort Law*, Cambridge (MA), Harvard University Press.
- Le Gall, Jean-Pierre (1976), 'Liability for Persons under Supervision', in André Tunc (ed.), *Torts, International Encyclopedia of Comparative Law*, Tübingen, Mohr.
- Limpens, Jean, Kruithof, Robert M. and Meinertzhagen-Limpens, Anne (1979), 'Liability for One's Own Act', in André Tunc (ed.), *Torts, International Encyclopedia of Comparative Law*, Tübingen, Mohr.
- Malloy, Robin Paul (1990), *Law and Economics: A Comparative Approach to Theory and Practice*, St. Paul, West Publishing Company.
- Marks, Stephen V. (1994), 'Discontinuities, Causation, and Grady's Uncertainty Theorem', **23** *Journal of Legal Studies*, 287-301.
- Mattei, Ugo (2000), *Comparative Law and Economics*, Ann Arbor, MI, University of Michigan Press.
- Mattei, Ugo and Pardolesi, Roberto (1991), 'Law and Economics in Civil Law Countries: A Comparative Approach', **11** *International Review of Law and Economics*, 265-275.
- Mattei, Ugo, Antonioli, Luisa and Rossato, Andrea (2000), 'Comparative Law and Economics', in Boudewijn Bouckaert and Gerrit De Geest (eds.), *Encyclopedia of Law and Economics*, vol. II, Cheltenham, Edward Elgar.
- McEwin, R. Ian (2000), 'No-Fault Compensation Systems', in Boudewijn Bouckaert and Gerrit De Geest (eds.), *Encyclopedia of Law and Economics*, vol. II, Cheltenham, Edward Elgar.

- McGregor, Harvey (1972), 'Personal Injury and Death', in André Tunc (ed.), *Torts, International Encyclopedia of Comparative Law*, Tübingen, Mohr.
- Miceli, Thomas J. (1997), *Economics of the Law: Torts, Contracts, Property, Litigation*, Oxford, Oxford University Press.
- Ogus, Anthony I. (1994), *Regulation: Legal Form and Economic Theory*, Oxford, Clarendon.
- Ollier, Pierre-Dominique and Le Gall, Jean-Pierre (1981), 'Various Damages', in André Tunc (ed.), *Torts, International Encyclopedia of Comparative Law*, Tübingen, Mohr.
- Orr, Daniel (1991), 'The Superiority of Comparative Negligence: Another Vote', **20** *Journal of Legal Studies*, 119-129.
- Parisi, Francesco (1994), 'Alterum non Laedere: An Intellectual History of Civil Liability', **39** *American Journal of Jurisprudence*, 317-351.
- Parisi, Francesco and Posner, Richard A. (1997), *International Library of Critical Writings in Economics, Volume II: Contracts, Torts and Criminal Law*, London, Edward Elgar.
- Png, Ivan Paak-Liang (1986), 'Optimal Subsidies and Damages in the Presence of Judicial Error', **6** *International Review of Law and Economics*, 101-105.
- Polinsky, A. Mitchell (1989), *An Introduction to Law and Economics*, 2nd ed., Boston: Little, Brown and Company.
- Polinsky, A. Mitchell and Che, Yeon-Koo (1991), 'Decoupling Liability: Optimal incentives for Care and Litigation', **22** *Rand Journal of Economics*, 562-570.
- Polinsky, A. Mitchell and Rogerson, William P. (1983), 'Product Liability, Consumer Misperceptions, and Market Power', **14** *Bell Journal of Economics*, 581-589.
- Polinsky, A. Mitchell and Shavell, Steven (1998), 'Punitive Damages: An Economic Analysis', **111** *Harvard Law Review*, 869-962.
- Polinsky, A. Mitchell and Shavell, Steven (2000), 'Punitive Damages', in Boudewijn Bouckaert and Gerrit De Geest (eds.), *Encyclopedia of Law and Economics*, vol. II, Cheltenham, Edward Elgar.
- Polinsky, Mitchell A. and Shavell, Steven (1994), 'Should Liability be Based on the Harm to the Victim or the Gain to the Injurer?', **10** *Journal of Law, Economics, and Organization*, 427-437.
- Posner, Richard A. (1973), 'Strict Liability: A Comment', **2** *Journal of Legal Studies*, 205-221.
- Posner, Richard A. (1998), *Economic Analysis of Law* (5th ed.), New York, Aspen Publishers, Inc..
- Posner, Richard A. (ed.) (1977), *Tort Law: Cases and Economic Analysis*, Boston, Little Brown.
- Rabin, Robert L. (1985), 'Tort Recovery for Negligently Inflicted Economic Loss: A Reassessment', **37** *Stanford Law Review*, 1513-1538.
- Rabin, Robert L. (1993), *Perspectives on Tort Law* (3rd ed.), Boston, Little Brown.

- Radin, Margaret J. (1987), 'Market Inalienability', **100** *Harvard Law Review*, 1849-1937.
- Rea, Samuel A., Jr. (1987), 'The Economics of Comparative Negligence', **7** *International Review of Law and Economics*, 149-162.
- Rizzo, Mario J. (1982), 'A Theory of Economic Loss in the Law of Torts', **11** *Journal of Legal Studies*, 281-310.
- Rose-Ackerman, Susan (1985), 'Inalienability and the Theory of Property Rights', **85** *Columbia Law Review*, 931-969.
- Rubinfeld, Daniel L. (1987), 'The Efficiency of Comparative Negligence', **16** *Journal of Legal Studies*, 375-394.
- Schäfer, Hans-Bernd (2000), 'Tort Law: General', in Boudewijn Bouckaert and Gerrit De Geest (eds.), *Encyclopedia of Law and Economics*, vol. II, Cheltenham, Edward Elgar.
- Schäfer, Hans-Bernd and Schönemberger, Andreas (2000), 'Strict Liability versus Negligence', in Boudewijn Bouckaert and Gerrit De Geest (eds.), *Encyclopedia of Law and Economics*, vol. II, Cheltenham, Edward Elgar.
- Shavell, Steven (1979), 'On Moral Hazard and Insurance', **93** *Quarterly Journal of Economics*, 541-562.
- Shavell, Steven (1980a), 'Strict Liability versus Negligence', **9** *Journal of Legal Studies*, 1-25.
- Shavell, Steven (1980b), 'An Analysis of Causation and the Scope of Liability in the Law of Torts', **9** *Journal of Legal Studies*, 463-516.
- Shavell, Steven (1982), 'On Liability and Insurance', **13** *Bell Journal of Economics*, 120-132.
- Shavell, Steven (1984), 'Liability for Harm versus Regulation of Safety', **13** *Journal of Legal Studies*, 357-374.
- Shavell, Steven (1985), 'Uncertainty over Causation and the Determination of Civil Liability', **28** *Journal of Law and Economics*, 587-609.
- Shavell, Steven (1986), 'The Judgment Proof Problem', **6** *International Review of Law and Economics*, 45-58.
- Shavell, Steven (1987), *Economic Analysis of Accident Law*, Cambridge (MA), Harvard University Press.
- Shavell, Steven (1992), 'Liability and the Incentives to Obtain Information about Risk', **21** *Journal of Legal Studies*, 259-270.
- Spence, A. Michael (1977), 'Consumer Misperceptions, Product Failure and Product Liability', **64** *Review of Economic Studies*, 561-572.
- Stoll, Hans (1972), 'Consequences of Liability: Remedies', in André Tunc (ed.), *Torts, International Encyclopedia of Comparative Law*, Tübingen, Mohr.
- Sykes, Alan O. (1981), 'An Efficiency Analysis of Vicarious Liability Under the Law of Agency', **91** *Yale Law Journal*, 168-206.

Sykes, Alan O. (1984), ‘The Economics Vicarious Liability’, *93 Yale Law Journal*, 1231-1282.

Trimarchi, Pietro (1959a), “Il caso fortuito quale limite della responsabilità per il danno da cose. Contributo and una teoria del rischio di impresa” (Fortuitous Event as a Limit of the Liability for Damage Caused by Things. Contribution to a Theory of Entrepreneurial Risk), *Rivista trimestrale di diritto e procedura civile*, 808 ff.

Trimarchi, Pietro (1959b), “La responsabilità per il fatto dei dipendenti. Contributo and una teoria del rischio di impresa” (Liability for Damage Caused by Employees. Contribution to a Theory of Entrepreneurial Risk), *Rivista di diritto civile*, 604 ff.

Trimarchi, Pietro (1961), *Rischio e responsabilità oggettiva* (Risk and Strict Liability), Milano, Giuffrè.

Trimarchi, Pietro (1967), *Causalità e danno* (Causality and Damage), Milano, Giuffrè.

Figures

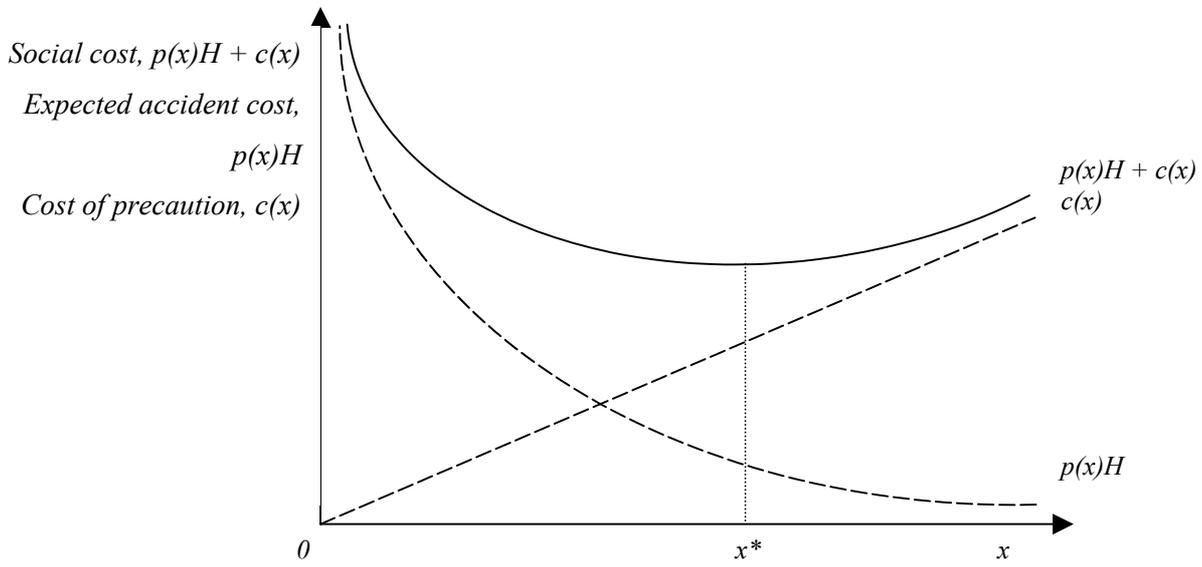


Figure 1: the basic model

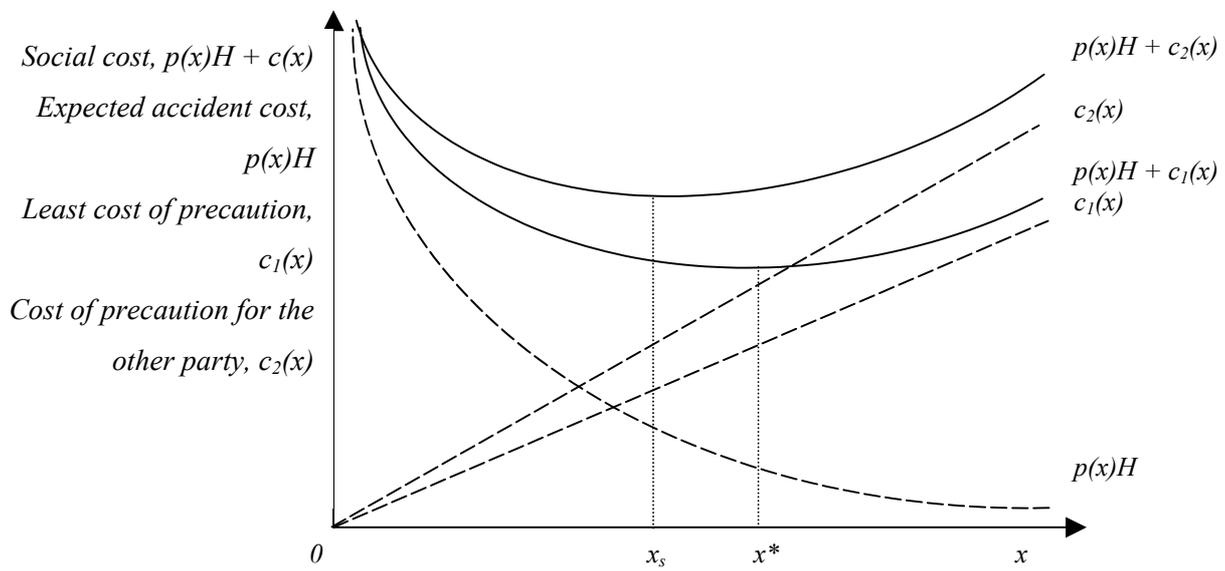


Figure 2: alternative precaution

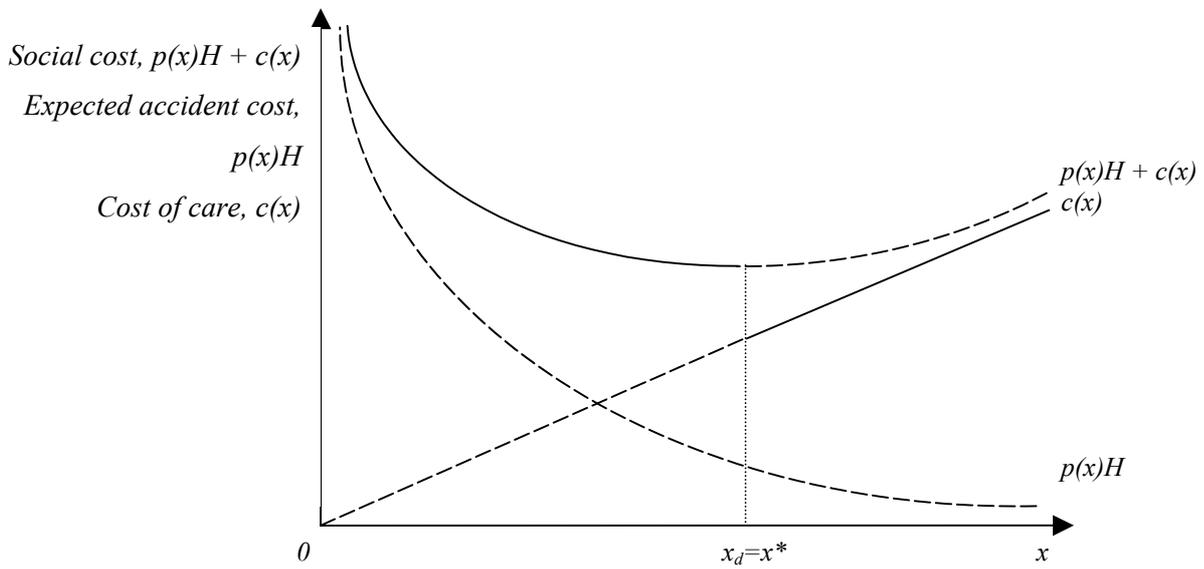


Figure 3: fault-based liability

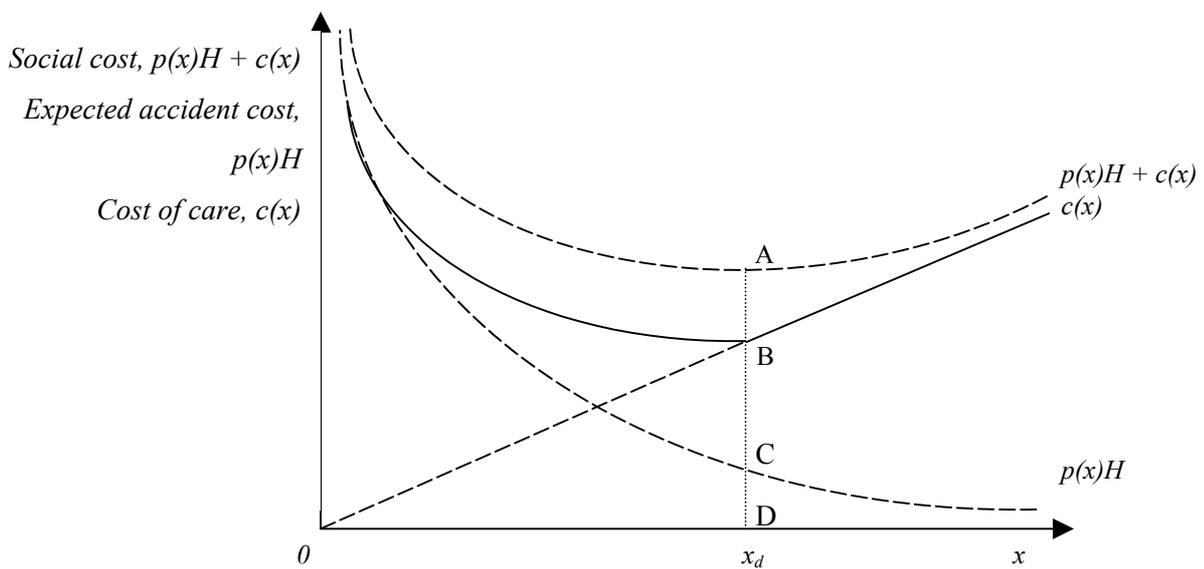


Figure 4: causation and negligence

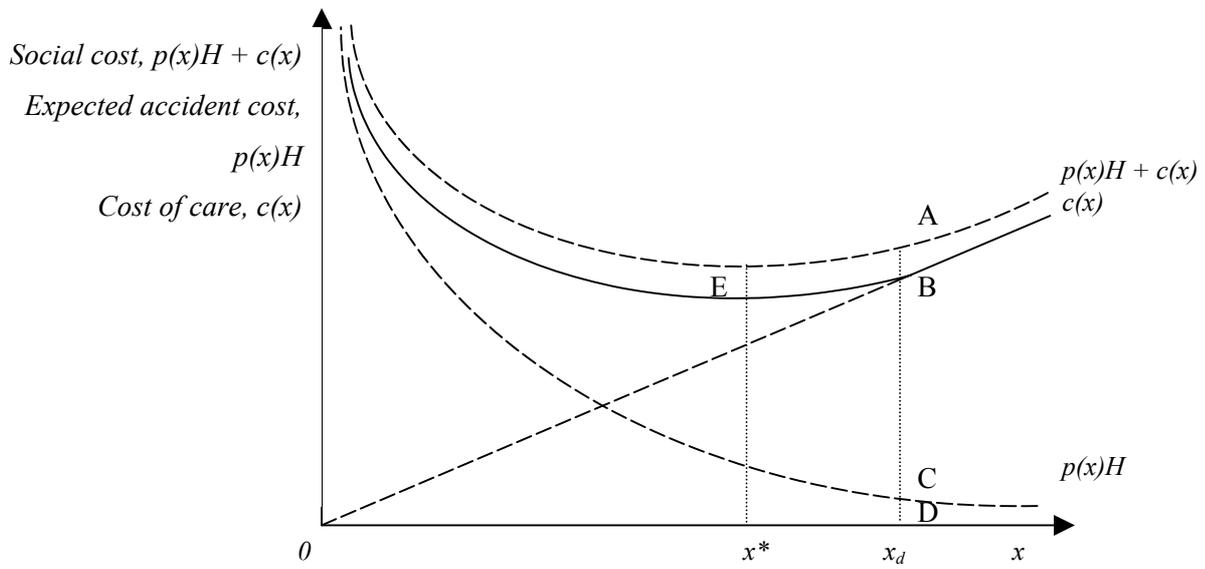


Figure 5: causation and negligence – due care higher than optimal care

ON THE DEFINITIONS OF CARE AND ACTIVITY LEVEL AND THE CHOICE OF LIABILITY RULES*

Giuseppe Dari Mattiacci

Utrecht University and Marie Curie Fellow at Hamburg University

ABSTRACT

Stat rosa pristina nomine, nomina nuda tenemus

Bernard of Morlay, *De contemptu mundi*¹

In this chapter, I show that the traditional definitions of care (as precaution) and activity level (as number or frequency of an action) are tautological and difficult to apply, as any precautionary measure always fits both categories. The correct definitions depend on the determination of negligence: precautionary measure included in the negligence criterion are to be defined as care, while precautionary measures that are not included in the negligence criterion are to be defined as activity level. As controlling parties' precaution through the negligence inquiry triggers an administrative cost, it is efficient to limit the extent of the negligence criterion to some precautionary measures only. This framework explains the evolution of liability systems over time and the choice of liability rules in different areas of modern tort law and permits us to reinterpret the two basic theorems of tort law and economics (the efficiency-equivalence theorem and the activity-level theorem) and the traditional argument concerning "strict liability vs. negligence".

JEL classification: K00, K13.

Keywords: tort, liability, activity level, care, precaution, negligence, historical development.

* Giuseppe Dari Mattiacci, Economic Institute/CIAV, Utrecht University, Kromme Nieuwegracht 22, 3512 HH Utrecht, the Netherlands; email: g.darimattiacci@econ.uu.nl. I am deeply indebted to Alessandra Arcuri, Gerrit De Geest and Tina Heubeck, with whom I discussed repeatedly the early ideas of this analysis. I thank Eimaer Coffie, Thomas Eger, Antony Ogus, Francesco Parisi, Hans-Bernd Schäfer, Wouter Wils, and the participants in the Marie-Curie workshop at Hamburg University and in the CASLE workshop at Ghent University, who provided helpful comments. Financial support for this research was provided by the European Commission, Marie Curie Fellowship Program 2001 at Hamburg University, that I gratefully acknowledge. Updated versions of this chapter will possibly be available on my web page: www.law.uu.nl/eco/gdm.htm.

¹ As quoted in Eco, Umberto (1980), *Il nome della rosa*, Milano, Bompiani.

3.1. Introduction

When both the injurer and the victim of an accident can reduce the expected accident loss, legal systems usually adopt a negligence rule in order to provide an appropriate incentive to both of them. One party is subject to a negligence inquiry and will pay the accident loss only if found negligent, while the other party bears the loss in the residual cases.

I will present a model that considers accidents between two parties who can take many different precautionary measures in order to reduce the expected accident loss. The optimal liability rule is defined as that rule which minimizes expected the accident loss, precaution expenditures and administrative costs. As controlling parties' precaution by means of the negligence criterion triggers administrative costs, the optimal extent of the negligence criterion is determined by a balance between those costs and the resulting advantage in terms of improved incentives to prevent accidents. I will demonstrate that it is efficient not to control those precautionary measures which trigger high administrative costs and have little impact on the reduction of the expected accident loss.

Such an approach permits us to redefine the concepts of care and activity level, to interpret the evolution of liability systems and to explain the choices made by modern legal systems between different liability rules in specific cases. These are my points:

- The decision of which precautionary measures should be controlled under the negligence inquiry generates the definitions of care and activity level: the precautionary measures included in the negligence criterion will be defined as care, while the precautionary measures excluded from the negligence criterion will be defined as activity level. Those definitions are endogenous to the model and are variable with the negligence criterion. The traditional exogenous definitions of care (precaution) and activity level (the number of repeated actions) are tautological and, hence, difficult to apply.
- At an initial stage, liability systems prefer strict rules to negligence rules, as the cost of applying the negligence criterion is too high. The choice between strict liability and no liability can be made by using a dangerousness criterion: the residual bearer should be that party whose actions affect the expected accident loss most.
- At a later stage, the development of the judiciary system and of society in general will result in a decrease of the costs associated with the negligence criterion. Legal systems tend to move towards a broader application of the negligence criterion. The choice between no-liability-based rules and strict-liability-based rules cannot be based only on the dangerousness criterion. As a negligence criterion is usually applied, the verifiability

of parties' behavior should also be taken into account: the residual bearer should be that party whose actions have a greater effect on the expected accident loss and whose precautionary measures are less verifiable before the court.

- In modern societies, we can observe a general application of no-liability-based negligence rules, such as simple negligence, as this solution economizes compensation costs; however, many areas of torts are still subject to strict liability. Those cases do not always concern dangerous activities and cannot be easily explained in terms of the dangerousness criterion only, as it has been attempted in the literature. They prove to be consistent with my approach and can be justified by the verifiability criterion mentioned above.

The next section criticizes the traditional definitions of care and activity level. Sections 3.2 to 3.4 present the reader with a formal model, which differs from the traditional model in two respects: first, it considers many different precautionary measures and not a unique and homogeneous form of precaution; second, it considers administrative costs while determining the extent of the negligence criterion. Section 3.5 discusses the new definitions of care and activity level, as generated by the model itself. Section 3.6 applies the model to the study of the evolution of liability systems over time, with two specific applications: the evolution of Roman and common law and a discussion of the industry-subsidy argument for early industrial negligence-based liability. Section 3.7 applies the same framework to an analysis of the choice between different liability rules in modern tort law. In section 3.8, I will provide a synthesis of my findings, their relation to the literature and some concluding remarks.

3.1.A. A criticism of the traditional definitions of care and activity level

In the Law and Economics literature on tort law, accident losses are said to be determined by parties' care and activity level. Care is defined as precaution reducing the expected accident loss at a certain cost, while activity level is defined as the number of repeated actions yielding a benefit but increasing the occurrence of accidents².

² The distinction between care and activity level was originally introduced by Shavell (1980a), and somewhat sketched by Posner (1973) at 208. Shavell (1987), at 5, provides the following definition: "The number of miles an individual drives, for instance, might be interpreted as his level of activity, and the precaution he takes when on the road (slowing for curves, paying attention to the presence of bicyclists) as his level of care. Similarly, how often a bicyclist rides where there is automobile traffic might be regarded as his level of activity, and his precaution when riding (staying close to the side of the road, using a brightly colored vest) as his level of care". Similar descriptions have been explicitly or implicitly adopted in the literature. Miceli (1997), at 27, defines activity level as "how frequently or intensively to engage in a risky activity", and writes, "For example, the driver of an automobile decides how carefully to drive, but also how often and how many miles". Landes and Posner (1981) at 851, 875-878, and (1987), at 61, and Posner (1998) adopt implicitly the same definition. Cooter and Ulen (2000), at 311, define activity level as the "amount" of one's action. The same approach can be found in Epstein (1999), at 97. The definition of care is generally a straightforward identification with the common concept of precaution.

The literature further assumes that the court cannot verify the activity level³ and concludes that, under any of the negligence rules, both parties take optimal care (efficiency-equivalence theorem⁴) but only the residual bearer chooses optimal activity level (activity-level theorem⁵).

Table 1: The traditional definitions of care and activity level and a criticism thereof

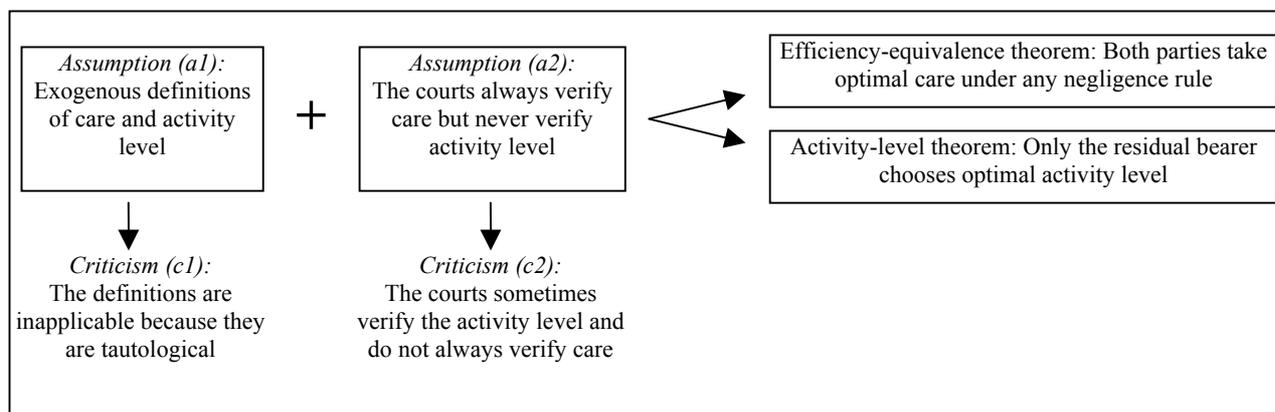


Table 1 summarizes such a line of reasoning: (a1) define care and activity level on the basis of the (exogenous) characteristics of the parties' actions (care as precaution and activity level as the number of repeated actions); (a2) assume that the court cannot verify⁶ the activity level. Hereafter, I will show the shortcomings of these two steps and will propose defining care and activity level endogenously, on the basis of the negligence criterion⁷.

³ The court cannot verify the activity level, or measure the cost of reducing it, or it is prevented from making a statement on activity level for some other reasons (Shavell, 1987, at 25, 50, 56, and 57, Landes and Posner, 1987, at 66-67, Miceli, 1997, at 28). The same justification applies to the measures of care excluded from the determination of negligence. See Gilles (1992) for a criticism of this assumption.

⁴ Landes and Posner (1980).

⁵ Shavell (1980a). Only the residual bearer, that is the party who bears the accident costs when both parties are non-negligent, has an incentive to internalize that cost in the decision on his activity level, and will decide optimally. The other party escapes liability by simply taking due care and will be likely to engage in too high a level of activity, for the reason that he does not bear any cost. Strict-liability-based negligence rules spur injurer's optimal activity levels, whereas no-liability-based negligence rules induce victim's optimal activity levels. Decoupling liability would give both parties the appropriate incentive with respect to care and activity level (Polinski and Che, 1991).

⁶ Latin (1987) and Donohue (1989) criticize the analysis of activity level in Shavell (1980a and 1987) and Landes and Posner (1987), but they do not dispute the plausibility of step (a2). Gilles (1992) criticizes step (a2) by demonstrating that the courts actually base their decisions not only on care, but also on precautionary measures, which are generally considered to fall within the definition of activity level. However, by doing so, he implicitly accepts that step (a1) can, in one way or another, be a valid basis for analysis. He recognizes, at 329 to 336, that activity level and care are not mutually exclusive concepts *when* there is a close substitute for the activity (e.g. using a bicycle instead of using a car); I go somewhat further and argue that they are *always* interchangeable. This permits me to discuss the problem put forward by Gilles (the courts sometimes consider precautions commonly understood as activity level), and its mirror twin (the courts do not always consider all those precautions, which are ordinarily regarded as care, see Shavell, 1987, at 9 and 17). Gilles argues that the assumption made in step (a2) is unrealistic; I argue that it is unnecessary, as the real drawback lies hidden in step (a1). This is the crucial point of my analysis as the endogeneity and variability of the definitions follow from the rejection of step (a1).

⁷ This point is not controversial, as the literature has always implicitly recognized that the distinction between care and activity level depends on the negligence criterion (Shavell, 1980a and 1987, at 9 and 17, Miceli, 1997, at 28, Kaplow and Shavell, 2002), but it has never investigated the effect of making such an implicit definition explicit. This is my aim in this chapter.

The first step (a1) is bound to fail. A decrease in the activity level reduces the expected accident loss at the cost of some foregone benefits; however, an action that reduces the occurrence of accidents and produces costs has been defined above as care. By starting with a case of activity-level reduction, we end up with care. We do not need to show that care can be reinterpreted similarly as activity level. In general, any precautionary measure, which is taken to reduce the expected accident loss, can be regarded both as care and as activity level⁸.

Example 1. Consider the case of a man returning home from work by bicycle. He can take a safe but long cycle track or can choose the shorter but more dangerous route along the highway. The frequency with which he chooses the second option can be interpreted both as a lack of care or as an excessive activity level. In fact, taking the shorter route is indeed an activity yielding a utility and increasing the occurrence of accidents; therefore, taking the longer path more frequently can be considered as a reduction of the activity level. However, taking the longer path is certainly a form of precaution reducing the expected accident loss at a certain opportunity cost (the additional time required); hence, it can be defined as care. The traditional definitions of care and activity level do not assist us in clarifying whether a particular action belongs to one category or the other.

It is clear that the definitions given above do not permit us to distinguish between care and activity level⁹. The reason is that they are tautological: they state that care is precaution (a synonym of care) and that the activity level is the number of repeated actions (a paraphrase for activity level).

Even if we assumed that step (a1) succeeds in distinguishing care from activity level, step (a2) would result in an ad hoc assumption to support the activity-level theorem. It is obvious that the court, on the one hand, can sometimes verify the activity level while, on the other hand, it is not always able to verify care¹⁰.

I argue that the real distinction between care and activity level depends on the negligence criterion. Both care and activity level are forms of precaution, which is an expensive way of reducing the expected accident loss. Care denotes those precautionary measures which are

⁸ See also Diamond (1974) at 110, “The distinction [between care and activity level] is somewhat artificial in that we could define negligent driving as a different activity from non-negligent driving”. Nevertheless, Diamond adopts slightly different concepts of care and activity level from those adopted by Shavell (1980a).

⁹ They are not mutually exclusive concepts, see Gilles (1992), supra note 6.

¹⁰ Shavell (1980a) at 354-355 answers the question of why “activity level isn’t usually considered in the formulation of the due care standard” by recognizing the courts’ difficulty in doing so and accounts for measures of precaution excluded from the standard of due care, which are to be considered as activity level. See Landes and Posner (1987) at 70-71, on the fact that the courts actually condition negligence on the number of repeated actions when it is possible to do so. Terry (1915) quotes the case of a man, who, after taking any (other) possible precaution, went upon the tracks to save a child but was killed by an oncoming train. The jury found him not guilty of contributory negligence (*Eckert vs. Long Island R. R. Co.*, 43 N.Y., 1871). An opposite decision would have been taken were the creature a kitten (Terry, 1915, at 43-44, see also Posner, 1992, at 252-253). Not attempting a rescue is indeed a form of precaution, at the cost of a life’s value. Going upon the tracks is hence considered in the negligence inquiry, although at first sight we would say that it is an activity level. Gilles (1992) constructs a strong argument on this point and gives many examples. He shows that “the courts should be able to regulate many activity-level choices by developing rules concerning reasonableness or unreasonableness of particular activities as well as their timing, place and extent” and indeed “modern American negligence law regulates activity levels to a considerably greater extent than has previously been recognized” (Gilles, 1992, at 320). See also Grady (1983).

included in the determination of negligence, while activity level denotes those precautionary measures which are omitted from the negligence inquiry¹¹.

Care \equiv precautionary measures included in the negligence criterion

Activity level \equiv precautionary measures omitted from the negligence criterion

Example 2. Let us consider again the situation described in example 1. The problem of defining activity level and care becomes easy to solve. If the court takes the fact of cycling along the highway into consideration in the determination of negligence, then that is to be considered as care. Otherwise, it is to be considered as activity level¹².

3.2. The model: assumptions and notation

We consider accidents between a victim and an injurer, strangers to each other, rational and risk neutral. They minimize the sum of precaution costs and expected accident loss that they bear under a given liability regime. The victim is the party that suffers harm; the injurer is the other party. Both parties can reduce the expected accident loss¹³ by taking many different precautionary measures, among which they allocate their expenditures optimally¹⁴.

Let:

P = vector of injurer's precautionary measures, $P = \{p_1, \dots, p_z\}$, $p_n \in [0, \infty)$,
 $n = 1, \dots, z$.

R = vector of victim's precautionary measures, $R = \{r_1, \dots, r_s\}$, $r_m \in [0, \infty)$,
 $m = 1, \dots, s$.

x_n = injurer's expenditure when the injurer takes the first n precautionary measures from the vector P , $x_n \in [0, \infty)$.

y_m = victim's expenditure when the victim takes the first m precautionary

¹¹ Gilles (1992) distinguishes between case-specific (ex post) and rule-based (ex ante) techniques to determine negligence, and argues that the choice between the two depends on the cost of evaluating precaution. As information costs become higher, "the courts are more likely to employ the techniques of rule-based negligence to make such inquiries possible" (Gilles, 1992, at 336). For this reason, as precautionary measures traditionally regarded as activity level are likely to be more costly to verify, they are usually dealt with through rule-based negligence (Gilles, 1992, at 337).

¹² Other examples might be easily thought of. Is the number of trains running between two cities to be considered as an activity level, or is the appropriate reduction of the train traffic on a specific track to be regarded as care? We could argue that this amounts to activity level, as it is clearly the case of a repeated action; but we could also argue that this is care, as the number of accidents can be reduced by reducing the frequency of trains, which is indeed a costly precaution to take. Similarly, is the frequency of traveling by car (instead of by bus) an activity level or is the decision whether to take a bus or to travel by car inherent to the level of care? An answer to those questions will always depend on the extent of the negligence criterion.

¹³ The model applies to the most general case in which both parties can take precautions (typically referred to as joint-care accidents). However, it is also suitable for describing unilateral-care and alternative-care accidents, in which one party's optimal precaution will always be zero.

¹⁴ We assume that the parties allocate the expenditure among a certain number of precautionary measures so that the expected accident loss cannot be further reduced, given the level of expenditure and the precautionary

measures from the vector R , $y_m \in [0, \infty)$.

$l(x_n, y_m)$ = expected accident loss; $l' < 0$, $l'' > 0$, with respect to x and y .

$L(\bullet)$ = allocative loss, $L' \leq 0$ ¹⁵: the difference between the total accident costs (expected accident loss plus precaution expenditures) when both parties take all their precautionary measures at the optimal level and the total accident costs when some precautionary measures are not taken. The more precautionary measures are taken, the more the allocative loss decreases. We will refer respectively to the injurer and to the victim by using the subscripts j and v .

$I(\bullet)$ = information costs; $I' \geq 0$: the administrative costs of ascertaining parties' precaution. Information costs increase the more precautionary measures the court has to verify. We will refer respectively to the injurer and to the victim by using the subscripts j and v .

K = compensation costs: the administrative costs of damage compensation.

3.3. Torts in a world with no administrative cost

The model is identical to the standard model adopted in the literature¹⁶, except for one aspect: it explicitly considers the possibility that injurers and victims might be able to take more than one precautionary measure, among which they allocate their precautionary expenditures optimally¹⁷. A precautionary measure is any action that parties can take in order to reduce the expected accident loss. We also refer, therefore, to those measures that under the traditional view would be regarded as activity level (as for example a reduction in the frequency of an activity).

The vector of the injurers' precautionary measures is denoted by P . The vector contains z different precautionary measures given in a certain order¹⁸. We make no assumption on the relationship between them: they may be substitutes, complements or simply independent measures. The injurer's precaution expenditure is denoted as x_n . The subscript n indicates how many precautionary measures are taken by the injurer: x_3 means, for example, that he spends a certain amount of money, x , and that he allocates x among precaution 1, precaution 2 and

measures taken.

¹⁵ $L' \leq 0$ follows from the proof of proposition 1.

¹⁶ See Landes and Posner (1980), Shavell (1980a), Miceli (1997) and Cooter and Ulen (2000).

¹⁷ See footnote 14.

¹⁸ The problem of how to order precautionary measures in the vector does not affect my results and it is easily dealt with in footnote 30.

precaution 3.¹⁹

The same applies to the victim, who selects his m precautionary measures from a different vector, R , composed of s precautionary measures, and allocates an expenditure y optimally among them.

3.3.A. The social cost in a world with no administrative cost

There are two choices to be made: *how many* precautionary measures to take, and *how much* to spend on precaution in order to minimize the social cost (the sum of both parties' precaution expenditures and expected accident loss).

Proposition 1. When there are no administrative costs, the total social cost is at its minimum if both parties choose the optimal level of precaution expenditures with respect to all their precautionary measures. The total social cost increases (or remains constant) if fewer precautionary measures are taken.

Proof.

The total social cost, when no administrative costs are taken into account, is the sum of expected accident loss and precaution expenditures. The proposition states

$$(1) \quad \min_{x,y,n,m} [l(x_n, y_m) + x_n + y_m] = \min_{x,y} [l(x_z, y_s) + x_z + y_s] = l(x_z^*, y_s^*) + x_z^* + y_s^*.$$

The first step (going from the first to the second part of the expression) refers to the choice of *how many* precautionary measures should be taken and states that the social cost is minimized when the parties' expenditures are allocated among *all* their precautionary measures. The reason is that, by reallocating the same expenditure among more precautionary measures, the parties can further reduce the expected accident loss²⁰.

The second step (going from the second to the third part of the expression) refers to the choice of *how much* should be spent on precautionary measures. The solution can be found by applying ordinary minimization techniques and denoted as (x_z^*, y_s^*) ²¹.

¹⁹ The allocation of the sum x among those three precautionary measures is left to the injurer and we assume that he will allocate it optimally, i.e. any different allocation will result in a higher (or the same) expected accident loss. Clearly, the outcome might be precaution 1 absorbing the whole expenditure x , while precautions 2 and 3 are taken at a level equal to zero. See also footnote 14.

²⁰ It is obvious that the expected accident loss never increases if the same precautionary expenditure is allocated among a greater number of precautionary measures, because if the new precautionary measures are not worth taking, then the parties will continue to spend only on the old ones. Taking an additional precautionary measure is equivalent to using better (or at least equivalent) accident-prevention technology.

²¹ The solution is internal and unique, as it follows from the assumptions made.

3.3.B. Liability rules in a world with no administrative cost

If there are no administrative costs, the optimal liability rule is that rule which induces both parties to take all their precautionary measures and to select the socially optimal level of expenditure. Here we consider four possible rules: two strict rules (strict liability and no-liability) and two negligence rules (strict liability with defense of contributory negligence, which is a strict-liability-based negligence rule, and simple negligence, which is a no-liability-based negligence rule). The analysis can be simplified as indicated in table 2.

Table 2: The choice of the optimal liability rule

		<i>Residual bearer</i>	
		Injurer <i>(strict-liability-based rules)</i>	Victim <i>(no-liability-based rules)</i>
<i>Extent of the negligence criterion for the non-residual bearer</i>	0	Strict liability	No liability
	s	Strict liability with defense of contributory negligence	Simple negligence

The legal system makes two choices. First, it decides whether to implement a strict rule or a negligence rule (vertical choice in the table). In the latter case, the legal system controls the behavior of the non-residual bearer through negligence and has to determine the extent of the negligence inquiry. It is a choice of the *degree* to which the legal system verifies the parties' behavior. The legal system can verify none (strict rule), one (a very limited negligence rule), two, or even all of the parties' precautionary measures (omni-comprehensive negligence rule). The analysis determines the optimal extent of the negligence criterion in the hypothesis that the victim is the residual bearer (optimal no-liability-based rule) and in the hypothesis that the injurer is the residual bearer (optimal strict-liability-based rule).

Second, the legal system selects either the victim (no-liability-based rules) or the injurer (strict-liability-based rules) as the residual bearer (horizontal choice in the table). Hence, it chooses between the optimal no-liability-based rule and the optimal strict-liability-based rule.

3.3.B.I. Vertical choice: allocative loss and determination of the optimal extent of the negligence criterion under no-liability-based rules (the victim is the residual bearer) in a world with no administrative cost

Under no liability, the legal system makes the victim bear the expected accident loss

irrespective of his (negligent) behavior. The victim is the residual bearer. A negligence criterion can be added in order to control the injurer's behavior (simple negligence²²). The negligence criterion lists the precautionary measures that the injurer has to take in order not to be found negligent as well as their level. We denote the vector of such precautionary measures as due care, C_n . The subscript n refers to *how many* precautionary measures are included in the vector. The superscript c refers to their due *level*.

$$(2) \quad C_n = \{p_1^c, \dots, p_n^c\}.^{23}$$

The non-negligent injurer is supposed to take all the precautionary measures listed at the required level²⁴. We denote as x_n^c the expenditure triggered by due care. A motorist, for example, has to fulfill a detailed list of requirements: speed, the condition of the car, the use of headlights and indicators, traffic lights, and so forth. For each of them a certain level is required. If the motorist fails to fulfill one of these requirements, he will be found negligent. Simple negligence results in the injurer being liable if he fails to fulfill the due-care requirements and non-liable otherwise.

For the sake of simplicity, we assume that the *level* the n precautionary measures included in the negligence criterion is set optimally²⁵ and that the injurer will take due care. The victim is the residual bearer and we denote the victim's precaution expenditure as y_s^c ²⁶. Therefore, under simple negligence, a negligence criterion set accordingly to Exp. (2) yields parties' precautionary expenditures equal to (x_n^c, y_s^c) ²⁷.

We define as the allocative loss of a liability rule the difference between the total social cost under that liability rule and the minimum social loss as defined in Exp. (1). Let us consider the allocative loss associated with the no-liability-based rule under examination. The

²² Although contributory negligence and comparative negligence also subject the victim's behavior to a negligence inquiry, the victim is the residual bearer, and he will hence take precaution with respect to all his precautionary measures (as shown by Shavell, 1980a). As a result, there is no difference with respect to simple negligence and only with respect to injurers does a distinction between activity level and care have a meaning.

²³ The negligence criterion targets some specific actions on the part of the injurer. In theory, the court could verify directly expenditure x , instead of verifying the injurer's behavior as regards many different precautionary measures. However, the cost of precaution is often non-monetary, and although it can sometimes be estimated, it is very difficult to verify. When it is possible to verify directly the costs, the problem becomes easier to solve.

²⁴ The problem of ordering the precautionary measures does not affect the results of the analysis. See footnote 30.

²⁵ We define as the optimal level that level which minimizes total accident loss (precaution expenditures plus expected accident loss), given n and the reaction of the victim. Hence, the due level of the n precautionary measures included in the negligence criterion is optimal if x_n^c minimizes $l(x_n, y_s^c) + x_n + y_s^c$.

²⁶ It follows for proposition 1 that the victim will take all s precautionary measures. Hence, the victim takes such a level of the s precautionary measures that $l(x_n^c, y_s) + y_s$ is minimized. Let y_s^c denote the solution to the former problem. That solution is internal and unique for the assumptions made on the functional form.

²⁷ This conclusion holds both under the traditional interpretation of the negligence criterion (a negligent injurer pays damages for any accident loss, which might occur) and under a causation-corrected model of negligence (a negligent injurer pays damages only for those accident losses, which would not have occurred had he been non-negligent, see Grady, 1983, and Kahan, 1989).

subscript j indicates that the loss occurs because the injurer does not take all his precautionary measures.

$$(3) \quad L_j(n) = [l(x_n^c, y_s^c) + x_n^c + y_s^c] - [l(x_z^*, y_s^*) + x_z^* + y_s^*].$$

From proposition 1, it follows that the first part of the right-hand side is greater than or equal to the second part and decreases or is constant in n . Hence, $L_j(n) \geq 0$, and $L_j' \leq 0$.

If $n=0$, then the rule is no liability and the allocative loss is maximal, as only the victim takes precaution. If $n>0$, then the rule is simple negligence, the injurer takes n precautionary measures and the allocative loss tends to decrease. The broader the negligence criterion (the more the injurer's precautionary measures included therein), the smaller the allocative loss tends to be. If $n=z$, the rule is an omni-comprehensive simple negligence rule, as all the injurer's precautionary measures are included in the negligence criterion; the allocative loss is zero.

Therefore, the task of the legal system is to set the negligence criterion as broad as possible, in order to reduce the allocative loss. The first best would be C_z , hence the optimal no-liability-based rule is omni-comprehensive simple negligence. The allocative loss is equal to zero, $L_j(z)=0$.

3.3.B.II. Vertical choice: allocative loss and the determination of the optimal extent of the negligence criterion under strict-liability-based rules (the injurer is the residual bearer) in a world with no administrative cost

The same analysis can apply to strict-liability-based rules. In this case, the negligence criterion targets the victim's behavior. Let us define the negligence criterion under strict liability with defense of contributory negligence as $G_m = \{r_1^g, \dots, r_m^g\}$.

The allocative loss takes the form of $L_v(m)$ and is a function of the number of the victim's precautionary measures included in the determination of negligence. If $m=0$, the rule is strict liability and $L_v(m)$ is maximal. If $m>0$, the rule is strict liability with defense of contributory negligence; $L_v(m)$ decreases if m increases. We can draw the same conclusion as before: the optimal strict-liability-based rule is strict liability with defense of omni-comprehensive contributory negligence, G_z . The allocative loss is equal to zero, $L_v(s)=0$.

3.3.B.III. Horizontal choice: the choice of the residual bearer in a world with no administrative cost

As the optimal no-liability-based rule and the optimal strict-liability-based rule trigger an

allocative loss equal to zero, they are both efficient²⁸.

This result can be summarized in the following proposition.

Proposition 2. When there are no administrative costs, the optimal negligence criterion includes all the non-residual bearer's precautionary measures. The optimal no-liability-based negligence rule is equivalent to the optimal strict-liability-based negligence rule.

3.4. Torts in a world with positive administrative cost

3.4.A. Administrative costs

In this section, we will introduce administrative costs into the model. We will consider two types of administrative costs: information costs and compensation costs.

3.4.A.I. Information costs

Information costs are those costs which result from the introduction of a negligence criterion on a strict-rule framework. They consist not only of the costs of gathering information concerning one party's behavior but also of the costs due to a higher likelihood of errors and increased litigation. They include the costs of judicial proceedings, lawyers' fees and any indirect costs borne by the parties.

We denote information costs as $I_j(n)$ for no-liability-based rules and as $I_v(m)$ for strict-liability-based rules. Information costs increase with n or m : if the negligence criterion expands, more information is needed, the likelihood of errors increases and further litigation might arise. $I' > 0$ in both cases.

3.4.A.II. Compensation costs

Compensation costs are the costs of transferring damage compensation from the injurer to the victim. In equilibrium, compensation costs arise only under strict-liability-based rules. In fact, under no-liability-based rules, the injurer's dominant strategy is to behave accordingly to the negligence criterion, and, hence, he never pays compensation. In addition to that, compensation costs are constant: they do not vary with m . Under strict-liability-based rules, the injurer always pays compensation to the victim, as the victim's dominant strategy is to

²⁸ Landes and Posner (1980) also proved that any sharing of the damages between the parties is equivalent. Hence, we can extend the result to other rules, such as contributory and comparative negligence, strict liability with defense of dual contributory negligence and strict liability with defense of comparative negligence. However, my formulation underlines that Landes and Posner's result completely holds only when the negligence criterion is fully comprehensive and when administrative costs are equal to zero.

behave non-negligently, given any appropriate negligence criterion, G_m . Let K denote compensation costs.

3.4.B. *Social cost and liability rules in a world with positive administrative cost. Genesis of the distinction between care and activity level*

In a world with positive administrative cost, the determination of the social cost cannot be disjoined from the choice of the liability rule, as administrative costs depend on the latter.

3.4.B.I. Vertical choice: determination of the optimal extent of the negligence criterion and genesis of the distinction between care and activity level under no-liability-based rules (the victim is the residual bearer) in a world with positive administrative cost

In section 3.3.B.I, we have defined the allocative loss, $L_j(n)$, as the measure of the relative performance of no-liability-based rules with respect to the minimization of the social cost. $L_j(n)$ decreases with n . In a world with no administrative cost, $n=z$ is optimal; in a world with positive administrative cost, the application of the negligence criterion triggers information costs $I_j(n)$, increasing with n .

The optimal extent of the negligence criterion for the injurer's behavior minimizes the total social cost under no-liability-based rules²⁹, and is denoted by n^* , which solves:

$$(4) \quad \min_n \left[L_j(n) + I_j(n) \right].$$

n^* can assume any value between 0 and z ³⁰.

- $n^*=0$. If information costs are high, the cost of applying a negligence rule might be too high when compared with the reduction in the allocative loss it would entail. The legal system might find it optimal to adopt no liability.
- $n^*=z$. When information costs are negligible (the injurer's behavior is easily verifiable), the optimal negligence criterion might be omni-comprehensive, all the injurer's precautionary measures are included in the determination of negligence. This is indeed an extreme case, and is unlikely to occur in practice³¹.
- $0 < n^* < z$. In intermediate cases, the negligence criterion includes only some of the

²⁹ The total social cost in a world with positive administrative cost is the sum of expected accident loss, precaution expenditures and administrative costs associated with the liability rule in use. Under no-liability-based rules, the minimization problem is $\min_{x,y,n,m} [l(x_n, y_m) + x_n + y_m + I_j(n)]$. Given a negligence criterion C_n , the problem is to define n so that $n = \operatorname{argmin} [l(x_n, y_s) + x_n + y_s + I_j(n)] = \operatorname{argmin} [l(x_n, y_s) + x_n + y_s - l(x^*_n, y^*_s) - x^*_n - y^*_s + I_j(n)] = \operatorname{argmin} [L_j(n) + I_j(n)]$.

³⁰ The solution is not necessarily internal and unique. Placing the injurer's precautionary measures in a different order might provide different level of n^* as a result. We take into consideration the order that triggers the lowest total cost $L_j(n^*) + I_j(n^*)$.

³¹ See Shavell (1987) at 30 on this point.

injurer's precautionary measures, while the rest are omitted as being too expensive to verify and/or not affecting the allocative loss to any substantial extent.

The set of the injurer's precautionary measures included in the negligence criterion, $C_{n^*} = \{p_1, \dots, p_{n^*}\}$, can be defined as care. The residual set of the injurer's precautionary measures excluded from the negligence criterion can be defined as the activity level, $A_{n^*} = \{p_{n^*+1}, \dots, p_z\}$. Care and activity level are just two subsets of the set of the injurer's precautionary measures, and their respective extent depends on the negligence criterion, $C_{n^*} \cup A_{n^*} = P$ and $C_{n^*} \cap A_{n^*} = \{\emptyset\}$. The injurer takes precaution only with respect to care, as once he takes due care, any additional precaution will increase his precautionary expenditure without reducing his exposure to liability³².

In the extreme case of $n^*=0$ (no liability), none of the injurer's precautionary measures is included in the negligence criterion. Hence, $C_{n^*} = \{\emptyset\}$ and $A_{n^*} = P$. In the other extreme case of $n^*=z$ (omni-comprehensive simple negligence), all the injurer's precautionary measures are included in the negligence criterion. Hence, $C_{n^*} = P$ and $A_{n^*} = \{\emptyset\}$.

In any case, no distinction between care and activity level can be claimed with respect to the victim's precautions, as the victim is the residual bearer and will allocate his precaution expenditure among all his precautionary measures³³.

3.4.B.II. Vertical choice: determination of the optimal extent of the negligence criterion and genesis of the distinction between care and activity level under strict-liability-based rules (the injurer is the residual bearer) in a world with positive administrative cost

Under strict-liability-based rules, the optimal extent of the negligence criterion for the victim's behavior is denoted by m^* , which solves:

$$(5) \quad \min_m [L(m) + I_v(m) + K].$$

m^* can assume any value between 0 and s . The logic is the same as in the former section.

- $m^*=0$: strict liability.
- $m^*=s$: strict liability with defense of omni-comprehensive contributory negligence.
- $0 < m^* < s$: strict liability with defense of contributory negligence.

The set of the victim's precautionary measures included in the negligence criterion,

³² See Shavell (1980a).

³³ See footnote 26 for a proof.

$G_{m^*} = \{r_1, \dots, r_{m^*}\}$, can be defined as care. The residual set of the victim's precautionary measures, which are excluded from the negligence criterion, can be defined as activity level, $B_{m^*} = \{r_{m^*+1}, \dots, r_s\}$. Care and activity level are just two subsets of the set of the victim's precautionary measures, and their respective extent depends on the negligence criterion, $G_{m^*} \cup B_{m^*} = R$ and $G_{m^*} \cap B_{m^*} = \{\emptyset\}$. The victim takes precaution only with respect to care³⁴.

In the extreme case of $m^*=0$ (strict liability), none of the victim's precautionary measures is included in the negligence criterion. Hence, $G_{m^*} = \{\emptyset\}$ and $B_{m^*} = R$. In the other extreme case of $m^*=s$ (strict liability with defense of omni-comprehensive contributory negligence), all the victim's precautionary measures are included in the negligence criterion. Hence, $G_{m^*} = R$ and $B_{m^*} = \{\emptyset\}$.

In any case, no distinction between care and activity level can be claimed with respect to the injurer's precautions, as the injurer is the residual bearer and will allocate his precaution expenditure among all his precautionary measures.

3.4.B.III. Horizontal choice: choice of the residual bearer in a world with positive administrative cost

In a world with positive administrative cost, the optimal no-liability-based rule is not necessarily equivalent to the optimal strict-liability-based rule, as the optimal extent of the negligence criterion might be less than omni-comprehensive in either or both cases.

The optimal liability rule is that rule which triggers the lowest social (allocative and administrative) cost. A no-liability-based rule, as determined in section 3.4.B.I, is the optimal liability rule if

$$(6) \quad L_j(n^*) + I_j(n^*) < L_v(m^*) + I_v(m^*) + K.$$

Otherwise, a strict-liability-based rule, as determined in section 3.4.B.II, is the optimal liability rule. The previous results can be summarized in the following propositions.

Proposition 3. Where administrative costs are positive, the optimal negligence criterion might not include all the non-residual bearer's precautionary measures: some precautionary measures are efficiently omitted from the negligence inquiry. The optimal no-liability-based negligence rule is typically not equivalent to the optimal strict-liability-based negligence rule.

³⁴ See footnote 32.

Proposition 4. A distinction between care and activity level can be claimed only with respect to the non-residual bearer's precaution: care is the subset of precautionary measures included in the negligence criterion and activity level is the subset of precautionary measures that are not included in the negligence criterion.

3.4.C. The dangerousness criterion and the verifiability criterion

Exp. (6) provides a criterion for the choice of the residual bearer, given an optimal determination of negligence. In this section, I will decompose such a criterion into three elements:

(i) The *dangerousness criterion*: the residual bearer should be that party, whose activity level triggers greater allocative loss (whose activity is more “dangerous”). $L_j(n^*)$ is the allocative loss due to the fact that some of the injurer's precautionary measures are excluded from the negligence criterion (the injurer's activity level). Therefore, $L_j(n^*)$ can be interpreted as the dangerousness of the injurer's activity. $L_v(m^*)$ is the dangerousness of the victim's activity. If $L_j(n^*) < L_v(m^*)$, the victim's activity is more dangerous: the dangerousness criterion requires choosing him as the residual bearer, and, hence, implementing a no-liability-based rule. Otherwise, a strict-liability-based rule would be desirable.

(ii) The *verifiability criterion*: the residual bearer should be that party, whose precautionary measures are more expensive to verify under the negligence inquiry. $I_j(n^*)$ the cost of verifying the injurer's precautionary measures under the optimal no-liability-based negligence rule, $I_v(m^*)$ is the cost of verifying the victim's precautionary measures under the optimal strict-liability-based negligence rule. If $I_j(n^*) < I_v(m^*)$, verifying the victim is more expensive, and hence a no-liability-based negligence rule is to be chosen. A strict-liability-based rule would be desirable otherwise.

(iii) No-liability-based rules do not trigger compensation costs, K , while strict-liability-based rules do so. Therefore, choosing the victim as the residual bearer saves administrative costs.

It is clear that the choice of the optimal liability rule has to take into account all the three criteria listed supra: neither if them is sufficient alone. Exp. (6) combines them in a unitary criterion. In section 3.6, I will exploit this framework in order to investigate the evolution of

liability systems over time. In section 3.7, I will show that the traditional explanation of the choice between strict-liability-based rules and no-liability-based rules, as provided in the literature, is incomplete because it is based on the dangerousness criterion only and does not account for the verifiability criterion.

3.5. Characteristics of the new definitions of care and activity level: endogeneity and variability

By recognizing that the definitions of care and activity level are a byproduct of the negligence criterion, we can dispense with two assumptions at the same time. We have an endogenous distinguishing criterion for care and activity level, therefore we do not need to search for a natural distinction between them *out of* the model. The negligence criterion is indeed a non-tautological and effective criterion to distinguish care from activity level.

In fact, the model produces mutually exclusive definitions. No precautionary measures can be defined at the same time as care and as activity level, because the intersection between the set of precautionary measures defined as care and the set of precautionary measures defined as activity level is always empty (see sections 3.4.B.I and 3.4.B.II). Hence, it completely solves the problem of defining those concepts and overcomes the shortcoming of step (a1) of the traditional definition (see table 1 and section 3.1).

In addition, we also avoid incurring a second theoretical problem, due to step (a2) of the traditional definition (see table 1 and section 3.1). Once we have defined as activity level those precautionary measures that are not taken into account by the court while deciding issues of negligence, the assumption that the court cannot verify the activity level becomes superfluous³⁵. The system turns out to be based on fewer assumptions, and this is indeed an improvement³⁶.

As the definitions of care and activity level are generated by the model itself and are, hence, endogenous to it, they depend on the negligence inquiry concerning the non-residual-bearer's precautions. Few consequences follow:

- It is logically incorrect to distinguish between care and activity level with respect to the

³⁵ Such an assumption would instead be indispensable in order to justify the activity-level theorem if activity level and care were defined exogenously. See Shavell (1980a) at 354 and Shavell (1987) at 25.

³⁶ For several hundred years, mathematicians tried to reduce the number of axioms on which the whole geometry is constructed from five (as listed by Euclid around 300 BC) to four, by demonstrating from the existing four that two parallel lines never intersect (a rough description of the fifth axiom of Euclidean geometry). Unfortunately, no one ever succeeded during more than 21 centuries, but, fortunately, after such a collective effort, in 1823 non-Euclidean geometry was discovered, and the number of axioms is still five. What I am trying to do is to eliminate two axioms from the economic theory of torts. By deriving the definitions of care and activity level from other existing concepts, I attempt demonstrating that they are definitions produced *by* the model and not assumptions *of* the model.

residual-bearer's precautions, as he always takes all his precautionary measures, and the negligence criterion applies to the other party³⁷.

- It is logically incorrect to distinguish between care and activity level under strict rules, such as strict liability and no liability. With respect to the residual bearer, no distinction can be made for the reason given supra. With respect to the non-residual bearer, no negligence criterion is applied in this case and, in theory, all his precautionary measures should be qualified as activity level, because the set of care measures is empty³⁸.

More in general, the definitions of care and activity level are variable with the extent of the negligence criterion and the choice of the residual bearer, which in turn respond to changes in the administrative costs and the prevention technology. This consideration is extremely important, as it implies that the concepts of care and activity level are relative to a specific legal, economic and juridical situation and can change over time with the extent of the negligence criterion.

When the negligence criterion expands (in the sense that more precautionary measures are verified by the court), the subset of precautionary measures that we define as care expands accordingly, while the subset of activity level shrinks. As the determination of the optimal extent of the negligence criterion depends on the balance between allocative loss and information costs, a change in either of them triggers an adjustment of the negligence criterion³⁹. The following sections exploit this characteristic.

3.6. A diachronic perspective: the evolution of liability rules

The foregoing analysis highlights some arguments that will be used here to suggest a theoretical framework for explaining the evolution of liability systems over time⁴⁰.

The exemplification made in table 2 permits us to decompose the choice of the optimal

³⁷ This result also holds if a negligence criterion is applied on the residual bearer as well, as for example under contributory negligence, comparative negligence and strict liability with defense of dual contributory negligence. See section 3.8 at (i).

³⁸ Shavell (1980b), while discussing issues of causation, bases his analysis on the distinction between activity level and care, but applies the distinction in a strict-liability framework. From my analysis, it follows that this approach is logically incorrect. I discuss this issue in a separate paper.

³⁹ See Landes and Posner (1987) at 71: "There is a trade off between the information cost of considering the injurer's activity level as an aspect of due care and the allocative cost of ignoring the activity level. And although [...] strict liability eliminates the problem of having to determine whether particular changes on the injurer's activity level would have been cost justified, it does so at the price of relaxing the victim's incentive to avoid the accident through a change in his own activity level."

⁴⁰ It is clear that allocative efficiency is not the only goal pursued in ancient (and modern) tort law systems, and sometimes it is not the principal one. Therefore, we might expect deviations from the pattern that we would draw in a theoretical evolution toward efficient rules. This is true not only for ancient law but also for more recent periods. My aim here is to suggest a possible additional way in which to examine the problem. For a law and economics analysis of the genesis of liability in ancient law see Parisi (2001).

liability rule into two components: the choice between a strict rule and the corresponding negligence variant (moving vertically in table 2, see sections 3.4.B.I and 3.4.B.II), and the choice of the residual bearer (moving horizontally in table 2, see section 3.4.B.III). We can describe two stages in the evolution of a liability system.

Proposition 5. At the first stage, liability systems tend to prefer (i) strict rules to negligence rules (vertical choice) and (ii) no-liability-based rules to strict-liability-based rules (horizontal choice).

Corollary: strict liability is chosen when the injurer's actions are more dangerous than the victim's actions and the allocative gain overcomes compensation costs.

Proof.

(i) Exp. (4) and (5) show that the choice of the extent of the negligence criterion under no-liability-based rules and strict-liability-based rules depends on the balance between information costs and allocative loss. In the early stage of the development of liability systems, information costs are likely to be extremely high for two reasons. Negligence is difficult to prove due to the lack of a good understanding of the laws of nature, systems of writing and records⁴¹ and a sufficient judicial expertise. The application of the negligence criterion increases litigation as it creates animosity⁴². Therefore, strict liability is likely to be preferred to strict liability with defense of contributory negligence as much as no liability is likely to be preferred to simple negligence.

(ii) Therefore, the choice of the residual bearer is mainly a choice between no liability and strict liability. Exp. (6) then becomes:

$$(7) \quad L_j(n=0) < L_v(m=0) + K. \quad ^{43}$$

The decision of the residual bearer depends on compensation costs. If the parties' precaution affects the expected accident loss to the same extent, no liability will be preferred as it saves compensation costs. In early liability systems compensation costs are likely to be particularly high for reasons which are similar to the ones already provided for information costs, and, therefore, the argument for no liability is strengthened. Strict liability will be adopted only in those areas of torts where the advantage of making the injurer take precaution instead of the victim overcomes compensation costs (when the difference between $L_j(n=0)$

⁴¹ Posner (1980).

⁴² Parisi (1992) at 33-34, and 56-58.

⁴³ Information costs are zero if strict rules are implemented.

and $L(m=0)$ is greater than K). This criterion can be referred to as the dangerousness criterion: the legal system targets the most dangerous party, the one who affects the accident loss the most⁴⁴.

Proposition 6. At a later stage, (i) liability systems tend to move from strict rules to negligence rules (vertical choice) but (ii) they still tend to prefer no-liability-based rules to strict-liability-based rules (horizontal choice).

Corollary: strict-liability-based rules are chosen when the injurer's actions are more dangerous and less controllable than the victim's actions and the allocative gain overcomes compensation costs.

Proof.

(i) Given the allocative superiority of negligence rules (they spur both parties' precaution), legal systems will move in their direction as soon as the information costs decrease below the allocative gains⁴⁵. Information costs decrease when literacy and knowledge develop, resulting in increased judicial expertise. This justifies a move away from strict rules. However, information costs also depend on the complexity of the parties' actions, which hinges upon social and technological development. In those areas of torts where the parties' actions become more complex (as for example in the production of certain goods), information costs might increase and thereby justify strict rules⁴⁶. Negligence rules become more common, although strict rules survive in some specific areas⁴⁷.

(ii) The reason for preferring no-liability-based rules is the same as under the former proposition: the legal system saves compensation costs. However, the choice of the optimal liability rule becomes more complex, as the negligence criterion is usually applied. The dangerousness of the parties' actions (i.e. the importance of their precaution in determining the allocative loss) should be balanced with the verifiability of their precautionary measures. Exp. (6) shows that both allocative loss and information costs now enter the comparison. Therefore, the dangerousness criterion alone does not provide an explanation for deciding between no-liability-based and strict-liability-based rules, and it has to be integrated with a verifiability criterion: the legal system considers both dangerousness of the activity (L) and

⁴⁴ See section 3.4.C.

⁴⁵ See also Landes and Posner (1987).

⁴⁶ Isaac (1918) singles out three periods of dominant strict liability in English law: (i) the 11th century, around the time of the Norman conquest, (ii) the 14th century, at the time of Edward I, and (iii) the beginning of the 20th century. Fault was the dominant criterion for liability in between those periods. He justifies such cyclical dynamics as a strive of law for approaching the goals of ethics.

⁴⁷ Isaac (1918) at 967 speaks of the "swinging of the pendulum between strict rules and negligence rules".

costs of controlling parties' precautionary measures (*I*)⁴⁸. In addition, even when the injurer happens to be more dangerous and less controllable (hence a strict-liability-based rule is optimal), the resulting rule does not necessarily have to be strict liability, as it might be profitable to embark on a negligence inquiry concerning the victim's behavior, which would trigger strict liability with defense of contributory negligence.

Legal systems tend to adopt a general no-liability-based negligence rule and to confine strict-liability-based rules to specific areas of torts and no liability to those instances in which compensation is denied.

3.6.A. *A historical case for proposition 5: Roman and English legal history*

Roman law confirms the theoretical pattern of evolution described above as it first developed a strict liability regime, gradually evolving towards negligence. In addition, at the beginning, the Roman tradition produced liability rules based on strict liability only for specified wrongs; the rest would fall under no liability. The *lex XII tabularum* and the *lex Aquilia* listed a series of wrongs that had to be restored through compensation. The requirement of *damnum iniuria datum* for those wrong which triggered liability at the beginning literally meant "a wrong committed against the law". Hence, only those wrongs admitted by the law could entitle the victim to compensation.

Limitations also followed from the application of the *corpore corpori* principle, which gave rise to compensation only for those wrongs materially committed by the injurer and resulting in material harm for the victim. These limitations on the types of wrongs triggering liability served the economic goal of reducing the administrative cost of the (not yet well developed) legal and judicial system, by *de facto* using strict liability as an exception to no liability⁴⁹.

English and American legal history seems to confirm this pattern too, as it developed from the strict-liability formant and slowly evolved towards negligence⁵⁰. Not until the nineteenth century was there any general acceptance of the fault principle. The early common law's main concern was with intentional torts and, even later, the attention was drawn to the nature of the victim's harm rather than to the injurer's behavior. Rudimental requirements of causation were used to select accident losses for which the victim should be entitled to compensation, regardless of the negligence of either party⁵¹.

The English writ system was similar to the early Roman system of *actiones*: a victim was entitled to compensation only for specified kinds of accident losses for which a procedural

⁴⁸ See section 3.4.C.

⁴⁹ See Parisi (1992) for a discussion of the problem and a vast bibliography.

⁵⁰ See Fleming (1983) at 300.

remedy existed and it was not necessary to prove fault on the part of the injurer.

3.6.B. A historical case for proposition 6 and a criticism of the application of the dangerousness criterion in order to explain the choice of liability regimes in modern societies: liability for industrial accidents and product liability from no-liability-based rules to strict-liability-based rules

The joint application of the dangerousness criterion and the verifiability criterion determines the choice of the optimal liability rule in modern societies, where the application of a negligence criterion is usually possible. This will immediately become clear with a theoretical example.

Example 3. Imagine that an accident occurs between an injurer carrying on a very dangerous activity, whose precautionary measures are all easily verifiable before the court (running a railway in the nineteenth century). On the contrary, victim's contribution to the accident is less dangerous but not as easily verifiable (crossing a railway in the nineteenth century).

A strict-liability-based rule will provide the injurer with an incentive to take all his precautionary measures (the injurer is the residual bearer), whereas it will provide the victim with only little incentives to take precaution, as the extent to which the court can consider the victim's precaution under the negligence criterion is limited, and therefore the extent of the victim's activity level will result large. On the contrary, simple negligence (a no-liability-based rule) will result in more efficient accident prevention. On the one hand, the victim is in fact led to take precaution with respect to all his precautionary measures, as he is the residual bearer. On the other hand, as the court can easily verify all the injurer's precautionary measures, the extent of his activity level is reduced to zero (more correctly, to the empty set), all his precautionary measures are to be considered as care, and thus he will take precaution with respect to all of them.

Although the injurer's actions influence largely the expected accident loss, the legal system attains a more efficient result by targeting the victim (simple negligence), as the victim's precautionary measures are less verifiable.

The example highlights two important points. First, dangerous injurer's activities do not necessarily require a strict liability regime, as also the verifiability test should be employed. Secondly, even injurer's activities that are not excessively dangerous can require a strict liability regime, if the injurer's precautionary measures are not readily verifiable (as would be evident by simply reversing the positions of the injurer and the victim in the example above)⁵².

⁵¹ White (1980) at 3 ff. and Fleming (1983) at 97.

⁵² Gilles (1992), at 362, after showing that the courts do consider precautionary measures traditionally regarded as

This argument might provide an efficiency justification for the adoption of a no-liability-based negligence regime for industrial accidents in the 19th century⁵³. It has been said that the adoption of a no-liability-based regime (mainly implemented with the application of a negligence criterion for the injurer) was meant to favor the development of industry in its early stages, by making the victims the residual bearers. However, my framework can provide a different interpretation of this problem.

In the nineteenth century, industrial activities were of rather simple a kind. The hazards created by them did not constitute any such subtle danger to escape anyone's notice⁵⁴. Products were also rather simple. Our previous discussion suggests that it was probably easier to control the injurer's behavior than victim's behavior in such situations, which could provide a justification for placing the residual loss on victims.

With the development of technology, industrial processes have become more complex and the cost of acquiring information over them has risen enormously. Hence, controlling injurers through negligence has become more expensive a task. The verifiability criterion suggests that the residual bearer should be the injurer⁵⁵.

Therefore, the parallel move in industrial and product liability from no-liability-based rules to strict-liability-based rules can be justified by the increasing information costs associated with the application of the negligence criterion to the injurer's behavior. No-liability-based rules are efficient in the early stages of industrial development, when production techniques are simple and therefore easily verifiable. During the following stages, when complexity increases, strict-liability-based rules become a more efficient solution⁵⁶.

In addition to that, another point should be clarified. If, in principle, primitive tort law systems tend to implement strict rules, we cannot state that modern tort law systems tend to negligence in all areas of torts. The persistence or the reappearance of strict rules (mainly strict liability) in specific areas of tort law can be explained by the relevant (or increasing)

activity level in the determination of negligence, concludes that "activity levels may be less important as a factor bearing on the choice between negligence and strict liability than has previously been thought". Although, this conclusion is correct, it is one-sided. In fact, the courts do not include many precautionary measures, traditionally regarded as care, in the negligence inquiry. A correct definition of the concepts of care and activity level, as linked to the negligence criterion, brings an informational argument into the discussion but does not necessarily undermine the importance of the activity level in the choice between strict-liability-based rules and no-liability-based rules. I simply argue that the activity level is something different from what the literature has always claimed. So is care.

⁵³ See Posner (1972 and 1998), Schwartz (1981), and Landes and Posner (1987) at 116 for a critical discussion.

⁵⁴ As noted by Posner (1992) at 257.

⁵⁵ See section 3.4.C. Posner (1992) at 256-258 recognizes that the hazards created by nineteenth century industrial production were of a simple and discernible kind and the cost of acquiring information thereon was particularly low. He applies this logic to accidents between parties to a contract, but not to accident between strangers, in which he generally accepts the subsidy argument. My analysis is somewhat broader, as I reinterpret the argument in both contractual and non-contractual settings and show that the verifiability argument can explain both situations.

⁵⁶ Landes and Posner (1987) at 284-285 suggest a similar solution with respect to product liability.

complexity of the parties' activities, which increase the costs and reduce the benefit of negligence, as in the case just discussed.

3.7. A synchronic perspective: “strict liability vs. negligence”, an analysis of liability rules in modern legal systems

The literature⁵⁷ has traditionally approached the problem of choosing between strict-liability-based rules (traditionally referred to simply as “strict liability”, the injurer is the residual bearer) and no-liability-based rules (traditionally referred to simply as “negligence”, the victim is the residual bearer) in modern tort law systems by applying the dangerousness criterion only. “If it is more important to control injurers' level of activity than victims'” the rules that result in greater social welfare are strict-liability-based rules⁵⁸. The legal system should otherwise adopt no-liability-based rules. In turn, the “importance” of controlling one party's activity can be seen in terms of riskiness or dangerousness of that party's activity⁵⁹.

This criterion has provided a classical and broadly accepted explanation for the choice between strict liability and negligence. However, not all strict liability regimes can be justified in terms of dangerousness. “Is the chance of a wild animal escaping from a zoo and doing harm, for which strict liability would probably result in the United States, greater than that of an automobile running down a pedestrian, for which the negligence rule would govern?”⁶⁰. The criterion is indeed “somewhat rough”⁶¹. Many actual legal solutions are unclear, as “the choices made between strict liability and negligence rules are not always easy to explain on the basis of differences in riskiness”⁶². I will provide here an alternative explanation, which elaborates upon our previous discussion in section 3.4.C: the verifiability criterion should also be considered.

3.7.A. The general application of negligence

Proposition 6 (i) shows that legal systems tend to move from strict rules towards negligence when parties' precautionary measures can be verified more easily. This leads to greater allocative gains (more precautionary measures can be included in the negligence criterion) and less cost (information costs due to the negligence criterion).

In addition, in proposition 6 (ii), we have been provided with a reason why legal systems

⁵⁷ Shavell (1980a) and Schäfer and Schönenberger (2000).

⁵⁸ Shavell (1987) at 29. See also Landes and Posner (1987) at 70.

⁵⁹ Shavell (1987) at 31-32.

⁶⁰ Shavell (1987) at 31.

⁶¹ Shavell (1987) at 32.

⁶² Shavell (1987), at 31-32. See also Landes and Posner (1980) and Schäfer and Schönenberger (2000).

tend to prefer no-liability-based negligence rules to strict-liability-based negligence rules, in order to reduce compensation costs.

These two findings, once combined, can explain why modern societies, where judicial expertise is quite well developed and the cost of acquiring information concerning many parties' precautionary measures tend to be low, adopt a general fault-based regime of tort liability, which holds the injurer responsible only if negligent. This choice saves the higher compensation costs due to a corresponding negligence regime based on strict liability, as formalized in the model by Exp. (6).

3.7.B. *The implementation of strict liability in specific areas of torts*

Let us first consider unilateral-precaution accidents, where the injurer is the only party who can take some precaution in order to reduce the expected accident loss and the domain of strict liability is incontestable, given the fact that this rule provides optimal incentives to injurers⁶³. The traditional argument is that a negligence rule will be defective in that it does not spur injurer's optimal activity level.

Nevertheless, I have already shown that this is precisely due to the cost associated with verifying some of the injurer's precautionary measures (those which are to be classified as activity level) and, therefore, if information costs are lower than compensation costs plus allocative costs, then a simple negligence rule might be preferable. In an extreme case, an omni-comprehensive negligence criterion (one in which all the precautionary measures are taken into consideration) will have the same allocative effect as a strict liability rule (providing the same incentives for the injurer). The choice between the two depends on the comparison between the higher compensation costs under strict liability and the information costs due to a simple negligence rule.

In bilateral precaution cases, we shall add that the introduction of strict liability triggers an allocative cost because of the reduction in victim's incentives to take precaution. Often the rule in force is not exactly a strict liability rule but a no-liability-based negligence rule⁶⁴ with inversion of the burden of proof⁶⁵. In such a case, the residual bearer is the victim⁶⁶. This

⁶³ See Cooter and Ulen (2000).

⁶⁴ The severity of the negligence criterion only slightly affects this conclusion, as it concerns only with the balance between costs and benefits of precaution, and such a balance could well lead to the conclusion that it would have been optimal for the injurer to take maximal precaution.

⁶⁵ The Italian civil code is a good example as it clearly indicates where the burden of proof lies in each case. It comprises a detailed list of provisions of presumed fault or presumed responsibility, in which the burden of proof is placed on the injurer, see artt. 2047, 2048, 2050, 2053 and 2054 cc. The French system has mostly reached the same results through jurisprudence and doctrine. German law also provides many examples of inversion of the burden of proof through the doctrine of *prima facie proof*, which is similar to the common law solution of *res ipsa loquitur*, except that the latter cannot be rebutted.

⁶⁶ The case of presumed fault should be distinguished from that of presumed responsibility, the latter being a form of strict liability as, in the case of both parties being non-negligent, the injurer is the residual bearer.

observation strengthens the information argument, as it shows that the result reached by a strict liability rule, can be sometimes obtained by a no-liability-based rule, if inverting the burden of proof can lower the cost of verifying the injurer's precautionary measures. Nevertheless, I will consider some of those cases in the following analysis together with cases in which a truly strict liability rule is regularly implemented, not only for their proximity to that rule, but in particular for they can be justified within the same informational perspective⁶⁷.

Ultra-hazardous activities

In order to support the dangerousness criterion for strict liability, the case of ultra-hazardous activities is used to prove that, at least in most cases, the dangerousness of the activity triggers strict liability. This is the hypothesis of the most general and open application of strict liability by American courts⁶⁸, and by the courts in civil law countries as well⁶⁹.

Strict liability, however, is not justified by dangerousness alone, but rather by the *abnormality* of the activity. The activity must not be *in common use*. Activities that imply a high but *typical* risk (as motoring in American law) are excluded⁷⁰. The criterion seems to be rather focused on what is well known as a distinguishable category from what is not well known and controllable. It seems that the injurer is "punished" by strict liability, not because he has embarked on a dangerous activity but rather because he has carried out an activity that goes beyond what all the others do, and hence he has to bear the *unknown consequences* of

⁶⁷ Landes and Posner (1987), at 107-122, consider strict liability regimes. They provide several efficiency justifications for strict liability, an analysis of which goes beyond the scope of this chapter. I will focus on two main aspects of their argumentation. Their claims are all based on the supposed natural dichotomy between care and activity level, which is criticized here. While Shavell (1987) limits his consideration to the dangerousness of the activity, they step forward and often consider the utility of the activity and compare it with the accident cost generated by the activity itself. Consequently, they derive a justification for strict liability when the utility is lower than the cost (as in the case of vicious dogs at 109). However, if costs and benefits can be confronted, the activity level will be easily considered in the determination of negligence, as shown by Gilles (1992). Hence, this cannot be a justification for strict liability. A second recurrent argument seems to confirm the thesis defended here, that the choice of the residual bearer depends on the verifiability of parties' behavior. With respect to employer's vicarious liability, at 121, they provide as the only argument for vicarious liability to be strict that "there are a number of activity measures (as distinct from care measures) that an employer can take to reduce accident behavior by his employees". The point is indeed the distinction between "care measures" (that also simple negligence would enhance) and "activity measures" (for which strict liability is needed). The only meaningful interpretation of such a distinction which I can think of is the one I provide in the text. However, once this step is taken, the argument becomes the one of verifiability that I discuss. They adopt the same justification at 115. "During the early stage of the development of a new product or activity, the legal system lacks sufficient experience to be able to determine whether the benefit of the product exceeds its full cost [...] One way to generate such information is to hold the producer or the user strictly liable". It is indeed a problem of information. Nevertheless, Landes and Posner (1987) do not consider one crucial consequence of such an approach: that the concepts of care and activity level expand and shirk depending on the cost of acquiring information concerning the parties' precaution. This is a central point of my analysis.

⁶⁸ Fleming (1983) at 302. Restatement Second of Torts, 520.

⁶⁹ See for example art. 2050 of the Italian civil code, which represents an advance over earlier continental codes, Stone (1972).

⁷⁰ In *Rylands v. Fletcher* (1866) L.R. 1 Ex. 265, affd. (1868) L.R. 3 H.L.330, the reference is to *non-natural use*.

having done so.

It does not seem strange to conclude that also the informational characteristics of the activity, and not only its dangerous potential, contribute to triggering strict liability⁷¹.

Further examples will corroborate this view, but first it is worth commenting on the legal definition of an *abnormally dangerous activity* as that activity which cannot be made safe by ordinary precaution. It is hard to find any economic justification for such a statement, as any activity, even those which are less risky, will result in some *efficient* accidents, given that the optimal level of precaution is not necessarily the level of precaution that results in no accidents at all. In these terms, no activity is safe. Moreover, the optimal level of precaution for a specific activity depends on the balance between its costs and benefits, and can well be higher or lower than the ordinary level (i.e. the optimal level of precaution for most of the activities).

Vicarious liability

Even under its many different applications (liability of the supervisor for the harmful actions of a child or an adult under supervision, liability of the employer for the actions of employees, within the scope of the employment), vicarious liability usually triggers either the application of strict liability⁷² on the defendant vicariously liable or the allocation to him of the burden of proof under the ordinary negligence regime⁷³.

In principle, vicarious liability is applied to a party (the principal) who can control the action of another party (the actual injurer). Such responsibility is usually considered as deriving from the principal's failure to select, direct, educate or bring up the other party, and could well be claimed under the general principle of negligence. Here it is even clearer that the application of the general negligence rule is rendered difficult by the high cost of verifying the injurer's precautionary measures, because it is inherently cumbersome to prove that the harm caused by a child is due to the bad education imparted by his parent, or that the harm caused by certain schoolchildren is due to a lack of teacher's discipline⁷⁴. The dangerousness criterion alone is clearly no appropriate explanation for these cases.

Damages caused by things

With respect to damages caused by things, a broad variety of legal solutions can be found,

⁷¹ See section 3.4.C.

⁷² Italian, French and American systems in general opt for strict liability in the case of liability of the employer for the actions of employees. The German system sometimes reaches the same result but only through a complicated interpretation of the general principle of negligence (Eörsi, 1975).

⁷³ Liability of the supervisor for the harmful action of a child or an adult under supervision is governed by a presumption of fault on the part of the supervisor in most legal systems (Le Gall, 1976).

⁷⁴ "Generally faulty supervision would be difficult of proof" (Le Gall, 1976, at 6).

partially due to national traditions and particular beliefs, as is evident in the case of damages caused by animals. Although, in some instances, the general principle of fault is applied, most legal systems make a broad use of strict liability or of the inversion of the burden of proof.

Damages caused by wild animals are generally dealt with under a strict rule, irrespective of the dangerous nature of the particular animal or of its domestication⁷⁵, and even non-dangerous species sometimes⁷⁶ fall under such a regime, merely because they are *unusual* in a particular country or region.

Ruinous buildings trigger the application strict liability in most cases⁷⁷, and so do thrown⁷⁸ or falling⁷⁹ things, as “in many cases, the activity or negligence of some person cannot be demonstrated”⁸⁰.

As a corollary, it is worth noticing that strict liability can also be used as an inexpensive negligence rule, which assumes that the optimal level of precaution is the maximal level possible. This occurs when the marginal cost of maximal precaution (not acting at all) is equal to or even smaller than the marginal benefit in terms of a reduction in the expected accident losses. On the one hand, an omni-comprehensive negligence criterion will serve the same purpose by making acting injurers liable, but it would bring along greater informational (and plausibly allocative) costs. On the other hand, banning the activity might be a feasible solution, at the price of forbidding some potentially welfare enhancing activities. Strict liability seems to be a good intermediate solution. In this case also, dangerousness alone does not provide a distinguishing criterion. An example is the application of strict liability to harm caused by a domesticated animal, if its keeper was aware of its dangerous disposition⁸¹.

Sometimes negligence defenses are allowed, but often they are not. As we have already noted, this depends on the balance between the information costs and the allocative gains of the negligence rule.

⁷⁵ This is the case under Italian, French, and German law, which employ a rigid distinction between wild (*ferae naturae*) and domesticated animals. On the contrary, common law, and, English law in particular, comprise a very refined and complex distinctive criterion. For domesticated animals the *scienter* is of general use. Stone (1972), Fleming (1983), Epstein (1999).

⁷⁶ In Italian law, for instance.

⁷⁷ French, Belgian and Italian law adopt a presumption of responsibility; common law makes general use of strict liability, although it employs fault for damages to persons in the premises. German law presumes injurer’s fault.

⁷⁸ Civil law countries follow the Roman tradition of dealing with these instances under a strict liability regime, while common law requires fault.

⁷⁹ In general, strict liability or presumed responsibility are largely applied.

⁸⁰ Stone (1976) at 32.

⁸¹ Stone (1976), Fleming (1983), Epstein (1999). The cost of the maximal precaution (not keeping the animal at the cost of a foregone benefit) seems plausibly lower than the expected personal injuries resulting from an accident. See Landes and Posner (1987) at 119.

3.7.C. *Room for no liability*

Although legal systems provide for a general principle of negligence and do not consider no liability as a liability rule, but rather as a non-rule, no liability is a liability rule in an economic sense, because it allocates the accident loss as any other liability rule does⁸². No liability results when the legal system does not entitle the victim to recover damages, irrespective of the injurer's level of precaution. No liability has a major advantage: it saves both information costs and compensation costs, as no reallocation of accident loss takes place.

Ancient law evolved by producing, at the beginning, tort rules based on strict liability only for specified wrongs, in order to serve the economic goal of reducing the administrative costs. Obviously, this is obtained at the (allocative) cost of no injurer's precaution.

Modern legal systems still limit the scope of their liability regimes in different and more sophisticated ways (by listing the categories of wrongs that give rise to liability or by setting minimum thresholds on the injurer's causal contribution to the accident), but tend to be broader than their ancient predecessors, as the overall cost of administering liability is supposed to be lower.

The discussion on no liability is not as rich as the one on strict liability, as, under the former, the legal system simply denies injurer's responsibility.

3.8. **Concluding remarks**

In this chapter, I have shown that the definitions of care (as precaution) and activity level (as number or frequency of actions), traditionally employed in the literature, are tautological and therefore difficult to apply, because any precautionary measure always fits both definitions. Hence, they do not provide for a criterion to decide whether a particular precautionary measure should be considered as belonging to one category or the other.

The reason is that the traditional concepts of care and activity level are defined out of the model, as assumptions. On the contrary, I have shown that the appropriate definitions thereof are produced by the model itself: care is the subset of precautionary measures included in the negligence criterion and activity level is the residual subset of precautionary measures that are not included in the negligence criterion.

The determination of the extent of the negligence criterion, i.e. the decision of which precautionary measures should be included in it, follows from a choice made by the legal system and based on efficiency considerations: the balance between the costs of inclusion (information costs of inquiring into parties' precaution) and its benefits (the parties take more

⁸² Landes and Posner (1987) at 62.

precaution). The activity-level theorem describes the effect of such a choice: precautionary measures excluded from the negligence criterion will remain untaken.

It has been said that the analysis of “activity level vs. care” and, consequently, the activity-level theorem⁸³ “undermined the confidence in the efficiency of tort law”⁸⁴. However, the opposite is true. The activity-level theorem does not concern any failure of the tort liability system, but it accounts for the allocative costs of excluding some precautionary measures from the determination of negligence. Since such exclusions can be justified on efficiency grounds as they save information costs, the whole picture should rather be regarded as the result of an efficient balance between costs and benefits.

The aim of this chapter has been indeed to show that the problem of determining the extent of the negligence inquiry, the boundaries of the concepts of care and activity level, and the relevance of the activity-level theorem can be fully understood only if the information costs of determining the parties’ precaution are taken into account.

The framework elaborated in this chapter also permits us to reinterpret the other fundamental theorem of tort law and economics: the efficiency-equivalence theorem by Landes and Posner (1980). They show that any negligence rule provides both parties with the same incentives with respect to care. My analysis shows that their result holds (and with respect not only to care, but to precaution in general) only in a world with no administrative cost (proposition 2). When administrative costs are considered, negligence rules yield different outcome, not only because they trigger different administrative costs, as the authors argue, but also because they trigger negligence criteria of different extent (proposition 3).

Besides the determination of the extent of the negligence criterion, the legal system makes a second fundamental choice: it decides whether the victim or the injurer is the residual bearer. This choice has traditionally been analyzed under the label of “strict liability vs. negligence”⁸⁵: under strict-liability-based rules the injurer is the residual bearer, under no-liability-based rules (usually simply referred to as negligence⁸⁶) the victim is the residual bearer. Such a choice has been explained by referring to the dangerousness criterion: the party, whose activity is more dangerous, is to be chosen as the residual bearer, in order to provide him with better incentives.

I have shown that the dangerousness criterion is not a sufficient explanation, as the cost of verifying parties precaution should also be taken into account (verifiability criterion): the party whose precautionary measures are more expensive to verify is to be chosen as the

⁸³ In their original formulation, by Shavell (1980a).

⁸⁴ Donohue (1988), at 1058.

⁸⁵ Shavell (1980a). The term “strict liability” refers to what we defined as strict-liability-based rules (with or without negligence defenses); the term “negligence” refers to no-liability-based negligence rules.

⁸⁶ See previous footnote.

residual bearer, in order to save administrative costs.

In addition to that, a third element is to be considered: strict-liability-based rules trigger compensation costs, while no-liability-based rules do not, because in the latter case a non-negligent injurer never pays compensation to the victim.

The choice between strict liability (strict-liability-based rules) and negligence (no-liability-based rules) is, therefore, to be analyzed under all those three lenses: neither of them is sufficient alone. I have shown that such an approach can explain the choice of liability rules in modern legal systems also in those cases that Shavell (1987)⁸⁷ considers as puzzling, as they cannot be justified by the dangerousness criterion. I have demonstrated that the same approach can explain the evolution of liability systems over time and the tendency of tort law to move from the localized use of strict liability towards the implementation of a general no-liability-based negligence regime.

The forgoing analysis raises some additional related points.

(i) The model considers a limited set of rules: no liability, strict liability and its negligence variants, simple negligence and strict liability with defense of contributory negligence. The latter have in common the negligence applies only to the non-residual bearer.

Legal systems provide other liability rules, by adding an additional negligence criterion to the residual bearer's behavior. Under contributory negligence, both the injurer (the non-residual bearer) and the victim (the residual bearer) are subject to the negligence inquiry: the injurer pays only if he is negligent and if the victim is non-negligent. However, in equilibrium, we expect both parties to behave non-negligently; hence, the victim always bears the accident loss.

This rule has no allocative advantage over simple negligence. The victim has an incentive to take all his precautionary measures (irrespective of the negligence criterion) as he bears the residual loss; the injurer has an incentive to comply with the negligence criterion, as under simple negligence. Nevertheless, contributory negligence triggers higher information costs, as it doubles the negligence inquiry. The application of comparative negligence would increase the information costs even further. My framework suggests that simple negligence should be preferred over contributory or comparative negligence, which may be justified on different grounds⁸⁸.

A similar reasoning applies to strict liability with defense of dual contributory negligence, a strict-liability-based negligence rule.

(ii) Risk allocation has not been taken into account. It is a well-established belief that the tort

⁸⁷ See Shavell (1987) at 31-32.

system is more expensive than the insurance system as a way to allocate risk. However, the cheapest-cost insurer criterion can play an important role in the decisions concerning the tort law system to implement⁸⁹, as the residual bearer is provided with incentives to buy insurance coverage.

(iii) Another dimension for the comparison between different legal rules is provided by an analysis of the incentives to acquire information about risk. In general, the residual bearer has an incentive to do so, and this could serve as a criterion for the choice of liability rules. Strict liability will provide incentives to injurers, simple negligence to victims⁹⁰.

(iv) The conclusions reached here for accidents between strangers could be profitably extended to analyzing accidents between parties of a contract, as in the case of product liability discussed in section 3.6.B.

(v) The negligence criterion can be defined ex post by the court, case by case, or be based on an ex ante rule of negligence precisely defined by a regulatory body, the legislature or judicial decisions. Gilles (1992) argues that the choice between case-specific (ex post) and rule-based (ex ante) techniques to determine negligence depends on the cost of evaluating precaution. As information costs become higher, “the courts are more likely to employ the techniques of rule-based negligence to make such inquiries possible” (Gilles, 1992, at 336). For this reason, as precautionary measures traditionally regarded as activity level are likely to be more costly to verify, they are usually dealt with through rule-based negligence (Gilles, 1992, at 337), while care measures are dealt with through case-based decisions. Such an approach is consistent with my framework and confirms that the distinction between care and activity level is endogenous to the model (and to the legal system).

3.9. References

- Cooter, Robert D. and Ulen, Thomas S. (2000), *Law and economics*, 3rd ed., Reading, Massachusetts, [etc.]: Addison-Wesley.
- Diamond, Peter A. (1974), ‘Single Activity Accidents’, **3** *Journal of Legal Studies*, 107-164.
- Donohue, John J. (1989), ‘The law and Economics of Tort Law: The Profound Revolution’ (Book Review of William M. Landes and Richard A. Posner, *The Economic Structure of Tort Law* and of Steven Shavell, *Economic Analysis of Accident Law*), **102** *Harvard Law Review*, 1047-1073.

⁸⁸ See Landes and Posner (1980). The literature on comparative negligence is quite extensive.

⁸⁹ See Calabresi (1970) and Trimarchi (1961).

⁹⁰ See on this point Shavell (1980a) and Shavell (1992).

- Eörsi, Gyula (1975), 'Private and Governmental Liability for the Torts of Employees and Organs', in André Tunc (ed.), *Torts, International Encyclopedia of Comparative Law*, Tübingen, Mohr.
- Epstein, Richard A. (1999), *Torts*, Gaithersburg, Aspen.
- Fleming, John G. (1983), *The Law of Torts* (6th ed.), Sydney, Law Book Company.
- Gilles, Stephen G. (1992), 'Rule-based Negligence and the Regulation of Activity Levels', **21** *Journal of Legal Studies*, 319-363.
- Grady, Mark (1983), 'A New Positive Theory of Negligence', **92** *Yale Law Journal*, 799-829.
- Isaac, Nathan (1918), "Fault and Liability", **31** *Harvard Law Review*, 954-979.
- Kahan, Marcel (1989), 'Causation and Incentives to Take Care under the Negligence Rule', **18** *Journal of Legal Studies*, 427-447.
- Kaplow, Louis and Shavell Steven (2002), 'Economic Analysis of Law', in Howerback, Alan J. and Feldstein, Martin S. eds., *Handbook of Public Economics*, Amsterdam, Elsevier Science.
- Landes, William M. and Posner, Richard A. (1980), 'Multiple Tortfeasors: An Economic Analysis', **9** *Journal of Legal Studies*, 517-555.
- Landes, William M. and Posner, Richard A. (1981), 'The Positive Economic Theory of Tort Law', **15** *Georgia Law Review*, 851-924.
- Landes, William M. and Posner, Richard A. (1987), *The Economic Structure of Tort Law*, Cambridge (MA), Harvard University Press.
- Latin, Howard (1987), 'Activity Levels, Due Care, and Selective Realism in Economic Analysis of Tort Law', **39** *Rutgers Law Review*, 487-514.
- Le Gall, Jean-Pierre (1976), 'Liability for Persons under Supervision', in André Tunc (ed.), *Torts, International Encyclopedia of Comparative Law*, Tübingen, Mohr.
- Levmore, Saul (1986), "Rethinking Comparative Law: Variety and Uniformity in Ancient and Modern Tort Law", **61** *Tulane Law Review*, 235-287.
- Miceli, Thomas J. (1997), *Economics of the Law*, New York, Oxford University Press.
- Parisi, Francesco (1992), *Liability for Negligence and Judicial Discretion* (2nd ed.), Berkeley, University of California Press.
- Parisi, Francesco (2001), 'Genesis of Liability in Ancient Law', **3** *American Law and Economics Review*, 82-124.
- Polinsky, A. Mitchell and Che, Yeon-Koo (1991), 'Decoupling Liability: Optimal Incentives for Care and Litigation', **22** *Rand Journal of Economics*, 562-570.
- Posner, Richard A. (1972), 'A Theory of negligence', **1** *Journal of Legal Studies*, 29-96.
- Posner, Richard A. (1980), 'A Theory of Primitive Society with Special Reference to Law', **23** *Journal of Law and Economics*, 1-53.
- Posner, Richard A. (1992), *Economic Analysis of Law* (4th ed.), New York, Aspen Publishers, Inc..
- Schäfer, Hans-Bernd and Schönenberger, Andreas (2000), 'Strict Liability versus Negligence', in Bouckaert, Boudewijn and De Geest, Gerrit (eds.), *Encyclopedia of Law and Economics*, Vol. II, Cheltenham, Edward Elgar, 597-624.
- Schwartz, Gary T. (1981), 'Tort Law and the Economy in Nineteen-Century America: A Reinterpretation', **90** *Yale Law Journal*, 1717-1775.

- Shavell Steven (1980a), 'Strict Liability versus Negligence', **9** *Journal of Legal Studies*, 1-25.
- Shavell Steven (1980b), 'An Analysis of Causation and the Scope of Liability in the Law of Torts', **9** *Journal of Legal Studies*, 463-516.
- Shavell, Steven (1987), *Economic Analysis of Accident Law*, Cambridge (MA), Harvard University Press.
- Shavell, Steven (1992), 'Liability and the Incentives to Obtain Information about Risk', **21** *Journal of Legal Studies*, 259-270.
- Stone, Ferdinand F. (1972), 'Liability for Damages Caused by Things', in André Tunc (ed.), *Torts, International Encyclopedia of Comparative Law*, Tübingen, Mohr.
- Terry, Henry T. (1915), 'Negligence', **29** *Harvard Law Review*, 40 ff.
- White, Edward G. (1980), *Tort Law in America. An Intellectual History*, Oxford, University Press.

COMPARATIVE NEGLIGENCE AS A FILTER MECHANISM WHEN COURTS CANNOT OBSERVE PRECAUTION COSTS*

Giuseppe Dari Mattiacci
Utrecht University

Gerrit De Geest
Utrecht University and Ghent University

ABSTRACT

This chapter presents the filtering properties of comparative negligence. In unilateral precaution accidents, if the court cannot observe which party could have avoided the accident, all-or-nothing rules fully direct the incentive stream towards one party, and randomly prevent only half the accidents. On the contrary, comparative negligence directs a weaker incentive stream towards both parties so to perform as a filter that prevents the most harmful accidents and let the less harmful ones occur. We show that comparative negligence is likely to generate a smaller social loss than all-or-nothing rules, and discuss the conditions for this result to hold. We also consider alternative precaution accidents and the possibility that courts over- or under-compensate the victim's harm.

JEL classification: K13

Keywords: contributory negligence, comparative negligence, all-or-nothing rule, joint precaution, alternative precaution, tort law, liability, filter.

* Giuseppe Dari Mattiacci, Economic Institute/CIAV, Utrecht University, Kromme Nieuwegracht 22, 3512 HH Utrecht, the Netherlands. Email: g.darimattiacci@econ.uu.nl; Gerrit De Geest, Economic Institute/CIAV, Utrecht University, and School of Law, Ghent University. Email: gerrit.degeest@rug.ac.be. The authors would like to thank Hans Bern Schäfer, Winand Emons, Robert Dur, Carlo Ferrigno, Jana McLean, Ben van Velthoven and the participants in the 17th annual conference of the European Association of Law and Economics, Ghent (2000), and in the annual meeting of the Italian Economists Society, Rome (2001), for helpful comments. Updated versions of this chapter will possibly be available on my web page: www.law.uu.nl/eco/gdm.htm.

4.1. Introduction

If both the injurer and the victim of an accident are negligent, the tort system has two fundamental options. First, it may direct the entire incentive stream towards either party, by holding him fully liable (all-or-nothing rules, ANRs). That party will be the injurer under the rules of simple negligence and strict liability with defense of dual contributory negligence and the victim under contributory negligence and strict liability with defense of contributory negligence. Second, it may direct a portion of the incentive stream towards the injurer and the remaining portion towards the victim (comparative negligence, CN). Old tort systems chose contributory negligence, an all-or-nothing rule that fully targets the victim. Modern tort systems generally prefer comparative negligence (Curran, 1992).

The shift from contributory negligence to comparative negligence is still largely unexplained in the law and economics literature. According to the ‘efficiency-equivalence theorem’ (for a survey see Schäfer and Schönenberger, 2000, 608-610), all negligence rules (simple negligence, contributory negligence and comparative negligence) lead to optimal incentives, if the due level of precaution is correctly defined (Landes and Posner, 1980, Haddock and Curran, 1985). Contributory negligence, however, generates less litigation costs (Low and Smith, 1985) and has been empirically shown to give parties stronger incentives to take precaution (White, 1989) than comparative negligence; it should therefore be preferred. The opinion of early writers about the efficiency of comparative negligence was largely negative (Brown, 1973, and Green, 1976, at 558) or was that the incentive effect of that rule strictly depends on the liability fraction (Diamond, 1974, at 140-145¹). Nevertheless, more recently and after the appearance of the efficiency equivalence theorem, some authors have developed economic arguments for comparative negligence, which seems to be advantageous when some of the assumptions of the standard model are not satisfied².

Cooter and Ulen (1986) argued that, if the courts make random errors in comparing the due level of precaution to the level of precaution actually taken by the parties (evidentiary uncertainty), they are induced to take an inefficient level of precaution in order to minimize their expected liability; comparative negligence, by distributing the loss between the parties, reduces the effect of errors and results in less distortions. Edlin (1994) argued instead that in these situations comparative and contributory negligence only differ with respect to the standard of negligence that is to be implemented for an efficient outcome to result.

Comparative negligence has been said to be more efficient than contributory negligence

¹ Diamond (1974) did not relate, however, the sharing rule under comparative negligence to the standard of due precaution, whereas Landes and Posner (1980) did so.

² See on this point Chung’s (1993) comment on Orr (1991).

when the due level of precaution is uniform for all parties, but the individual costs of precaution differ (Rubinfeld, 1987), and when judges err regarding the precaution costs actually borne by the parties (Haddock and Curran, 1985). Shavell (1987) showed that comparative negligence could be superior in non-equilibrium situations, when (for some reasons) victims or injurers are found negligent. Nevertheless, a counter-argument against comparative has been recently proposed by Bar-Gill (2001).

Under a different perspective, comparative negligence can enhance a more efficient outcome when individuals are not identical (Emons, 1990a, 1990b, Emons and Sobel, 1991) and when parties make precaution decisions sequentially, so that one party can observe the other party's level of precaution before choosing his (Rea, 1987, Grady, 1990). A completely different approach eventually leads to fairness and equity arguments in favor of comparative negligence, insofar as sharing losses when parties are similarly at fault is socially regarded as just.

All the studies mentioned above focused on the use of comparative negligence as a criterion to share the accident loss when both parties are at fault in order to improve the parties' incentives to take care. Parisi and Fon (2001) highlight instead a different point, as they argue that loss sharing can be profitably implemented even in those situations in which both parties have been found non-negligent, in order to provide them with improved incentive to take optimal activity level. In those cases the accident loss is apportioned on the basis of the parties' relative causal contribution to the loss rather than on the basis of their relative degree of negligence.

This chapter will add a new efficiency argument. We will develop a model in which courts cannot observe precaution costs at all³ and argue that in this case the best solution is to share the accident costs between the parties (comparative negligence). Consequently, our argument applies both to those precautionary measures traditionally regarded as care and to the ones usually considered as an activity level. The application to the latter is even more evident, as the accepted view is that the activity level is difficult or impossible to verify before the court⁴.

Suppose, for instance, that after a great conference dinner, a law and economics scholar is walking through a small dark street in Ghent while discussing new theories with a colleague. He hardly notices what is happening around him. When he suddenly goes one step to the left, a car hits him. A judge has to decide who should bear the accident costs. However, he cannot determine the optimal level of precaution because he has no information about the

³ Informational constraints on the courts' ability to determine the parties' costs of precaution was already considered as a source of inefficiency by Green (1976), who however only focused his analysis on the choice of the best all-or-nothing rule; so did Diamonds (1974b) and Diamonds and Mirrlees (1975).

⁴ See Shavell (1980 and 1987), I have analyzed the definitions of care and activity level in chapter 3.

precaution costs. All he knows is that, in theory, the victim could have avoided the accident by looking around all the time, and that, in theory, the injurer could have avoided the accident by reducing his speed. What is the pedestrian's cost of always looking around? What is the motorist's cost of reducing his speed?⁵

Although the judge cannot appropriately determine due levels of precaution, he has to decide the case. He has three possibilities: letting the injurer pay the accident costs, letting the victim pay or sharing the costs between them. In addition, the judge has to find legal arguments (doctrines) to justify his decision.

Most modern courts would probably share the costs. Although many legal techniques exist to reach this result, the most straightforward one is *comparative negligence*⁶. In the past, the courts would probably have let the victim bear the accident costs. The most straightforward doctrine to justify this decision is *contributory negligence*⁷. A third possibility is that the injurer bears all accident costs, and it can be justified by the doctrine of *simple negligence*⁸.

In the analysis we will focus on accidents of the alternative precaution type (where the accident can be avoided by either party) as distinct from the joint precaution type (where both parties need to be careful in order to prevent the accident). The alternative precaution type will be further divided in three types: unilateral precaution accidents (only one party can avoid the accident at a reasonable cost), alternative precaution accidents strictly defined (each of the parties can avoid the accident at a reasonable cost), and efficient accidents (no party can avoid the accident at a reasonable cost). Since we assume that the courts cannot observe the precaution costs, they know neither who is the (least cost) avoider nor whether the accident is of unilateral, alternative or efficient type⁹.

If the accident is in reality a unilateral precaution case, comparative negligence operates as a filtering mechanism which prevents the most harmful accidents, that is, it prevents all

⁵ The difficulties to determine levels of due precaution may be increased by a lack of information about probabilities. What is the probability of causing an accident if a pedestrian is not paying attention to the cars in that particular street of Ghent at that particular time? What is the probability of causing an accident if a car driving is driving 50 km/h in that particular street of Ghent at that particular time? Therefore, it is not unrealistic to assume that in some cases courts cannot determine optimal level of precaution.

⁶ Another legal technique to obtain the same result is to declare that only the injurer was negligent and to deliberately underestimate the damages. That there can be difference between the legal doctrines as they are defined in the economic literature and the names of the doctrines used by courts to reach the same results has been remarked by other writers, see Rubinfeld (1987, p. 367).

⁷ Other doctrines which reach the same result are strict liability with defense of contributory negligence and no liability.

⁸ Other doctrines which would attain the same result are strict liability with defense of dual contributory negligence and strict liability. Under all-or-nothing rules the outcome might also depend on the burden of proof: if the burden of proof is on the victim and he fails in proving the injurer's fault, he will be the cost-bearer; to the contrary, if the burden of proof is on the injurer and he fails in proving his carelessness, he will be the cost-bearer. Nevertheless, the crucial point is still that all-or-nothing rules sanction only one party and do not allow any cost sharing.

⁹ Nevertheless it is plausible that the court will be able to distinguish between joint precaution and alternative

those accidents in which the precaution cost was low and the expected harm was high.

Under all-or-nothing rules, such as simple or contributory negligence, the judge only targets one party, and has a 50% chance to provide the party which could have avoided the accident with incentives to take precaution. So only in 50% of the cases the accident will be prevented.

On the contrary, if each party bears 50% of the accident costs, the avoider will surely be targeted, but the crucial question is whether the incentive is powerful enough to induce him to prevent the accident. The outcome depends on the relative magnitude of his precaution cost, compared to the expected harm. If the precaution cost is lower than 50% of the expected harm (if, for instance, the precaution cost is 30%), the accident will be prevented. If the precaution cost is higher than 50% (for instance, 70%), the party will not have an incentive to prevent the accident, as the cost of preventing the accident is higher than the portion of the expected harm that he will have to bear.

Therefore, comparative negligence does not prevent all accidents either. But *comparative negligence will allow a loss sharing which prevents the most harmful accidents*, that is, it prevents the accidents in which the expected harm is much higher than the precaution cost. *Comparative negligence performs as a clever filter*. For instance, if one group of accidents can be avoided at a cost equal to 30% of the expected harm and another group at 70%, comparative negligence prevents all accidents of the first group and allows all accidents of the second group. *All-or-nothing rules (such as simple negligence or contributory negligence) are blind selection mechanisms*: they simply have a 50% chance to target the avoider and thus prevent some of the most harmful as well as some of the least harmful accidents.

The overall superiority of comparative negligence depends, however, on the distribution of the precaution costs. Suppose that all accidents can be avoided at a cost equal to 70% of the expected harm. In this case, comparative negligence prevents no accidents at all, while an all-or-nothing rule prevents 50% of the accidents. This makes the analysis more complex. In section 3, we will first demonstrate the superiority of comparative negligence under a uniform distribution of accident types (there are as many accidents that can be prevented at 10% of the expected harm, as at 20%, ..., 80%, 90%, ...), then we will derive a general criterion to choose between comparative negligence and all-or-nothing rules under non-uniform distributions of accident types. In *appendix 1* we will show by means of a computer simulation that comparative negligence generates more efficient results in more cases than all-or-nothing rules.

When both parties can avoid the accident at a reasonable cost (alternative precaution strictly defined), the results become ambiguous, as shown in *section 4*. In this case, efficiency

precaution accidents, as it will be clarified in section 4.2.

requires that only the least cost avoider takes precaution. All-or-nothing rules provide only one party with incentives to prevent the accident, and risk to target the wrong one: the accident will be avoided anyway but at too high a cost. Comparative negligence risks to give incentives to both parties (one party's precaution cost could have been saved) or no party at all. Whether this risk outweighs the benefits of the filtering effect depends on the distributions of the precaution costs among individuals.

When accidents are efficient (*section 5*), comparative negligence is superior, for another reason: it suffers less from overdeterrence (caused by overcompensation) than all-or-nothing rules.

4.2. The model

Two parties are in a situation that might result in an accident. The victim is the party that will suffer harm; the injurer is the other party.

Discrete precaution levels. For reasons of simplicity, we assume that each party has a discrete choice between being careful and careless.

Varying precaution costs. Each party can prevent the accident at a cost that is lower than infinite. We denote the precaution costs as c and more specifically the precaution costs of the injurer as c_I and the precaution costs of the victim as c_V . The two costs may have different values, and the chance that the precaution costs have a certain value may differ between the parties. In other words, the distribution of the accident types with respect to precaution costs (accidents that are relatively inexpensive to prevent, accident that are relatively more expensive to prevent, ...) can have any form and is not necessarily the same for both parties. In the appendix a computer simulation for different distribution curve shapes will be discussed. Precaution costs are said to be *reasonable* if they are lower than (or equal to) the harm ($c \leq 1$). For instance if $c_I = 0.25$ and $c_V = 1.3$, we will say that the injurer can avoid the accident at a reasonable cost, while the victim cannot.

Harm normalized to one. For mathematical simplicity, we assume that if both parties are careless, an accident will certainly occur. This avoids the possible confusion between accident costs and expected accident costs (= accident costs times the probability of an accident if both are careless). This assumption is relaxed in the appendix. The harm that the victim suffers is denoted by h . For mathematical simplicity, we will assume that h always equals 1 ($h = 1$). The normalization of h allows us to express the precaution costs as a percentage of the harm. For instance if $c_I = 0.25$ and $c_V = 1.3$, this means that the injurer can avoid the accident at a precaution cost that is 25% of the accident costs and that the victim can avoid the accident at a

precaution cost that is 130% of the accident costs.

Types of accidents. Alternative precaution family: the accident is prevented if at least one party is careful. *Further division of the alternative precaution family:* if only one of the parties can avoid the accident at a reasonable cost ($c_I \leq 1$ and $c_V > 1$, or $c_I > 1$ and $c_V \leq 1$), we will speak of a *unilateral precaution accident*. If each of the parties can avoid the accident at a reasonable cost ($c_I \leq 1$ and $c_V \leq 1$), we will speak of an *alternative precaution accident (strictly defined)*. If none of the parties can avoid the accident at a reasonable cost ($c_I > 1$ and $c_V > 1$), we will speak of an *efficient accident*¹⁰.

(Least-cost) avoider, non-avoider. In the case of unilateral precaution accidents, we will refer to the party who can avoid the accident at a reasonable cost as the *avoider* and to the other party as the *non-avoider*. In other words, the avoider can be the injurer as well as the victim. In alternative precaution cases (strictly defined), we will use the term *least cost avoider* for the party whose precaution costs are the lowest.

Courts cannot observe precaution costs. The crucial assumption of our model is that courts cannot observe precaution costs at all. Since they do not know c_I and c_V , they do not know whether they are dealing with a unilateral precaution accident, an alternative precaution accident (strictly defined), or an efficient accident. If it is a unilateral precaution accident, they do not know what party (the injurer or the victim) is the avoider and if it is an alternative precaution accident, they do not know who is the least cost avoider. Courts do not know the likelihood of the distribution of precaution costs either. They have no reason, for instance, to assume that injurers in general bear lower precaution costs than victims or vice versa.

Can courts observe precaution levels? As explained above, both parties have a discrete choice between being careful or careless. Can courts observe who has been careful and who has been careless? If accidents are of the alternative precaution type, this question is irrelevant. If a case is brought before court, this means that an accident has taken place and that both parties have been careless. If one party was careful, or both were careful, an accident would not have occurred.

Do courts know whether the accident is of the alternative precaution family or of the joint precaution family? In our analysis we assume that courts do know of what type the accident is. This is quite realistic an assumption. Even if courts have no information on the precaution

¹⁰ In the literature there is some confusion between the concepts *unilateral precaution* and *alternative precaution*. In the archetype of unilateral precaution accidents, one party can avoid the accident at a cost < 1 , while the other party's precaution costs are infinite. In the archetype of alternative precaution accidents, each of the parties can prevent the accident at a cost < 1 . Terminological confusion may arise when one party can avoid the accident at a cost < 1 , while the other party's precaution costs are > 1 but not infinite and when both parties' precaution costs are > 1 but not infinite.

costs, they can in most cases figure out whether only either party (alternative precaution cases) or both parties (joint precaution cases) should have changed their behavior in order to avoid the accident. We focus the analysis only on the alternative precaution family.

Legal rules. We assume that courts have only 3 choices: (i) letting the injurer bear all accident costs, (ii) letting the victim bear all accident costs, or (iii) letting the parties share the accident costs. For (i) we will use the term *simple negligence*, for (ii) *contributory negligence* and for (iii) *comparative negligence* (CN), although courts can also use other labels and legal techniques to reach the same results. *All-or-nothing rule* (ANR) is the term that we will employ to refer to simple negligence and contributory negligence. For simplicity, we assume that under comparative negligence, the losses will be equally shared between the parties (a 50/50 division), although in *appendix 4* we will look at the filtering characteristics of different sharing rules (such as 60/40, 70/30,). We will denote as d the damages paid by the injurer. Under simple negligence and perfect compensation, the injurer pays $d=1$. Under contributory negligence, the injurer pays $d=0$. Under comparative negligence (with a 50/50 division), the injurer pays $d=0.5$.

4.2.A. *Standard assumptions*

In our model, we will make a number of additional assumptions that can be considered as standard in the economic literature on tort law:

- (i) Only the tort law system generates an incentive stream. Criminal sanctions (a second incentive stream) do not exist. The incentives created by the tort law are not undermined by insurance (insurance is not available);
- (ii) Parties are perfectly informed on each other precaution cost and on the harm (parties' *perfect information*) and know exactly who will be declared responsible in the case of an accident and to what extent (*perfect predictability of the judicial decisions*, this assumption will be relaxed in *section 4.3.F*);
- (iii) No under- or overcompensation by the court (*perfect compensation*, this assumption will be relaxed in *section 4.3.D*);
- (iv) Parties never escape liability (the apprehension rate is 100% and there is no judgment proofness);
- (v) Parties are rational and utility maximizing.
- (vi) Since we solely focus on incentives and not on transaction costs or the costs of a suboptimal risk allocation, we assume transaction costs away and we assume that parties are risk neutral.

4.2.B. *The graph*

To clarify the analysis we use a simple graph, which shows on the y-axis the relative costs of

precaution (calculated as a percentage of the harm). As already explained, the harm h is normalized to 1. As a consequence, the costs of precaution are also normalized by the expected harm. For instance, if the expected harm is \$100,000 and the precaution cost is \$70,000, in our normalized model the harm h is 1 and the precaution cost is 0.7.

The x -axis shows the ranked order of accidents, from the accidents which are (relatively) inexpensive to avoid to the accidents which are too expensive to avoid. Thus on the x -axis all accidents are ranked on the basis of their y -value. For mathematical simplicity, the ranked order of accidents is also normalized by the number of accidents to be avoided (those accident the precaution cost of which is smaller than the expected harm), so that $x=1$ means 100% of the accidents to be avoided. By definition, the precaution cost curve slopes upward ($dy/dx \geq 0$): because of the ranked order of accidents, an accident cannot have a smaller y than the preceding one.

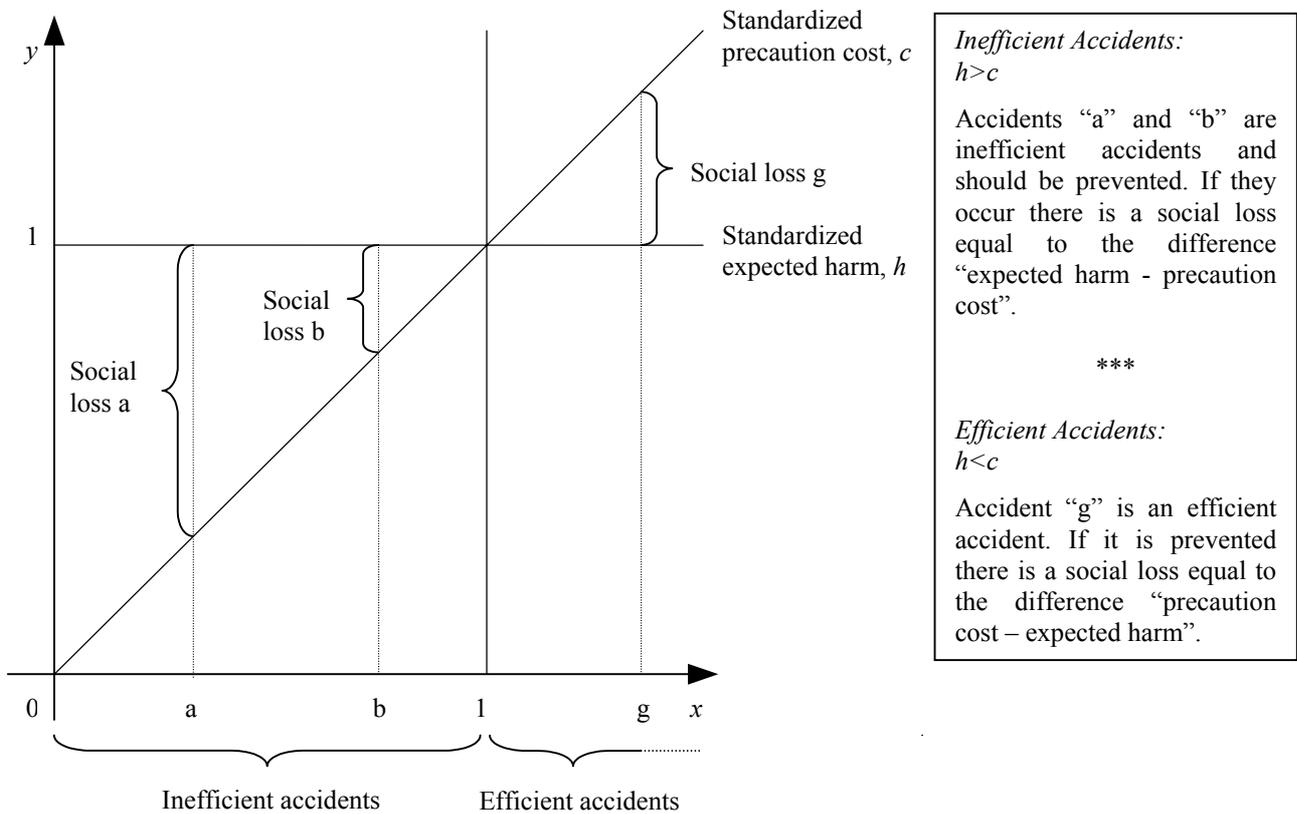


Figure 1. The basic unilateral-precaution model

In figure 1, the precaution cost curve $y(x)$ is a straight line, which means assuming a uniform distribution of accident types (there are as many accidents that can be prevented at 10% of the expected harm as at 20%, 30%,... 90%, 100%, ...) so that the costs of precaution $y=x$ (for further details see *appendix 1*).

4.2.C. *The social welfare function*

Efficiency requires the minimization of the social loss. If $y < I$ ($c < h$) the accident is inefficient; if it occurs, the social cost is the difference between the harm and the costs of precaution. If $y = I$ it is indifferent. If $y > I$ ($c > h$) letting the accident occur is better than preventing it; if the accident is prevented, there is a social loss of a different type: too high a precaution cost. The social loss in that case is the difference between the precaution cost and the expected harm. In the graph the accidents to be avoided are in the region delimited by the $x = I$ line, out of that region a good rule is a rule that lets the accidents occur.

Accidents “a” and “b” in *figure 1* are accidents to be avoided ($c < h$), while accident “g” is an efficient accident ($c > h$) and the legal rule should let it occur. Therefore, the social losses “a” and “b” occur if the accidents “a” and “b” are *not* prevented, the social loss “g” occurs if the accident “g” is prevented.

4.3. Unilateral precaution (only one party can avoid the accident at a reasonable cost)

In this section we will analyze the consequences of judicial decisions concerning accidents of the unilateral precaution type.

As explained in the former section, the crucial assumption in our model is that courts have no information about the precaution costs in an individual case. They also have no information on the distribution curves. Consequently, there is no reason why they would presume that the injurers have on average lower precaution costs than the victims or vice versa. This makes the analysis for both all-or-nothing rules the same: whether the court fully targets the injurer (simple negligence) or the victim (contributory negligence), the court has a 50% chance to provide the wrong party (the non-avoider) with incentives to take precaution. Therefore, the comparison will be made between comparative negligence (CN) and all-or-nothing rules (ANRs).

4.3.A. *Uniform distributions*

Under all-or-nothing rules, the legal system fully targets only one party. That party can be the avoider or the non-avoider. If the legal system targets the avoider (50% chance), the avoider will bear 100% of the harm, which is hard enough a sanction to make him prevent all inefficient accidents (the incentive stream is equal to 1 in *figure 1*). But if the legal system targets the wrong party, the avoider bears 0% of the harm, and he will prevent no accidents at all (the incentive stream is equal to 0 in *figure 1*). In that case, the social loss is indicated by the dotted area in *figure 2*.

Under comparative negligence (as defined above) the judge will equally share the cost

between the parties. In this case the incentive stream is simultaneously directed toward both the victim and the injurer. The avoider always bears 50% of the cost of the accident that he caused. If the injurer has to compensate the victim for half harm, the victim can recover from the injurer only half. Both the victim and the injurer have an incentive stream equal to $\frac{1}{2}$ in *figure 2*.

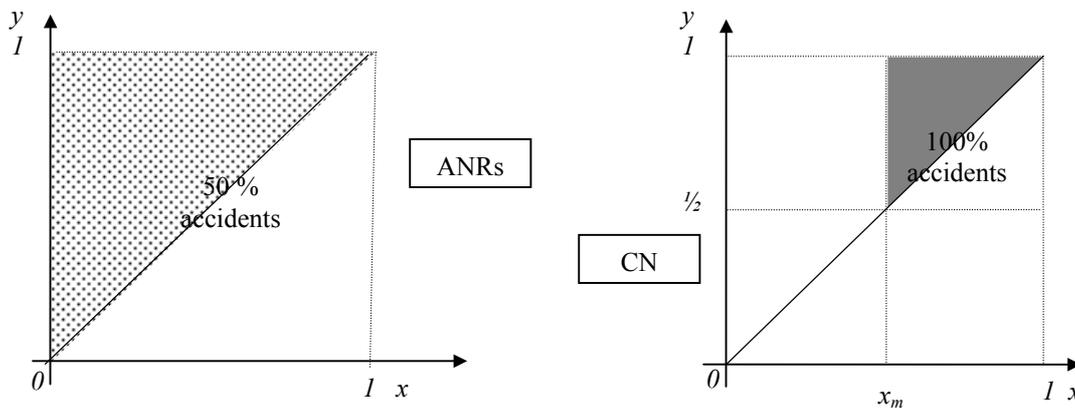


Figure 2. Social loss under ANRs and CN

Under comparative negligence, the avoider will prevent those accidents the precaution cost of which is lower than $\frac{1}{2}$ harm. If the precaution cost is higher, it is cheaper for the avoider to let the accident occur and pay half the accident cost than to pay the full precaution cost. The gray area in *figure 2* shows the social loss.

Comparing the social loss under all-or-nothing rules with the social loss under comparative negligence, it is easy to notice that the dotted area in *figure 2* (ANRs) is 4 times bigger than the gray area in the same figure (CN)¹¹. Since the loss depicted by the big dotted area of the all-or-nothing rules occurs in only 50% of the cases (the legal system only targets the wrong party in 50% of the cases) the social loss is only $\frac{1}{2}$ of that area; so *under all-or-nothing rules the social loss is twice as large as under comparative negligence in the case of a uniform distribution function*.

How can we explain this difference? The fundamental cause of the difference is that comparative negligence operates as a clever *filter mechanism*. Both CN and ANRs prevent 50% of the accidents. But CN prevents the most harmful accidents, i.e. the accidents that can be prevented at a lower cost than half harm. ANRs, on the other hand, provide a blind mechanism, which does not make any selection between more harmful and less harmful accidents.

¹¹ The vertical side of the dotted triangle is twice as high as the vertical side of the gray triangle and the horizontal side is also twice as long.

4.3.B. Non-uniform distributions

In the former analysis, we have assumed a uniform distribution (i.e. there are as many accidents that can be prevented at 10% of the expected harm, as at 20%, ... 80%, 90%, ...). In reality this does not need to be the case. If the precaution cost curve $y(x)$ is not a straight line (i.e. under non-uniform distributions), the social loss caused by comparative negligence may change.

Under a general precaution cost curve, $y(x)$, comparative negligence prevents an accident if $y(x) \leq \frac{1}{2}$ and let the accident occur otherwise. The number of accidents prevented is given by $x_m = f^{-1}(\frac{1}{2})$.

Under an all-or-nothing rule $\frac{1}{2}$ of the accident will randomly occur without any selection. The social loss can be easily calculated by following the same reasoning as for the uniform distribution case. Comparative negligence will enhance a lower social loss if:

$$1 - x_m - \int_{x_m}^1 f(x) dx < \frac{1}{2} \left(1 - \int_0^1 f(x) dx \right),$$

comparative negligence will lead to a greater social loss than under all-or-nothing rules otherwise.

In *appendix 1* we provide the reader with the results of a computer simulation, which was performed in order to test whether or not comparative negligence remains superior if the distribution change. To avoid overly complex mathematics, we restricted the simulation to a specific distribution function: $y(x) = x^q$ with varying q levels. This formula does not allow us to capture all possible distribution forms, but it describes well the essentials of what happens when there are more/less relatively wasteful accidents.

In the former section, we assumed a uniform distribution; this means that $q=1$ and $y(x)=x$. *Figure 3* shows the pattern of the social loss under comparative negligence if q becomes smaller or greater than 1. *Figure 4* shows the pattern of the social loss under all-or-nothing rules.

If we compare the figures, we can notice that the argument pro comparative negligence becomes stronger for $q > 1$ and weaker for $q < 1$. In the appendix, we show that, even for q values lower than 1, comparative negligence remains superior, except for q values lower than $\frac{1}{2}$. And even in those cases, the superiority (in terms of social cost saving) of all-or-nothing rules is never as clear as the superiority of comparative negligence for other q values.

If the precaution cost curve is known, then there is a clear criterion to decide between comparative negligence and all-or-nothing rules. From the opposite standpoint, if the distribution of the precaution cost among parties is not known, comparative negligence should be preferred, for it yields smaller social costs for a wider range of precaution cost

distributions than all-or-nothing rules¹².

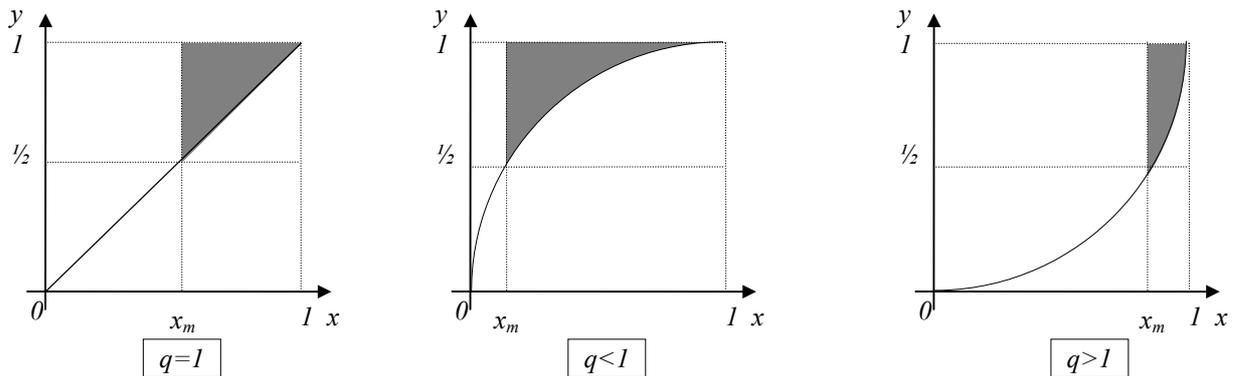


Figure 3. Social loss under CN – three different distributions

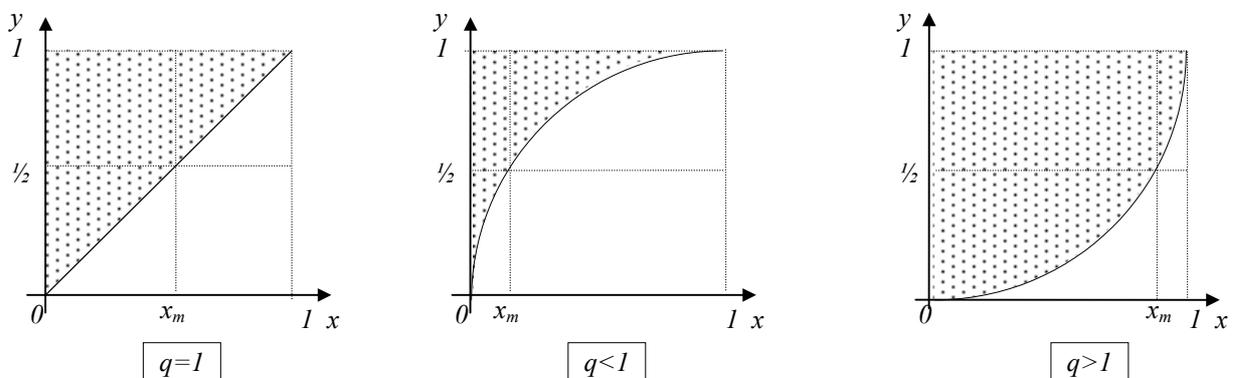


Figure 4. Social loss under ANRs – three different distributions

4.3.C. Sharing rules other than 50/50. All-or-nothing rules as limit-cases of comparative negligence.

What happens to the filtering effect if costs are not equally shared (50/50), but unequally, from 0/100, 10/90, 20/80, ... up to 100/0? Simple negligence (100% of the harm to be borne by the injurer, 0% by the victim) and contributory negligence (0/100) can here be seen as limiting cases of comparative negligence¹³.

Let us consider, for instance, a comparative negligence rule that always lets the injurer bear 80% and the victim 20%. Suppose for simplicity that the distribution of precaution costs is uniform ($q=1$ so that $y(x)=x$). Compared to a 50/50 rule, more accidents where the injurer is the avoider will be prevented but more accidents where the victim is the avoider will occur. The 80/20 still functions as a filter. The 80% accidents prevented at the injurer's side, are the

¹² This result holds if we consider all distribution curves as equally probable.

¹³ The idea that contributory negligence can be seen as a special limiting case of comparative negligence can also be found in Edlin (1994), although applied in a different context.

80% most harmful ones (since the precaution costs was less than 80% of the harm). The 20% accidents prevented at the victim's side, are the 20% most harmful ones (since the precaution costs were less than 20% of the harm). But the accidents that are no longer prevented at the victim's side (the range between 20% and 50%) are more harmful than the accidents that are now prevented at the injurer's side (the range between 50% and 80%). Therefore, an 80/20 rule is less efficient than a 50/50 rule.

The further the division deviates from 50/50, the greater the social loss, or, to put it differently, the smaller the filtering effect. In the extreme cases of (100/0) and (0/100) there is no longer any filtering effect. In *appendix 4* we will discuss different sharing percentages also for non-uniform ($q \neq 1$) distributions of accident costs.

4.3.D. *Under or overcompensation*

The conclusions reached so far have been based on the assumption of perfect compensation of the harm. In this section we will relax this assumption.

Before we start analyzing what happens in case the court over or underestimates the harm, we should remark that this problem is not relevant for one type of all-or-nothing rules: contributory negligence. Since that rule does not award any compensation to the victim when both parties are found negligent, errors in measuring compensation do not change the outcome, which remains 0/100.

Under or overcompensation changes the division of a comparative negligence rule. If the judicial estimate of the harm is 80% of the true harm, a 50/50 division is changed into a 40/60 division: the injurer pays less; thus the victim bears more. A similar underestimate would transform a simple negligence rule (100/0) into a kind of comparative negligence rule (80/20). If the judicial estimate of the harm is 50% of the true harm, a simple negligence rule (100/0) is transformed into a 50/50 comparative negligence rule and a 50/50 comparative negligence rule is transformed into a 25/75 comparative negligence rule. Since 50/50 yields a better filtering effect than 25/75, the original 'simple negligence rule' becomes more efficient than the original 'comparative negligence rule'. But as a matter of fact, a 50/50 comparative negligence rule is just another name for a 50% undercompensatory all-or-nothing rule (that intends to target fully the injurer)¹⁴.

A similar change takes place when courts overcompensate the harm. While reducing the magnitude of damage compensation means shifting the incentive stream from the injurer to the victim, increasing the magnitude of damage compensation means increasing the incentive

¹⁴ In that respect, a badly applied all-or-nothing rule can sometimes be superior to a badly applied comparative negligence rule.

stream to the injurer¹⁵. Yet, the analysis of overcompensation is more complex, because a second source of social loss arises: overdeterrence (too few accidents occur because too much precaution is taken). To put it differently, the legal system may fail not only because (i) an accident that should have been prevented ($c < h$) occurs, but also because (ii) an accident that should occur ($c > h$) is prevented. Since the latter effect is not a problem for ‘unilateral precaution cases’ (in our narrow definition), but for efficient accidents, we will discuss it in *section 5*. In *appendix 2* and *3* we will show that comparative negligence is generally superior to all-or-nothing rules both in the case of over and undercompensation also under non-uniform distribution functions ($q \neq 1$).

4.3.E. *Relaxing the assumption of systematic and predictable errors in compensation*

So far we have analyzed the case in which the parties can predict the judicial decision concerning both the cost-bearer and the amount of damages. In addition, we have assumed that the courts always make the same type of errors to the same extent.

Now we relax this assumption and consider the case in which the parties can still foresee who will be declared responsible but they can no longer know in advance the amount of the compensation. The courts randomly make errors of different types (either over or undercompensation). Thus the parties have to substitute “ d ” (the damages that the victim is entitled to recover) with “ de ” in their cost-benefit analysis; de is the statistical expectation of the judicial decisions with respect to damages to be paid by the injurer; therefore, de is the mean of all possible d weighted by the probability of the judicial error. Those data come from the past decisions of the courts. It is important to note that the result of such a calculation is an expectation of the compensation. The only difference between d and de is the risk involved in the latter. The parties can base on de their behavior as they do on d . Thus, the analysis just made for errors in compensation can be used also in the case of unpredictable random errors.

4.3.F. *Relaxing the assumption of perfect predictability of the judicial decisions*

The conclusions of the previous sections relied on the assumption that the parties know exactly who will be declared responsible in the case of an accident. This assumption was used to build up the analysis and implies that the parties can perfectly predict the judicial outcome, which means that each party knows exactly whether or not the judge will sanction him. What happens if courts use all-or-nothing rules but parties do not know who will be declared liable in a specific case? An extreme example could be a rule that permits the judge to flip a coin in order to decide the case when he cannot observe the precaution costs.

¹⁵ In addition, as the victim gains from the occurrence of accidents, he might intentionally place himself in the condition of being injured.

In that case, the parties do not know who will have to bear the cost of an individual accident. They just know what the probability to be sanctioned and the probability to escape the tort liability are.

$$a = s * h + (1-s) * 0$$

“*a*” is the expected cost of accident which results from the sum of the possible outcomes multiplied by the probability to occur.

“*h*” is the expected harm, which is equal to “*d*” under perfect compensation.

“*s*” is the probability to bear *h* and “*1-s*” is the probability that the other party will bear it.

The judge makes mistakes in 50% of the cases, thus the probability that the avoider is sanctioned is exactly 50%. If $s=50\%$ then $a=50\%*h$

Unpredictable all-or-nothing rules thus lead to the same results as comparative negligence. Also under comparative negligence the expected cost of an accident is $\frac{1}{2}h$ for all the avoiders.

One conclusion is that unpredictable all-or-nothing rules are better than predictable all-or-nothing rules if courts cannot observe precaution costs. That means that, if the court is unable to determine the cost of parties' precaution and if the rule is an all-or-nothing rule, the decision should be taken using a completely casual and unpredictable system. In other words the court had better use a lottery system instead of a legal criterion. From a different point of view, a consistent practice in the judiciary while deciding doubtful cases, such as a criterion adopted by the supreme court and followed by all the judges for the civil law countries or the precedence rule for the common law countries, makes the judicial decision predictable and in this sense it causes more losses than the lack of such a general practice, which allows different judges to decide in different ways and it is closer to random decisions.

In addition and more important, the risk is the only difference between comparative negligence and all-or-nothing rules under non-predictability. Comparative negligence is a risk-free situation. Therefore, under the non-predictability assumption, the advantage of comparative negligence consists of removing the risk from the judicial outcome, by making the outcome that the parties would rationally expect sure¹⁶.

4.4. Alternative precaution in a strict sense (each party can prevent the accident at a reasonable cost)

We will now analyze alternative precaution accidents (in a strict definition). While in unilateral precaution cases only one party can avoid the accident at a reasonable cost, in

¹⁶ The conclusions do not change if we introduce errors in compensating the harm. Again all-or-nothing rules under the non-predictability assumption can be interpreted using the analysis already made for comparative negligence.

alternative precaution cases *either* party can avoid the accident at a reasonable although different cost, so there are two avoiders but only one is the *least cost avoider*. (While in joint precaution situations the accident can be avoided only if *both* parties are careful, in alternative precaution cases *either* party's precaution is enough to avoid the accident.)

As in unilateral precaution cases the legal system should provide the parties with the incentive to prevent an accident only if the precaution cost is smaller than the expected cost of the accident; in addition, efficiency requires only the least cost avoider to take precaution. We denote as c_1 the least precaution cost and as c_2 the other precaution cost, so that $c_1 \leq c_2$. If the wrong party takes precaution there is a social loss due to the fact that the accident could have been avoided at a lower cost. That loss is given by the difference between the two costs of precaution, $c_2 - c_1$.

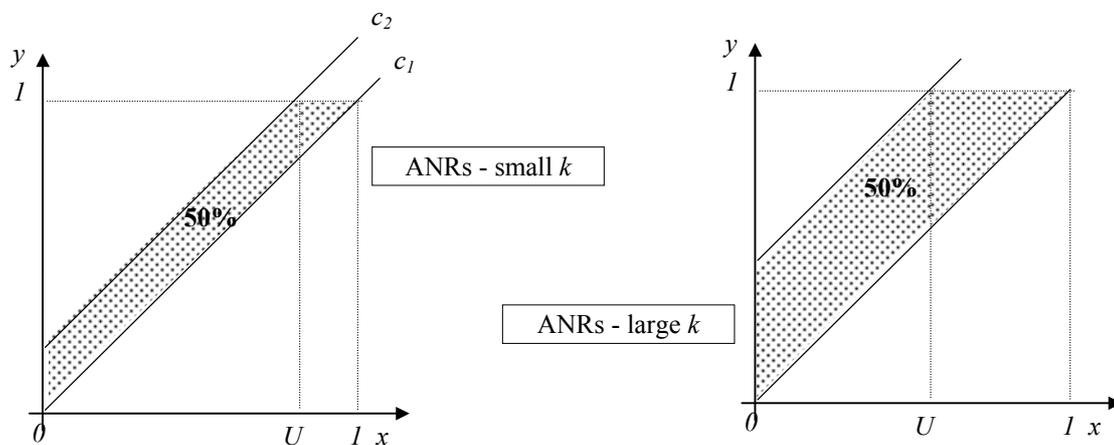


Figure 5. The unknown alternative precaution model - ANRs

For a similar reason, in case both parties take precaution there is a loss equal to the higher cost c_2 , because the least cost avoider's precaution is already enough to avoid the accident.

If no one takes precaution, and an inefficient accident occurs, the loss is given by the difference "expected harm - least precaution cost" ($h - c_1$), because the least cost avoider is always the one who should be careful.

Basically, we can model alternative precaution situations by introducing a second precaution cost in the unilateral precaution model. For simplicity, we assume that there is a constant relationship between the two costs of precaution. This is quite heavy an assumption but allows us to illustrate some crucial points of the theory. More specifically we assume that whatever the shape of the least cost function, the highest cost is equal to the least cost plus a constant (k), so that the vertical distance between the two curves is always equal to k .

$$c_2 = c_1 + k, \quad k \geq 0.$$

In *figure 5* we use straight lines, but the conclusions can be verified for all kinds of distributions. We first consider small values of k , which implies that the two precaution cost functions are close to each other, then we account for larger values of k .

The area containing the accidents to avoid is determined by the least precaution cost, because it is efficient to avoid an accident as far as the least precaution cost is lower than the expected harm ($c_1 < h$). However, it is clear that, because $c_2 > c_1$, there is a point at which c_2 becomes bigger than h while c_1 is still smaller than h (accident to avoid). We denote this point as U .

Before U the situation can be described as an alternative-precaution situation strictly defined, but to the right of that point the least cost avoider becomes the only avoider, because he is the only one who could take precaution at a reasonable cost (which is a cost smaller than the expected harm). Strictly speaking we are now in a unilateral precaution situation.

Under all-or-nothing rules, alternative precaution has a major advantage with respect to unilateral precaution. If the judge makes an error and does not direct the incentive stream towards the least cost avoider, the other party will avoid the accident anyway. The loss is just the difference between the two costs of precaution. If the cost functions are close to each other the loss will be small. In unilateral precaution cases the loss would be half the triangle above the c_1 function, but in alternative precaution cases, the loss is reduced at half the area between the two cost functions.

We want to underline that the loss increases or decreases as k increases or decreases. If k is zero (i.e. the two cost functions are equal), also the loss is zero, because the accident will always be avoided at the optimal cost, irrespective of the party that the court targets. On the contrary, if k increases, the area between the two curves tends to fill the whole triangle; " U " moves to the left and the unilateral precaution zone becomes bigger. Thus, under all-or-nothing rules the social loss is directly related to k .

Under comparative negligence the relationship between k and the social loss is completely different. Because the court shares the damages between the parties, it gives an incentive to avoid the accident only if $c < \frac{1}{2}h$.

When $c_1 > \frac{1}{2}h$, also $c_2 > \frac{1}{2}h$, so to the right of x_m no party will take precaution (zone III in *figure 6*). The loss is exactly the same as in unilateral precaution cases. The U -line does not influence at all that case.

When $c_2 > \frac{1}{2}h$ and $c_1 < \frac{1}{2}h$ (zone II), only the least cost avoider will take precaution and the outcome is efficient.

So far there is not any difference with respect to unilateral precaution, but if we look at

zone I, where both c_2 and c_1 are smaller than $\frac{1}{2}h$, a game theoretical problem arises. Each party has the incentive to prevent the accident but he prefers the other party to take precaution. Such a game has a mixed-strategy solution, which attaches a probability to each party's careful behavior. The outcome is efficient only in a small portion of the accidents, when only the least cost avoider takes precaution. In the rest of the cases it can happen that both parties, no one or the wrong party takes precaution, which generates a social loss. So in zone I the waved area indicates the social loss due to the parties' strategic behavior.

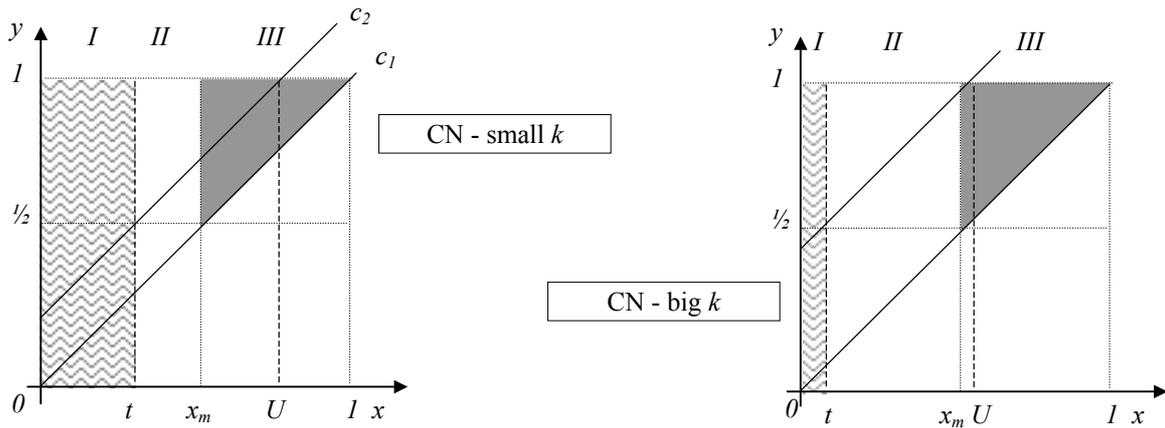


Figure 6. The unknown-alternative-precaution model - CN

Again, we want to underline that there is a relationship between k and the social loss also in the case of comparative negligence. The loss in zone III does not hinge on k , and therefore remains constant for any value of k . On the contrary, if k increases, c_2 goes quickly above $\frac{1}{2}h$ and the number of cases in which both parties receive an incentive stream (zone I) becomes smaller, and thus also the social loss due to strategic behavior decreases. Thus, under comparative negligence the loss is (approximately) inversely related to k .

If k reaches an extremely high value, so that c_2 is always above $\frac{1}{2}h$, the loss due to strategic behavior completely disappears, because only the least cost avoider has the incentive to take precaution. The total social loss is the same as in unilateral precaution cases.

In conclusion, alternative precaution generates a more complex situation than unilateral precaution, mostly because of the strategic behavior of the parties under comparative negligence. However, our simple model shows that if the difference between the two precaution costs is small, all-or-nothing rules are to be preferred over comparative negligence. Under all-or-nothing rules the accident will be always avoided even if at little excessive a precaution cost.

On the contrary, when k assumes high values, the outcome is different, because large

values of k tend to eliminate the game theoretical problem due to comparative negligence. As k increases, an alternative-precaution situation slowly becomes closer to a unilateral-precaution situation in which comparative negligence is superior.

4.5. Efficient accidents (each party can prevent the accident, but no party at a reasonable cost)

Now let us analyze efficient accident accidents ($c_I > 1$ and $c_V > 1$). If there is any social loss here, it is due to overdeterrence (contrary to unilateral precaution cases, in which inefficient outcomes are due to underdeterrence).

It is easy to see that if there is no overcompensation, there can be no overdeterrence and hence the outcome must be efficient. A party will take precaution only if his precaution cost is lower than the damages that he would have to pay. As precaution costs are $c > 1$ for each party, damages d must be higher than 1 too.

Since overcompensation is impossible for contributory negligence (that lets the victim bear the accident costs), only one all-or-nothing rule is vulnerable to overdeterrence: simple negligence. Of course comparative negligence is vulnerable to overdeterrence too. Suppose, for instance, that the court's estimate of the harm is 300% of the real harm. A 50/50 rule will let the injurer pay 150% of the harm. If the injurer's precaution costs $c_I = 1.2$, the injurer will prevent the accident. In *appendix 2* we will formally demonstrate that comparative negligence remains superior to simple negligence in the case of overcompensation. We will now just briefly explain the underlying intuitions. As long as the court's estimate is lower than 200% of the real harm, comparative negligence remains also superior to contributory negligence.

The intuition can be explained as follows. Overcompensation increases the incentive stream above the level of the harm and stimulates the avoider to take precaution even in those cases in which the precaution cost is greater than the harm, and thus avoiding the accident is uneconomical. This problem arises for simple negligence already at minor forms of overcompensation: instead of 100 for the injurer and 0 for the victim, it is now 110/0, 120/0, 130/0, ... For comparative negligence, this overdeterrence effect only starts at high levels of overcompensation. For low levels, it simply changes the balance between the cost to the injurer and the cost to the victim: 50/50 becomes 55/45, 60/40, 65/35, ... Comparative negligence remains superior, although the superiority becomes less and less explicit since overcompensation gives comparative negligence more and more the characteristics of all-or-nothing rules. Overdeterrence only takes place when the damages are more than twice as much as the real harm to the victim (what we call 'large errors'). But even for large errors, the overdeterrence problem is always more serious for simple negligence than for comparative

negligence.

Comparative negligence remains superior also to contributory negligence as long as the court's estimate is lower than 200% of the real harm. Suppose that the court's estimate is 180% of the true harm. A 50/50 sharing rule becomes then a 90/10 sharing rule. But this is still better than the 0/100 division of contributory negligence, since 90/10 still creates some filtering effect, while 0/100 creates no filtering effect at all.

4.6. Concluding remarks

This chapter has shown that if courts cannot observe the precaution costs, the best rule is comparative negligence. The main reason is that only comparative negligence creates a filtering effect that prevents the most harmful accidents. The case for comparative negligence is very clear in unilateral precaution situations. In alternative precaution cases, strictly defined, results are ambiguous, but one cannot say that all-or-nothing rules are superior, since comparative negligence keeps its filtering effect.

In the case of efficient accidents, comparative negligence is superior, although for a different reason: it is less exposed to the risk of overdeterrence.

The strength of these arguments depends of course on the plausibility of our assumptions. The crucial assumption that courts cannot observe precaution costs at all may be unrealistic in a number of real life situations. We can conclude, however, that the greater the difficulties that the courts encounter while acquiring information on the precaution costs, the stronger the argument for comparative negligence.

4.7. References

- Bar-Gill, Oren (2001), 'Does Uncertainty Call for Comparative Negligence?', *John M. Olin Center for Law, Economics and Business, Discussion Paper No. 346*, Harvard Law School.
- Chung, Tai-Yeong (1993), 'Efficiency of Comparative Negligence: A Game Theoretic Analysis', *22 Journal of Legal Studies*, 395-404.
- Cooter, Robert D. and Ulen, Thomas S. (1986), 'An Economic Case for Comparative Negligence', *61 New York University Law Review*, 1067-1110.
- Curran, Christopher (1992), 'The Spread of the Comparative Negligence Rule in the United States', *12 International Review of Law and Economics*, 317-332.
- Diamond, Peter A. (1974a), 'Single Activity Accidents', *3 Journal of Legal Studies*, 107-164.
- Diamond, Peter A. (1974b), 'Accident Law and Resource Allocation', *5 Bell Journal of Economics*, 366-406.
- Diamond, Peter A. and Mirrlees, James A. (1975), 'ON the Assignment of Liability: The Uniform Case', *6 Bell Journal of Economics*, 487-516.

- Edlin, Aaron S. (1994), 'Efficient Standards of Due Precaution: Should Courts Find More Parties Negligent under Comparative Negligence', **14** *International Review of Law and Economics*, 21-34.
- Emons, Winand (1990a), 'Efficient Liability Rules for an Economy with Non-identical Individuals', **42** *Journal of Public Economics*, 89-104.
- Emons, Winand (1990b), 'Some Recent Developments in the Economic Analysis of Liability Law: An Introduction', **146** *Journal of Institutional and Theoretical Economics*, 237-248.
- Emons, Winand and Sobel, Joel (1991), 'On the Effectiveness of Liability Rules when Agents Are Not Identical', **58** *Review of Economic Studies*, 375-390.
- Grady, Mark (1990), 'Multiple Tortfeasors and the Economy of Prevention', **19** *Journal of Legal Studies*, 653-678.
- Green, Jerry (1976), 'On the Optimal Structure of Liability Laws', **7** *Bell Journal of Economics*, 553-574.
- Haddock, David D. and Curran, Christopher (1985), 'An Economic Theory of Comparative Negligence', **14** *Journal of Legal Studies*, 49-72.
- Landes, William M. and Posner, Richard A. (1980), 'Joint and Multiple Tortfeasors: An Economic Analysis', **9** *Journal of Legal Studies*, 517-555.
- Low, Stuart and Smith, Janet Kiholm (1995), 'Decision to Retain Attorneys and File Lawsuits: An Examination of the Comparative Negligence Rule in Accident Law', **24** *Journal of Legal Studies*, 535-557.
- Orr, Daniel (1991), 'The Superiority of Comparative Negligence: Another Vote', **20** *Journal of Legal Studies*, 119-129.
- Parisi, Francesco and Fon, Vincy (2001), 'Comparative Causation', *Law and Economics Working Paper Series 01-30*, George Mason University, School of Law.
- Rea, Samuel A., Jr. (1987), 'The Economics of Comparative Negligence', **7** *International Review of Law and Economics*, 149-162.
- Rubinfeld, Daniel L. (1987), 'The Efficiency of Comparative Negligence', **16** *Journal of Legal Studies*, 375-394.
- Schäfer, Hans-Bernd and Schönenberger, Andreas (2000), 'Strict Liability versus Negligence', in Bouckaert, Boudewijn and De Geest, Gerrit (eds.), *Encyclopedia of Law and Economics*, II, Aldershot, Edward Elgar, 597-624.
- Shavell, Steven (1980), 'Strict Liability versus Negligence', **9** *Journal of Legal Studies*, 1-25.
- Shavell, Steven (1987), *Economic Analysis of Accident Law*, Cambridge (MA), Harvard University Press.
- Schwartz, Gary T. (1978), "Contributory and Comparative Negligence: A Reappraisal", **87** *Yale Law Journal*, 697-727.
- White, Michelle J. (1989), 'An Empirical Test of the Comparative and Contributory Negligence Rules in Accident Law', **20** *Rand Journal of Economics*, 308-330.

APPENDIX

Appendix 1. Simulation of non-uniform distributions of the costs of precaution in unilateral precaution cases

In the appendix we will simulate the effect that comparative negligence and all-or-nothing rules have on the social loss under non-uniform distributions of the costs of precaution. We will use a more complex model than the one employed in *section 1*. The computer simulations were implemented in the MATHEMATICA environment.

We will refer to the graph in *figure 1*.

- For each accident we consider the precaution cost $c=[0, \infty)$ needed to avoid it and the expected harm $h=(0, \infty)$; “ h ” is given by “ Hp ”, where “ p ” is the probability that, if the avoider is careless, an accident will occur and “ H ” is the harm;
- Calculate for each accident $y=c/h$ ($0 \leq y < \infty$), the standardized precaution cost, and rank all the accidents on the base of the value of y ; if $y < 1$ (precaution cost < expected harm) the accident is socially undesirable; if $y > 1$ (precaution cost > expected harm) it is socially advantageous that the accident occurs; if $y=1$ it is indifferent.
- Let $x=N/N_s$, N_s = number of accidents with $y \leq 1$ (accidents which should not occur), N = position of the accident in the raked order described above, notice that $x=1$ when $y=1$, x and y are standardized measures for the precaution cost and the ranked order of the accidents; for simplicity we will consider x as a continuous variable, even if in fact it has discrete values;
- Consider a function $y=f(x)$, which provides a relationship between the standardized precaution cost and the ranked order of the accidents; $dy/dx \geq 0$ (because of the ranked order, an accident cannot have a smaller y than the preceding one); $f(0)=0$ the first accident to avoid is the one with the smallest y (c infinitely smaller than h); $f(1)=1$ the last accident to avoid has $y=1$ ($c=h$);
- We will test the model by employing the function $y=x^q$, $q=(0, \infty)$.

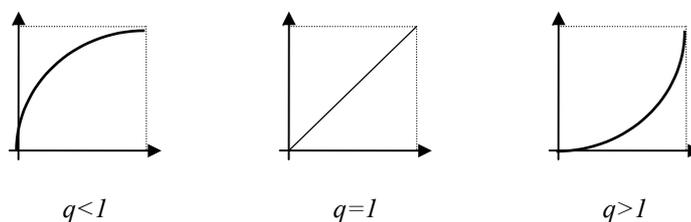


Figure 7. Different shapes of the function $y=x^q$.

By employing this model, we will calculate the social loss under ANRs and the social loss under CN as functions of q , and then we will compare them in order to test whether or not CN yields smaller social losses than ANRs for any possible value of q . This will provide us with a criterion to choose between ANRs and CN in unilateral precaution cases described in the previous sections.

The results of that comparison will be plotted. We will use a graph showing on the x -axis the value of q and on the y -axis the value of the difference between the social loss under ANRs and the social loss under CN. The value depicted on the y -axis is to be interpreted as the comparative advantage of CN in reducing the social loss due to a judicial lack of information for different precaution cost distributions.

Because the central case is the uniform distribution of precaution costs, given by $q=1$, the graphs will be centered on this value. To the right-hand side of the graph, $q>1$ -distributions will be shown ($1 < q < \infty$), and to the left $q<1$ -distributions will be presented ($0 < q < 1$) using the inverse of the values indicated on the right-hand side (see *figure 8*). This solution accounts for the fact that, with respect to the central uniform distribution of accidents (the straight line in *figure 7*), $q=2$ -distribution is the mirror image of $q=1/2$ -distribution, $q=4$ -distribution is the mirror image of $q=1/4$ -distribution, and so forth.

Perfect compensation

We will now refer to *figure 2*. The social loss is given by the sum of the differences $h-c$ for each single accident that the legal rule under analysis fails to prevent.

For ANRs, we denote as " A_1 " the area of the dotted triangle in *figure 2*, and as $\frac{1}{2}A_1$ the social loss due to ANRs (50% of the accidents will occur).

$$A_1 = 1 - \int_0^1 f(x) dx; \quad f(x) = x^q.$$

For CN, we denote as A_2 the gray area in *figure 2*; in this case the social loss is equal to the whole area A_1 , because all the accidents between $x=x_m$ and $x=1$ occur.

$$A_2 = 1 - 1 * x_m - \int_{x_m}^1 f(x) dx; \quad f(x) = x^q; \quad x_m = \left(\frac{1}{2}\right)^{\frac{1}{q}}.$$

In order to choose between ANRs and CN, we need to determine which rule yields the smaller social loss. For this purpose, we can compare $\frac{1}{2}A_1$ (social loss under ANRs) and A_2 (social loss under CN). CN is preferable over ANRs if $\frac{1}{2}A_1 - A_2 > 0$. ANRs are preferable over CN if $\frac{1}{2}A_1 - A_2 < 0$. The two rules are equivalent if $\frac{1}{2}A_1 - A_2 = 0$. Let $L(q)$ denote the former difference.

So, CN is preferable over or equivalent to ANRs if:

$$\frac{1}{2}A_1 - A_2 > 0$$

$$\frac{1}{2}A_1 - A_2 = \frac{1}{2} \left[1 - \int_0^1 f(x) dx \right] - \left[1 - 1 * x_m - \int_{x_m}^1 f(x) dx \right] > 0$$

After replacing $f(x) = x^q$; $x_m = \left(\frac{1}{2}\right)^{\frac{1}{q}}$; we obtain:

$$\frac{1}{2}A_1 - A_2 = \frac{-q + \left(\frac{1}{2}\right)^{\frac{1}{q}}(2q+1)}{2(q+1)} = L(q) > 0 \quad a1)$$

$$\min L(q) = 0.065, \text{ for } q = 0.206$$

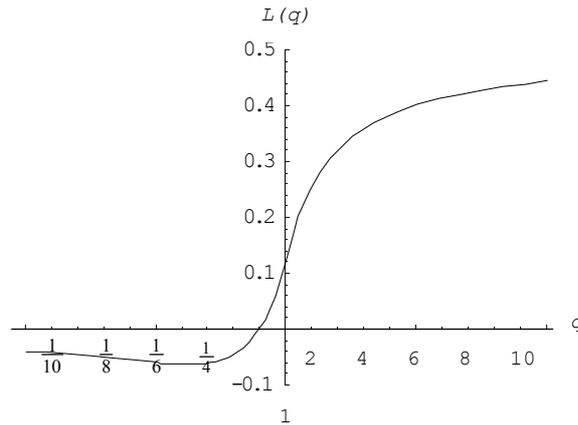


Figure 8: Perfect compensation

$$\lim_{q \rightarrow 0} L(q) = 0; \quad L(1) = \frac{1}{8}; \quad \lim_{q \rightarrow \infty} L(q) = \frac{1}{2}$$

Figure 8 shows the value of the function $\frac{1}{2}A_1 - A_2 = L(q)$ for different values of q , i.e. for different precaution cost distributions. The social loss under CN is smaller than the social loss under ANRs if $L(q)$ is positive; in this case CN performs as a filter and yields a smaller social loss than ANRs. A

negative value of $L(q)$ indicates that the loss under CN is greater than the loss under ANRs and thus ANRs are the best solution. From *figure 8* it is clear that CN is better than ANRs for any value of $q > 1/2$. ANRs have a better performance only for $0 < q < 1/2$. However, the gain in social loss saving (given by the vertical distance between the curve and the x -axis) of comparative negligence is largely superior than the one due to ANRs. It can be also noticed that the social loss does not depend on the number of accidents each rule prevents. ANRs always prevent 50% of the accidents to avoid, while CN prevents a variable number of accidents depending on the value of q .

The graph on the left-hand side of *figure 9* shows the previous statement. The greater q , the larger the number of accidents which will be prevented. On the contrary, the number of accidents prevented under ANRs is constant. The graph on the right-hand side of *figure 9* shows the shape of the precaution cost distribution when $q = 1/2$. In this case CN prevents only 25% of the accidents and the social loss is equal to the social loss under ANRs.

The number of accidents prevented under CN (x_m) can be calculated as follows:

$$x = f^{-1}(y)$$

$$x_m = f^{-1}\left(\frac{1}{2}\right) = \left(\frac{1}{2}\right)^{\left(\frac{1}{q}\right)}$$

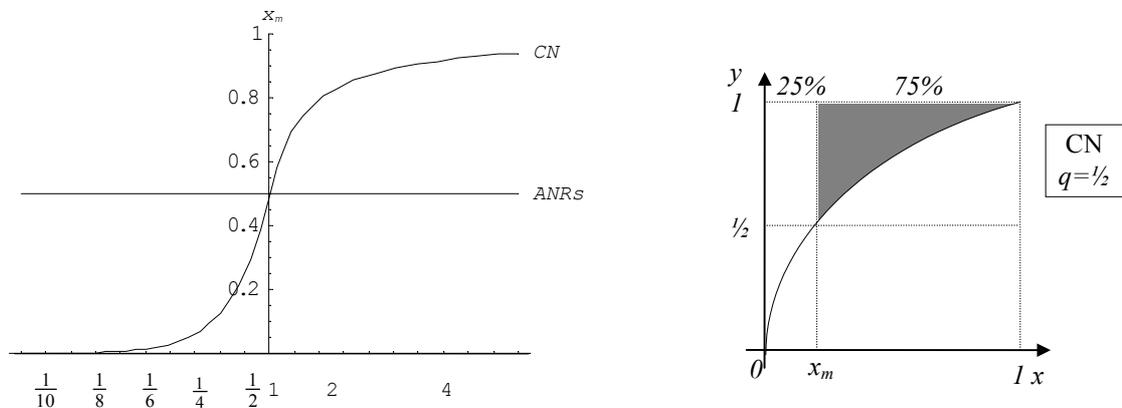


Figure 9: Number of accidents prevented under CN and under ANRs

Figure 9 shows also that CN tends to prevent more accidents than ANRs when $q > 1$. The following table resumes the previous results.

- $q > 1$: n° of accidents under CN < n° of accidents under ANRs
social loss under CN < social loss under ANRs
- $q = 1$: n° of accidents under CN = n° of accidents under ANRs
social loss under CN < social loss under ANRs
- $1/2 < q < 1$: n° of accidents under CN > n° of accidents under ANRs
social loss under CN < social loss under ANRs
- $q = 1/2$: n° of accidents under CN > n° of accidents under ANRs
social loss under CN = social loss under ANRs
- $q < 1/2$: n° of accidents under CN > n° of accidents under ANRs
social loss under CN > social loss under ANRs

Appendix 2. Overcompensation

In this section we will relax the assumption of perfect compensation. Under ANRs, on the one hand, if the court makes the victim bear the cost of the accident, errors in compensation do not have any effect on the incentive stream simply because no compensation takes place. On the other hand, if the court

always targets the injurer, overcompensation will create an extra incentive for him, while undercompensation will shift part of the incentive stream towards the victim. In other words, the court could minimize the effect of errors by always targeting the victim. Under CN, the court always directs the incentive stream towards both parties, and the effect of errors cannot be eliminated in the same way.

We will run two analyses. First we will compare ANRs targeting the injurer (for example, simple negligence) to CN. Second, we will compare ANR targeting the victim (for example, contributory negligence) to CN.

In both cases D (the damages that the injurer has to pay) are no longer equal to H (the harm that the victim suffers), but $D=H(1+e)$, where e is the error committed by the court while estimating H , consistently also the expected damages change, $d=h(1+e)$. The errors are systematic and predicted by the parties. The court can overcompensate or undercompensate the victim for the harm.

- overcompensation: $e>0$
- perfect compensation: $e=0$
- undercompensation: $e<0$.

For overcompensation we denote:

- overcompensation with small errors $0<e<1$
- overcompensation with large errors $e\geq 1$ (D is twice or more than twice as much as H)

Overcompensation - small errors ($0<e<1$) – ANRs targeting the injurer

Under ANRs targeting the injurer, overcompensation of the harm increases the incentive stream directed toward the injurer and induces the injurer to take overprecaution. The injurer will avoid an accident if the expected damages (d) are greater than the precaution costs (c), even if $c>h$. In this case there is a loss given by the difference $c-h$ (zone II in figure 10). Because we assume that the court always targets the injurer and that only in 50% of those cases the injurer is the avoider, that extra social loss arises only in 50% of the cases. Therefore, the social loss is equal to the social loss under perfect compensation plus $\frac{1}{2}$ of the stripped triangle in figure 10.

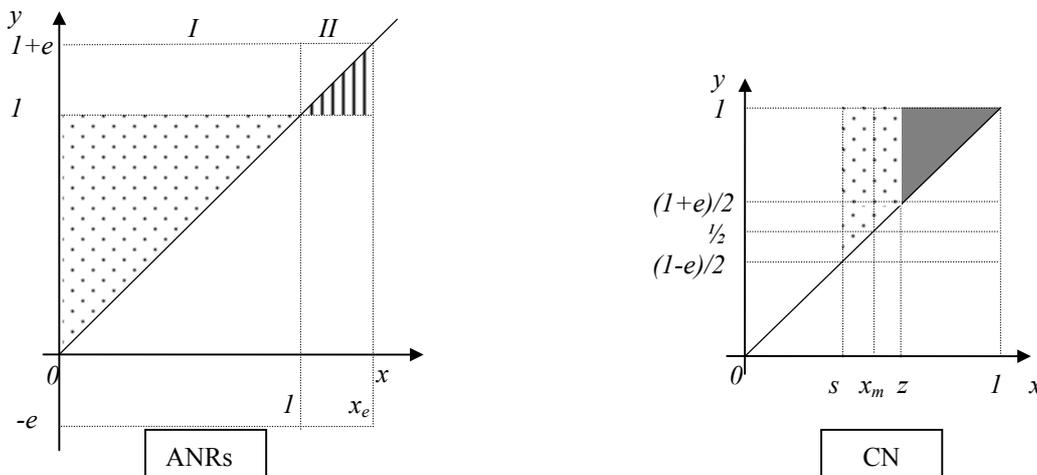


Figure 10. Overcompensation - small errors

Also in the case of overcompensation, under CN the court over-valuates the damages, but the effect is lower because CN contemporarily targets the victim and the injurer with half the maximum incentive stream, so that the injurer pays $D=H(1+e)/2$, and the victim bears the remaining $H(1+e)/2$. The incentives for the injurer increase: the injurer will pay D to the victim, so he will be careful as far as the precaution cost = d (which is greater than $h/2$). The incentives for the victim are undermined: the victim will expect to bear just $h-d = h - h(1+e)/2 = h(1-e)/2$ (which is smaller than $h/2$).

In figure 10's right-hand side graph: the injurer will be careful till the point z (more than under perfect compensation, x_m), while the victim will be careful only till the point s (less than under perfect compensation, x_m). Before s both will be careful, after z both will be careless. The result is that under overcompensation 100% of the accidents in the gray area and 50% of the accidents (only the injurer is careful) in the dotted area will occur.

It is to be noticed that the analysis is exactly the same as for undercompensation. The only

difference is that in that case the victim will be over-careful and the injurer under-careful, *ceteris paribus*.

In order to test our model in the case of small errors in overcompensating damages, we compare the total (ordinary loss + additional loss due to errors) social loss under ANRs and the total social loss under CN. If the difference “social loss under ANRs targeting the injurer (L_1)” – “social loss under CN (L_2)” > 0 then CN yields to a smaller social loss in the case of errors.

$$L_1 - L_2 > 0$$

$$L_1 - L_2 = \left\{ \frac{1}{2} \left[1 - \int_0^1 f(x) dx \right] + \frac{1}{2} \left[\int_1^{x_e} f(x) dx - (x_e - 1) \right] \right\} - \left\{ \left[(1-z) - \int_z^1 f(x) dx \right] + \frac{1}{2} \left[(z-s) - \int_s^z f(x) dx \right] \right\} > 0$$

After replacing

$$f(x) = x^q; \quad x_e = (1+e)^{\frac{1}{q}}; \quad z = \left(\frac{1+e}{2}\right)^{\frac{1}{q}}; \quad s = \left(\frac{1-e}{2}\right)^{\frac{1}{q}};$$

we obtain:

$$L_1 - L_2 = \frac{(1+e)^{\frac{1}{q}}(2e-2q) + \left(\frac{1+e}{2}\right)^{\frac{1}{q}}(2q-e+1) + \left(\frac{1-e}{2}\right)^{\frac{1}{q}}(2q+e+1)}{4(q+1)} = L(q) > 0 \tag{a2}$$

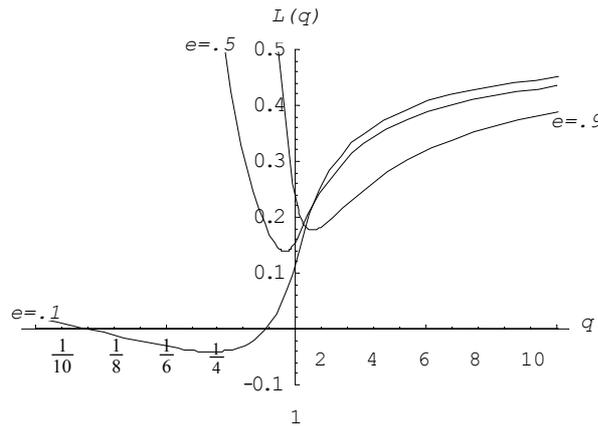


Figure 11: CN against ANRs targeting the injurer – Overcompensation - small errors

Figure 11 shows the performance of CN for any value of q and for different values of e ($0 < e < 1$, $e=.1$, $e=.5$, $e=.9$), and shows that CN is still to be preferred on ANRs because yields a smaller social loss. The result becomes clearer when the errors become larger.

Overcompensation - large errors ($e \geq 1$) – ANRs targeting the injurer

We now consider the case in which large overcompensation ($e \geq 1$) occurs. That is, the injurer has to pay damages that are (twice or) more than twice as much as the harm suffered by the victim. The analysis differs from the cases already examined only with respect to CN. With respect to ANRs the analysis is exactly the same as in the case of small errors.

Under CN the analysis is completely different. In the big dotted triangle 50% of the accidents occur, because only the injurer will be careful, as the victim is entirely compensated for the harm. The small stripped triangle depicts the loss due to over-precaution of the injurer, regarding again 50% of the accidents.

The injurer will be careful as far as $c < h(1-e)/2$ (precaution cost $<$ expected damages to pay to the victim), which corresponds to the point x_f on the x -axis. The victim does not have any incentive to be careful because he will recover the whole harm from the injurer, because $(1-e)/2 < h$. If only the injurer

is careful 50% of the accidents will occur both in zone I and in zone II. However, the social loss in zone I is given by the difference $h-c$ (accidents we should prevent), while the social loss in zone II is given by the difference $c-h$ and is due to the injurer's overprecaution (accidents which should occur). If loss under ANRs (L_1) - loss under CN (L_2) > 0 , then CN is more efficient than ANRs in minimizing the

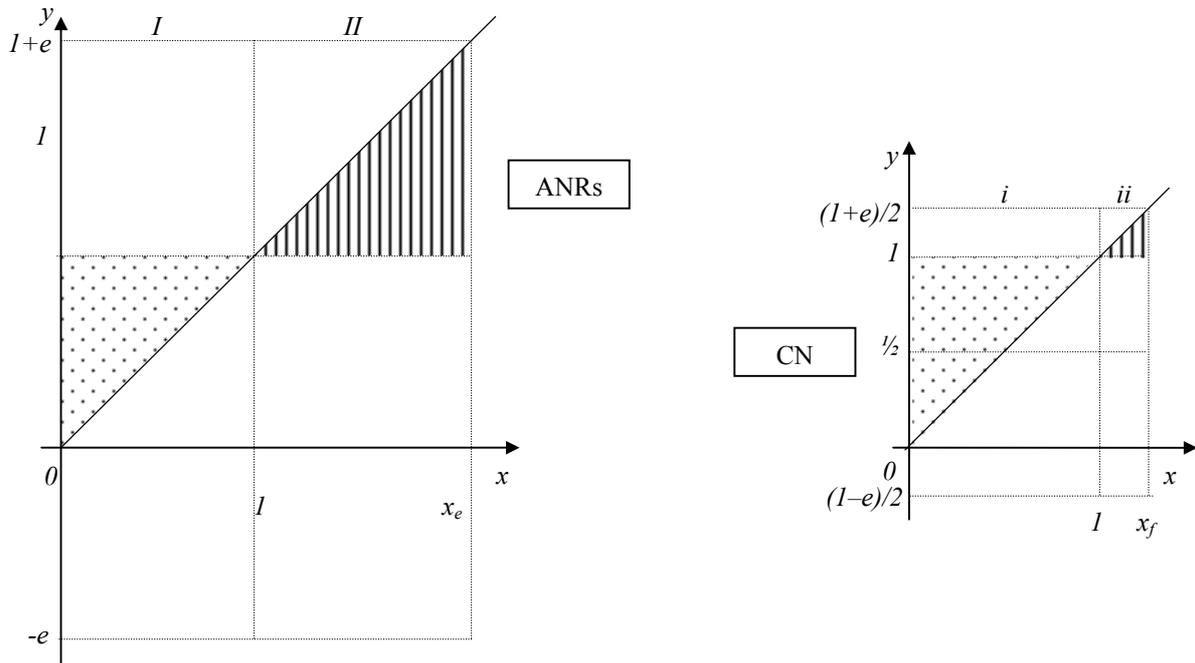


Figure 12. Overcompensation - large errors

social loss in the case of large errors.

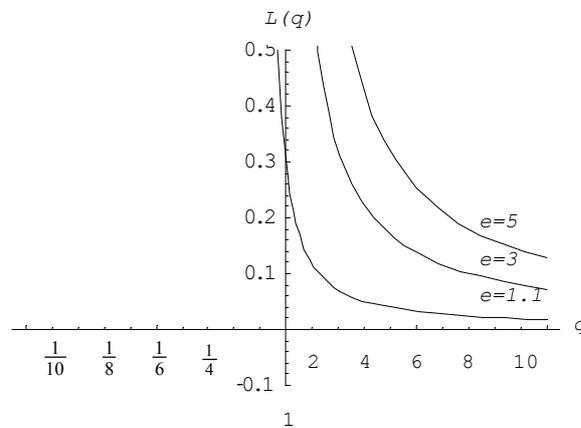


Figure 13: CN against ANRs targeting the injurer – Overcompensation - large errors

$$L_1 = \frac{1}{2} \left[1 - \int_0^1 f(x) dx \right] + \frac{1}{2} \left[\int_1^{x_e} f(x) dx - (x_e - 1) \right]$$

$$L_2 = \frac{1}{2} \left[1 - \int_0^1 f(x) dx \right] + \frac{1}{2} \left[\int_1^{x_f} f(x) dx - (x_f - 1) \right]$$

$$L_1 - L_2 = \left\{ \frac{1}{2} \left[1 - \int_0^1 f(x) dx \right] + \frac{1}{2} \left[\int_1^{x_e} f(x) dx - (x_e - 1) \right] \right\} - \left\{ \frac{1}{2} \left[1 - \int_0^1 f(x) dx \right] + \frac{1}{2} \left[\int_1^{x_f} f(x) dx - (x_f - 1) \right] \right\} > 0$$

$$L_1 - L_2 = \frac{1}{2} \left[\int_1^{x_e} f(x) dx - (x_e - 1) \right] - \frac{1}{2} \left[\int_1^{x_f} f(x) dx - (x_f - 1) \right] > 0$$

After replacing

$$f(x) = x^q; \quad x_e = (1+e)^{\frac{1}{q}}; \quad x_f = \left(\frac{1+e}{2} \right)^{\frac{1}{q}}$$

we obtain:

$$L_1 - L_2 = \frac{(1+e)^{\frac{1}{q}}(2e-2q) + \left(\frac{1+e}{2} \right)^{\frac{1}{q}}(1+2q-e)}{4(q+1)} = L(q) \quad a3)$$

Figure 13 shows even a clearer pattern than in the previous case. The function $L(q)$ is always positive for any value of q and any value of e (we have tested $e=1.1$, $e=3$, $e=5$). That means that CN always yields a smaller social loss.

Overcompensation - small errors ($0 < e < 1$) – ANRs targeting the victim

We now consider the case in which under ANRs the court always targets the victim. In this case no problem of over or undercompensation arises, because, by making the victim bear the entire cost of the accident, the court does not need to estimate the harm. However, errors still affect the performance of CN in the same way as we have already seen in the former two sections.

Therefore, in order to test the relative efficiency of the two rules, we will compare ANRs without errors and CN with errors. We will start with small errors in overcompensating damages ($0 < e < 1$).

We can still refer to figure 10. The loss due to ANRs targeting the victim is just the big dotted triangle (if the court targets the victim errors in compensation have no effect) in the left-hand side graph. The loss under CN is the same as in the case tested above for overcompensation, small errors (see figure 10, the right-hand side graph). If the difference “social loss under ANRs targeting the victim (L_1)” – “social loss under CN (L_2)” > 0 then CN yields a smaller social loss in the case of errors.

$$L_1 - L_2 > 0$$

$$L_1 - L_2 = \frac{1}{2} \left[1 - \int_0^1 f(x) dx \right] - \left\{ \left[(1-z) - \int_z^1 f(x) dx \right] + \frac{1}{2} \left[(z-s) - \int_s^z f(x) dx \right] \right\} > 0$$

After replacing

$$f(x) = x^q; \quad z = \left(\frac{1+e}{2} \right)^{\frac{1}{q}}; \quad s = \left(\frac{1-e}{2} \right)^{\frac{1}{q}};$$

we obtain:

$$L_1 - L_2 = \frac{-2q + \left(\frac{1+e}{2}\right)^{\frac{1}{q}}(2q-e+1) + \left(\frac{1-e}{2}\right)^{\frac{1}{q}}(2q+e+1)}{4(q+1)} = L(q) > 0 \quad a4)$$

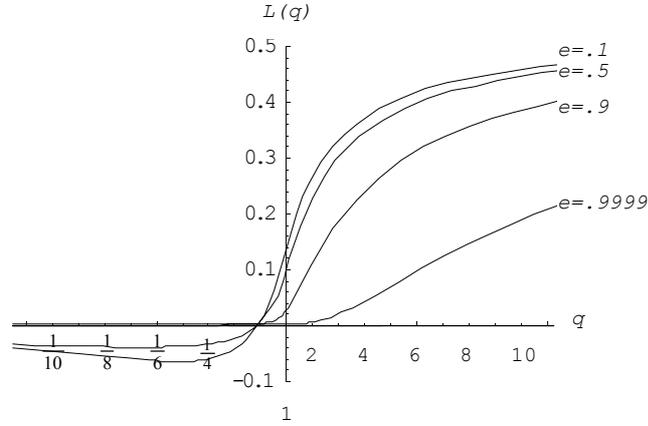


Figure 14: CN against ANRs targeting the victim – Overcompensation - small errors

It is to be noted that small errors tend to transform a CN rule in an ANR, so that if the error is equal to 1 there is no difference between the two rules. That is why, even if we introduce small errors CN is still better than ANRs (as shown in the graph) and the difference between the two rules slowly disappears as e becomes closer to 1.

Overcompensation - large errors ($e \geq 1$) – ANRs targeting the victim

We have already said that if the error is equal to 1, CN and ANRs targeting the victim give the same results. We are now going to test the effect of large errors in overcompensation. Again ANRs targeting the victim are not affected by errors. On the contrary, under CN the injurer will be over-careful and the victim will not be careful at all. As we have already explained, the loss is depicted by the two dotted triangles in the right-hand side graph of figure 12.

If “social loss under ANRs (L_1)” – “social loss under CN (L_2)” > 0 , then CN is more efficient than ANRs as it yields a smaller social loss in the case of large errors.

$$L_1 - L_2 = \frac{1}{2} \left[1 - \int_0^1 f(x) dx \right] - \left\{ \frac{1}{2} \left[1 - \int_0^1 f(x) dx \right] + \frac{1}{2} \left[\int_1^{x_f} f(x) dx - (x_f - 1) \right] \right\} > 0$$

$$L_1 - L_2 = -\frac{1}{2} \left[\int_1^{x_f} f(x) dx - (x_f - 1) \right] > 0$$

After replacing

$$f(x) = x^q; \quad x_f = \left(\frac{1+e}{2}\right)^{\frac{1}{q}}$$

we obtain:

$$L_1 - L_2 = \frac{-2q + \left(\frac{1+e}{2}\right)^{\frac{1}{q}}(1+2q-e)}{4(q+1)} = L(q) \quad a5)$$

In this case ANRs are always better than CN. The graph shows that the value of the function $L(q)$ is negative for any value of q . When the court entitles the victim to recover more than twice as much as the harm that he actually suffered, CN becomes worse than ANRs targeting the victim.

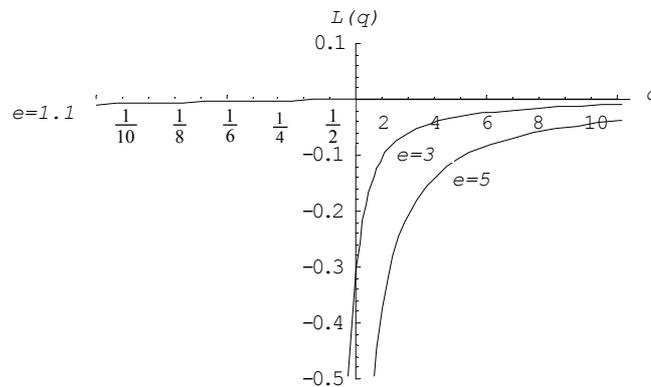


Figure 15: CN against ANRs targeting the victim – Overcompensation - large errors

Concluding remarks for overcompensation

The simulation shows that CN generally yields a smaller social loss than ANRs also in the case of overcompensation. In the case of small errors (both with respect to ANRs targeting the injurer, equation $a2$, and figure 11, and ANRs targeting the victim, equation $a4$, and figure 14) the superiority of CN is due to the filtering effect. In the case of large errors CN is also superior to ANRs targeting the injurer because of a lighter over-deterrence effect. On the contrary, compared to ANRs targeting the victim in the case of large overcompensation CN is no longer the best rule. This is due to the fact that large errors erode CN's filtering effect. However, because this is the only case, we can conclude for a general superiority of CN.

Appendix 3. Undercompensation

If the court underestimates the harm suffered by the victim there is an effect on the incentive streams that the tort-liability system gives both to the victim and to the injurer.

For undercompensation we denote:

- undercompensation with small errors $-\frac{1}{2} < e < 0$;
- undercompensation with large errors $-1 < e \leq -\frac{1}{2}$ (D is half or less than half H , if $e = -1$ the rule becomes equivalent to no liability).

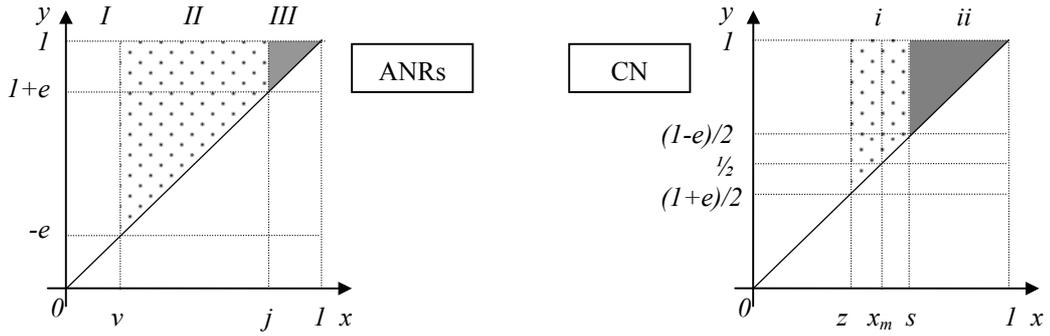


Figure 16. Undercompensation - small errors

Undercompensation – small errors ($-\frac{1}{2} < e < 0$) – ANRs targeting the injurer

If the court errs in setting the damages that the injurer has to pay, and if those damages are set at a lower level than under the perfect-compensation assumption, $D=H(1+e)$, and $e < 0$. The incentive stream that the injurer faces is undermined in the same measure, $d=h(1+e)$. However, undercompensation creates an additional incentive stream for the victim. This is due to the fact that he has to bear the portion of the harm not internalized by the injurer ($H-D=H-H*[1-e]=H-H+e=e$; thus the victim bears a portion of the harm equal to the error that the court makes).

Table 1. ANRs targeting the injurer. Undercompensation, small errors

	avoider	Zone I	Zone II	Zone III
1.	injurer	careful	careful	careless
2.	victim	careful	careless	careless
		no accidents	50% of the accidents	100% of the accidents

The effect of errors in (under)compensation is actually ambiguous: on the one hand, the injurer is less careful, on the other, the victim has an incentive to be careful also in the case of the injurer being liable. The table shows the parties' behavior and its effect on the number of accidents. In zone I (in figure 16) both the victim and the injurer have an incentive to avoid the accident because both bear a portion of the harm which is greater than the precaution cost. There is no social loss. In zone II only the injurer has an incentive to be careful. In zone III neither the victim nor the injurer has an incentive to be careful, because the precaution cost exceeds the portion of the harm that each party has to bear.

Under CN, the case of undercompensation is the same as the case of overcompensation. The only difference is that in this case the victim is over-careful and the injurer under-careful, thus also the points s and z are inverted. The analysis already made in the case of overcompensation with small errors can be employed for the cases of both undercompensation with small errors and undercompensation with large errors.

CN is preferable if “social loss ANRs (L_1)” – “social loss CN (L_2)” > 0 . We denote as A_{II} , A_{III} (for ANRs) and A_i , A_{ii} (for CN) the social-loss areas corresponding to the zones II and III and to the zones i and ii in figure 16. For each area we calculate a fraction consistently to table 1.

$$L_1 - L_2 > 0$$

$$L_1 - L_2 = \left\{ \frac{1}{2} A_{II} + A_{III} \right\} - \left\{ \frac{1}{2} A_i + A_{ii} \right\} > 0$$

$$L_1 - L_2 = \left\{ \frac{1}{2} \left[j - v - \int_v^j f(x) dx \right] + \left[1 - j - \int_j^1 f(x) dx \right] \right\} - \left\{ \frac{1}{2} \left[(s - z) - \int_z^s f(x) dx \right] + \left[(1 - s) - \int_s^1 f(x) dx \right] \right\} > 0$$

After replacing

$$f(x) = x^q; \quad v = (-e)^{\frac{1}{q}}; \quad j = (1+e)^{\frac{1}{q}}; \quad s = \left(\frac{1-e}{2}\right)^{\frac{1}{q}}; \quad z = \left(\frac{1+e}{2}\right)^{\frac{1}{q}};$$

we obtain:

$$L_1 - L_2 = \frac{(1+e)^{\frac{1}{q}}(2e-2q) - (-e)^{\frac{1}{q}}(2q+2+2e) + \left(\frac{1+e}{2}\right)^{\frac{1}{q}}(2q+1-e) + \left(\frac{1-e}{2}\right)^{\frac{1}{q}}(2q+1+e)}{4(q+1)} = L(q) > 0 \quad a6)$$

In the case of undercompensation the result is ambiguous. CN maintains an advantage for small values of e , while when e increases ANRs targeting the injurer seems to be better. The reason behind this result is that a systematic undercompensation and the fact that the injurer is always held liable make ANRs close to CN. When the error is equal to $\frac{1}{2}$ (the victim bears half the harm) ANRs (with errors) give exactly the same result as CN (without errors). If we introduce errors also under CN the performance becomes worse.

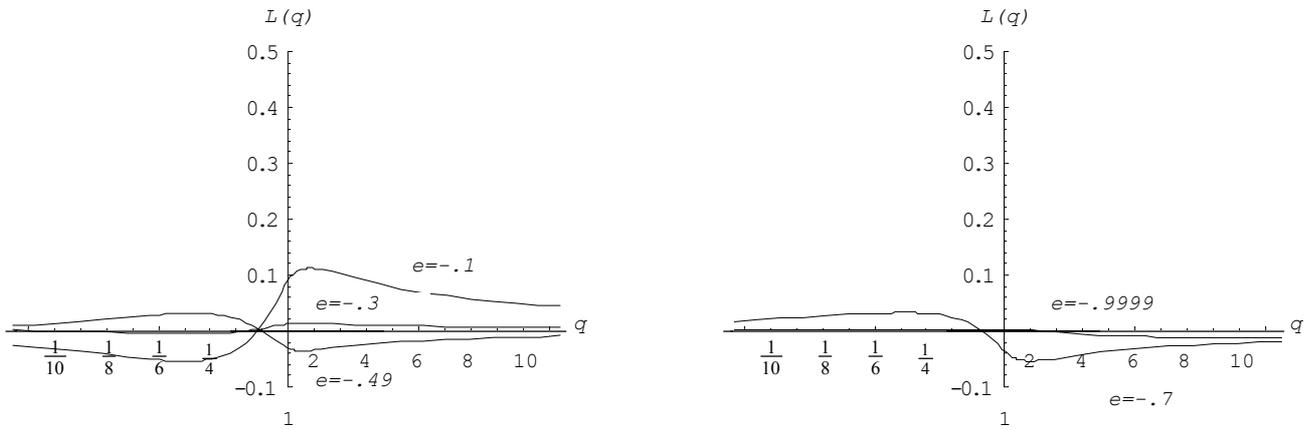


Figure 17: CN against ANRs shooting on the injurer - Undercompensation – small and large errors

Undercompensation, large errors ($-1 < e \leq -\frac{1}{2}$) – ANRs targeting the injurer

We now turn to the case of large errors in undercompensation ($-1 < e \leq -\frac{1}{2}$), where the damages that the

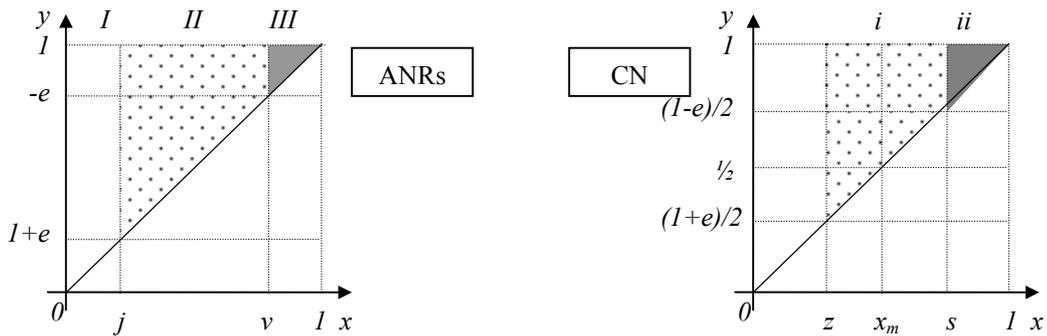


Figure 18. Undercompensation – large errors

victim is entitled to recover are equal to or less than half harm. In these cases the analysis differs from scenarios involving small errors only with respect to ANRs, but the result is the same. The portion of the harm that the injurer bears is less than the portion of the harm that the victim bears. Because $-1 < e \leq -\frac{1}{2}$, thus $(1+e) \leq -e$. The point j and v are inverted with respect to small undercompensation as shown in figure 16.

Table 2. ANRs. Undercompensation, large errors

	avoider	Zone I	Zone II	Zone III
1.	injurer	careful	<i>careless</i>	<i>careless</i>
4.	victim	careful	careful	<i>careless</i>
		no accidents	50% of the accidents	100% of the accidents

Table 2 illustrates the parties' behavior. Nothing changes in zone I and III with respect to small errors. On the contrary, in zone II only the victim is careful, even if the injurer is the cost-bearer.

The rest of the analysis is exactly the same as under small errors. In the specific case of $e=-\frac{1}{2}$, $v=j$ and $l+e=-e$, zone II disappears and only zones I and III are important for the analysis. However, the analysis can be run in the same way so that the simulation made for small errors applies also in the case of large errors and in the particular case of $e=\frac{1}{2}$. The result is given by equation a6). Therefore, figure 17 shows the comparative performance of CN also in the case of large errors (see the right-hand side graph).

Undercompensation – small errors ($-\frac{1}{2} < e < 0$) – ANRs targeting the victim

If the court targets the victim under ANRs errors make no effect. Moreover, under CN the effect of undercompensation is exactly the same as the effect of overcompensation. Therefore, we can refer to the analysis already made for small errors in overcompensation and ANRs targeting the victim (see figure 14, and note that under CN a negative errors has the same effect as a positive error). The only difference with respect to overcompensation is that the points z and s are inverted, the victim is more careful and the injurer is less careful than under perfect compensation. The result of the analysis is given by equation a4).

Undercompensation, large errors ($-1 < e \leq -\frac{1}{2}$) – ANRs targeting the victim

The case of large errors in undercompensation ($-1 < e \leq -\frac{1}{2}$) is analogous to the case of small errors just discussed. Again we consider no effect under ANRs. Under CN the effect of small and large errors in undercompensating damages is the same, thus we can still refer to figure 14 for the graph of the results. Also in this case negative errors produce the same effect as positive errors. The result of the analysis is given again by equation a4).

Concluding remarks for overcompensation

The simulation shows that CN generally yields a smaller social loss than ANRs also in the case of undercompensation. In the case of ANRs targeting the injurer, the results are interesting. On the one hand, errors weaken the filtering effect of CN, on the other hand, they tend to transform a ANR targeting the injurer into a CN rule. When the court underestimates the damages by half, ANRs give the same result as CN. Those two effects combined give the ultimate result. If errors are small the filtering effect is not seriously undermined and CN is still the best rule. When errors become larger (already for $e=.4$), CN loses a vast part of the filtering effect while ANRs improve, because errors create a similar filtering effect (see equation a6 and figure 17).

On the contrary, when compared to ANRs targeting the victim, CN yields better results, because in this case ANRs are not improved by errors. The result can be interpreted using equation a4) and figure 14, both for small and large errors. When errors increase, CN progressively loses the filtering effect and at $e=-1$ the outcome of CN is the same as ANRs.

We can conclude also in this case for a general superiority of CN.

Appendix 4. Sharing rules other than 50/50

We assumed that under CN both the victim and the injurer bear $\frac{1}{2}$ of the harm. We will now consider the filtering effect of different sharing rules. Let $\frac{1}{2}+w$ denote the portion of the harm borne by the injurer and let $\frac{1}{2}-w$ denote the portion of the harm borne by the victim, $w=[-\frac{1}{2}, \frac{1}{2}]$. It is clear that the sum of the portions borne by each party is always 1, the whole harm, and that if $w=0$ the rule is a 50/50 sharing rule. Moreover if $w>0$ the injurers bears more than the victim does, while if $w<0$ the victim bears more than the injurer. When $w=\pm\frac{1}{2}$ the rule is an ANR. We will denote a general sharing rule as

SR and a 50/50 sharing rule as $\frac{1}{2}$ SR.

We are now going to test whether a 50/50 division is the best sharing rule to minimize the social loss under different precaution cost distributions.

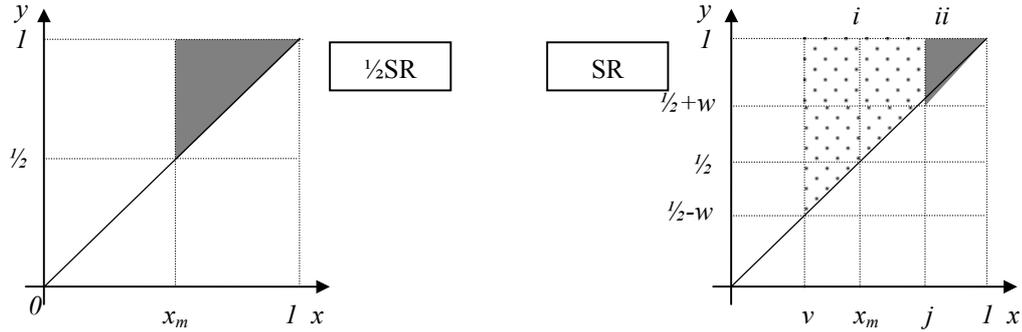


Figure 19. Sharing rules

Figure 19 shows that in the case of a 50/50 sharing rule there is a pure filtering effect, both the victim and the injurer have the same incentives to take precaution; before x_m no accident will occur, after x_m no accident will be prevented. A SR other than 50/50 will provide the victim and the injurer with different incentive streams, so that the victim will be careful till the point v and the injurer till the point j . Note that the graph pictures the case when $w > 0$. Indeed if $w < 0$ the points v and j will be inverted, but nothing will change in the results of the analysis. It is to be noted also that we are assuming that the victim and the injurer have the same probability to be the avoider. The analysis will be different if we relax this assumption.

Let us test whether a 50/50 division yields minimum social loss.

$$L(w, q) = \frac{1}{2} \left[j - v - \int_v^j f(x) dx \right] + 1 - z - \int_z^1 f(x) dx$$

After replacing

$$f(x) = x^q; \quad x_m = \left(\frac{1}{2}\right)^{\frac{1}{q}}; \quad j = \left(\frac{1}{2} + w\right)^{\frac{1}{q}}; \quad v = \left(\frac{1}{2} - w\right)^{\frac{1}{q}};$$

we obtain:

$$L(w, q) = \frac{4q - \left(\frac{1}{2} + w\right)^{\frac{1}{q}} (1 + 2q - 2w) - \left(\frac{1}{2} - w\right)^{\frac{1}{q}} (1 + 2q + 2w)}{4(q + 1)} \tag{a7}$$

Equation a7) gives the magnitude of the social loss as a function of w and q . For $w = 0$ the result is the social loss under a 50/50 sharing rules, that we have used in the former analysis. It is useful to plot the results with respect to w . Figure 20 gives the graph of the function for any value of w and for selected values of q ($q = 5, q = 1, q = .5, q = .3, q = .1$). The $L(w, q)$ function is convex if $q > .5$, is linear if $q = .5$, and is concave if $q < .5$.

Moreover when $q > .5$ the function has its minimum at $w = 0$, i.e. when the victim and the injurer bear $\frac{1}{2}$ of the harm each, which means that a 50/50 division yields the least social loss. When $q = .5$ any sharing rule gives the same result, including the two extreme sharing rules ANRs targeting the victim ($w = -\frac{1}{2}$) and ANRs targeting the injurer ($w = \frac{1}{2}$). The result is consistent to the one we have had from the comparison between CN and ANRs under the perfect compensation assumption. In that case, CN yields a smaller social loss for $q \geq \frac{1}{2}$. On the contrary, if $q < .5$ $L(w, q)$ is a concave function and $w = 0$ is the maximum of that function, so a 50/50 division is the worst choice. Figure 20 shows those results.

However, unless the court knows exactly the value of q , the 50/50 division guarantees the best result in a large number of cases ($\frac{1}{2} < q < \infty$) and, more importantly, in the value of q which are in the neighborhood of the central value $q=1$. That means that the filtering effect reduces the social loss not only when the most-wasteful accident are more than the least-wasteful accidents ($q > 1$), but also when the most-wasteful accident are as many as the least-wasteful accidents ($q=1$), and when the least-wasteful accidents are slightly more than the most-wasteful accident ($\frac{1}{2} < q < 1$). Which makes us conclude that if the q value is unknown the 50/50 division has to be preferred. As we have already noted, if we assume that the injurer is more often the avoider than the victim is (or vice versa), an asymmetric division could be better.

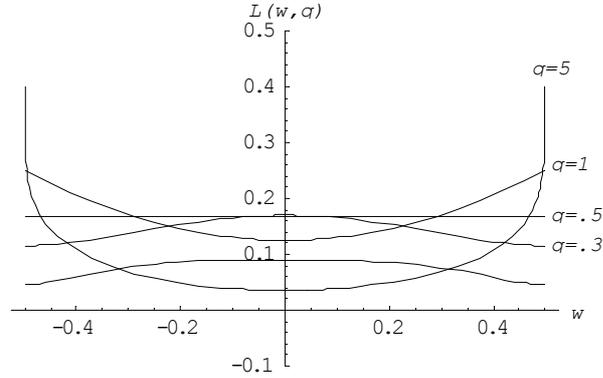


Figure 20: Sharing rules

AN ANALYSIS OF THE JUDGMENT-PROOF PROBLEM UNDER DIFFERENT TORT MODELS*

Giuseppe Dari Mattiacci

Utrecht University and Marie Curie Fellow at Hamburg University

Gerrit De Geest

Utrecht University and Ghent University

ABSTRACT

The judgment-proof problem arises from a threshold imposed by nature (limited assets) or by a rule maker (liability caps) on the maximum amount of damages that the injurer has to pay in the case of an accident. This chapter analyzes the effects of the judgment-proof problem on the injurer's incentive to take precaution. It attempts to draw a more generalized picture than previous contributions in the literature, which tended to focus on specific solutions under specific models. It is demonstrated that the effects of the judgment-proof problem depend on the type of accident or externality. Four tort models are analyzed: the probability model (where more precaution on the part of the injurer only reduces the probability of an accident), the magnitude model (where more precaution only reduces the magnitude of the harm), the joint probability-magnitude model (where more precaution at the same time reduces the probability of the accident and the magnitude of the harm) and the separate probability-magnitude model (where one precautionary measure reduces the probability of an accident and another precautionary measure reduces the magnitude of the harm). In addition, a stochastic variant for each of the four models is discussed. Different legal solutions to the problem are analyzed: punitive damages, average compensation of the harm, under-compensatory damages, and negligence. It is shown that in general subsidizing potentially judgment-proof injurers yields a more desirable outcome than increasing the damages that they have to pay.

JEL classification : K13, K32

Keywords: tort, judgment proof, capping liability, punitive damages, under-compensatory damages.

* Giuseppe Dari Mattiacci, Economic Institute/CIAV, Utrecht University, Kromme Nieuwegracht 22, 3512 HH Utrecht, the Netherlands; email: g.darimattiacci@econ.uu.nl. Gerrit De Geest, Economic Institute/CIAV, Utrecht University, and School of Law, Ghent University; email: g.degeest@econ.uu.nl. The authors would like to thank Jef De Mot, Henrik Lando, Nuno Garoupa and the participants in Copenhagen Business School Seminar and 18th Annual Conference of the European Association of Law and Economics, Vienna 2001, for helpful comments. Giuseppe Dari Mattiacci gratefully acknowledges the financial support provided by the European Commission, the Marie Curie Fellowship Programme in Law and Economics at Hamburg University. Updated versions of this chapter will possibly be available on my web page: www.law.uu.nl/eco/gdm.htm.

5.1. Introduction

A maximum upper threshold is an upper limit to the amount of damages that the injurer actually pays in the case of an accident. There are two main sources of maximum upper thresholds. First, the limit can be due to the total injurer's assets, and thus we will consider such a limit as set by nature. Second, it can be set by the legislature in several ways, such as by liability caps or corporate limited liability. It looks obvious that, whatever the source of the maximum upper threshold is, it can undermine the injurer's incentives to take precautions, as the accident costs are no longer fully internalized. In such a case, the injurer is said to be judgment-proof.

We will show, however, that the incentive effects largely depend on the assumptions of the model. The literature generally uses a specific type of model that we call the *probability model*. A good example may be provided by aircraft accidents, where more precaution of the pilot leads to fewer accidents, but where the harm in case of an accident remains constant (total loss). However, many if not most accidents are of a different type. By reducing his speed, a car driver reduces both the probability of an accident and the magnitude of the harm (*joint probability-magnitude model*). In most nuisance cases, less precaution (e.g. more noise at night) leads to bigger externalities to the neighbors (*magnitude model*). Finally, injurer can quite often take one precautionary measure in order to reduce the probability of an accident and another precautionary measure to reduce the possible harm (*separate (3-dimensional) probability-magnitude model*). For instance, a ship owner can install radar in order to avoid shipwrecks and place lifeboats on the deck in order to reduce the harm to the passengers in the case of a shipwreck occurring.

Those simple examples suggest the need of taking a look at the whole picture. We will analyze the incentives to take precaution in the presence of maximum upper thresholds under these four models.

Probability model

Under a probability model, maximum upper thresholds always induce too little precaution. The level of precaution depends directly upon the threshold: the lower the threshold, the lower the level of precaution. This model closely follows Shavell (1986).

Magnitude model

A magnitude model gives another picture. If the threshold is high enough, it does not affect the level of precaution chosen by the injurer, and he will take optimal precaution. If the threshold is too low, the injurer will take no precaution at all. So, the level of precaution

depends directly on the threshold also in this case, but we observe only a binary outcome: either optimal precaution or no precaution at all. In this case, the injurer can influence the magnitude of the harm, and therefore he can decide whether or not to be potentially judgment-proof (by causing a greater harm than the threshold). If he decides not to be potentially judgment-proof, he internalizes the full harm and, hence, he will take optimal precaution.

On the contrary, if he decides to be potentially judgment-proof, in the case of an accident he will pay an amount of money equal to the threshold (his assets or the cap set by law). In the judgment-proof zone, he will choose no precaution at all, as precaution would reduce the harm but not his exposure to liability. Therefore, the injurer's decision is simply whether to be potentially judgment-proof (no precaution) or not (optimal precaution). The choice depends on the level of the threshold.

Joint probability-magnitude model

In joint probability-magnitude models the outcome is binary as well for the same reason as before. If the threshold is high enough, it does not affect the level of precaution chosen by the injurer, and he will take optimal precaution, but if the threshold is too low, the injurer will take too little precaution (although not necessarily no precaution at all). In this case, the injurer can reduce both the magnitude and the probability of the accident. Even in the judgment-proof zone, precaution reduces his exposure to liability as it reduces the probability of an accident occurring.

Therefore, the decision is again whether or not to be potentially judgment-proof, but in this case underprecaution results instead of no precaution. The choice depends on the level of the threshold. The lower the threshold, the lower the level of precaution taken in the judgment-proof zone.

Separate probability-magnitude model

In the fourth model, the injurer can discriminate between the expenditure on reducing the probability of an accident and the expenditure on reducing the magnitude of the harm. By choosing the level of his magnitude-reducing precaution, the injurer can decide whether or not to be potentially judgment-proof as in the second model. If he decides not to be judgment-proof, he internalizes the full harm; hence, he will take optimal precaution not only with respect to the magnitude-reducing precaution, but also with respect to the probability-reducing precaution.

On the contrary, if he decides to be potentially judgment-proof, he will take no magnitude-reducing precaution (as it would not reduce his exposure to liability) and an inefficient level of probability-reducing precaution. If the threshold is lower than the magnitude of the harm that would result given optimal precaution, in the judgment-proof zone

the injurer internalizes less than the optimal harm, and therefore will take lower probability reducing precaution than optimal. On the contrary, the threshold might be equal to or higher than the magnitude of the harm that would result given optimal precaution, but not high enough to induce the injurer to take optimal magnitude-reducing and probability-reducing precaution. In such a case, in the judgment-proof zone, the injurer internalizes the optimal harm or more than the optimal harm, and therefore will take optimal or higher than optimal probability-reducing precaution.

In synthesis, the analysis shows two results. First, a maximum upper threshold does not always undermine the incentive to take precaution. On the one hand, the legislature can, for some purposes, set a threshold on the maximum amount of damages that the victim is entitled to recover from the injurer, without causing an increase in the number or the harmfulness of accidents. On the other hand, a little judgment proof problem will have in general no effect at all on the level of precaution taken by the injurer.

The second important result is that, in all those cases in which the injurer can reduce the magnitude of the harm, the effect of too low a threshold might be much greater than foreseen under the traditional approach. The injurer might choose no precaution at all with respect to the magnitude of the harm, and an inefficient level of precaution with respect to the probability of an accident.

We analyze four possible tort-law solutions to the judgment-proof problem: punitive damages, average compensation of the harm, undercompensation, and negligence. An approach similar to ours has only been provided by Boyd and Ingberman (1994). They concluded that punitive damages should be used in cases governed by the first model, while they would cause even more serious underprecaution when the injurer can reduce the harm; instead, in these cases undercompensation should be applied. Here we account for these results and illustrate, to the contrary, that average compensation is a better solution than punitive damages, because it results in a lower bankruptcy rate and it is easier to apply for the court. In addition to that, we argue that average compensation might in theory be applied to some cases in which also the magnitude can be reduced, but it is to be put aside for another reason: it is cumbersome to apply. We run then a broader analysis and consider undercompensation and negligence under the four models.

We find that:

Average compensation is in most cases superior to *punitive damages*; average compensation (as punitive damages) only functions in stochastic models and it is in theory applicable not only when injurers can reduce the probability of an accident, but also when

injurers can reduce the magnitude of the harm. In general, practical difficulties discourage the use of average compensation in the latter situations.

Undercompensation can reduce the judgment-proof problem only if the injurer has some influence on the harm; hence, it is not a good solution to the probability model. However, while in the magnitude model undercompensation can be applied as a fixed reduction in the damage award for any level of precaution, in the mixed models it has to be tailored on the level of precaution taken by the injurer and, hence, triggers higher administrative costs.

Negligence functions as a harm subsidy for non-negligent injurers, in a similar way as undercompensation. However, since it creates a discontinuity in the harm function, it reduces the effect of judgment-proofness in all models, also in the probability model, where undercompensation fails.

Those findings may furnish with the following policy indications:

(i) In probability models both negligence and average compensation are feasible solutions; however, they operate under different conditions, hence either of them might be preferable given the specific circumstances of the case. Average compensation corrects underprecaution if the harm is stochastic and unpredictable and if the average is lower than the threshold. Negligence corrects underprecaution in non-stochastic models as well, but requires the due level of precaution to be set at the optimal level and triggers some administrative costs due to the application of the negligence inquiry.

(ii) In magnitude models the best solution is a fixed undercompensation of the harm, as it corrects judgment-proofness without triggering the administrative costs due to the application of the negligence criterion.

(iii) In mixed models (joint-probability-magnitude models and separate-probability-magnitude models), the best solution is negligence, as the administrative costs due to its application are lower than those yielded by undercompensation. Undercompensation cannot be applied as a fixed reduction in the damage award, but has to be variable with the level of precaution; thus, it triggers the administrative costs due to the inquiry into the injurer's precaution plus some compensation costs due to the payment of (reduced) damage awards. Negligence triggers administrative costs as well, but it creates a clear cut: negligent injurers pay damages, while non-negligent injurers do not pay damages. Therefore, it saves some compensation costs.

In synthesis, all our solutions to the judgment-proof problem are based on a partial subsidy of the harm, which is 'paid' to the injurer by the victim. Undercompensation is the clearest case: the injurer always pays damages equal to a portion of the harm; the rest is

externalized to (subsidized by) the victim.

Negligence relies on a more complex system: the harm-subsidy is only paid to non-negligent injurers, as they do not have to compensate the victim.

Average compensation is a mixed solution as it creates a subsidy in all those cases, in which the harm happens to be higher than average (the injurer pays under-compensatory damages), and a sort of “fine” when the harm happens to be lower than average (the injurer pays over-compensatory damages). However, in average, damages are compensatory.

The harm-subsidy created by those solutions tends to counterbalance the harm-subsidy created by the judgment-proof problem (as judgment-proof injurers do not pay the full expected accident loss). The reason is that the judgment-proof subsidy attracts the injurer towards low precaution levels as it increases when precaution decreases (the lower precaution, the greater the probability to receive the subsidy). In contrast, the harm subsidy created by the solutions that we analyze attracts injurers toward optimal precaution, as it reaches its maximum when precaution is optimal, as it will be shown in the analysis.

Punitive damages are in general an inferior solution, as by punishing a potentially insolvent injurer by means of overcompensation, they might amplify the judgment-proof problem¹. On the one hand, a judgment-proof injurer is not able to pay compensatory damages and, therefore, cannot pay for overcompensation either; hence, punitive damages cannot provide him with better incentives. On the other hand, some potentially judgment-proof injurers might become judgment-proof precisely because of punitive damages; hence, punitive damages might increase the rate of bankruptcy. In addition, if improperly applied, punitive damages might lead to overprecaution.

5.1.A. Applicability of the analysis

Our analysis applies to situations in which the expenditure on precaution does not reduce the maximum upper threshold: (i) thresholds set by nature, when the threshold is due to limited injurer’s assets, but precaution is of non-monetary kind; (ii) thresholds set by law (liability caps)².

¹ See sections 4.A. and 4.B. for a more complete analysis of the condition of this result.

² The case of disappearing defendants (where there is a certain probability that the injurer will not be sued, see Summers, 1983) is not a judgment-proof problem, as it does not generate a maximum upper threshold on the amount of damages that the injurer has to pay in the case of an accident, but simply dilute the probability that the injurer will actually pay damages. Under a probability model, the two problems have the same effect on the injurer’s incentives, as both reduce the portion of the harm that the injurer internalizes (see Shavell 1984, who considers the two together). However, under the other three models that we discuss in this chapter, where the injurer can also (or only) reduce the magnitude of the accident, the case of disappearing defendants is different from the case of a maximum upper threshold on liability. In the latter, the injurer pays the harm, $h(x)$, if this is lower than or equal to the threshold, $h(x) \leq t$ (precaution, x , sufficiently high), and pays the threshold, t , otherwise: the threshold cuts the harm function: the portion of the harm which exceeds the threshold is completely externalized. On the contrary, if there is a probability a that the injurer will not be sued, he internalizes an accident

We discuss the remaining case of thresholds set by nature, when the threshold is due to limited injurer's assets, and precaution is of monetary kind in a separate paper³. In such a case, the amount of money spent for precautions reduces the amount of money available for paying damages, and thus the threshold decreases when the level of precaution increases. This might change the incentives to take precaution and might lead to either under or overprecaution .

With respect to maximum upper thresholds set by law, we do not discuss the reasons why such thresholds are introduced, but only their effects on the injurer's precaution.

This chapter is structured as follows. After providing with a note on the literature (section 5.1.B), we will describe the four models (section 5.2). In section 0, we will introduce the possibility that harm vary stochastically. In section 5.4, we will examine the possible tort-law solutions to the judgment-proof problem. Section 5.5 will address some concluding remarks.

5.1.B. A note on the literature

Several aspects of maximum upper thresholds have already been studied in the law and economics literature, though under different labels (insolvency, liability caps, corporate limited liability, bankruptcy, inability to pay damages). In this chapter, we consider all these different issues as being of the same nature and label them as maximum upper threshold for the sake of generality. Our analysis applies to each of these problems and can be extended to contract liability as well, for the parties' incentives to take precaution are also of main concern in contract settings.

More importantly, this chapter tries to draw a more complete picture than previous contributions in the literature, which were focused on specific solutions under specific models. The first formal analyses of the judgment-proof problem were made by Summers (1983) and Shavell (1986), but both only used a (non-stochastic) probability model. Boyd and Ingberman (1994) analyzed the judgment proof problem by employing a stochastic probability model, and showed that punitive damages can be a solution. We attempt to make a more general analysis, by inquiring into four different tort models (each of which with a stochastic variant) and into four possible legal solutions: punitive damages, average compensation, under-compensatory damages, and negligence.

loss lower than the actual harm, $ah(x)$, for any level of x . Hence, the case of disappearing defendants has the same effect on all models, and cannot be considered as a maximum upper threshold on liability. A punitive-damage coefficient equal to $1/a$ solves the problem under all models.

³ See chapter 7 of this book; Beard (1990) was the first to analyze this problem in a (stochastic) probability model, in which the injurer's precaution costs are deducted from the injurer's remaining assets in case of an accident. Garoupa and Gravelle (2002) apply a similar refinement to the case of law enforcement with legal expenditures.

5.2. The model

We make the following assumptions:

1. Unilateral precaution situation, only the injurer can take precaution, only the victim suffers harm;
2. The rule in force is strict liability⁴;
3. Parties are perfectly informed and the courts entitle victims to perfect compensation;
4. Parties are risk-neutral⁵, rational and utility maximizing.

Notation:

- x = the expenditure on precaution⁶ or its monetary equivalent (if non-monetary),
 $x=[0, \infty)$;
- s, z = the expenditures in probability-reducing and magnitude-reducing precaution,
 $s=[0, \infty), z=[0, \infty)$;
- p = the probability of an accident, $p=[0, 1]$;
- h = the harm if an accident occurs, $h=[0, \infty)$;
- J = the injurer's total expenses (expected liability plus precaution);
- SC = the social cost
- t = the maximum upper threshold (maximum amount of damages that the injurer has to pay).

The social cost function can generally be described as the sum of the expected accident loss (the probability of an accident times the magnitude of the harm) plus the expenditure on precaution. Depending on the model, the expected accident loss will be defined differently (either the probability or the magnitude of the harm or both will be considered as a function of the expenditure on precaution, x).

We first describe and analyze the four models in their non-stochastic (or 'deterministic') variant: we assume that there are not random variations around h , i.e. the injurer exactly knows the harm resulting in the case of an accident. Later we relax this assumption and consider stochastic models.

⁴ We will consider simple negligence in section 4.

⁵ Our analysis will be performed under a risk-neutrality assumption, which makes the model easier to follow. In general, risk-averse individuals might take a different level of precaution than risk-neutral individuals.

⁶ In some models x stands for the *level* of precaution and the *cost* of precaution is denoted by $c(x)$. Our solution (also adopted, for instance, in Shavell, 1987) is not qualitatively different and the results would be the same under either approach. We prefer the expenditure on precaution to be denoted simply by x in order to simplify the notation and the expressions. We will use indifferently the terms "expenditure on precautions", "cost of precaution" and "level of precaution" to refer to x . This approach could also be seen as an additional assumption that the cost of precaution is constant and unitary, and therefore the level of precaution coincides to the cost of

5.2.A. First model: probability model

In this model, the probability of an accident is endogenous; it is a function of the expenditure on precaution. Let:

$$p=p(x)=(0,1); p'(x)<0, p''(x)>0.$$

h is exogenous. Plausible examples of accidents in which the injurer can only influence the probability are airplane accidents and nuclear accidents. The social cost function is given by:

$$(1) \quad SC(x) = p(x)h + x$$

Let x^* ⁷ denote the level of precaution that minimizes the social cost function. We consider the introduction of a maximum upper threshold, t , so that the injurer pays damages that he causes only if $h \leq t$, and pays t if $h > t$. The injurer's cost function is given by:

$$(2) \quad \begin{cases} J(x) = p(x)h + x & \text{if } h \leq t \\ J_t(x) = p(x)t + x & \text{if } h > t \end{cases}$$

*Proposition 1: in a probability model, a potentially judgment-proof injurer ($t < h$) takes underprecaution systematically. The lower the threshold, the lower the level of precaution that the injurer takes*⁸.

Proof.

Since the injurer cannot influence the magnitude of damages at all, the introduction of a threshold reduces h for any value of x . $J(x)$ is minimized by x^* ; let x_t ⁹ denote the level of x that minimizes $J_t(x)$.

x_t is clearly lower than x^* and decreases with t , as it can be easily shown by confronting the first order conditions for the two functions in Exp. (2)¹⁰. The minimum level for the threshold being innocuous is:

$$(3) \quad t_{min} = h.$$

Figure 1 shows the former statements.

precaution.

⁷ The solution is internal and unique for the assumptions made.

⁸ The probability-model follows closely Shavell (1986).

⁹ The solution is internal and unique for the assumptions made.

¹⁰ The two first order conditions are $p'(x^*)h = -1$ for the first equation in Exp. (2), and $p'(x_t)t = -1$, for the second equation. We can write $p'(x^*)h = p'(x_t)t$. Because $t < h$, and $p'(x) < 0$, then $p'(x_t) < p'(x^*)$. Recalling that $p''(x) > 0$, we can conclude that also $x_t < x^*$, that is the injurer will take a sub-optimal level of precaution.

5.2.B. Second model: magnitude model

The magnitude of the harm is now endogenous ($h=h(x)$) and the probability p exogenous. More precaution only leads to a decrease in the magnitude of the harm in case of an accident. Let:

$$h=h(x)=(0, \infty); h'(x)<0, h''(x)>0.$$

Clear examples are nuisance cases (as already mentioned) where the probability is not only fixed but in many cases also close to 1 (if I make too much noise at night it is certain that my neighbors will suffer). As examples where p is lower than 1 we could think off a safety web for acrobats in a circus put by the circus owner in order to reduce the injuries in the case of a fall. The social cost function is given by:

$$(4) \quad SC(x) = ph(x) + x$$

Let x^{*11} denote the level of precaution that minimizes the social cost function. If a threshold t is present, the injurer's expenses function will be

$$(5) \quad \begin{cases} J(x) = ph(x) + x & \text{if } h(x) \leq t \\ J_t(x) = pt + x & \text{if } h(x) > t \end{cases}$$

Proposition 2: in a magnitude model, if the maximum upper threshold is sufficiently high the injurer decides not to be potentially judgment-proof and takes optimal precaution. Otherwise he takes no precaution at all.

Proof

$J(x)$ is minimized by x^* , hence the level of precaution chosen by the injurer is socially optimal. $J_t(x)$ is minimized by $x=0$, as the function is linear in x . The logic of this result is that the injurer has no incentive to take precaution in the judgment-proof zone, since he has to pay t anyway.

In the magnitude model there is not simply underprecaution but a danger that the injurer will choose no precaution at all. The injurer will choose the optimal level of precaution if

$$(6) \quad J_t(0) = pt \geq ph(x^*) + x^* = J(x^*).$$

He will take no precaution otherwise.

From Exp. (6), the level of the threshold that does not undermine the incentive for the injurer to take optimal precaution is:

¹¹ The solution is internal and unique for the assumptions made.

$$(7) \quad t_{\min} = h(x^*) + \frac{x^*}{p}^{12}.$$

If $t < t_{\min}$ the injurer will choose no precaution and the outcome will be largely inefficient. On the contrary, if t is set at a level which is greater than or equal to t_{\min} , the presence of a threshold does not have any effect on the injurer's level of precaution.

It is to be noted that this case is different from the case of proposition 1 in two respects. First, the minimum level at which the threshold has no effect is different (in the case described by proposition 1, $t_{\min} = h$). Second, the effect on incentives does not vary with the threshold, once the threshold is below the minimum level, the injurer takes no precaution and further diminutions do not bring any consequence.

Figure 2 shows the former statements.

5.2.C. Third Model: joint probability-magnitude model

In this case $h = h(x)$ and $p = p(x)$. The injurer can reduce both the probability and the magnitude of the harm, but cannot discriminate between them. An increase in the expenditure on precaution, x , causes a simultaneous reduction in $h(x)$ and $p(x)$. Let:

$$\begin{aligned} p &= p(x) = (0, 1); \quad p' < 0, \quad p'' > 0; \\ h &= h(x) = (0, \infty); \quad h'(x) < 0, \quad h''(x) > 0. \end{aligned}$$

A good example is the speed of a car: the higher the speed (= less precaution), the higher the chance that an accident will happen and the higher the harm in case of an accident. The social costs function is given by:

$$(8) \quad SC(x) = p(x)h(x) + x.$$

Let x^{*13} denote the level of precaution that minimizes the social cost function. If a threshold t is present the injurer's cost function becomes:

$$(9) \quad \begin{cases} J(x) = p(x)h(x) + x & \text{if } h(x) \leq t \\ J_t(x) = p(x)t + x & \text{if } h(x) > t \end{cases}$$

$J(x)$ is minimized by x^* , so that the level of precaution chosen by the injurer is socially optimal (the accident costs are fully internalized). Let x_t^{14} denote the lowest point of the function $J_t(x)$.

¹² It is to be noticed that Exp. (7) implies that $t > h(x^*)$, i.e. it requires the injurer being able to pay damages resulting from his taking optimal precaution.

¹³ The solution is internal and unique for the assumptions made.

¹⁴ The solution is internal and unique for the assumptions made.

Proposition 3: in a joint probability-magnitude model, if the maximum upper threshold is sufficiently high the injurer decides not to be potentially judgment-proof and takes optimal precaution. Otherwise he takes underprecaution. The lower the threshold, the lower precaution that the injurer takes in the judgment-proof zone.

Proof.

First we define the minimum level of the threshold for the injurer choosing the optimal level of precaution, and then we prove that if the threshold is at a lower than minimum level, the injurer will choose too low a level of precaution, $x_t < x^*$.

The injurer will choose the optimal level of precaution, x^* , if:

$$(10) \quad J_t(x_t) = tp(x_t) + x_t \geq h(x^*)p(x^*) + x^* = J(x^*).$$

Thus, the threshold could be set at a level which does not undermine the incentive for the injurer to take optimal precaution. From Exp. (10), the minimum level of the threshold is:

$$(11) \quad t_{\min} = \frac{p(x^*)h(x^*) + x^* - x_t}{p(x_t)}.$$

If t is set at a level greater than or equal to t_{\min} the presence of a threshold does not have any effect on the injurer's level of precaution. On the contrary, if $t < t_{\min}$ the injurer will choose too low a level of precaution, x_t and the outcome is inefficient, as x_t will always be lower than x^* .

In order to prove this point, let x_c denote the level of precaution that satisfies $h(x_c) = t$. Graphically, this level of precaution divides the figure in two regions: on the left of it $J(x) > J_t(x)$ [judgment-proof zone], while on the right of it $J(x) < J_t(x)$ [solvency zone], as it can be easily verified from Exp. (9) and from figures 3 and 4.

If the threshold is so that $t < h(x^*)$, then $x^* < x_c$, which follows from the definition just given. In this case, described in figure 3, the socially optimal level of precaution lays in the region to the left of x_c . Obviously in this region Exp. (11) is never satisfied as $J(x) > J_t(x)$. In other words the injurer, by taking the optimal level of precaution would cause a harm that he cannot compensate, as $h(x^*) > t$, and will face a total expected cost of $p(x^*)t + x^* = J_t(x^*)$. However, this is not optimal for the injurer since we have defined as x_t the level of precaution which minimizes $J_t(x)$. The injurer can, therefore, reduce his cost by moving from x^* to x_t . Now we have to prove that x_t is a lower than optimal level of precaution, that is $x_t < x^*$.

For $J(x)$ to be minimized, x^* has to fulfill the following first order condition:

$$(12) \quad p'(x^*) = -\frac{1 + h'(x^*)p(x^*)}{h(x^*)}.$$

For $J_t(x)$ to be minimized, x_t has to be such that

$$(13) \quad p'(x_t) = -\frac{1}{t}.$$

By noting that $h'(x)$ is negative and, therefore, that $1+h'(x^*)p(x^*) < 1$, we can conclude that, as $h(x^*) > t$, $p'(x_t) < p'(x^*)$. As $p''(x) > 0$, then $p'(x_t) < p'(x^*)$ implies $x_t < x^*$, which is what we wanted to prove. Therefore, if $t < h(x^*)$, the injurer will choose a lower than optimal level of precaution.

Let us now consider the case of the threshold being greater than or equal to the level of the harm due to optimal precaution, $t \geq h(x^*)$, described in figure 4. In this case $x^* \geq x_c$. Obviously, if $x_t > x_c$, then both x^* and x_t lay to the right of x_c , where the threshold excides the actual harm, but we know that the injurer never pays more than the harm, and therefore this situation is not to be considered. Mathematically, $x > x_c$ implies $J(x) < J_t(x)$ and therefore Exp. (11) is satisfied, and the injurer chooses x^* .

If $x_t < x_c$, then the injurer will take x^* if Exp. (11) is satisfied and x_t otherwise. In any case, because $x_t < x_c$, and $x^* > x_c$, then $x_t < x^*$, also here. We can conclude that, if $h(x^*) \leq t$, the injurer will choose x^* if the threshold is sufficiently high, and a lower than optimal level of precaution, x_t , otherwise.

In addition, it is worth noting that the cost borne by the injurer, given that he chooses x_t , is increasing with the level of the threshold. Let x_t^1 denote the level of precaution that minimizes $J_{t^1}(x)$ given a threshold t^1 , and x_t^2 the level of precaution that minimizes $J_{t^2}(x)$ given a threshold $t^2 > t^1$; then it is clear from Exp. (13), that $x_t^2 > x_t^1$. It follows that also $J_{t^2}(x) > J_{t^1}(x)$ for any x . Therefore, also $p(x_t^2)t^2 + x_t^2 > p(x_t^1)t^1 + x_t^1$, since x_t^1 is by definition the value of x that minimizes $p(x)t^1 + x$.

Therefore, when the level of the threshold increases the cost associated to choosing x_t increases as well, and becomes more convenient to choose x^* , as already shown by Exp. (11).

In the analysis we have considered the possibility that t is below $h(x)$ for low values of x , but at higher precaution levels, t rises above $h(x)$. Two more situations may occur. First, t can be so low that it is always below $h(x)$, even if x is infinitely (or maximally) high. In this case, the injurer's expenses are always $p(x)t + x$. The model becomes identical as our first model (probability model) and the results (the injurer choosing a too low precaution level) are identical as well.

Second, t can be so high that it is always above $h(x)$. Here, the injurer's expenses are always $p(x)h(x) + x$. The effects are perfectly internalized, and the injurer always chooses the optimal precaution level x^* .

Figure 3 shows the former statements.

5.2.D. *Fourth model: separate (3-dimensional) probability-magnitude model*

In this model, the injurer can discriminate between the expenditure on reducing the probability and the expenditure on reducing the magnitude of the harm in case of an accident.

A good example is the shipwreck case. The owner of a ship can invest in controlling the shore and the engines, in accurate weather forecasts, or in navigation devices, in order to reduce the probability of a shipwreck. However, once a shipwreck has occurred, the presence of lifeboats on board will enormously reduce the magnitude of the damages to the passenger and to the crew (though lifeboats are completely irrelevant for the occurrence of the accident). The two expenditures are independent from each other.

In this case $h=h(z)$ and $p=p(s)$, where z and s are expenditures in two different kinds of precautions for reducing respectively h and p . The injurer can reduce both the probability and the magnitude of the harm, and can freely discriminate between them. The social costs function is given by:

$$(14) \quad SC(s, z) = p(s)h(z) + s + z$$

Let s^* and z^{*15} denote the levels of probability-reducing precaution and magnitude-reducing precaution respectively that minimize the social cost function. If a threshold t is present the injurer's cost function becomes:

$$(15) \quad \begin{cases} J(s, z) = p(s)h(z) + s + z & \text{if } h(z) \leq t \\ J_t(s, z) = p(s)t + s + z & \text{if } h(z) > t \end{cases}$$

Proposition 4: in a separate (3-dimensional) probability-magnitude model, if the maximum upper threshold is sufficiently high, the injurer decides not to be potentially judgment-proof and takes optimal precaution. For intermediate levels of the threshold, the injurer takes no precaution with respect to the magnitude-reducing measure and overprecaution (or optimal precaution) with respect to the probability-reducing measure. For lower levels of the threshold, the injurer takes no precaution with respect to the magnitude-reducing measure of the harm and underprecaution with respect to the probability-reducing measure. The lower the threshold, the lower the probability-reducing precaution that the injurer takes in the judgment-proof zone.

Proof.

$J(s, z)$ is minimized by (s^*, z^*) ; thus, the level of precaution chosen by the injurer is socially optimal. Let $(s, z=0)$ denote the levels of precaution that minimize $J_t(s, z)$, as the function is

¹⁵ The solutions are internal and unique for the assumptions made.

linear in z and convex in s ¹⁶. $s_t \neq s^*$, so that s_t is an inefficient level of precaution. Unless $z^*=0$, also z_t is an inefficient level of precaution. For $J(s, z)$ being minimized, the first order conditions are:

$$(16) \quad \begin{cases} p'(s^*)h(z^*) = -1 \\ p(s^*)h'(z^*) = -1 \end{cases};$$

and, given $z=0$, for $J_t(s, z)$ being minimized, the first order condition is:

$$(17) \quad p'(s_t)t = -1.$$

Therefore:

$$(18) \quad p'(s_t) = \frac{-1}{t} \quad \text{and} \quad p'(s^*) = \frac{-1}{h(z^*)}$$

From Exp. (18) it follows that $p'(s_t) \leq p'(s^*)$, if $t \leq h(z^*)$. Because $p''(s) > 0$, also $s_t \leq s^*$. The injurer will under-invest in precaution to reduce the harm ($z_t=0$) and under-invest in precaution to reduce the probability of an accident ($s_t \leq s^*$). On the contrary, if $t > h(z^*)$, the injurer will under-invest in precaution to reduce the harm ($z_t=0$) but will over-invest in precaution to reduce the probability of an accident ($s_t > s^*$).

The injurer will choose $z_t=0$ and s_t if

$$(19) \quad J_t(s, 0) = p(s_t)t + s_t < p(s^*)h(z^*) + s^* + z^* = J(s^*, z^*).$$

The threshold, therefore, might be set at a level that does not undermine the incentive for the injurer to take optimal precaution. From Exp. (19), the minimum level of the threshold is:

$$(20) \quad t_{\min} = \frac{p(s^*)h(z^*) + s^* + z^* - s_t}{p(s_t)}.$$

Therefore, the results depend on whether or not the injurer will choose the efficient level of precaution (s^* and z^*) and, in the latter case, whether the chosen level of probability reducing precaution will be lower than, equal to or greater than the efficient level (the inefficient level of magnitude reducing precaution being always zero).

The results can be summarized in the table below. In the case of the injurer being able to reduce both the probability of an accident and the magnitude of the harm by investing in different kinds of precaution, a maximum upper threshold will result in no precaution with respect to magnitude-reducing precaution if the threshold is below the minimal level. With respect to probability-reducing precaution different outcomes can arise: over, under, and optimal precaution can result. The logic is that in some cases it is advantageous for the injurer

¹⁶ The solution with respect to s only is internal and unique for the assumptions made.

to spend in reducing the probability of an accident rather than the magnitude as only the first expenditure reduces his expected liability. If the threshold is above the minimal level efficient precaution will result with respect to both precaution expenditures.

Table 1. Precaution in separate probability-magnitude models

Threshold	Probability-reducing precaution	Magnitude-reducing precaution
$t < h(z^*)$	$s_t < s^*$	$z_t = 0$
$t = h(z^*)$ ¹⁷	$s_t = s^*$	$z_t = 0$
$h(z^*) < t < \frac{p(s^*)h(z^*) + s^* + z^* - s_t}{p(s_t)}$	$s_t > s^*$	$z_t = 0$
$t \geq \frac{p(s^*)h(z^*) + s^* + z^* - s_t}{p(s_t)}$	s^*	z^*

5.3. Stochastic variation of the harm

In the foregoing sections we have described a deterministic model, that is a model in which the harm is exactly determined either exogenously, h , or endogenously as a function of the level of precaution, $h(x)$ or $h(z)$. In this section we introduce a stochastic variation of the harm for each of the four models, and discuss possible variations of the conclusions reached so far. As we are assuming risk-neutrality, the increase in the risk borne by the injurer does not affect *per se* the outcome.

We find that stochasticity does not change the conclusions reached for the first model, because the injurer's expectation will be the same. In contrast, under the other three models, the level of the harm depends on the injurer's precaution. In the deterministic model, when the injurer decreases his level of precaution, and the harm becomes too great, he will pay just the threshold. In the stochastic model this sharp effects fades, because when the injurer decreases his level of precaution, he will not be able to pay damages if the harm is high, but he will pay for them if the harm is low.

Therefore, there is an intermediate zone between paying the harm and paying only the threshold; in this zone the injurer pays sometimes the harm (if low) and sometimes the threshold (if the harm is high). This fact might change the results: the injurer might take an inefficient level of precaution, but different from the one the deterministic model would

¹⁷ In this case, as $s_t = s^*$, then t_{min} becomes equal to $h(z^*) + z^*/p(s^*)$, which is clearly not satisfied by $t = h(z^*)$. The injurer takes optimal precaution-reducing precaution but still opts for no magnitude-reducing precaution.

predict.

First (stochastic) model: probability model

With respect to the first model nothing changes if we introduce stochasticity. In fact, the harm is independent from the level of precaution and the injurer, instead of paying a determined harm, pays an expected harm.

We introduce a stochastic variation for h between a lower limit, \underline{h} , and an upper limit, \bar{h} ; $\pi(h)$ will be the density function for h , and $E(h)$ the expected harm.

$$(21) \quad E(h) = \int_{\underline{h}}^{\bar{h}} \pi(h)h dh.$$

The injurer's expenses function is $p(x)E(h)+x$, and it is minimized by x^* , the optimal level of precaution¹⁸. In the presence of a threshold, $t < \bar{h}$, the expected harm becomes:

$$(22) \quad E(t) = \int_{\underline{h}}^t \pi(h)h dh + \int_t^{\bar{h}} \pi(h)t dh < E(h).^{19}$$

The injurer's expenses function is $p(x)E(t)+x$, and it is minimized by x_t , a lower than optimal level of precaution, as he expects to pay lower damages for any level of precaution²⁰. Therefore, as in the deterministic model, if the threshold is lower than (some realizations of) the harm, the injurer does not fully internalize the harm that he can cause and hence takes less than optimal precaution.

Second (stochastic) model: magnitude model

In the magnitude model the harm is a function of the level of precaution, $h(x)$. In order to capture the stochastic component of the harm, we make a simplifying assumption: we assume that nature randomly select a magnitude function, $h_i(x)$, among all possible magnitude functions, with probability π_i , ($\sum \pi_i = 1$), and we assume that all possible magnitude functions have the same characteristics of being decreasing in x at a decreasing rate and that they do not intersect. As the core of our argument does depend on the presence of a stochastic variation and not on the number of different possible harm functions, it will not change if we consider the simplest case of only two magnitude functions, $\underline{h}(x) < \bar{h}(x)$, for any x , occurring respectively with probability $\underline{\pi}$ and $\bar{\pi}$, ($\underline{\pi} + \bar{\pi} = 1$).

¹⁸ As under the deterministic model x^* is optimal because it yields the minimum in the social cost function, $p(x)E(h)+x$.

¹⁹ Obviously, if t is lower than the lower limit, the first addendum disappears, and the injurer always pay t .

²⁰ This point can be easily verified by confronting the first order conditions for $E(h)$ and for $E(t)$.

The difference with respect to the deterministic model, Exp. (5), is that in this case we can single out three cases instead of only two: $J(x)$ occurs when the injurer can always compensate the victim for the harm, $t \geq \bar{h}(x)$; $J_t(x)$ occurs when the injurer cannot pay full compensation for any harm, $t \leq \underline{h}(x)$; in addition to those, there is an intermediate case, $J_m(x)$, occurring when the injurer can pay the harm only if the lower harm results, $\underline{h}(x) < t < \bar{h}(x)$.

$$(23) \quad \begin{cases} J(x) = p[\pi \underline{h}(x) + \bar{\pi} \bar{h}(x)] + x & \text{if } \bar{h}(x) \leq t \\ J_m(x) = p[\pi \underline{h}(x) + \bar{\pi} t] + x & \text{if } \underline{h}(x) < t < \bar{h}(x) . \\ J_t(x) = pt + x & \text{if } \underline{h}(x) \geq t \end{cases}$$

The injurer fully internalizes the harm in $J(x)$ and his level of precaution will be the optimal level of precaution x^* , which minimizes his expenditure function, as in the deterministic model. Similarly, $J_t(x)$ is minimized by $x=0$, as in the deterministic model. The only difference is therefore the second case. $J_m(x)$ is clearly convex²¹, and is minimized by a level of precaution, denoted by x_m , which is lower than optimal, $x_m < x^*$ ²².

After confronting $J(x^*)$, $J_t(0)$ and $J_m(x_m)$, the injurer will select the level of precaution which yields the lowest cost. Therefore, the feasible outcomes are not just two as in the deterministic model, but three: x^* , $x=0$, and x_m . In other words, the outcome can be optimal precaution, no precaution, or underprecaution.

There are however two criteria to be used to limit the range of possible outcomes:

- (i) If $t > \underline{h}(0)$, then $J_m(x_m) < J_t(0)$ ²³, otherwise either $J_m(x_m) < J_t(0)$ or $J_m(x_m) > J_t(0)$ can result;
- (ii) If $t < \bar{h}(x^*)$, then $J_m(x_m) < J(x^*)$ ²⁴, otherwise either $J_m(x_m) < J(x^*)$ or $J_m(x_m) > J(x^*)$ can result.

Therefore, we can predict some of the outcomes, by just looking at the level of the threshold. By combining the two former results we obtain:

- (a) If $t < \underline{h}(0)$ and $t < \bar{h}(x^*)$, the outcome might be either $x=0$ or x_m ;
- (b) If $\underline{h}(0) < t < \bar{h}(x^*)$, the outcome will be x_m ;
- (c) If $\bar{h}(x^*) < t < \underline{h}(0)$, the outcome could be any of $x=0$, x_m , and x^* .

²¹ It is worth recalling that $h''(x) > 0$.

²² By confronting the first order conditions for $J(x)$ and $J_m(x)$ we can verify this result. For $J(x)$ we have $\underline{h}'(x^*) = -\frac{1 + p\pi h'(x^*)}{p\pi}$, and for $J_m(x)$ we have $\underline{h}'(x_m) = -\frac{1}{p\pi}$. As $p\pi h'(x) < 0$, then $\underline{h}'(x^*) > \underline{h}'(x_m)$, and as

$h''(x) > 0$, necessarily also $x^* > x_m$.

²³ In fact in this case: $J_t(0) > J_m(0) > J_m(x_m)$.

²⁴ In fact in this case: $J(x^*) > J_m(x^*) > J_m(x_m)$.

(d) If $t > \underline{h}(0)$ and $t > \bar{h}(x^*)$, the outcome could be either x_m or x^* .

More precise answer to the question of what level of precaution that the injurer will choose may be obtained by adopting a more refined model for the stochastic variation of the harm. The point we want to make here is that the stochastic variation might sometimes change the conclusions of our model. However, a little stochastic variation around the expected value of the harm is not likely to affect our conclusions.

Third (stochastic) model: joint probability-magnitude model

In the third model, the injurer can reduce the probability of an accident as well as the magnitude of the harm. Nevertheless, the same kind of reasoning as before can be applied here also.

$$(24) \quad \begin{cases} J(x) = p(x) [\underline{\pi} \underline{h}(x) + \bar{\pi} \bar{h}(x)] + x & \text{if } \bar{h}(x) \leq t \\ J_m(x) = p(x) [\underline{\pi} \underline{h}(x) + \bar{\pi} t] + x & \text{if } \underline{h}(x) < t < \bar{h}(x) \\ J_t(x) = p(x)t + x & \text{if } \underline{h}(x) \geq t \end{cases}$$

The four criteria described above ought to be rephrased as follows:

- (a) If $t < \underline{h}(0)$ and $t < \bar{h}(x^*)$, the outcome could be either x_t or x_m ;
- (b) If $\underline{h}(0) < t < \bar{h}(x^*)$, the outcome will be x_m ;
- (c) If $\bar{h}(x^*) < t < \underline{h}(0)$, the outcome could be any of x_t , x_m , and x^* .
- (d) If $t > \underline{h}(0)$ and $t > \bar{h}(x^*)$, the outcome could be either x_m or x^* .

Fourth (stochastic) model: separate (3-dimensional) probability-magnitude model

The fourth model results as a combination of the first two; therefore, the same considerations already made can be applied here.

5.4. Correcting inefficient deterrence

We will analyze several methods to correct inefficient deterrence due to a maximum upper threshold and precisely punitive damages (and in general overcompensation), average compensation, undercompensation and negligence. It will be shown that the measure to be adopted ought to be tailored on the specific model which we are presented with. A synthesis of the results is given in the introduction.

5.4.A. Punitive damages (overcompensation)

Punitive damages could be introduced as a way to cope with a maximum upper threshold. This method has been proposed by Boyd and Ingberman (1994) as a way to reduce the

underprecaution problem in a stochastic probability model. Under a deterministic probability model such a system would produce the opposite effect.

In fact, as the injurer is subject to a threshold t on the damages that he can pay, he will not be able to pay punitive damages either; therefore, the underprecaution problem remains in place. In addition, if applied to non-judgment-proof injurers, punitive damages might generate judgment-proofness, if they increase over the threshold the damages to be paid to the victim.

In the second model (when the injurer can only influence the magnitude of the harm) punitive damages would increase the comparative advantage of taking no precaution at all, amplifying instead of solving the problem. A similar outcome would result in the remaining two models as well.

On the contrary, punitive damages can enhance efficient precaution under model 1, with stochastic harm, the reason being that under such a model the injurer faces an expected harm that is reduced by the threshold; punitive damages, by making the injurer pay more than the harm that he actually caused when the latter is below the threshold tend to restore a correct expectation.

To understand how such a mechanism functions it is necessary to go back to the stochastic variation for model 1 described in section 4, Exp. (22), and introduce punitive damages, in order to correct the expectation:

$$(25) \quad E(t) + PD = E(h).$$

The potentially judgment-proof injurer's cost function is again $p(x)E(h) + x$, and it is minimized by x^* , the optimal level of precaution. The underprecaution problem is solved.

There are, however, three problems. First, the application of punitive damages increases the number of cases in which the injurer will be insolvent, which can be an undesirable outcome, as already noted by Boyd and Ingberman (1994)²⁵.

Second, the application of punitive damages in order to correct a maximum upper threshold problem is difficult in practice. The court has to apply a punitive measure to solvent individuals or firms on the basis of the hypothesis that if the harm had happened to be greater the injurer would not have been able to pay. In the case of natural maximum upper threshold (limited injurer assets) the court has to ascertain exactly the injurer's wealth and determine in which cases he would not be able to compensate the victim. On the basis of this information the court sets punitive damages.

In addition, the solution proposed by Boyd and Ingberman (1994) leads to the conclusion that poor injurers should pay more than rich ones. Besides the difficulty of accepting such an

²⁵ In Boyd and Ingberman (1999), the authors argue that punitive damages might yield an additional distortion: they might induce injurers to reduce their exposure to liability through reducing their assets, t . They assume that

undeeep-pocket rule, this solution proves to be economically undesirable. In addition to the difficulty inherent to the judicial determination of the correct amount of punitive damages to award, such an amount should be easily predictable *ex ante* in order to induce efficient behavior. Moreover, individuals and firms' assets vary over time, which makes it more difficult to calculate the magnitude of punitive damages. Fewer problems might arise when the threshold is legally determined (liability caps, and corporate limited liability), for its determination is easier.

Third, punitive damages cannot solve every maximum upper threshold problem, since they are subject to a limit. The maximum amount of punitive damages that can be applied to an injurer is in fact determined by the difference between the threshold, t , and the harm, h , actually realized in the accident, and only if h is lower than t . *Ex ante*, the maximum *PD* an injurer may expect to be charged with is:

$$(26) \quad PD \max = \int_h^t \pi(h)(t-h)dh.$$

Such an amount of *PD* has to be sufficient for $E(t)+PD \max \geq E(h)$. Therefore:

$$(27) \quad \int_h^t \pi(h)h dh + \int_t^{\bar{h}} \pi(h)t dh + \int_h^t \pi(h)(t-h)dh \geq E(h).$$

The former Exp. Can be rewritten as:

$$(28) \quad t \geq E(h),$$

which is the condition for punitive damages to solve completely the underprecaution problem. Similar concerns would apply to other means of increasing the cost of the accident for the injurer, such as fines to be imposed if an accident occurs or other monetary sanctions.

5.4.B. Average compensation of the harm

This solution consists of making the injurer pay always $E(h)$, regardless of whether the actual realization of the harm, h , is greater (undercompensation) or lower (overcompensation) than its mean, $E(h)$. Such a compensation rule, leading to an inaccurate determination of damages, is not attractive when the injurer can determine the actual magnitude of the harm in advance, and therefore discriminate between accident to avoid (overcompensated) and accident to let occur (undercompensated)²⁶. However, when the outcome of the accident is *ax-ante* unpredictable, average compensation solves some problems due to the application of punitive

injurers can control t ; in our analysis we assume that t be exogenously determined.

²⁶ Accuracy in the determination of damages has been analyzed in a different contest than judgment-proofness by

damages.

First, the court has to gather less information, for average compensation can be applied both for solvent and potentially insolvent injurers, without any distinction. The court, therefore, does not need to ascertain the injurer's assets and to perform any hypothetical reasoning. The measure of the compensation can be determined per categories of accidents in the form of tables (already in use in some courts for different purposes), and not on an individual basis.

Second, the rate of bankruptcy does not increase for the application of average compensation. On the contrary, it is reduced, for little harms are overcompensated but great harms are undercompensated, and therefore injurers will be found insolvent less frequently. If $t \geq E(h)$ for all potentially insolvent injurers, none of them will go bankrupt under an average compensation rule, while some of them would under perfect compensation.

Last, the limit for average compensation being sufficient to solve completely the underprecaution problem is $t = E(h)$ as in the case of punitive damages.

Besides the probability model, average compensation could lead to positive results also in stochastic magnitude or mixed models. Indeed by making the injurer always pay the average harm, a stochastic model is converted into the corresponding deterministic one. Therefore, in theory, average compensation could be applied in all those cases in which the outcome of the stochastic model is x_m , while the outcome of the correspondent deterministic model would be x^* . However, this solution brings along non easily surmountable applicative problems. In fact, since the magnitude of the harm depends on the level of precaution taken by the injurer, the average should be calculated as a function of the level of precaution. This is obviously not an easy task for the court, given the amount of information about the magnitude functions that it would require. On the contrary, a fixed average, not based on the level of precaution, would reduce the incentives to take precaution, because the injurer would receive no advantage from taking more precaution, as he would pay always the same.

Even if theoretically appealing also when the injurer can reduce the magnitude of an accident, this method should be restricted to cases in which the magnitude is exogenous.

5.4.C. Undercompensation of the harm

Average compensation fails to restore an appropriate incentive stream in the case of the injurer being able to influence the magnitude of the harm. In those cases, the judgment-proof problem creates a harm-subsidy for injurers who take low levels of precaution: if precaution is low, the expected harm will be great and probably greater than the threshold; hence, a part of

Kaplow (1994) and Kaplow and Shavell (1996).

the harm will not be paid, and it is as if it will be subsidized by the victim. The judgment-proof subsidy attracts the injurer towards low levels of precaution.

Undercompensation of the harm creates a similar harm-subsidy for injurers at *any level* of precaution: even if the level of precaution is high, a portion of the harm will be subsidized. For this reason, the undercompensation subsidy can partially contrast the judgment-proof subsidy²⁷.

In the next sections we will show how undercompensation can be used as a possible solution to the judgment-proof problem when the injurer can only or also influence the magnitude of the harm.

Magnitude model

If the injurer can influence only the magnitude of damages and fixed undercompensation, g , is introduced, Exp. (5) becomes:

$$(29) \quad \begin{cases} J(x) = ph(x) - pg + x & \text{if } h(x) - g \leq t \\ J_t(x) = pt + x & \text{if } h(x) - g > t \end{cases}$$

The introduction of undercompensation does not change the minimizing level of precaution for $J(x)$ and $J_t(x)$, which are still minimized by x^* and $x=0$ respectively.

However, Exp. (7) becomes

$$(30) \quad t_{\min}^u = h(x^*) + \frac{x^*}{p} - g,$$

Therefore, for some level of t it is possible to set a fixed level of $g \leq h(x^*)$ so that $t = t_{\min}^u < t_{\min}$, and the outcome is efficient. The optimal g is $g = J(x^*) - J_t(x_t) \leq h(x^*)$.

Joint probability-magnitude model

If the injurer can influence jointly the probability of an accident and the magnitude of the harm, the introduction of fixed undercompensation would create under-deterrence. However,

if variable undercompensation, $g(x) = \frac{g}{p(x)}$ is introduced, Exp. (9) becomes:

$$(31) \quad \begin{cases} J(x) = p(x)[h(x) - g(x)] + x & \text{if } h(x) - g(x) \leq t \\ J_t(x) = p(x)t + x & \text{if } h(x) - g(x) > t \end{cases}$$

and Exp. (11) becomes:

$$(32) \quad t_{\min}^u = \frac{p(x^*)[h(x^*) - g(x^*)] + x^* - x_t}{p(x_t)}$$

²⁷ See also Boyd and Ingberman (1994).

Therefore, for some level of t it is possible to set a variable level of g such that $\frac{g}{p(x^*)} \leq h(x^*)$, so that $t = t_{\min}^u < t_{\min}$, and the outcome is efficient. It is noteworthy that $g'(x) > 0$, and therefore the level of undercompensation varies positively with the level of precaution. This can be seen as a sort of comparative negligence rule, as the compensation depends on the level of precaution. A major problem is that the application of variable undercompensation requires precise information about the cost of precaution, which might lead to practical difficulties in its applicability.

Separate probability-magnitude model

If the injurer can influence separately the probability of an accident and the magnitude of the harm, the introduction of fixed undercompensation would create under-deterrence. However, variable undercompensation, $g(s) = \frac{g}{p(s)}$, can be introduced and induce an efficient outcome as above; the proof can be easily derived in the same way.

It is significant that undercompensation can be applied also when the harm is stochastic. Without need of going into modeling, it is clear that the effect of undercompensation is in practice one of lifting the threshold, and therefore it brings along a reduced incidence on the injurer expenses functions, as in the deterministic models.

5.4.D. Negligence (harm subsidy to non-negligent injurers)

Under a rule of negligence²⁸, the legal system sets a due level of precaution²⁹. An injurer pays damages to the victim only if he is found negligent, that is if his level of precaution was below the due level. On the contrary, a non-negligent injurer pays no damages.

Shavell (1986) only analyzed the effects of a negligence rule in probability models and found that it reduces the judgment-proof problem, compared to strict liability. Shavell attributed his findings to the “sharpness of the incentives”; under a negligence rule, negligent injurers pay $p(x)h+x$, while non-negligent injurers pay only x . This creates a discontinuity in the injurer’s expenses function.

Another explanation might be that the negligence rule provides non-negligent injurers with a harm-subsidy (equal to $p(x)h$). On the contrary, the judgment-proof problem gives a harm-subsidy to negligent injurers, which increases the more the level of precaution decreases. Under strict liability the judgment-proof subsidy only operates and reduces the

²⁸ In this section we discuss simple negligence. Contributory negligence and comparative negligence are uninteresting solutions in our setting as we assumed that only the injurer can take precaution. Hence, the victim can never be found at fault.

²⁹ We assume that the due level of precaution set by the legal system equals the optimal level of precaution.

injurers' incentives to take precaution. Under a negligence rule, the judgment-proof subsidy to negligent injurer's is contrasted by the negligence-subsidy to non-negligent injurer, which tends to rebalance the incentives toward optimal precaution.

For this reason, a negligence rule is less vulnerable to the judgment-proof problem than a strict liability rule, but is not totally immune from it: a judgment-proof problem may still undermine the injurer's incentives even under a negligence rule. Hereafter, we will give a formal interpretation of the judgment-proof problem under a negligence rule in the four models. We will show that negligence counterbalances the judgment-proof problem not only in the probability model (as noticed by Shavell, 1986), but also under the magnitude and the two mixed probability-magnitude models. We compare negligence to the other means of giving a harm-subsidy to the injurer that we have already analyzed: under-compensatory damages.

Under a negligence rule, the accident loss caused by non-negligent injurers is actually paid by victims, while injurers bear only the cost of precaution. As a consequence, the injurer's total expenses, when the injurer takes optimal precaution, are reduced to $J_n(x^*)=x^*$ ³⁰ only, while the victim pays the expected accident loss (i.e. he subsidizes the injurer).

Clearly, the injurer's total expenses under negligence are lower than the injurer's total expenses under a strict liability rule: $J_n(x^*)=x^* < p(x^*)h+x^*=J(x^*)$.³¹

In our analysis, we have shown that the decision whether or not to be potentially insolvent depends on the comparison between the injurer's total expenses given optimal precaution, $J(x^*)$ ³², and the injurer's total expenses given the inefficient level of precaution due to insolvency, $J_i(x_i)$ ³³.

The judgment-proof problem does not undermine the injurer's incentives toward optimal precaution if $J(x^*) \leq J_i(x_i)$. Under negligence, the criterion becomes $J_n(x^*) \leq J_i(x_i)$; since $J_n(x^*) < J(x^*)$, it is clear that the judgment-proof problem arises for much lower levels of the threshold than under strict liability. Taking optimal precaution becomes more convenient under negligence, because the cost of doing so is partially subsidized by the victim.

Under such an interpretation, negligence reduces the judgment-proof problem in the same way as undercompensation of the harm: the cost associated with optimal precaution is partially externalized to the victim. Nevertheless, the main difference between the two systems is that undercompensation reduces the damage compensation continuously all over the harm function for any level of precaution, while the negligence criterion creates a

³⁰ $J_n(s^*, z^*)=s^*+z^*$, in our fourth model.

³¹ The formalization refers to our first model. However, the inequality holds for all our four models. $J(x^*)=ph(x^*)+x^*$, in our second model; $J(x^*)=p(x^*)h(x^*)+x^*$, in our third model; $J(s^*, z^*)=p(s^*)h(z^*)+s^*+z^*$ in our fourth model.

³² $J(s^*, z^*)$ in our fourth model.

discontinuity: damage compensation is due entirely if the injurer is negligent ($x < x^*$) and it is completely subsidized if the injurer is non-negligent ($x \geq x^*$). This difference leads to some further considerations.

First, in the probability model undercompensation of the harm does not solve the judgment-proof problem, while negligence is also a feasible solution in such a case, as it will be clear in the next section.

Second, in the magnitude model, undercompensation is easier to apply as it requires less information. Undercompensation simply requires the damage award to be reduced by a fixed extent, without any inquiry into the level of precaution taken by the injurer, while the application of the negligence criterion triggers higher administrative costs due to such an inquiry.

Third, in the joint-probability-magnitude model and in the separate-probability-magnitude model, undercompensation is rather complex a solution, as it requires that the reduction in the damage award should be variable with the level of precaution, which seems to be a similar solution to comparative negligence. A negligence rule would trigger the same administrative costs of inquiring into parties' precaution, but offers an easier solution, as its application creates a clear cut on the injurer's precaution: either he pays damages or he does not. In the latter case, negligence saves compensation costs.

A more accurate analysis of the four models under negligence follows.

*Probability model*³⁴

Under strict liability, a judgment-proof problem reduces the injurer's incentives if $t < h$. On the contrary, if the rule in force is negligence, $t < h$ might not result in underprecaution. In fact a non-negligent injurer pays $J(x^*) = x^*$, while a negligent (judgment-proof) injurer pays $J_t(x_t) = p(x_t)t + x_t$. The injurer takes optimal precaution if $J_n(x^*) \leq J_t(x_t)$, hence we can write the minimum level for the threshold being innocuous as:

$$(32) \quad t_{\min}^n = \frac{x^* - x_t}{p(x_t)}.$$

Negligence reduces the effect of the judgment-proof problem in some cases, in which judgment proofness would have undermined the injurer's incentives under strict liability, in fact $t_{\min}^n < t_{\min} = h$, where the latter is the minimum level for the threshold being innocuous under strict liability. The proof is simple. When t is close to but smaller than h (a judgment-

³³ $J_t(s_t, 0)$ in our fourth model.

³⁴ The use of negligence as a solution to the judgment-proof problem in probability models was proposed by Shavell (1986).

proof problem arises under strict liability), also x_t is close to x^* , hence t_{\min}^n must be close to zero and, hence, smaller than t (no judgment-proof problem arises under negligence).

Figure 5 describes this result.

Magnitude model

In a magnitude model also, the injurer decides not to be potentially judgment-proof (and takes optimal precaution) if $J_n(x^*)=x^* \leq pt=J_t(x_t)$. The minimum level of the threshold can be written as:

$$(33) \quad t_{\min}^n = \frac{x^*}{p}.$$

In this case too, negligence reduces the effect of the judgment-proof problem in some cases, in which judgment proofness would have undermined the injurer's incentives under strict liability. In fact, $t_{\min}^n = \frac{x^*}{p} < h(x^*) + \frac{x^*}{p} = t_{\min}$. A threshold lower than t_{\min} , but higher than t_{\min}^n yields no precaution under strict liability, but optimal precaution under negligence.

Figure 6 depicts this result.

Joint-probability-magnitude model

In a joint-probability-magnitude model the logic is the same, the injurer decides not to be potentially judgment-proof (and takes optimal precaution) if $J_n(x^*)=x^* \leq p(x_t)t+x_t=J_t(x_t)$. The minimum level of the threshold is the same as in the probability model:

$$(34) \quad t_{\min}^n = \frac{x^* - x_t}{p(x_t)}.$$

In this case also, negligence reduces the effect of the judgment-proof problem in some cases, as $t_{\min}^n = \frac{x^* - x_t}{p(x_t)} < \frac{p(x^*)h(x^*) + x^* - x_t}{p(x_t)} = t_{\min}$.

Separate-probability-magnitude model

In a separate-probability-magnitude model, the injurer decides not to be potentially judgment-proof (and takes optimal precaution) if $J_n(s^*, z^*)=s^*+z^* \leq p(s_t)t+s_t=J_t(s_t, 0)$. The minimum level of the threshold is the same as in the probability model:

$$(35) \quad t_{\min}^n = \frac{s^* + z^* - s_t}{p(s_t)}.$$

In this case also, negligence reduces the effect of the judgment-proof problem, as

$$t_{\min}^u = \frac{s^* + z^* - s_t}{p(s_t)} < \frac{p(s^*)h(z^*) + s^* + z^* - s_t}{p(s_t)} = t_{\min}.$$

5.5. Concluding remarks

We have discussed both the theoretical and practical applicability of different solutions which can be employed within tort law in order to correct the inefficient deterrence problem due to a maximum upper threshold on liability. We have found that in general in a probability model the best solution is to make the injurer pay the average harm or to implement a negligence rule; in a magnitude model it is better to let the injurer pay less than the harm (through fixed undercompensation); in mixed models, negligence offers instead the easiest solution to implement.

5.6. References

- Beard, Rudolph T. (1990), 'Bankruptcy and Care Choice', **21** *Rand Journal of Economics*, 626 ff.
- Boyd, James and Ingberman, Daniel (1994), 'Noncompensatory Damages and Potential Insolvency', **23** *Journal of Legal Studies*, 895-910.
- Boyd, James and Ingberman, Daniel (1999), 'Do Punitive Damages Promote Deterrence?', **19** *International Review of Law and Economics*, 47-68.
- Dari Mattiacci, Giuseppe and De Geest, Gerrit (2002), 'When Will Judgment-Proof Injurers Take Too Much Precaution?', Utrecht University, Economic Institute/CIAV, *working paper*.
- Garoupa, Nuno and Gravelle, Hugh (2002), 'Optimal Deterrence with Legal Defense Expenditure', *Economic Inquiry*.
- Kaplow, Louis (1994), 'The Value of Accuracy in Adjudication: An Economic Analysis', **23** *Journal of Legal Studies*, 307-401.
- Kaplow, Louis and Shavell, Steven (1996), 'Accuracy in the Assessment of Damages', **39** *Journal of Law and Economics*, 191-209.
- Shavell, Steven (1984), 'A Model for Optimal Use of Liability and Safety Regulation', **15** *Rand Journal of Economics*, 271-280.
- Shavell, Steven (1986), 'The Judgment Proof Problem', **6** *International Review of Law and Economics*, 45-58.
- Summers, John (1983), 'The Case of the Disappearing Defendant: An Economic Analysis', **132** *Pennsylvania Law Review*, 145-185.

Figures

The dashed lines describe the injurer's expenses when different from the social cost.

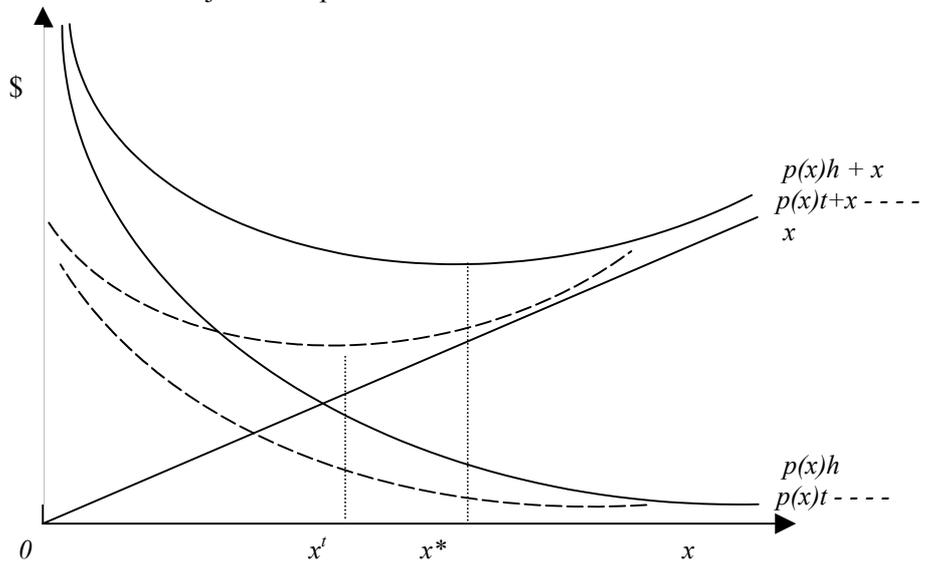


Figure 1: probability model (the injurer can reduce only the probability of an accident)

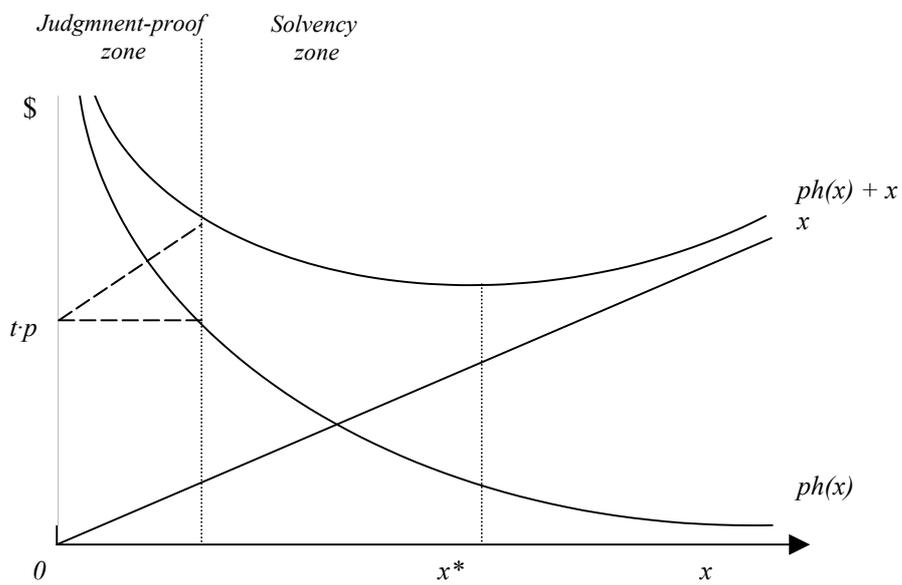


Figure 2: magnitude model (the injurer can reduce only the magnitude of the harm)

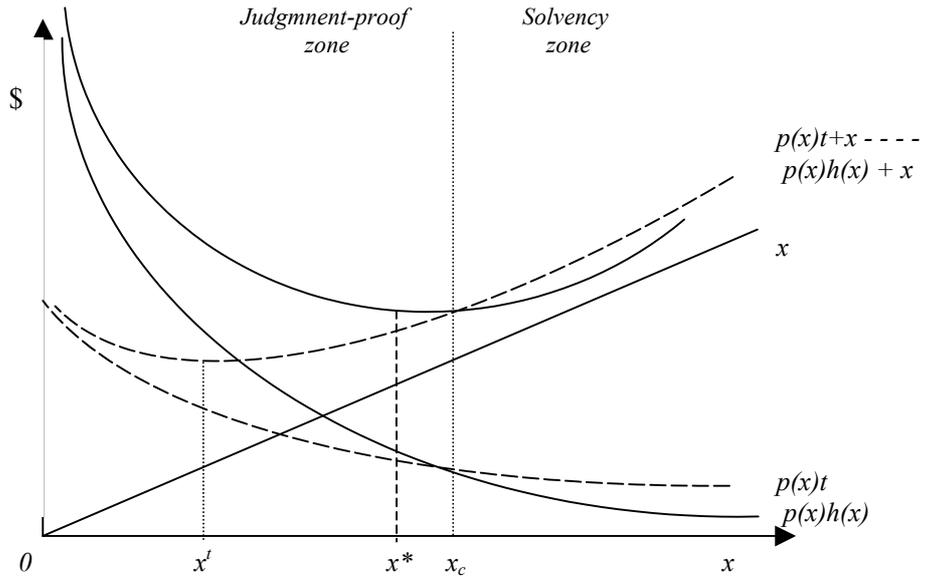


Figure 3: joint probability-magnitude model (the injurer can reduce jointly both the probability and the magnitude of the harm)
If $t < h(x^*)$

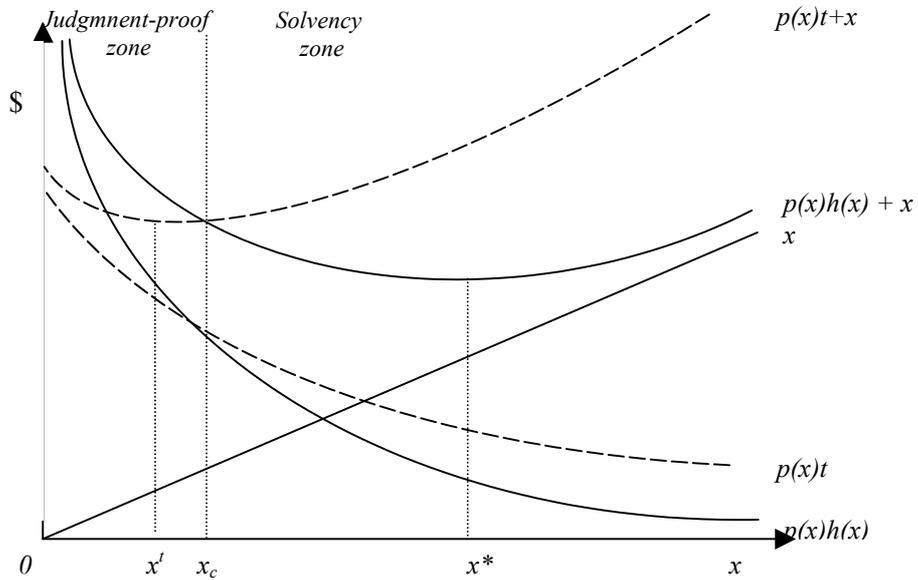


Figure 4: joint probability-magnitude model (the injurer can reduce jointly both the probability and the magnitude of the harm)
If $t \geq h(x^*)$

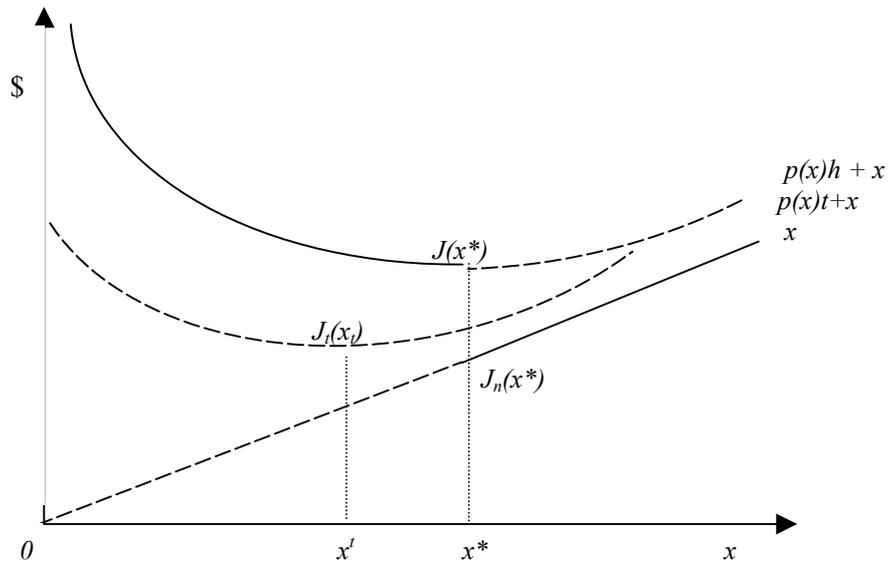


Figure 5: probability model (the injurer can reduce only the probability of an accident) under negligence (standard of negligence equal to the optimal level of precaution)

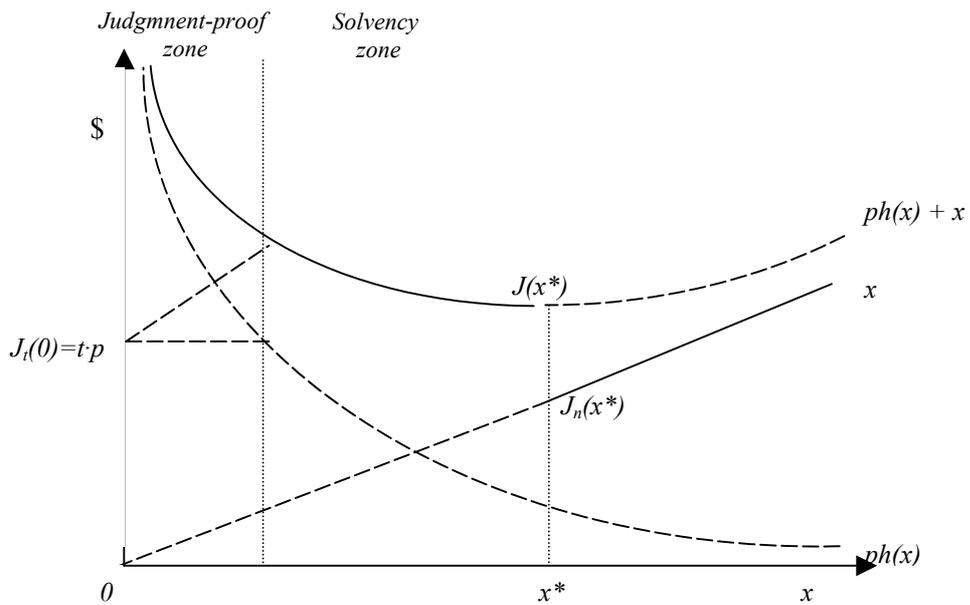


Figure 6: magnitude model (the injurer can reduce only the magnitude of the harm) under negligence (standard of negligence equal to the optimal level of precaution)

ON THE COMBINED USE OF ANTI-JUDGMENT-PROOF REGULATION AND TORT LAW*

Gerrit De Geest
Utrecht University and Ghent University

Giuseppe Dari Mattiacci
Utrecht University

ABSTRACT

Tort law and regulations are two alternative means of controlling externalities. Nevertheless, tort liability is often combined with minimum regulatory standards. The reason for their joint use is that the incentives created by tort law might be diluted by the judgment-proof problem, while regulators might lack appropriate information on risk. Thus, they may be used to complement each other. We demonstrate that minimum regulatory standards, which are insufficient to give injurers an incentive to take optimal precaution, might be sufficient to completely remove the effect of judgment-proofness. Once the judgment-proof problem has been corrected, tort law can provide incentives toward optimal precaution, and hence remedy the information deficit of regulatory bodies.

JEL classification: K13, K32.

Keywords: insolvency, judgment-proof problem, liability, bankruptcy, overprecaution, regulation.

* Gerrit De Geest, Economic Institute/CIAV, Utrecht University and School of Law, Ghent University; email: g.degeest@econ.uu.nl. Web: www.gerritdegeest.net. Giuseppe Dari Mattiacci, Economic Institute/CIAV, Utrecht University, Kromme Nieuwegracht 22, 3512 HH Utrecht, the Netherlands; email: g.darimattiacci@econ.uu.nl. Updated versions of this chapter will possibly be available on my web page: www.law.uu.nl/eco/gdm.htm.

6.1. Introduction

Tort law and ex ante regulation are two means of controlling negative externalities. Motorists may be provided with incentives to drive at a reasonable speed by letting them bear the accident loss (tort law) or by organizing speed controls and sanctioning those who violate the limit (regulation).

In spite of the fact that tort law and regulation are substitutes, legal systems often make joint use of them. Such a combination is at first sight puzzling (and still insufficiently explained in the literature). If tort law is already employed to induce parties to take precaution, then why should regulation be implemented for the same purpose? Similarly, if incentives are provided by ex ante regulation, why do legal systems maintain tort liability in place?

In the legal literature there also seems to be concern about how the courts should set the level of due care under fault-based liability when a regulatory standard is already enforced by an administrative body. If the injurer has violated the standard (e.g. the speed limit) should he also be found automatically (*per se*) negligent? If the injurer has complied with the regulatory standard, should he then be relieved of liability (compliance defense)?¹.

To illustrate, suppose that the regulatory standard is 40. An accident has occurred and the injurer's precaution level was 39, thus he might be found negligent under the *per se* rule. Nevertheless, even if his precaution level had been 40, he might still be found negligent, as the courts might require a higher due care level, for instance 50, i.e. might not accept a compliance defense. The fact that regulatory standards are often set at a suboptimal level raises two additional questions. If regulation is deployed to control injurer's precaution, why is not the regulatory standard set at the socially optimal level? Moreover, if the regulatory standard is so low that it has to be corrected by tort law, what is the advantage of laying down regulation in the first place?

This chapter addresses the former questions and identifies a number of cases in which the combined use of regulation and tort law may be desirable and in which it may be advantageous to set the regulatory standards below the optimal level of due care. First, we will furnish with a justification of why tort law and regulation might not be effective if used alone. Then we will inquire into their joint use.

6.1.A. A reason why tort law and regulation might fail if employed alone.

Tort law is an effective system with respect to the acquisition of information about risk, as it

operates *ex post*, after the accident has occurred. However, it might fail in those cases in which injurers are potentially judgment proof, for they have limited assets or because the law sets a limit on the maximum amount of damages the victim is entitled to recover in the case of an accident (liability cap). Judgment-proof injurers might take an inefficient level of precaution for they internalize only a portion of the accident loss. Thus, tort law might result in suboptimal accident prevention.

Regulation does not suffer from the judgment-proof problem, because regulatory standards are enforced *ex ante*, before and irrespective of the actual occurrence of an accident, hence they also induce compliance in relation to those injurers who might be unable to compensate the harm fully. Nevertheless, regulators might experience difficulty in setting standards at the socially optimal level of precaution, because they have *ex ante* imperfect information about risk².

6.1.B. How to combine regulation and tort law: regulation of minimum levels of precaution

We have provided the reader with a reason why tort law and regulation might not result in the socially optimal accident prevention if used alone. We will hence illustrate how they can complement each other in order to attain a socially optimal outcome. We will consider strict liability as the liability rule in force.

The kernel of the argument to be made resides in the fact that in most situations injurers can reduce not only the probability of the accident but also the magnitude of the harm by taking precaution. Therefore, they will be insolvent only if their level of precaution is so low that the harm might exceed their assets (or the liability cap). In essence, by controlling the magnitude of the harm through their precaution, injurers may decide whether or not to be judgment-proof. They might find it advantageous to be insolvent because in the judgment-proof zone (low levels of precaution) they can reduce their precaution to very low levels (even zero) without having to pay for the increment in the harm, as their liability only equals their assets. The greater the portion of the harm that remains unpaid, the greater the advantage of being judgment-proof.

Consequently, we will demonstrate that preventing injurers from taking the very low levels of precaution by means of minimum regulatory standards is often sufficient to remove the advantage of being insolvent. Hence, minimum regulatory standards are sufficient to deliver tort law from the judgment-proof problem. The reason why they should not be set too high is that regulatory agencies might not guess the optimal level of precaution and overprecaution may result (given the *ex ante* information deficit of regulators).

¹ See Burrows (1999).

Once the judgment-proof problem has been corrected by regulation, tort law can operate efficiently and exploit its informative advantage by providing solvent injurers with appropriate incentives to take the optimal (higher than the minimum regulatory standard) level of precaution. The numerical example of *table 1* might be helpful.

Consider a potential injurer who is only able to pay damages up to \$79,000; the probability that an accident occurs is 10%, while the harm depends on the level of precaution. If he takes a level of precaution equal to or lower than 3, the harm will exceed \$79,000, and hence he will be insolvent. On the contrary if he takes the optimal level of precaution, 4*, or a higher level of precaution, the harm will be lower than \$79,000, and therefore he will be solvent. Although the injurer would be solvent if he took optimal precaution (4*), he finds it advantageous to take low precaution, as his total costs are lower in the judgment-proof zone than in the solvency zone. Moreover, the table shows that in the judgment-proof zone, the total injurer's costs decrease if precaution decreases. Therefore, the injurer will decide to be insolvent and will take a level of precaution equal to 0, rather dramatic an outcome (as indicated in the second but last column).

Table 1. Anti-judgment-proof regulation: the minimum regulatory standard is set at 2^R, while the optimal level of precaution is equal to 4. Bold numbers indicate minimum values*

	Precaution level and precaution cost	Harm	Exp. harm (10%)	Total social costs (precaution plus exp. harm)	Injurer's exp. liability (10%)	Injurer's total costs (precaution plus exp. liability)	Injurer's total costs with regulation
Judgment-proof zone (pay \$79,000)	$0 = 0$	150,000	15,000	15,000	7,900	7,900	--
	$1 = 1,000$	100,000	10,000	11,000	7,900	8,900	--
	$2^R = 2,000$	80,000	8,000	10,000	7,900	9,900	9,900 ^(R)
Solvency zone (pay the harm)	$3 = 3,000$	65,000	6,500	9,500	6,500	9,500	9,500
	$4^* = 4,000$	52,000	5,200	9,200	5,200	9,200	9,200 ^(*)
	$5 = 5,000$	45,000	4,500	9,500	4,500	9,500	9,500

Consider now the introduction of anti-judgment-proof regulation as proposed above. A minimum regulatory standard equal to 2^R is laid down and enforced, so that the injurer, in order to avoid the regulatory sanction is induced to take at least a level of precaution equal to 2^R (as indicated in the last column). It is useful to observe that the regulatory standard is lower than the optimal level of precaution and it lies still within the judgment-proof zone. Nevertheless, once the injurer is forced to take at least precaution 2^R, the advantage of being judgment-proof is diluted, as at 2^R the total (insolvent) injurer's costs are \$9,900 and are higher than any value in the solvency zone, for instance \$9,500 at 3. Consequently, such a minimum standard is sufficient to remove the judgment-proof problem, as it keeps the injurer out of the judgment-proof zone. At this stage, tort law is enabled to provide the injurer with appropriate incentives to take optimal precaution. In fact, in the solvency zone, the normal

² See Shavell (1984a and 1984b).

functioning of tort law (i.e. the injurer pays the accident loss) induces the injurer to take optimal precaution 4^* , which minimizes both the social costs and the injurer's total costs (last column).

The former example should have clarified the reason to implement only minimum regulatory standard. However, such a solution is subject to two conditions. First, the injurer must be able to reduce the magnitude of the harm by taking precaution. If the magnitude of the harm is exogenous, the injurer finds himself either solvent or insolvent and has no influence on his judgment-proofness. In such cases, minimum regulation will not remove the judgment-proof problem, and a regulatory standard equal to the optimal level of precaution will be necessary in order to induce optimal accident prevention.

Second, the injurer must be solvent at the optimal level of precaution. If he is insolvent at the optimal level of precaution (he is still in the judgment-proof zone) the regulatory standard will not have the desired effect. Again, a regulatory standard equal to the optimal level of precaution will be necessary.

6.1.C. How to combine regulation and tort law: regulation of magnitude reducing precaution

We will now illustrate our second point. Selective regulatory standards that only focus on magnitude reducing precaution will often be sufficient to correct the judgment-proof problem, hence agencies might be dispensed with the need of regulating probability reducing precaution. The reason is that, when the injurer is induced to take the optimal level of magnitude precaution, he internalizes the same magnitude of the harm, as he would bear in the absence of a judgment-proof problem. Therefore, the injurer will also choose the optimal level of the probability precaution, for this is the level of precaution that minimizes his total costs. *Table 2* clarifies this issue.

Consider an injurer who can take two different precautionary measures in order to reduce the expected accident loss. One precautionary measure reduces the probability of the accident (e.g. radars reduce the probability of shipwrecks) and can be taken at three different levels, another precautionary measure reduces the magnitude of the harm (e.g. lifeboats reduce the harm to passengers in the case of a shipwreck) and can be taken at three different levels as well. The social costs are the sum of the magnitude precaution cost, the probability precaution cost and the expected harm; for example \$11,000 (the minimum level of social costs) is given by 3,000 (the cost of the magnitude precaution 3) + 1,500 (the cost of the probability precaution b) + 6,500 (the expected harm if precautions 3 and b are taken).

The injurer is potentially judgment proof as he can only pay damages up to \$85,000. Therefore, the injurer is insolvent only if he takes the magnitude precaution at level 1 . The injurer's expected liability and the injurer's total costs when he is insolvent are indicated in

brackets. From the numerical example in *table 2* it can be observed that the injurer will decide to be insolvent and will take the magnitude precaution at level *1* (instead of 3) and the probability precaution at level *a* (instead of *b*).

It is interesting to notice that the decision to be insolvent always implies a lower than optimal level of magnitude precaution. However, with respect to the probability precaution the result may vary. In our example the injurer takes an excessive level of probability precaution. The reason is that when he is bankrupt he actually pays higher damages (\$85,000) than he would pay if he took the optimal level of magnitude precaution (\$65,000). Nevertheless, his total costs are lower (\$10,800 instead of \$11,000), and hence he decides to be bankrupt. Examples can be easily constructed yielding lower than optimal or optimal probability precaution³.

Table 2. Anti-judgment-proof regulation: the separate probability - magnitude model and selective regulatory standard. Bold numbers indicate minimum values

	Precaution level and precaution cost	Injurer's liability	Exp. liability (the harm times the probability)			Total social (injurer's) costs: magnitude precaution plus probability precaution plus exp. harm (liability)		
			<i>a</i> =8% 3,000	<i>b</i> =10% 1,500	<i>c</i> =14% 1,000	<i>a</i>	<i>b</i>	<i>c</i>
Judgment-proof zone	<i>1</i> = 1,000	100,000 (85,000)	8,000 (6,800)	10,000 (8,500)	14,000 (11,900)	12,000 (10,800)	12,500 (11,000)	16,000 (13,900)
Solvency zone	<i>2^R</i> = 2,000	80,000	6,400	8,000	11,200	11,400	11,500	14,200
	<i>3*</i> = 3,000	65,000	5,200	6,500	9,100	11,200	11,000^(*)	13,100

Consider a regulatory standard set at a suboptimal level and only concerning the magnitude precaution, *2^R*. The injurer is induced to take at least that level of magnitude precaution which is enough to take him out of the judgment-proof zone as before. Again the normal functioning of tort law induces him to take both the optimal magnitude precaution and the optimal probability precaution, simply because once out of the judgment-proof zone, the injurer internalizes the full harm.

The condition is the same as before: regulation of magnitude precautions can be implemented only if the injurer is solvent at the optimal level of precaution.

6.1.D. When regulatory compliance should not relieve of liability for negligence

Negligence rules, such a simple negligence, reduce the incidence of judgment proofness. Since a non-negligent injurer never pays damage compensation and only bears the cost of precaution, the costs associated with the solvency zone decrease. Therefore, the judgment-

³ It should be noticed that, since regulation can correct the judgment-proof problem only if the injurer's assets or the liability cap is higher than the harm resulting at the optimal level of precaution, anti-judgment-proof regulation applies only to those situations in which the insolvent injurer pays higher damages than at the optimal level of precaution, thus he takes higher probability precaution than optimal.

proof problem arises less often and only if the limit on the injurer's liability is particularly low. Nevertheless, even under simple negligence injurers might find it advantageous to be insolvent.

Anti-judgment-proof regulation can be implemented in combination with a simple negligence rule in the same way as with strict liability. Moreover, when a negligence criterion is applied the question arises of whether or not regulatory compliance should be accepted as a negligence defense. As we have illustrated above (and we will demonstrate in the next several sections) that a regulatory standard can be efficiently set at a suboptimal level, we are now furnished with a reason why, whenever anti-judgment-proof regulation is employed, regulatory compliance should not relieve injurers of liability for negligence.

As the only function of the anti-judgment-proof regulatory standard is to deliver tort law from the judgment-proof problem and the function of due care levels is to provide incentives toward optimal precaution, then both should be enforced; if the former are not enforced tort law incentives are undermined by the judgment-proof problem, if the latter are not employed suboptimal precaution might result.

6.1.E. Structure of this chapter

After providing with a short note on the literature (section 6.1.F), in section 6.2 we will present the reader with the basic magnitude model in which the injurer can reduce only the magnitude of the harm, while the probability is exogenous⁴. We will always discuss two variants of the models that we will propose: the one-pocket variant (in which the expenditure on precaution reduces the amount of money available for compensation, as in the case of injurer's limited assets) and the two-pocket variant (in which the expenditure on precaution does not affect the limit on injurer's liability, as in the case of liability caps). We will define the judgment-proof problem in a formal way under strict liability and under simple negligence.

In section 6.3 we will analyze anti-judgment-proof regulation in the magnitude model and prove that a suboptimal regulatory standard might be sufficient to remove the judgment-proof problem. In section 6.4 we will analyze the separate-magnitude-probability model (in which the injurer can reduce the probability and the magnitude by taking two different precautionary measures)⁵ and prove that focusing regulation only on magnitude precaution might be sufficient. In section 6.5 we will discuss our results in probability models (the

⁴ Many safety measures in workplaces (helmets for example) are designed to reduce the harm to workers in the case of an accident; in nuisance cases, parties usually affect the magnitude and not the probability of the harm.

⁵ This is for example the case discussed in the text in section 6.1.C: radars reduce the probability of a shipwreck while lifeboats placed on the deck reduce the magnitude of the harm to passengers.

injurer can reduce only the probability of an accident)⁶, and joint-magnitude-probability models (the same precautionary measure reduces both the probability of the accident and the magnitude of the harm)⁷. Section 6.6 will finally provide with some concluding remarks.

6.1.F. A note on the literature

The literature has analyzed regulation and tort law mainly as alternative ways of controlling externalities. The only formal analyses of these two means as complements are provided by Shavell (1984b), Kolstad, Ulen and Johnson (1990) and Burrows (1999), and are based on the fact that neither instrument might be sufficient alone and the combination of the two might yield a second best solution. We attempt to prove that a first best outcome is possible.

Shavell (1984b) is the closest approach to ours, as he discusses the joint use of regulation and tort law in the presence of judgment-proofness (as a limit for tort law) and limited information on risk (as a limit for regulation). He finds that regulation can be used to correct partially the failure of tort law due to insolvency.

Our analysis is different in two respects: we consider not only probability models but also three other models in which the injurer can influence the magnitude of the accident. Our findings apply in the latter situations. In a pure probability model the injurer finds himself either solvent or insolvent, since both the magnitude of the harm and the limit on his liability are exogenous; therefore, regulation cannot lead him out of the judgment-proof zone either.

In addition to that, we consider also one-pocket variants of those models (when the expenditure on precaution reduces the injurer's assets) and discuss our results with respect to them. Under certain conditions, anti-judgment-proof regulation applied in magnitude models removes the judgment-proof problem and therefore enables tort law to induce a first best level of precaution.

In addition to that, Shavell (1984a) considers also the problem of disappearing defendant (where there is a certain probability that the injurer will not be sued) as a cause for the dilution of incentives under tort law. Nevertheless, the case of disappearing defendants is not a judgment-proof problem, as it does not generate a maximum upper threshold on the amount of damages the injurer has to pay in the case of an accident, but simply dilutes the probability that the injurer will actually pay damages. Under a probability model, the two problems have the same effect on injurer's incentives, for both reduce the portion of the harm that the injurer internalizes.

However, under the other three models that we discuss in this chapter, where the injurer

⁶ In the case of aircraft accident, for example, more precaution reduces the probability of an accident, although the magnitude of the harm is more or less exogenous.

⁷ This is the case of precaution while driving. A lower speed reduces both the probability of an accident and the

can reduce also (or only) the magnitude of the accident, the case of disappearing defendants is different from the judgment-proof problem. In the latter, the injurer pays the harm, $h(x)$, if this is lower than or equal to the threshold on his liability, $h(x) \leq t$ (precaution, x , sufficiently high), and pays the threshold, t , otherwise: the threshold cuts the harm function: the portion of the harm that exceeds the threshold is completely externalized. On the contrary, if there is a probability a that the injurer will not be sued, he internalizes an accident loss lower than the actual harm, $ah(x)$, for any level of x . Hence, the case of disappearing defendants cannot be considered as similar to a judgment-proof problem in magnitude models.

Our analysis builds upon previous contributions. For a complete discussion of the four two-pocket models see Dari Mattiacci and De Geest (2001)⁸. For an analysis of the four one-pocket models see Dari Mattiacci and De Geest (2002)⁹. The differences between a one-pocket approach and a two-pocket approach were first underlined by Beard (1990), who only focused on probability models.

Kolstad, Ulen and Johnson (1990) discuss the combined use of regulation and tort law in a different context. Tort law is said to generate incentive problems because of uncertainty in the application of the negligence rule. Regulation is used to repair such shortcoming.

Burrows (1999) focuses his analysis on instrumental uncertainty: a situation in which the injurer does not know in advance the severity of either or both the negligence rule and the regulatory standard.

6.2. The one-pocket and the two-pocket magnitude models

We make the following assumptions:

1. Unilateral precaution situation, only the injurer can take precaution, only the victim suffers harm;
2. Parties are perfectly informed and the courts entitle victims to perfect compensation;
3. Parties are risk-neutral, rational and utility maximizing;
4. The rule in force is alternatively strict liability or simple negligence.

Let:

x = the injurer's precaution cost;

p = the probability of an accident occurring (exogenous), $0 < p < 1$;

$h(x)$ = the magnitude of the harm (i.e. the accident loss), $h > 0$, $h' < 0$, $h'' > 0$;

magnitude of the harm.

⁸ See chapter 5.

⁹ See chapter 7.

t = a maximum upper threshold on injurer's liability (injurer's assets or liability cap);

$J(x)$ = the injurer's expected total expenditure;

$SC(x)$ = the social cost

The socially optimal level of precaution is the level of precaution that minimizes the sum of the precaution cost and the expected harm.

The social cost function is:

$$(1) \quad SC(x) = ph(x) + x.$$

We denote as x^* the level of precaution that minimizes Exp. (1). The assumptions made guarantee convexity.

6.2.A. Judgment-proofness under strict liability

The injurer seeks to minimize his total expenditure, which is the sum of his precaution cost x and his expected liability expenses in the case of an accident. We first consider the case of an injurer being judgment-proof because the magnitude of the harm, $h(x)$, might be greater than his total assets, t . The amount of money spent in taking precaution reduces the injurer's asset; therefore, after taking precaution, the total assets available for paying damages are $t-x$. If the harm $h(x)$ is greater than the residual assets, the injurer will not be able to compensate the victim fully for the harm and will pay only $t-x$.

Therefore, the injurer's expenditure function under strict liability is¹⁰:

$$(2) \quad \begin{cases} J(x) = ph(x) + x & \text{if } h(x) + x \leq t \\ J_{t1}(x) = pt + (1-p)x & \text{if } h(x) + x > t \end{cases} \quad [\text{One-pocket magnitude model}].$$

We call this model one-pocket model, as the injurer has only one pocket from which he pays both the precaution expenditure and damage compensation.

Secondly, we consider also those situations, in which the judgment-proof problem is due to a liability cap (a legal upper limit on the amount of damages that the victim is entitled to recover) or to limited injurer's assets when precaution is non-monetary. In these cases, neither does the expenditure on precaution reduce the assets available for damage compensation nor it affects the legal limitation on them. If the harm, $h(x)$, is greater than the threshold, t , the injurer pays t to the victim.

Therefore, the injurer's expenditure function under strict liability is:

¹⁰ Note that $p(t-x)+x$ can be rewritten as $pt+(1-p)x$.

$$(3) \quad \begin{cases} J(x) = ph(x) + x & \text{if } h(x) \leq t \\ J_{i2}(x) = pt + x & \text{if } h(x) > t \end{cases} \quad [\text{Two-pocket magnitude model}].$$

We call this model two-pocket model, as the injurer behaves as if he had two pockets: one to pay his precaution expenditures and another one to pay damage compensation.

$J(x)$ is clearly minimized by x^* , in both Exp. (2) and Exp. (3). $J_{i1}(x)$ in Exp. (2) and $J_{i2}(x)$ in Exp. (3) are both minimized by $x=0$, as they are linear in x . Note that $J_{i1}(0)=J_{i2}(0)=pt$.

The injurer will choose to take no precaution if his total expenditure at $x=0$ is lower than his total expenditure at x^* , that is if $J_{i1}(0)=J_{i2}(0)<J(x^*)$. He will choose x^* , otherwise. This condition can be rewritten as:

$$(4) \quad t < h(x^*) + \frac{x^*}{p} \quad [\text{Judgment-proofness under strict liability}].$$

The solution obtained by applying Exp. (4) always satisfies the conditions imposed by Exp. (2): $h(0)>t$, if $x=0$ is the solution, and $h(x^*)+x^*/p \leq t$, if x^* is the solution. It satisfies the conditions imposed by Exp. (3) also: $h(0)>t$, if $x=0$ is the solution, and $h(x^*) \leq t$, if x^* is the solution¹¹. Exp. (4) constitutes a condition for the judgment-proof problem to be relevant: if the level of the threshold is high enough it does not affect the precaution decisions of the injurer, both in the one-pocket and in the two-pocket model. Therefore, it is of our concern in this chapter to analyze situations in which the threshold is too low (as in Exp. (4)) and anti-judgment-proof regulation might be socially advantageous.

6.2.B. Judgment-proofness under simple negligence

Under simple negligence, if the injurer takes due care, he does not pay damage compensation to the victim and bears only the precaution cost, x . Let us assume that the due-care level is equal to the optimal level of precaution, x^* . $J(x)$ becomes:

$$(5) \quad J(x) = \begin{cases} ph(x) + x & \text{if } x < x^* \\ x & \text{if } x \geq x^* \end{cases}.$$

It can be easily shown that the injurer will take x^* ¹²; hence, $J(x^*)=x^*$. The condition in Exp. (4) becomes:

¹¹ If $t \geq h(x^*) + x^*/p$, then x^* is the solution and both $h(x^*) + x^*/p \leq t$ in Exp. (2) and $h(x^*) \leq t$ in Exp. (3) are satisfied, as $0 < p < 1$. If $t < h(x^*) + x^*/p$, then $x=0$ is the solution and $h(0) > t$ in Exp. (2) and in Exp. (3) is satisfied, as $h(x^*) + x^*/p > t$ can be rewritten as $ph(x^*) + x^* > pt$; given the definition of x^* , we can write $ph(0) + 0 > ph(x^*) + x^* > pt$, which yields $h(0) > t$.

¹² The first expression in (5) is minimized by x^* , the second expression is linear in x and has its minimum at x^* , hence the injurer will choose x^* and bear only the cost of precaution.

$$(6) \quad t < \frac{x^*}{p} \quad [\text{Judgment-proofness under simple negligence}]^{13}.$$

If the threshold is too low, a judgment-proof problem arises in both the one-pocket and the two-pocket model under simple negligence. Note that the level of the threshold at which a judgment-proof problem arises is higher under simple negligence than under strict liability¹⁴.

6.3. Anti-judgment-proof regulation

If the liability rule in force is an ineffective means for inducing optimal precaution because of the judgment-proof problem, regulation can be introduced in order to remove the effect of judgment-proofness on injurer's incentives.

Let:

x_r = the level of precaution required by the regulatory standard.

We assume certainty in the enforcement of the regulatory standard: if a standard x_r is set by the regulatory authority, the injurer will take at least x_r .

By setting x_r , the regulatory authority prevents the injurer from taking less precaution than x_r . Hereafter we prove that the optimal regulatory standard can be set at a lower level than the optimal level of precaution, $x_r \leq x^*$. Nevertheless, the optimal level of precaution will result.

In fact, as a potentially judgment-proof injurer is induced to take at least $x_r > 0$, his total expenditure increases: $J_{II}(x_r) > J_{II}(0)$, as $J_{II}(x)$ is linear in x . A regulatory standard, x_r , such that it makes injurer's total expenditure in the case of judgment-proofness higher than the total injurer's expenditure when he takes optimal precaution, $J_{II}(x_r) \geq J(x^*)$ ¹⁵, is sufficient to induce the injurer to take optimal precaution, as stated in proposition 1.

Proposition 1. If the injurer is not bankrupt at the optimal level of precaution: (i) a regulatory standard lower than the optimal level of precaution is sufficient to remove the judgment-proof problem; (ii) once regulation has removed the judgment-proof problem, tort law

¹³ Under simple negligence, the model in Exp. (2) changes slightly. As a non-negligent injurer never pays damages, if the injurer takes due care, the first condition $h(x)+x \leq t$ becomes $x^* \leq t$. If $t \geq x^*/p$, the solution is x^* and the first condition in Exp. (2) is satisfied, as $0 < p < 1$. If $t < x^*/p$, the solution is $x=0$, and the second condition in (2) is satisfied as $t < x^*/p$ can be rewritten as $x^* > pt$; it follows that $ph(x^*)+x^* > pt$, as $ph(x^*) > 0$, and $ph(0)+0 > ph(x^*)+x^* > pt$, by definition of x^* . Therefore, $h(0) > t$. The model in Exp. (3) changes in a similar way. As a non-negligent injurer never pays damages, if $t \geq x^*/p$, the solution is x^* , and the first condition is always satisfied. If $t < x^*/p$, the solution is $x=0$, and the second condition is always satisfied, the proof being the same as the one just described for Exp. (2).

¹⁴ See Shavell (1986) and Dari Mattiacci and De Geest (2001) on this point.

induces optimal precaution.

The next 2 sections will provide with a formal proof of proposition 1 under strict liability and simple negligence.

6.3.A. Anti-judgment-proof regulation under strict liability

Under strict liability, in the one-pocket model $J_{11}(x)=pt+(1-p)x$. A regulatory standard, x_r , may be introduced such that $J_{11}(x_r)\geq J(x^*)$ ¹⁶. In the latter case, the condition in Exp. (4) is not satisfied any more and optimal precaution x^* results. x_r must satisfy $pt+(1-p)x_r\geq ph(x^*)+x^*$. In addition to that, x_r must not be greater than x^* , in order not to induce overprecaution. These conditions can be rewritten as

$$(7) \quad x^* - \frac{p}{1-p} [t - h(x^*) - x^*] \leq x_r \leq x^*$$

[Regulation under strict liability in the one-pocket magnitude model].

The lower limit is smaller than x^* if and only if $t > h(x^*) + x^*$, i.e. if and only if the injurer is not bankrupt at x^* . In such case, any regulatory standard set between these two limits prevents the injurer from deciding to be potentially bankrupt by taking too low a level of precaution. Given such an injurer's decision, tort law is delivered from the judgment-proof problem and enhances optimal precaution x^* .

Under strict liability, in the two-pocket model $J_{12}(x)=pt+x$. A regulatory standard, x_r , may be introduced such that $J_{12}(x_r)\geq J(x^*)$. x_r must satisfy $pt+x_r\geq ph(x^*)+x^*$ and must not be greater than x^* , in order not to induce overprecaution. These conditions can be rewritten as

$$(8) \quad x^* - p[t - h(x^*)] \leq x_r \leq x^*$$

[Regulation under strict liability in the two-pocket magnitude model].

See figure 1.

The lower limit is smaller than x^* if and only if $t > h(x^*)$, i.e. if and only if the injurer is not bankrupt at x^* . Otherwise $x_r = x^*$ is optimal.

Therefore, when the injurer is not bankrupt at the optimal level of precaution, a regulatory standard lower than the optimal level of precaution may remove the judgment-proof problem, both in the one-pocket and in the two-pocket magnitude model. It can be easily shown that the lower limit in Exp. (7) is higher than the lower limit in Exp. (8)¹⁷.

¹⁵ The expression refers to the one-pocket model, but the reasoning is the same for the two-pocket model as well.

¹⁶ We assume that, if $J_{11}(x_r)=J(x^*)$, the injurer chooses optimal precaution x^* . The same will be assumed for the two-pocket model under strict liability and for the one-pocket and two-pocket models in the case of simple negligence.

¹⁷ $x^* - p[t - h(x^*) - x^*] / (1-p) > x^* - p[t - h(x^*)]$ can be rewritten as $ph(x^*) + x^* > pt$, which is always verified in the case

Hence, anti-judgment-proof regulatory standards might be higher in one-pocket situations.

On the contrary, if the injurer is bankrupt at the optimal level of precaution, $t < h(x^*) + x^*$ (one-pocket model) or $t < h(x^*)$ (two-pocket model), a regulatory standard equal to the optimal level of precaution, $x_x = x^*$, is optimal. In such case, regulation as the only means of reducing risk is optimal.

6.3.B. Anti-judgment-proof regulation under simple negligence

Under simple negligence, $J(x^*) = x^*$. In the one-pocket model, $J_{i1}(x) = pt + (1-p)x$. A regulatory standard, x_r , must satisfy

$$(9) \quad x^* - \frac{p}{1-p} [t - x^*] \leq x_r \leq x^*$$

[Regulation under simple negligence in the one-pocket magnitude model].

The lower limit is smaller than x^* if and only if $t > x^*$, i.e. if and only if the injurer is not bankrupt at x^* .

Under simple negligence, in the two-pocket model $J_{i2}(x) = pt + x$. A regulatory standard, x_r , may be introduced such that $J_{i2}(x) \geq J(x^*)$. x_r must satisfy

$$(10) \quad x^* - pt \leq x_r \leq x^*$$

[Regulation under simple negligence in the two-pocket magnitude model].

See figure 2.

The lower limit is always smaller than x^* ; hence, a regulatory standard lower than the optimal level of precaution may always solve the judgment-proof problem; therefore, both the regulatory standard and the due care standard ought to be separately enforced, which proves proposition 2.

Similarly to what we have previously noted, regulation might set an higher standard in the one-pocket case¹⁸.

Proposition 2. Compliance with anti-judgment-proof regulation should not relieve the injurer of liability for negligence.

A lower than optimal regulatory standard might be enough to remove the judgment-proof problem but tort law is needed in order to provide optimal incentive. Both the regulatory standard and the negligence standard have to be enforced.

of judgment-proofness as Exp. (4) must be satisfied.

¹⁸ $x^* - p[t - x^*]/(1-p) > x^* - pt$ can be rewritten as $x^* > pt$, which is satisfied in the case of judgment-proofness as Exp. (6) must be satisfied.

6.3.C. Imperfectly informed regulators

The foregoing analysis shows that if the injurer is not bankrupt at the optimal level of precaution, tort law can be complemented by a regulatory standard, which can be set at a lower level than the optimal level of precaution, but at a higher level than the minimum indicated in the analysis, Exp. (7) to (10). Therefore, regulators do not need to have perfect information on the optimal level of precaution and on the actual level of the threshold, as any standard set between those two limits will enable tort law to induce optimal precaution.

As injurers might have different assets the minimal level of the regulatory standard might change among them. Therefore, the regulatory standard should be set as close as possible to the optimal level of precaution in order to correct judgment-proofness in relation to as many injurers as possible.

6.4. The separate probability-magnitude model

This section considers those situations, in which the injurer can take one precautionary measure in order to reduce the magnitude of the accident and another precautionary measure (independent from the first) in order to reduce the probability of the accident occurring.

Let:

- s = the injurer's probability-precaution cost;
- z = the injurer's magnitude-precaution cost;
- $p(s)$ = the probability of an accident occurring, $0 < p < 1$, $p' < 0$, $p'' > 0$;
- $h(z)$ = the magnitude of the harm (i.e. the accident loss), $h > 0$, $h' < 0$, $h'' > 0$;
- z_r = the level of magnitude precaution required by the regulatory standard

The social cost function is:

$$(11) \quad SC(s, z) = p(s)h(z) + s + z .$$

We denote as (s^*, z^*) the level of precaution that minimizes Exp. (11). The assumptions made guarantee convexity. In the presence of a threshold, t , the one-pocket separate-probability-magnitude model is:

$$(12) \quad \begin{cases} J(s, z) = p(s)h(z) + s + z & \text{if } h(z) + s + z \leq t \\ J_{11}(s, z) = p(s)t + (1 - p(s))s + (1 - p(s))z & \text{if } h(z) + s + z > t \end{cases}$$

[One-pocket separate-probability-magnitude model].

The two-pocket separate-probability-magnitude model is:

$$(13) \quad \begin{cases} J(s, z) = p(s)h(z) + s + z & \text{if } h(z) \leq t \\ J_{12}(s, z) = p(s)t + s + z & \text{if } h(z) > t \end{cases}$$

[Two-pocket separate-probability-magnitude model].

Our concern here is only to show that regulation focused only on the magnitude precaution might enhance optimal precaution both with respect to the magnitude precaution and to the probability precaution¹⁹.

Proposition 3. If the injurer is not bankrupt at the optimal level of precaution: a regulatory standard focused only on the magnitude precaution induces optimal precaution with respect to both the magnitude and the probability precaution.

In the one-pocket model, $J(s, z)$ is minimized by (s^*, z^*) , the first order condition being:

$$(14) \quad \begin{cases} \frac{\partial J(s, z^*)}{\partial s} = 0 \\ \frac{\partial J(s^*, z)}{\partial z} = 0 \end{cases}$$

$J_{11}(s, z)$ and $J_{12}(s, z)$ are minimized by $z=0$ and s_{11} and s_{12} ²⁰ that we define respectively as the level of probability precaution which minimizes $J_{11}(s, 0)$ and as the level of probability precaution which minimizes $J_{12}(s, 0)$. The injurer chooses to be potentially insolvent if $J_{11}(s_{11}, 0) < J(s^*, z^*)$ in the one-pocket model and if $J_{12}(s_{12}, 0) < J(s^*, z^*)$ in the two-pocket model.

If a regulatory standard, z_r , is only introduced on the magnitude precaution, so that $J_{11}(s_{11}, z_r) > J(s_{11}, z^*)$, in the one-pocket model the judgment-proof problem is removed under the same condition as in proposition 1: the injurer must not be bankrupt at the optimal level of precaution. Once the injurer takes z^* , he minimizes $J(s, z^*)$. The first order condition is the same as in Exp. (14), hence s^* results, which proves proposition 3.

The same holds for the two-pocket model. The results hold true both under strict liability and under simple negligence.

6.5. The probability model and the joint-probability-magnitude model

In this section, we discuss the two remaining possibilities: the probability model (the injurer reduces only the probability of an accident, the magnitude is exogenous), both in its two-pocket version and in its one-pocket version, and the joint-probability-magnitude model, also

¹⁹ An analysis of the model depicted in Exp. (12) is provided in Dari Mattiacci and De Geest (2002), in chapter 7. An analysis of the model described by Exp. (13) is elaborated in Dari Mattiacci and De Geest (2001), in chapter 5.

²⁰ Both $s_{11} > s^*$ and $s_{11} \leq s^*$ might result: see Dari Mattiacci and De Geest (2002), in chapter 7. Similarly both $s_{12} > s^*$ and $s_{12} \leq s^*$ might result: see Dari Mattiacci and De Geest (2001), in chapter 5.

with respect to one-pocket and two-pocket variant.

6.5.A. The probability model

Let us use the notation of section 6.2, with only one difference. Let:

$$\begin{aligned} p(x) &= \text{the probability of an accident occurring, } 0 < p < 1, p' < 0, p'' > 0; \\ h &= \text{the magnitude of the harm (exogenous), } h > 0. \end{aligned}$$

The social cost function is:

$$(15) \quad SC(x) = p(x)h + x.$$

In the one-pocket version, the injurer's expenditure function under strict liability is:

$$(16) \quad \begin{cases} J(x) = p(x)h + x & \text{if } h + x \leq t \\ J_{11}(x) = p(x)t + [1 - p(x)]x & \text{if } h + x > t \end{cases} \quad [\text{One-pocket probability model}].$$

Proposition 1 never holds because if the injurer is not bankrupt when he takes optimal precaution, $t > h + x^*$, then the solution can never be $x_{11} < x^*$; hence, the regulatory standard has to be set at the optimal level of precaution. The proof is as follows. The injurer chooses x_{11} if $J_{11}(x_{11}) < J(x^*)$, he chooses x^* otherwise. Assume that, when $t > h + x^*$, $x_{11} < x^*$ minimizes $J_{11}(x)$, then $t > h + x_{11}$ is also true, hence $J_{11}(x_{11}) = p(x_{11})[t - x_{11}] + x_{11} > p(x_{11})h + x_{11} > p(x^*)h + x^* = J(x^*)$ (by definition of x^*); hence, x^* would be the solution, and a judgment-proof problem would not arise. Therefore, $x_{11} < x^*$ cannot be a solution of injurer's minimization problem.

In the two-pocket version, the injurer's expenditure function under strict liability is:

$$(17) \quad \begin{cases} J(x) = p(x)h + x & \text{if } h \leq t \\ J_{12}(x) = p(x)t + x & \text{if } h > t \end{cases} \quad [\text{Two-pocket probability model}].$$

In this case proposition 1 never applies because, if the judgment-proof problem arises, $h > t$, the injurer is always unable to pay losses at any level of precaution, including the optimal level; hence, the condition of proposition 1 will never be satisfied. The underlying logic is that the decision whether to be potentially judgment-proof is not available: the injurer finds himself either potentially bankrupt ($h < t$) or not. Anti-judgment-proof regulation is meant to correct such decisions.

6.5.B. The joint-probability-magnitude model

Let us use again the notation of section 6.2, but let:

$$p(x) = \text{the probability of an accident occurring, } 0 < p < 1, p' < 0, p'' > 0;$$

$h(x)$ = the magnitude of the harm, $h > 0$, $h' < 0$, $h'' > 0$.

The social cost function is:

$$(18) \quad SC(x) = p(x)h(x) + x.$$

In the one-pocket version, the injurer's expenditure function under strict liability is:

$$(19) \quad \begin{cases} J(x) = p(x)h(x) + x & \text{if } h(x) + x \leq t \\ J_{11}(x) = p(x)t + [1 - p(x)]x & \text{if } h(x) + x > t \end{cases}$$

[One-pocket joint-probability-magnitude model].

Proposition 1 holds only when the magnitude component of the model is particularly relevant, the injurer has large control on $h(x)$, and this situation can be interpreted by a large absolute value of $h'(x^*)$. Proof: assume that, when $t > h(x^*) + x^*$, $x_{11} < x^*$ minimizes $J_{11}(x)$, then the first order condition is:

$$(20) \quad p'(x_{11}) = -\frac{1 - p(x_{11})}{t - x_{11}}.$$

The first order condition of $J(x)$ is

$$(21) \quad p'(x^*) = -\frac{1 + p(x^*)h'(x^*)}{h(x^*)}.$$

Because $t > h(x^*) + x^*$ and $x_{11} < x^*$, then $t > h(x^*) + x_{11}$ is also true, hence $t - x_{11} > h(x^*)$. As $1 - p(x_{11}) < 1 + p(x^*)h'(x^*)$ only for $h'(x^*)$ sufficiently small: in this case $p'(x_{11}) > p'(x^*)$, and since $p'' > 0$, then also $x_{11} > x^*$, which contradicts our assumption that if $t > h(x^*) + x^*$ then $x_{11} < x^*$ might result, and proposition 1 does not hold. However, if $h'(x^*)$ is sufficiently large, then the comparison between $p'(x_{11})$ and $p'(x^*)$ becomes indeterminate and proposition 1 might hold²¹.

In the two-pocket version, the injurer's expenditure function under strict liability is:

$$(22) \quad \begin{cases} J(x) = p(x)h(x) + x & \text{if } h(x) \leq t \\ J_{12}(x) = p(x)t + x & \text{if } h(x) > t \end{cases}$$

[Two-pocket joint-probability-magnitude model].

In this case proposition 1 applies, because $x_{12} < x^*$, and $J_{12}(x)$ is increasing for $x > x_{12}$. The proof of the former is that if $x_{12} \geq x^*$ then $J_{12}(x_{12}) = p(x_{12})t + x_{12} > p(x_{12})h(x_{12}) + x_{12} > p(x^*)h(x^*) + x^* = J(x^*)$ (the first inequality follows from the condition of proposition 1 that

²¹ For a more complete discussion of the one-pocket joint-probability-magnitude model see Dari Mattiacci and De

$t > h(x^*)$ and the second follows from the definition of x^*). Hence, if $x_{i2} \geq x^*$, then $J_{i2}(x_{i2}) > J(x^*)$, the solution would be x^* and the judgment-proof problem would not arise. The judgment-proof problem arises only if $x_{i2} < x^*$ ²² and in this case it can be solved by anti-judgment-proof regulation.

6.6. Concluding remarks.

Anti-judgment-proof regulation has the sole purpose of removing the judgment-proof problem. It operates in two ways.

(i) Regulation of minimum levels of precaution. When the injurer can influence only or also the magnitude of the accident, he has actual control on whether he will be potentially bankrupt or not. If he takes low levels of precaution, the harm will be great, and it might be greater than his assets (or the legal limit on his liability). If he takes more precaution the harm decreases and he will be able to pay for it. Anti-judgment-proof regulation only increases the cost associated to judgment-proofness and operates by targeting those levels of sub-optimal precaution that would be chosen by judgment-proof injurers. This is enough to prevent injurers' from deciding to expose themselves to insolvency.

(ii) Regulation of magnitude precaution only. In many cases, injurers can take a precautionary measure to reduce the probability and another to reduce the magnitude of the accident. They will be either potentially judgment-proof or not depending on the level of their magnitude precaution. Therefore, a regulatory standard that only targets magnitude precaution might be enough to prevent injurers from deciding to expose themselves to insolvency.

In the analysis, we have underlined two major conditions for this solution to be effective: injurers must be solvent at the optimal level of precaution and they must be able to reduce the magnitude of the harm by taking precaution.

Once the judgment-proof problem has been removed by anti-judgment-proof regulation, the problem remains of how to induce injurers to take optimal precaution. The normal functioning of tort law attains this objective. Both strict liability and simple negligence have been considered. In the latter case, the analysis suggests that compliance with the regulatory standard should not relieve injurers of liability for negligence. The reason is that the anti-judgment-proof regulatory standard might be suboptimal and, hence, might not be sufficient to induce optimal precaution. Therefore, both the regulatory standard and the negligence

Geest (2002), in chapter 7.

²² For a more complete discussion of the two-pocket joint-probability-magnitude model see Dari Mattiacci and De Geest (2001), in chapter 5.

criterion should be independently implemented.

In addition, the analysis suggests that the level of the regulatory standard may be set at any level between two limits: the optimal level of precaution (as an upper limit) and the minimum level necessary to prevent the injurer from being judgment-proof (as a lower limit). Therefore, the regulatory standard must be sufficiently high to keep the injurer out of that zone, but it should not be set at what the regulator believes to be the optimal precaution level, since the regulator operates *ex ante* and has less information than the courts, which operate *ex post*.

In a world with only one injurer, regulation can be set high enough to avoid judgment-proofness, and low enough to avoid overprecaution. In a more realistic model with more injurers (with a varying assets, so that some are more likely to become judgment proof than others), the regulator needs to balance the costs due to the judgment-proof problem (associated to too low a regulatory standard) and the costs of overprecaution (associated to too high a regulatory standard). Here we suggest two extreme cases. If judgment-proof problems are relatively unimportant, and the danger of overprecaution due to the lack of information of the regulator is more relevant, the optimal solution may be no regulation at all. On the contrary, if judgment-proof problems are likely to be enormous, the optimal solution may consist of completely relying on regulation, without recurring to tort law. In many situations, the optimal solution is in between these two extremes, where regulation and tort law need to be combined, and where the regulatory standards are set at a lower level than optimal precaution.

6.7. References

- Beard, Rudolph T. (1990), 'Bankruptcy and Care Choice', *21 Rand Journal of Economics*, 626 ff.
- Burrows, Paul (1999), 'Combining Regulation and Legal Liability for the Control of External Costs', *19 International Review of Law and Economics*, 227-244.
- Dari Mattiacci, Giuseppe and De Geest, Gerrit (2001), 'An Analysis of the Judgment-Proof Problem under Different Tort Models', Utrecht University, Economic Institute/CIAV, *working paper* (presented at the Annual Conference of the European Association of Law and Economics, Vienna, September 2001).
- Dari Mattiacci, Giuseppe and De Geest, Gerrit (2002), 'When Will Judgment-Proof Injurer Take too Much Precaution?', Utrecht University, Economic Institute/CIAV, *working paper*.
- Kolstad, Charles D., Ulen, Thomas S., and Johnson, Gary V. (1990), 'Ex Post Liability for Harm vs. Ex Ante Safety Regulation: Substitutes or Complements?', *80 American Economic Review*, 888-901.

- Shavell, Steven (1984a), 'Liability for Harm Versus Regulation for Safety', **13** *Journal of Legal Studies*, 357-374.
- Shavell, Steven (1984b), 'A Model for Optimal Use of Liability and Safety Regulation', **15** *Rand Journal of Economics*, 271-280.
- Shavell, Steven (1986), 'The Judgment Proof Problem', **6** *International Review of Law and Economics*, 45-58.
- Summers, John (1983), 'The Case of the Disappearing Defendant: An Economic Analysis', **132** *University of Pennsylvania Law Review*, 145-185.

Figures

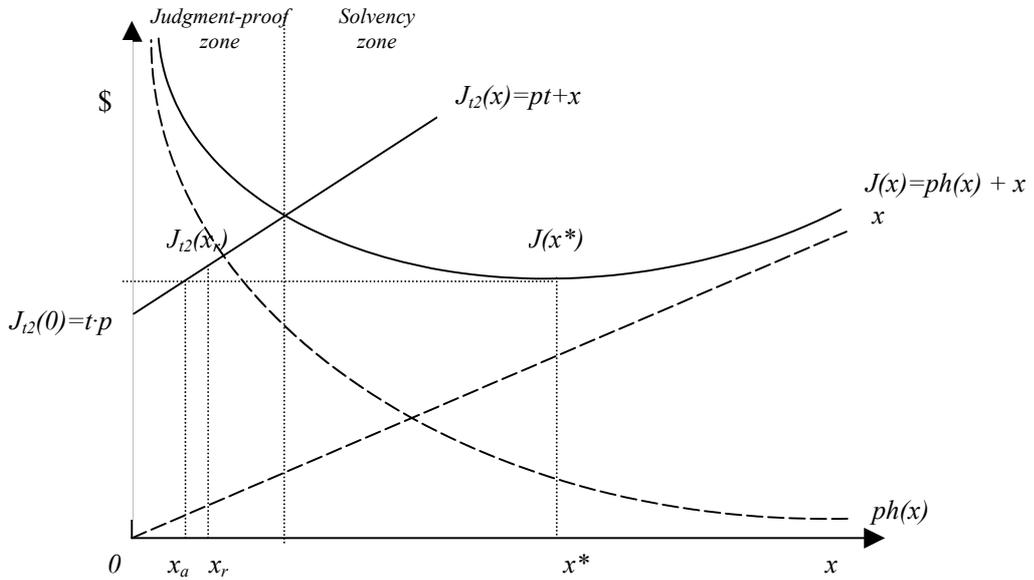


Figure 1: Anti-judgment-proof regulation in the two-pocket magnitude model under strict liability
 - x_a is the lower limit in Exp. (8), x_r can be set at any level between x_a and x^* -

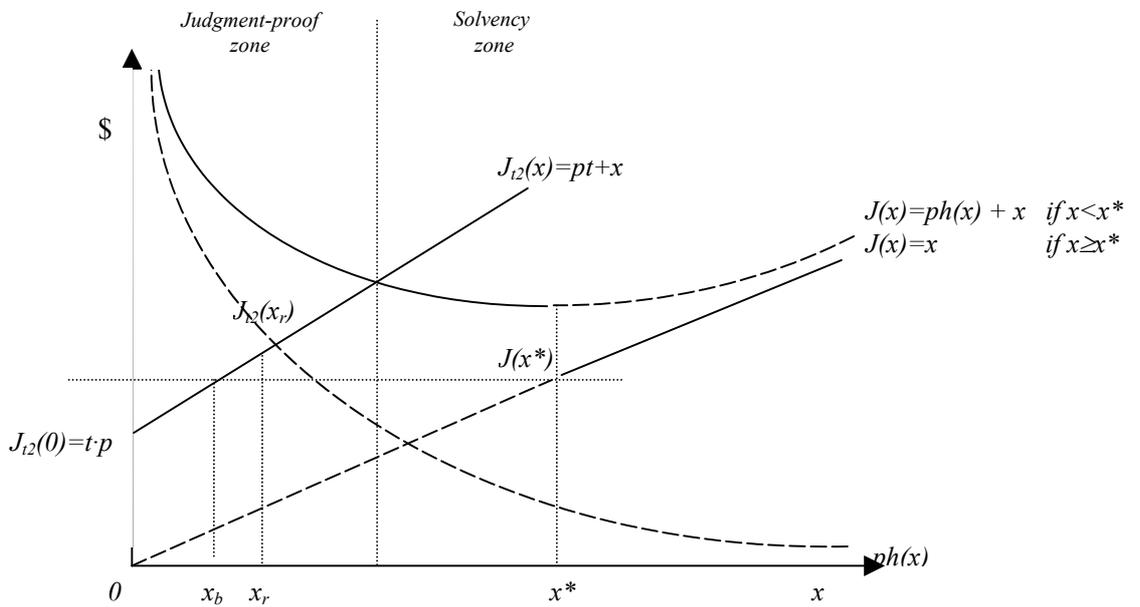


Figure 2: Anti-judgment-proof regulation in the two-pocket magnitude model under simple negligence
 - x_b is the lower limit in Exp. (10) x_r can be set at any level between x_a and x^* -

WHEN WILL JUDGMENT-PROOF INJURERS TAKE TOO MUCH PRECAUTION?*

Giuseppe Dari Mattiacci

Utrecht University and Marie Curie Fellow at Hamburg University

Gerrit De Geest

Utrecht University and Ghent University

ABSTRACT

An injurer is said to be judgment proof when his total assets are less than the harm he may cause. In principle, judgment-proof injurers can be expected to take less than optimal precaution, as they bear only a part of the accident loss (as formally proven by Shavell, 1986, using a probability model, i.e. a model where more precaution results in a lower probability of accidents). Beard (1990), however, showed that under certain conditions the judgment-proof problem can lead to overprecaution. Beard used a (stochastic) one-pocket probability model (where it was assumed that the more the injurer spends on precaution, the less assets will remain available to pay damages). While probability models can be appropriate to analyze some accident types (such as aircraft accidents), magnitude models (where more precaution reduces the magnitude of the loss and not the probability of the accident) are more appropriate to analyze other externality problems (such as nuisance, many types of environmental pollution or safety measures). In this chapter it is argued that overprecaution can never occur in a magnitude model. In addition, we consider a non-stochastic probability model and show that overprecaution is possible. We also analyze mixed models. Finally, we discuss the policy implications of our analysis.

JEL classification: K13, K32.

Keywords: insolvency, judgment proof problem, liability, bankruptcy, overprecaution.

* Giuseppe Dari Mattiacci, Economic Institute/CIAV, Utrecht University, Kromme Nieuwegracht 22, 3512 HH Utrecht, the Netherlands; email: g.darimattiacci@econ.uu.nl. Web: www.law.uu.nl/eco/gdm.htm. Gerrit De Geest, Economic Institute/CIAV, Utrecht University; email: g.degeest@econ.uu.nl. We would like to thank Nuno Garoupa, Vicente Mendes and Jef de Mot for helpful comments. Giuseppe Dari Mattiacci gratefully acknowledges the financial support provided by the European Commission, the Marie Curie Fellowship Programme in Law and Economics at Hamburg University. Updated versions of this chapter will possibly be available on my web page: www.law.uu.nl/eco/gdm.htm.

7.1. Introduction. Two-pocket probability model and two-pocket magnitude model.

An injurer is said to be judgment proof when his total assets are less than the harm he may cause. The injurer's total assets can be regarded as a maximum upper threshold on his liability.

Summers (1983) and Shavell (1986) showed that judgment-proof injurers tend to take less than optimal precaution. The reason is that, as not all accident losses are internalized, injurers bear the full marginal cost of any additional precaution but receive less than the full marginal benefits thereof (a reduction in the expected harm). In order to prove this result, Shavell (1986) used a probability model in which injurers can reduce the probability of an accident, $p(x)$, by taking more precaution, x , but they cannot influence the magnitude of the harm, h , which is considered to be exogenous.

In addition, he made the simplifying assumption that precaution expenses do not reduce the assets available for compensation in the case of an accident. In most real-world cases, however, the more the injurer spends in precaution, the less he will be able to pay in the case of an accident. This simplifying assumption is realistic only in two cases: non-monetary precautionary measures (which do not reduce the injurer's assets) and legal thresholds (where the law creates an artificial cap on the damages to be paid, and such a cap is set at a lower level than the injurer's total assets). These situations result in what could be called a *two-pocket* model: the injurer behaves as if he had two separate pockets. The first, limited by t , consists of the assets available for victim compensation; the second, unlimited, consists of the resources to be used to take precaution. The money spent on precaution, x , does not reduce the assets available for compensation, t .

In a two-pocket probability model, the injurer seeks to minimize the following function:

$$(1) \quad \begin{cases} J(x) = p(x)h + x & \text{if } h \leq t \\ J_t(x) = p(x)t + x & \text{if } h > t \end{cases} \quad [\text{Two-pocket probability model}].$$

While probability models can be appropriate to analyze some accident types (such as aircraft accidents), magnitude models (where more precaution reduces the magnitude of the loss and not the probability of an accident) are more appropriate to analyze other externality problems (such as nuisance, many types of environmental pollution and safety measures). Magnitude models, however, can lead to different analytical results. In chapter 5 it has been showed that the judgment-proof effect in a two-pocket magnitude model is different from the effect in a two-pocket probability model. While the latter leads to systematic underprecaution, a magnitude model yields either optimal precaution or no precaution at all.

$$(2) \quad \begin{cases} J(x) = ph(x) + x & \text{if } h(x) \leq t \\ J_t(x) = pt + x & \text{if } h(x) > t \end{cases} \quad [\textit{Two-pocket magnitude model}].$$

The reason is that, as the magnitude of the harm depends directly on precaution, the injurer can actually decide whether or not to go bankrupt by selecting his level of precaution. If he does not go bankrupt, $h(x) \leq t$ [solvency zone], he will be able to pay the full harm, hence he will choose the optimal level of precaution. If he goes bankrupt, $h(x) > t$ [judgment-proof zone], any precaution will be worthless, as he will pay anyway all his assets; therefore, he will choose no precaution at all. The judgment-proof effect generates therefore a binary outcome: the injurer decides either to be insolvent (no precaution) or to be solvent (optimal precaution). See figure 1.

In two-pocket models in general, precaution is optimal or lower than optimal: overprecaution never results.

This chapter analyzes the remaining two possibilities, the one-pocket probability model and the one-pocket magnitude model: the injurer has only one pocket to pay both precaution expenses and damages. The more he spends on precaution, the less will be available to pay damages.

The one-pocket probability model has already been studied by Beard (1990), who showed that under certain conditions the judgment-proof problem might lead to overprecaution. The fundamental intuition is the following. The judgment-proof problem distorts the injurer's incentives in two ways. First, it provides the injurer with an implicit harm subsidy: the greater the expected accident loss, the greater the portion thereof which will remain uncompensated in the case of an accident (an incentive to take less precaution in order to increase the expected accident loss). Second, it provides the injurer with an implicit precaution subsidy: the more the injurer spends on precaution, the greater the portion of the harm which will remain uncompensated (an incentive to take more precaution). In some cases, the precaution subsidy may dominate the harm subsidy, which induces the injurer to take too much precaution. Beard (1990) used a stochastic model. We show that his result holds also in a non-stochastic model: in a one-pocket probability model, overprecaution might result.

Nevertheless, Beard's model was still a probability model, although more complex than the one employed by Shavell (1986), because it took into consideration a one-pocket stochastic variant thereof.

We also analyze the one-pocket magnitude model and find that, when the injurer can reduce the magnitude instead of the probability of the accident, overprecaution will never occur, because the harm subsidy (an incentive towards less precaution) always offsets the

precaution subsidy (an incentive towards more precaution). If an accident occurs, a judgment-proof injurer is left with nothing, regardless of his level of precaution; to the contrary, if an accident does not occur, a judgment-proof injurer is left with his assets minus what he spent on precaution: therefore, it is convenient to take as little precaution as possible.

Section 7.2 will present the reader with the one-pocket probability model, and show that overprecaution is possible. In section 7.3, we will consider the one-pocket magnitude model, and demonstrate that overprecaution never results. Section 7.4 will explain the logic behind our results. In section 7.5, we will analyze mixed one-pocket probability-magnitude models. In the concluding section (section 7.6), we will discuss the policy implications of our findings.

7.2. The one-pocket probability model

In sections 7.2 to 7.5, we will consider accidents between a victim (the party which suffers a loss) and an injurer (the party which does not suffer any loss). They are strangers to each other. For the sake of simplicity, we assume unilateral accidents: only the injurer can take precaution in order to reduce the expected harm. The rule in force is strict liability.

Let:

x = the injurer's precaution costs, $x \in [0, t]$;

t = the injurer's assets (maximum upper threshold on injurer's liability);

$J(x)$ = the injurer's expected total expenditure;

$SC(x)$ = social cost

The injurer seeks to minimize his total expenditure, which is the sum of his precaution costs, x , and his expected liability expenses in the case of an accident. The socially optimal level of precaution is the level which minimizes the sum of the precaution costs and the expected harm. The injurer has limited assets t , therefore his exposure to liability may be less than the harm. In addition, the injurer's precaution costs x reduce the assets that are available to pay compensation to $t-x$.

In a one-pocket *probability* model, the injurer can reduce the probability of an accident by spending more on precaution, but he cannot reduce the magnitude of the harm, which is exogenous.

Let:

$p(x)$ = probability of an accident, $p \in (0, 1)$, $p' < 0$, $p'' > 0$;

h = accident loss, i.e. magnitude of the harm (exogenous).

The social cost function is:

$$(3) \quad SC(x) = p(x)h + x.$$

Let x^* denote the level of precaution which minimizes Exp. (3). The assumptions made guarantee convexity.

In a probability model, the harm h remains constant (irrespective of x). In case an accident occurs, the injurer's total expenditure is $h + x$ if he is not judgment proof. If he is judgment proof (that is, unable to compensate for the full harm h) his expenditure in the case of an accident is $(t - x) + x = t$. In the case of an accident, the injurer pays $t-x$ (all assets minus what he spent in precaution) to the victim. Therefore, the injurer's expenditure function is:

$$(4) \quad \begin{cases} J(x) = p(x)h + x & \text{if } h + x \leq t \\ J_t(x) = p(x)[t - x] + x & \text{if } h + x > t \end{cases} \quad [\text{One-pocket probability model}].$$

$J(x)$ is clearly minimized by x^* . Let x_t denote the level of precaution that minimizes $J_t(x)$. The assumptions made guarantee convexity. The injurer will choose to take optimal precaution if his total expenditure is lower at x^* than at x_t , i.e. if $J(x^*) \leq J_t(x_t)$. He will choose x_t otherwise. This condition can be rewritten as:

$$(5) \quad t \geq \frac{p(x^*)h + x^* - [1 - p(x_t)]x_t}{p(x_t)}.$$

The solution obtained by applying Exp. (5) always satisfies the conditions imposed by Exp. (4): $h + x^* \leq t$ if x^* is the solution, and $h + x_t > t$ if x_t is the solution¹.

It is easy to prove that x_t can be lower than, equal to or greater than x^* . Therefore, a judgment-proof injurer ($t < h + x_t$) might take underprecaution ($x_t < x^*$), optimal precaution ($x_t = x^*$) or overprecaution ($x_t > x^*$)².

If $h > t$, the injurer is always judgment proof, and only the second equation in Exp. (4) applies. However, if $h < t$, the injurer might go bankrupt just because he has spent a portion of his assets on precaution. Even an injurer, who is not judgment proof if he takes optimal

¹ The condition in Exp. (5) implies that the solution obtained satisfies the conditions imposed by Exp. (4), in fact the opposite is impossible. Let us assume that the solution is x^* ; if $h + x^* > t$, then $p(x^*)h + x^* > p(x^*)[t - x^*] + x^* > p(x_t)[t - x_t] + x_t$ (by definition of x_t), hence $J(x^*) > J_t(x_t)$ and the solution would be x_t , which contradicts the premise. Therefore, if x^* is the solution for Exp. (5), then $h + x^* \leq t$ in Exp. (4) must be satisfied. Let us now assume that the solution is x_t ; if $h + x_t \leq t$, then $p(x_t)[t - x_t] + x_t > p(x_t)h + x_t > p(x^*)h + x^*$ (by definition of x^*), hence $J(x_t) > J(x^*)$ and the solution would be x^* , which contradicts the premise. Therefore, if x_t is the solution for Exp. (5), then $h + x_t > t$ in Exp. (4) must be satisfied. In addition to that, for the same reason at least either $h + x^* \leq t$ or $h + x_t > t$ must be satisfied.

² Let us assume that $h + x^* > t$, then x_t is the solution (see footnote 1 for a proof). x^* satisfies $p'(x^*)h = -1$, x_t satisfies $p'(x_t)[t - x_t] - p(x_t) = -1$. The second derivative is positive in both cases. If $x_t > x^*$, then $p'(x^*)[t - x^*] - p(x^*) < -1$. By substituting $p'(x^*) = -1/h$ in the former we obtain $t > h + x^* - p(x^*)h$. Hence if $h + x^* - p(x^*)h < t < h + x^*$, the solution is $x_t > x^*$; if $h + x^* - p(x^*)h = t < h + x^*$ the solution is $x_t = x^*$; if $t < h + x^* - p(x^*)h$ the solution is $x_t < x^*$. In addition to that, $x_t > x^*$ might also be the solution when $h + x^* \leq t$, i.e. when the injurer would be able to pay compensation to the victim at the optimal level of precaution, but decides to take overprecaution in order to be judgment proof. In fact, $h + x^* \leq t$ implies $t > h + x^* - p(x^*)h$, hence $x_t > x^*$. If $J(x^*) > J_t(x_t)$, then x_t is the solution.

precaution ($h+x^*\leq t$), might decide to take overprecaution ($x_t>x^*$) in order to be judgment proof ($h+x_t>t$)³.

Beard (1990) showed that overprecaution is possible in a (one-pocket) stochastic probability model. Beard (1990, p.634) attributed his findings to a number of features of his model, including the stochastic elements. In this section, we have shown that overprecaution can also occur in a (one-pocket) *non-stochastic* probability model.

7.3. The one-pocket magnitude model

In this section, we will consider a one-pocket *magnitude* model. The injurer can reduce the magnitude of the harm, but not its probability, which is exogenous.

Let:

$$\begin{aligned} p &= \text{probability of an accident (exogenous), } 0 < p < 1; \\ h(x) &= \text{accident loss, i.e. magnitude of the harm, } h' < 0, h'' > 0; \end{aligned}$$

The social cost function is:

$$(6) \quad SC(x) = ph(x) + x.$$

Let x^* denote the level of precaution which minimizes Exp. (6). The assumptions made guarantee convexity.

The injurer pays the full harm to the victim only if $h(x)+x\leq t$ (the assets are large enough to pay both the precaution costs and the harm). Otherwise, the injurer pays $t-x$ (all assets minus what he spent in precaution).

Therefore, the injurer's expenditure function is⁴:

$$(7) \quad \begin{cases} J(x) = ph(x) + x & \text{if } h(x) + x \leq t \\ J_t(x) = pt + (1-p)x & \text{if } h(x) + x > t \end{cases} \quad [\text{One-pocket magnitude model}].$$

$J(x)$ is clearly minimized by x^* , while $J_t(x)$ is minimized by $x=0$. The injurer will choose to take optimal precaution if his total expenditure is lower at x^* than at $x=0$ ⁵, that is, if $J(x^*)\leq J_t(0)$. He will chose $x=0$ otherwise. This condition can be rewritten as:

³ See footnote 2 for a proof.

⁴ Note that $p(t-x)+x$ can be rewritten as $pt+(1-p)x$.

⁵ Note that, if $h(0)<t$, $x=0$ cannot be a solution as $J_t(x)$ is minimized by $x^>0$ such that $h(x^)+x^=t$. Nevertheless, $J_t(x^)=ph(x^)+x^$ is always greater than $J(x^)=ph(x^)+x^$ (by definition of x^*) and therefore the injurer will always choose x^* , unless he is bankrupt at x^* , i.e. if $t<h(x^)+x^$. Hence, $x^$ could be the outcome only if both $h(0)<t$ and $t<h(x^)+x^$ were simultaneously true. However, this is impossible, as it can be proven by the following simple algebra. If $h(0)<t$, then $ph(0)+0<pt$. By definition of x , $ph(x^)+x^<ph(0)+0$. Therefore, we can write $ph(x^)+x^<pt$, which yields $t>h(x^)+x^/p$. As $p<1$, then we can write $t>h(x^)+x^$. The latter proves that if $h(0)<t$, then $t<h(x^)+x^$ can never result. Therefore, $x^$ can never be a solution of the injurer's minimization problem: the

$$(8) \quad t \geq h(x^*) + \frac{x^*}{p}.$$

The solution obtained by applying Exp. (8) always satisfies the conditions imposed by Exp. (7): $h(x^*) + x^* \leq t$ if x^* is the solution, and $h(0) > t$ if $x=0$ is the solution⁶.

If the injurer's assets are sufficiently large to satisfy Eq. (8), the injurer will take optimal precaution x^* . Note that, as $p < 1$, the condition requires the injurer's assets to be higher than (not simply equal to) the sum of optimal precaution costs and optimal harm, hence the injurer will not go bankrupt if he takes optimal precaution. However, the condition does not require the assets to be large enough to pay any possible harm. The assets might be quite limited and the injurer might be potentially insolvent at low precaution levels. If condition (8) is not satisfied, the injurer will opt for no precaution at all, $x=0$. See figure 1.

7.4. The logic behind the different findings for the one-pocket magnitude model and the one-pocket probability model

Overprecaution is only possible in one-pocket models, since only these models allow an implicit precaution subsidy to the injurer: the more the injurer spends on precaution, the greater the portion of the harm that will remain unpaid in case he causes an accident which renders him insolvent.

In one-pocket models, the judgment-proof effect creates, therefore, two contrasting implicit subsidies: first, it subsidizes a portion of the harm (an incentive toward less precaution); second, it subsidizes precaution costs (an incentive toward more precaution). These implicit subsidies operate in opposite directions. The lower the precaution level, the lower the precaution subsidy and the higher the harm subsidy (the latter holds in the magnitude model where less precaution leads to a higher harm, as well as in the probability model where less precaution leads to a more frequent occurrence of accidents). The stronger subsidy (in marginal terms) will prevail. If the marginal harm subsidy prevails, underprecaution will result; if the marginal precaution subsidy prevails, the outcome will be overprecaution; if they perfectly set-off each other, optimal precaution will be taken.

In the one-pocket probability model any of those outcomes might result, while in the

only two possibilities are x^* and $x=0$. Figure 1 clarifies this issue.

⁶ The proof is similar to the one given in footnote 1. The condition in Exp. (8) implies that the solution obtained satisfies the conditions imposed by Exp. (7), in fact the opposite is impossible. Let us assume that the solution is x^* ; if $h(x^*) + x^* > t$, then $ph(x^*) + x^* > p[t - x^*] + x^* > pt$, hence $J(x^*) > J_t(0)$ and the solution would be $x=0$, which contradicts the premise. Hence if x^* is the solution for Exp. (8), then $h(x^*) + x^* \leq t$ in Exp. (7) must be satisfied. Let us now assume that the solution is $x=0$; if $h(0) \leq t$, then $pt + 0 > ph(0) + 0 > ph(x^*) + x^*$ (by definition of x^*), hence $J(x^*) \leq J_t(0)$ and the solution would be x^* , which contradicts the premise. Hence if $x=0$ is the solution for Exp. (8), then $h(0) > t$ in Exp. (7) must be satisfied. In addition to that, for the same reason, at least either $h(x^*) + x^* \leq t$ or

one-pocket magnitude model, the precaution subsidy can never prevail over the harm subsidy. We will provide the reader first with an intuitive reason for this result to hold and then with a formal explanation.

Injurers can reduce their total expenditure by reducing precaution (and, hence, by saving some precaution costs) or by increasing precaution (and, thus, by reducing the expected harm). The optimal level of precaution provides the efficient balance of these opposite forces; however, the judgment-proof subsidies alter such a balance.

On the one hand, the precaution subsidy renders precaution less expensive, but does not reduce the cost of precaution to zero. In fact, the subsidy is “paid” to the injurer only if an accident occurs, as only in this case a reduction in the injurer’s assets is perceived as a benefit, i.e. as a decrease in the damages that will actually be paid to the victim. On the contrary, the cost of precaution is borne by the injurer even if an accident does not occur. Consequently, the precaution subsidy simply reduces the incentives to save precaution costs (i.e. somewhat increases the incentives to take precaution), but does not reduce them to zero. The effect is similar under one-pocket probability and magnitude models.

On the other hand, the harm subsidy decreases expected liability, but has different effects under the two models. In the one-pocket probability model the harm subsidy only reduces the expected damages to be paid to the victim. If, for instance, the harm is equal to 100 and the injurer’s assets are equal to 85, only 15% of the expected harm is subsidized. The injurer maintains an incentive to spend on precaution and to reduce the probability to pay 85. Therefore, the effect of the harm subsidy competes against the effect of the precaution subsidy and the outcome will depend on their relative weight.

In the magnitude model, to the contrary, an insolvent injurer pays his total assets, 85, with a given probability and has no incentive to reduce the magnitude of the harm from 100 to, say, 90. 100% of the harm above the injurer’s assets is subsidized. Since increasing the level of precaution yields no benefit to an insolvent injurer, he maintains only some incentives to decrease precaution in order to reduce his total expenditure. The latter incentives are never annihilated by the precaution subsidy and, hence, they induce the injurer to reduce precaution to zero. As a consequence, the harm subsidy always prevails over the precaution subsidy in one-pocket magnitude models. A formal interpretation thereof follows.

The second Exp. in (7) depicts the cost function of a bankrupt injurer and can be rewritten as follows:

$$(9) \quad J_t(x) = ph(x) + x - p[h(x) - t] - px$$

[Harm subsidy and precaution subsidy in the one-pocket magnitude model].

$h(0) > t$ must be satisfied.

The first two terms in (9) represent the social cost function of Exp. (6); the third term, $p[h(x)-t]$, describes the expected harm subsidy: the portion of the harm which will remain uncompensated in the case of an accident. The fourth term, px , describes the expected precaution subsidy: the portion of the precaution costs which are subsidized as they reduce the assets available for compensation in the case of an accident.

The first derivative of Exp (9) depicts the marginal values of the four components just described and is given by the following Exp.:

$$(10) \quad ph'+1-ph'-p$$

[Marginal harm subsidy and marginal precaution subsidy in the one-pocket magnitude model].

The first term in (10) depicts the optimal incentive to reduce the social cost by means of increasing precaution (marginal reduction in the expected accident loss); the second term depicts the optimal incentive to reduce the social cost by means of reducing precaution (marginal reduction in the precaution costs). If the injurer were solvent, these two contrasting incentives would yield the optimal level of precaution, which optimally balances costs and benefits of precaution.

The judgment-proof subsidies alter such an optimal balance. The third term refers to the harm subsidy, which equals and completely neutralizes the optimal incentive to increase precaution (first term). Because of the harm subsidy, the injurer has no incentive to increase his level of precaution.

The fourth term, the marginal precaution subsidy, is equal to the marginal precaution costs in terms of absolute values (and thus it is equal to 1): if the injurer spends one more dollar in precaution, he will be able to pay one dollar less in the case of an accident. In terms of expected values, however, the marginal precaution subsidy is p , because the expenditure on precaution reduces the damage payment only if an accident occurs. Contrary to what we have noticed in relation to the marginal harm subsidy, the marginal precaution subsidy is never powerful enough to counteract the optimal incentive to reduce precaution ($p < 1$): the result is that the injurer maintains some incentives to reduce precaution.

Therefore, the combined actions of the judgment-proof subsidies only provide the injurer with incentives to reduce precaution, as the harm subsidy has a stronger effect than the precaution subsidy. Thus, his optimal choice will always be no precaution.

In the one-pocket probability model the harm subsidy is much weaker than in the magnitude model, and it is no longer sufficient to completely remove the effect of the (marginal) harm on injurer's expenditure. The second Exp. in (4) can be rewritten as follows

$$(11) \quad J_t(x) = p(x)h + x - p(x)[h - t] - p(x)x$$

[Harm subsidy and precaution subsidy in the one-pocket probability model].

The first two terms in (11) represent the social cost function of Exp. (3), the third term, $p(x)[h-t]$, describes the expected harm subsidy and the fourth term, $p(x)x$, the expected precaution subsidy. The first derivative of Exp. (11) is

$$(12) \quad p'h + 1 - p'[h-t] - [p'x + p(x)]$$

[Marginal harm subsidy and marginal precaution subsidy in the one-pocket probability model].

In the probability model, the marginal harm subsidy (third term) is $p'[h-t]$, while the marginal harm is $p'h$ (first term, the optimal incentive to increase precaution). The harm subsidy only reduces the incentive to increase precaution, because the marginal harm subsidy is always smaller than the marginal harm. Therefore, the injurer maintains some incentives to increase precaution.

The marginal precaution subsidy (fourth term) is $p'x+p(x)$. Since p' is negative, the first term $p'x$ reduces the second, which means that the marginal precaution subsidy is lower than 1 (second term, the optimal incentive to reduce precaution). In the probability model, the precaution subsidy reduces but does not completely remove the optimal incentives to reduce precaution.

Neither the harm subsidy nor the precaution subsidy is powerful enough to completely neutralize the optimal incentives to increase and to reduce precaution respectively. Consequently, the injurer maintains some incentives to increase precaution and some incentives to reduce precaution.

The result of the combined actions of the judgment-proof subsidies is indeterminate: either of the two might prevail. Therefore, the outcome might be an increase in the incentives to take precaution over the optimal level (overprecaution) as well as a decrease therein (underprecaution). It is also possible that the two subsidies balance each other perfectly and do not alter the optimal incentives (optimal precaution results).

7.5. Mixed one-pocket probability-magnitude models

So far, we have analyzed two stereotypical situations: a pure probability model and a pure magnitude model. In reality, injurers can often control through precaution both the probability of the accident and the magnitude of the harm. It is, therefore, worth analyzing briefly two mixed cases:

$$(13) \quad SC(x) = p(x)h(x) + x,$$

in which the injurer can reduce both the probability and the magnitude with the same precautionary measure (joint probability-magnitude model), and

$$(14) \quad SC(s, z) = p(s)h(z) + s + z ;$$

in which the injurer can reduce the probability by using a precautionary measure s and the magnitude by using a different precautionary measure z (separate probability-magnitude model).

In the first case, Exp. (13), the injurer’s expenditure function is⁷:

$$(15) \quad \begin{cases} J(x) = p(x)h(x) + x & \text{if } h(x) + x \leq t \\ J_t(x) = p(x)t + [1 - p(x)]x & \text{if } h(x) + x > t \end{cases}$$

[One-pocket joint-probability-magnitude model].

$J(x)$ is clearly minimized by x^* . Let x_t denote the level of precaution that minimizes $J_t(x)$. The injurer will choose to take optimal precaution if his total expenditure is lower at x^* than at x_t , that is, if $J(x^*) \leq J_t(x_t)$. He will chose $x=0$ otherwise. This condition can be rewritten as:

$$(16) \quad t \geq \frac{p(x^*)h(x^*) + x^* - [1 - p(x_t)]x_t}{p(x_t)} .$$

The solution obtained by applying Exp. (16) always satisfies the conditions imposed by Exp. (15): $h(x^*) + x^* \leq t$ if x^* is the solution, and $h(x_t) + x_t > t$ if x_t is the solution⁸. As in the one-pocket probability model, x_t can be lower than, equal to or greater than x^* . Therefore, a judgment-proof injurer ($t < h(x_t) + x_t$) might take underprecaution ($x_t < x^*$), optimal precaution ($x_t = x^*$) or overprecaution ($x_t > x^*$)⁹.

A precaution subsidy exists. Whether this leads to overprecaution will depend on whether the magnitude component of the model prevails over the probability one.

In the second case, Exp. (14), the injurer can take two separate precautionary measures. It is important to notice that the threshold t affects directly precaution z as in the pure magnitude model.

⁷ Note that $p(x)(t-x) + x$ can be rewritten as $p(x)t + [1-p(x)]x$.

⁸ The proof is similar to the one given in footnote 1. The condition in Exp. (16) implies that the solution obtained satisfies the conditions imposed by Exp. (15), in fact the opposite is impossible. Let us assume that the solution is x^* , if $h(x^*) + x^* > t$, then $p(x^*)h(x^*) + x^* > p(x_t)[t-x_t] + x_t > p(x_t)[t-x_t] + x_t$, (by definition of x_t) hence $J(x^*) > J_t(x_t)$ and the solution would be x_t , which contradicts the premise. Hence if x^* is the solution for Exp. (16), then $h(x^*) + x^* \leq t$ in Exp. (15) must be satisfied. Let us now assume that the solution is x_t ; if $h(x_t) + x_t \leq t$, then $p(x_t)[t-x_t] + x_t > p(x^*)h(x^*) + x^*$ (by definition of x^*), hence $J(x_t) \leq J_t(x_t)$ and the solution would be x^* , which contradicts the premise. Hence if x_t is the solution for Exp. (16), then $h(x_t) + x_t > t$ in Exp. (15) must be satisfied. In addition to that, for the same reason at least either $h(x^*) + x^* \leq t$ or $h(x_t) + x_t > t$ must be satisfied.

⁹ The proof is analogous to the one already given for the one-pocket probability model in footnote 2. Let us assume that $h(x^*) + x^* > t$, then x_t is the solution (see footnote 8 for a proof). x^* satisfies $p'(x^*)h(x^*) + p(x^*)h'(x^*) = -1$, x_t satisfies $p'(x_t)[t-x_t] - p(x_t) = -1$. The second derivative is positive in both cases. If $x_t > x^*$, then $p'(x^*)[t-x^*] - p(x^*) < -1$. By substituting $p'(x^*) = -[1+p(x^*)h'(x^*)]/h(x^*)$ in the former we obtain $t > \{[h(x^*) - p(x^*)h(x^*)]/[1+p(x^*)h'(x^*)]\} + x^*$. Hence, if $\{[h(x^*) - p(x^*)h(x^*)]/[1+p(x^*)h'(x^*)]\} + x^* < t < h(x^*) + x^*$, the solution is $x_t > x^*$, if $\{[h(x^*) - p(x^*)h(x^*)]/[1+p(x^*)h'(x^*)]\} + x^* = t < h(x^*) + x^*$ the solution is $x_t = x^*$, if $t < \{[h(x^*) - p(x^*)h(x^*)]/[1+p(x^*)h'(x^*)]\} + x^*$ the solution is $x_t < x^*$.

$$(17) \quad \begin{cases} J(s, z) = p(s)h(z) + s + z & \text{if } h(z) + s + z \leq t \\ J_t(s, z) = p(s)t + (1 - p(s))s + (1 - p(s))z & \text{if } h(z) + s + z > t \end{cases}$$

[One-pocket separate-probability-magnitude model].

$J(s, z)$ is clearly minimized by (s^*, z^*) . Let $(s_b, z=0)$ denote the level of precaution that minimizes $J_t(s, z)$ ¹⁰. The injurer will choose to take optimal precaution if his total expenditure is lower at (s^*, z^*) than at $(s_b, z=0)$, that is, if $J(s^*, z^*) \leq J_t(s_b, 0)$. He will choose $(s_b, z=0)$ otherwise. This condition can be rewritten as:

$$(18) \quad t \geq \frac{p(s^*)h(z^*) + s^* + z^* - [1 - p(s_t)]s_t}{p(s_t)}.$$

The solution obtained by applying Exp. (17) always satisfies the conditions imposed by Exp. (17): $h(z^*) + s^* + z^* \leq t$ if (s^*, z^*) is the solution, and $h(0) + s_t > t$ if $(s_b, z=0)$ is the solution¹¹.

As in the one-pocket magnitude model, the level of z will either be optimal, z^* , or equal to zero, hence no overprecaution takes place. On the contrary s_t might be lower, equal or higher than s^* , but this is not only due to the precaution subsidy (which reduces the cost of precaution by $1-p(s)$), but also to a sort of substitution effect between z and s . In fact, t might be higher than $h(z^*)$, and hence an insolvent injurer may face a higher expected harm than a solvent one, and be led towards more precaution s . As in the one-pocket probability model, overprecaution might result with respect to s ¹². In some cases, s_t might even be higher than $s^* + z^*$ ¹³: a bankrupt injurer spends in total for precaution more than a solvent injurer.

Consequently, when probability and magnitude depend on two different precautionary measures, judgment-proofness might yield overprecaution only with respect to the probability-reducing precaution, s . However, the result might be more relevant than in the pure probability model, as two forces push precaution forward: the precaution subsidy and the substitution effect. On the contrary, a magnitude-reducing measure z never experiences overprecaution.

¹⁰ The second equation is increasing in z , hence $z=0$ minimizes it. See also footnote 5.

¹¹ The proof is similar to the one given in footnote 1. The condition in Exp. (17) implies that the solution obtained satisfies the conditions imposed by Exp. (17), in fact the opposite is impossible. Let us assume that the solution is (s^*, z^*) ; if $h(z^*) + s^* + z^* > t$, then $p(s^*)h(z^*) + s^* + z^* > p(s^*)[t - s^* - z^*] + s^* + z^* > p(s_t)[t - s_t] + s_t$, (by definition of s_t) hence $J(s^*, z^*) > J_t(s_b, 0)$ and the solution would be $(s_b, 0)$, which contradicts the premise. Hence if (s^*, z^*) is the solution for Exp. (17), then $h(z^*) + s^* + z^* \leq t$ in Exp. (17) must be satisfied. Let us now assume that the solution is $(s_b, 0)$; if $h(0) + s_t \leq t$, then $p(s_t)[t - s_t] + s_t > p(s_t)h(0) + s_t > p(s^*)h(z^*) + s^* + z^*$ (by definition of s^* and z^*), hence $J(s^*, z^*) \leq J_t(s_b, 0)$ and the solution would be (s^*, z^*) , which contradicts the premise. Hence if $(s_b, 0)$ is the solution for Exp. (17), then $h(0) + s_t > t$ in Exp. (17) must be satisfied. In addition to that, for the same reason at least either $h(z^*) + s^* + z^* \leq t$ or $h(0) + s_t > t$ must be satisfied.

¹² The proof is analogous to the one given in footnote 2.

¹³ Let us assume that $h(z^*) + s^* + z^* > t$, then $(s_b, 0)$ is the solution (see footnote 11 for a proof). s_t satisfies $p'(s_t)[t - s_t] - p(s_t) = -1$. The second derivative is positive. If $s_t > s^* + z^*$, then $p'(s^* + z^*)[t - s^* - z^*] - p(s^* + z^*) < -1$. By rearranging, we obtain $t > \{[1 - p(s^* + z^*)] / [-p'(s^* + z^*)]\} + s^* + z^*$. Hence, if $[1 - p(s^* + z^*)] / [-p'(s^* + z^*)] < t < h(z^*) + s^* + z^*$, the solution is $s_t > s^* + z^*$.

7.6. Concluding remarks. Policy implications

A complete policy analysis would require empirical data, in order to determine whether in real situations injurers can affect the probability of an accident, the magnitude of the harm or both. Therefore, this section can only highlight some general policy implications of our analysis.

Many categories of accidents are subject to regulation. In most of the cases, the justification for regulatory intervention is the concern that tort law alone would fail to enhance optimal precaution, as injurers are judgment proof. Our analysis shows that it is important to distinguish between different categories of accidents.

In one-pocket probability and joint-probability-magnitude models, regulators should be concerned not only with underprecaution, but also with overprecaution, which might result as a consequence of the precaution subsidy created by bankruptcy. A regulatory standard coupled with tort liability will solve the underprecaution problem but will not prevent injurers from taking too much precaution. A maximum limit on parties' precaution should be set for that purpose by regulation, but it would be difficult to enforce.

In one-pocket magnitude models, overprecaution never results, but no precaution at all might be the outcome; therefore, the main concern of the regulator should be to force injurers to take precaution. In this case, regulation of minimum required level of precaution might suffice.

In one-pocket separate-probability-magnitude models, if injurer's assets are not particularly low, $t > h(z^*)$, it might be sufficient to regulate the magnitude-reducing precaution alone. If injurers are forced to choose z^* , then efficient precaution results automatically also with respect to s , which does not need to be regulated.

Many safety measures are likely to be pure magnitude measures, and hence should be analyzed under our approach. Fire escapes, lifeboats, helmets and safety belts, for instance, reduce the magnitude of the harm, and do not affect at all the probability of an accident occurring.

7.7. References

- Beard, Rudolph T. (1990), 'Bankruptcy and Care Choice', **21** *Rand Journal of Economics*, 626 ff.
- Dari Mattiacci, Giuseppe and De Geest, Gerrit (2001), 'An Analysis of the Judgment-Proof Problem under Different Tort Models', Utrecht University, Economic Institute/CIAV, *working paper* (presented at the Annual Conference of the European Association of Law and Economics, Vienna, September 2001).
- Shavell, Steven (1986), 'The Judgment Proof Problem', **6** *International Review of Law and Economics*, 45-58.

Summers, John (1983), 'The Case of the Disappearing Defendant: An Economic Analysis',
132 *University of Pennsylvania Law Review*, 145-185.

Figures

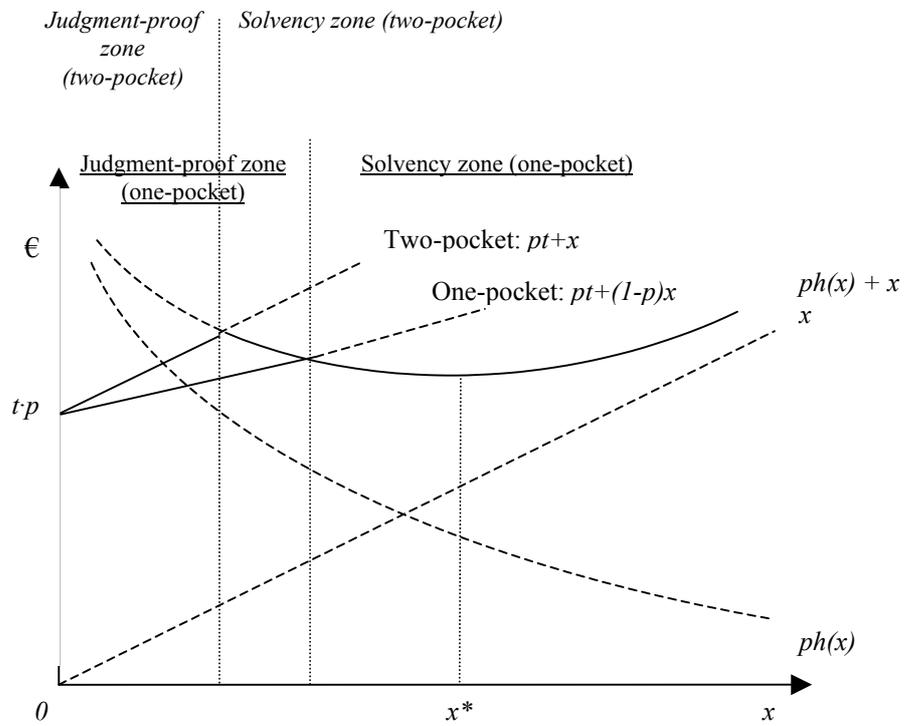


Figure 1: magnitude model (the injurer can reduce only the magnitude of the harm)

**NOXAE DEDITIO AND NEGLIGENCE:
ROMAN AND MODERN SOLUTIONS TO INEFFICIENT MONITORING
UNDER VICARIOUS LIABILITY**

*Giuseppe Dari Mattiacci**
Utrecht University

ABSTRACT

Why is vicarious liability strict for employers and duty-based for parents and supervisors? Why did Roman law limit *paterfamilias*' vicarious liability to the monetary value of the tortfeasor by means of *noxae deditio*? The answer to these two questions is the same: parents, supervisors and *patresfamilias* do only partially bear the cost of their agent's precaution. Under an unrestricted strict-liability rule, they would tend to over-prevent accidents. This chapter discusses these and some other related questions in a formal framework.

JEL classification: K13.

Keywords: vicarious liability, negligence, liability cap, threshold, negligence, noxae deditio, Roman law.

* Giuseppe Dari Mattiacci, Economic Institute/CIAV, Utrecht University, Kromme Nieuwegracht 22, 3512 HH Utrecht, the Netherlands; email: g.darimattiacci@econ.uu.nl. I would like to thank Gerrit De Geest, Willemijn Lindenhovius, Jef de Mot and Francesco Parisi for their helpful comments. Updated versions of this chapter will possibly be available on my web page: www.law.uu.nl/eco/gdm.htm.

8.1. Introduction

In the Law and Economics literature, vicarious liability is regarded as a system to provide judgment-proof agents with incentives to take optimal precaution by means of principals' monitoring or to enhance the efficient allocation of risk between the parties.

After the early contributions of Sykes (1981, 1984 and 1988) and Kornhauser (1982), who clarified the logic of vicarious liability and delimited its range of application, the literature has mainly been interested in corporate liability, and, in general, in vicarious liability of firms or employers¹, the main concern being the desirability of a vicarious liability rule. Little attention has been devoted to another aspect of the problem: the cost borne by the vicariously liable party in order to monitor the potential injurer.

Under personal liability, the optimal level of precaution is determined by the equality between the marginal benefit of precaution (the reduction in the expected accident loss) and the marginal cost thereof, which yields the minimization of total costs.

Under vicarious liability, the agent's precaution positively depends on the principal's monitoring expenditures. A reduction in the expected accident loss is attained at higher marginal costs, comprising both the precaution cost and the monitoring cost. Therefore, the optimal level of precaution under vicarious liability is lower than under personal liability.

Those principals, who are only liable for the negligent conduct of their agents, might be induced to choose too high a level of monitoring and their agents to take overprecaution, if the agents' negligence standard is set at a level which would be optimal under personal liability but is excessive under vicarious liability. This problem can be overcome by setting an appropriate negligence standard to the principal's level of monitoring also or by taking into account the principal's monitoring cost while setting the agent's standard of negligence.

Employers usually bear the precaution costs of their agents, both because they have to reward the agents for their efforts and because taking precaution results in an opportunity cost (for instance, agents could produce more instead of preventing accidents). On the contrary, parents (and supervisors in general) might not completely bear the precaution cost of their children (and of persons under their supervision), and they might be too concerned with reducing the expected accident loss, regardless of the cost of taking precaution. Again, an appropriate negligence standard for the principal's level of monitoring can solve the problem.

Similarly, in Roman family clans the *paterfamilias* did not bear the full cost of his slaves and family members' precaution. As a solution, Roman law provided a maximum upper

¹ See Arlen (1994), Chapman (1992 and 1996), Choi (1998), Croley (1996), Fischel and Sykes (1996), Khanna (1996), Kraakman (1984a, 1984b and 1986), Parker (1996), Schwartz (1996a), Shavell (1997).

threshold on the damages to be paid by the *paterfamilias*. This mechanism reduced the incentive for precaution, in a similar way to judgment-proofness². The *noxae deditio* provided in fact the *paterfamilias* with the possibility to avoid damage compensation by handing over the wrongdoer to the victim's family. The wrongdoer would then become the property of the victim's family. Consequently, *paterfamilias*' liability was limited to the monetary value of the wrongdoer.

I will also take into account the case that parents are too indulgent to their children and overestimate their cost of precaution, while trying to take that into account. Negligence can also be a solution to this problem.

Section 8.2 will formalize these findings and prove them in a simple model, which generalizes the results in the case of bilateral-precaution accidents also. Section 8.3 will account for different combinations of the negligence criterion on the principal and/or the agent in employers' liability. Section 8.4 will explore the use of negligence as a solution to over-monitoring and overprecaution when principals do not bear their agents' precaution cost, and the actual implementation thereof in different legal systems. Section 8.5 will focus on *noxae deditio* in Roman law. Section 8.6 will address some concluding remarks.

8.1.A. A numerical example

Under personal liability, the optimal level of precaution minimizes the sum of the precaution cost and the expected accident loss.

Table 1: personal liability

Expected accident loss: ph	Precaution cost: x	Total costs: $x+ph$
100	10	110
82	20	102
68	30	98
60	40	100

Under vicarious liability, the optimal level of precaution minimizes the sum of the precaution cost, the expected accident loss and the monitoring cost, and is therefore lower than the optimal level of precaution under personal liability (20 instead of 30). If the principal bears all these costs, he will enforce the optimal (under vicarious liability) level of precaution (I).

Table 2: vicarious liability when the principal bears the cost of the agent's precaution.

Expected	Monitoring	Precaution	Social cost:	Negligence on the	Negligence on the
----------	------------	------------	--------------	-------------------	-------------------

² Shavell (1986).

accident loss: ph	cost: m	cost: x	$x+ph+m$	agent: $x+ph+m$ if $x < 30$ $x+m$ if $x \geq 30$	principal: $x+ph+m$ if $m < 5$ $x+m$ if $m \geq 5$
100	2	10	112	112	112
82	5	20	107 ^(I)	107	25 ^(III)
68	10	30	108	40 ^(II)	40
60	17	40	117	57	57

However, if a negligence standard is set for the agent's precaution, the principal only pays damages if the agent is found negligent. A negligence standard set at the personal-liability optimal level (30) does not take the monitoring cost into account and might induce the principal to require excessive precaution on the part of his agent (II). In fact, he bears the precaution cost plus the expected accident loss plus the monitoring cost if the agent takes less than 30, but he only pays the precaution cost and the monitoring cost if the agent takes at least 30. The problem can be solved either by taking the monitoring cost into account while setting the standard of agent's negligence or by setting an appropriate standard to the principal's monitoring also (III).

If the principal does not bear the precaution cost, he will opt for too high a level of precaution (IV, in table 3), as he minimizes only the sum of the expected accident loss and the monitoring cost. Again, the problem can be solved by a negligence standard concerning the principal's monitoring level (V).

Table 3: vicarious liability when the principal does not bear the cost of the agent's precaution.

Expected accident loss: ph	Monitoring cost: m	Precaution cost: x	Principal's costs: $ph+m$	Principal's costs: $ph+m$ if $m < 5$ m if $m \geq 5$
100	2	10	102	102
82	5	20	87	5 ^(V)
68	10	30	78	10
60	17	40	77 ^(IV)	17

Roman law adopted a different solution: a maximum upper threshold on the damages to be paid by the principal (*noxae deditio*). In order to clarify this point, let us assume that if the agent takes precaution the probability of an accident decreases, while the harm is exogenous and remains constant. An appropriate upper limit on the damages (300 instead of 1000) makes the principal internalize a portion of the harm only and reduce his level of monitoring, and consequently the agent's level of precaution (VI instead of IV in table 4).

Table 4: *Noxae deditio*

Expected accident loss: ph	Monitoring cost: m	Precaution cost: x	Principal's costs: $ph+m$	Expected liability: pt	Principal's costs: $pt+m$
$10\% \cdot 1000=100$	2	10	102	$10\% \cdot 300=30$	32
$8.2\% \cdot 1000=82$	5	20	87	$8.2\% \cdot 300=24.6$	29.6 ^(VI)
$6.8\% \cdot 1000=68$	10	30	78	$6.8\% \cdot 300=20.4$	30.4
$6\% \cdot 1000=60$	17	40	77 ^(IV)	$6\% \cdot 300=18$	35

8.2. The model

8.2.A. Unilateral accidents

We will employ a simple model in order to describe the monitoring decision of a party (the principal), who is vicariously liable for accident losses caused by another party (a judgment-proof agent)³. We assume that parties are risk-neutral, rational and utility maximizing.

Derivatives are denoted by subscripts. Let:

x = agent's level of precaution, $x=[0, \infty)$;

$m(x)$ = principal's monitoring expenditure, $m=[0, \infty)$; $m_x > 0$; $m_{xx} > 0$ ⁴;

$p(x)$ = probability of an accident occurring, $p=[0, 1]$; $p_x < 0$; $p_{xx} > 0$;

h = harm (exogenous), $h > 0$;

a = agent's wealth, $a < h$ ⁵;

w = principal's payment to the agent in a contractual setting;

t = maximum upper threshold on the damages to be paid by the principal, $t < h$.

The probability of the accident is a function of the agent's level of precaution only; the principal and the victim cannot reduce it by taking precaution on their own (bilateral accidents will be considered in section 8.2.B). The principal can induce the agent to take a certain level of precaution either by simply conditioning the agent's payment to that level, or by enforcing

³ We do not discuss here whether vicarious liability is preferable to personal liability. We only examine those situations in which this is the case. This question has been the main concern of the literature on vicarious liability.

⁴ The causal relationship between x and m is the opposite as a matter of fact. Monitoring induces precaution and not the other way around. However, writing $m(x)$ simplifies the model and does not alter the substance of our reasoning at all. The principal decides how much precaution he wants the agent to take, and invests in monitoring so that that level of precaution will result. It is simply another way of looking at the same relationship. The positive second derivative depicts the diminishing returns (in terms of precaution) of the investment in monitoring.

⁵ We do not consider $a \geq h$ for the reason that in such a case the agent is not judgment-proof and vicarious liability is not a necessary device to induce optimal precaution: personal liability would provide the agent with perfect incentives.

it directly by means of monitoring the agent's level of precaution and sanctioning non-compliance. Given the magnitude of the sanction, the monitoring cost increases with the level of precaution that the principal enforces at an increasing rate.

We assume certainty in the principal's enforcement: given a combination of monitoring and sanctions, the required level of the agent's precaution results; as the agent always complies, sanctions are never applied and we do not consider them into the model⁶.

8.2.A.I. Social optimum

Proposition 1. The optimal level of precaution under vicarious liability is lower than the optimal level of precaution under personal liability.

When liability is vicarious, the reduction in the expected accident loss involves not only the agent's precaution cost (as under personal liability) but also the principal's monitoring cost, and hence the same level of precaution will be attained at a higher marginal cost.

If liability is personal, the socially optimal level of precaution, x^p , minimizes the sum of the precaution cost and the expected accident loss.

$$(1) \quad \min_x [x + p(x)h].$$

As Exp. (1) is convex⁷, from the first order condition we have:

$$(2) \quad p_x(x^p) = -\frac{1}{h}$$

If liability is vicarious the socially optimal level of precaution, x^* , minimizes the sum of the precaution cost, the expected accident loss and the monitoring cost.

$$(3) \quad \min_x [x + p(x)h + m(x)].$$

As Exp. (3) is convex⁸, from the first order condition we have:

$$(4) \quad p_x(x^*) = -\frac{1 + m_x}{h}.$$

$p_x(x^*)$ in Eq. (4) is clearly less than $p_x(x^p)$ in Eq. (2)⁹. As $p_{xx} > 0$, it follows that $x^* < x^p$: the optimal level of precaution under vicarious liability is lower than the optimal level of precaution under personal liability.

⁶ This assumption can be relaxed without undermining the results of the analysis.

⁷ First order condition: $1 + p_x h = 0$. Second order condition $p_{xx} h > 0$.

⁸ First order condition: $1 + p_x h + m_x = 0$. Second order condition $p_{xx} h + m_{xx} > 0$.

⁹ h and m_x are both greater than zero.

8.2.A.II. Contractual settings¹⁰

The principal and the agent are parties to a contract; they may be, for example, employer and employee respectively. The agent has limited wealth, a , and is judgment proof whatever payment he receives from the principal, $a+w < h^{11}$. The principal is instead solvent. The contract is Pareto optimal; therefore, it maximizes one party's utility, given the utility of the other party. Let the agent's utility be \tilde{w} , and let the principal maximize his utility by minimizing the payment to the agent plus his expected liability. Let w_n denote the payment in the case of no accident, and w_h the payment in the case of an accident occurring. The principal sets

$$(5) \quad \min_{w_n, w_h, x} [(1-p(x))w_n + p(x)(w_h + h) + m(x)],$$

subject to the constraint of constant agent's utility:

$$(6) \quad a + (1-p(x))w_n + p(x)w_h - x = \tilde{w}$$

Substituting Eq. (6) in Exp. (5), we obtain:

$$(7) \quad \min_x [\tilde{w} - a + x + p(x)h + m(x)] \quad \text{or} \quad \min_x [x + p(x)h + m(x)]$$

Exp. (7), the principal's minimization problem, is the same as in Exp. (3), the social cost minimization problem in the presence of vicarious liability; therefore, the principal will enforce x^* and bear $m(x^*)$ monitoring costs.

Proposition 2. If the principal bears the agent's precaution cost, strict liability for both the principal and the agent achieves the social optimum.

In a contractual setting, the principal bears the agent's precaution costs, as he has to compensate the agent for his effort. Strict liability on the principal implies principal's liability regardless of his level of monitoring, $m(x)$. Strict liability on the agent implies that the principal is liable regardless of the agent's level of precaution. The proof of Proposition 2 immediately follows from the fact that the principal bears exactly the social cost, Exp. (7).

Remark. The monitoring costs are likely to be zero for the first range on x . If the agent's assets plus the payment he receives if the accident does not occur are greater than zero, he will take some precaution also if the principal does not monitor¹². In such a case, the agent

¹⁰ The first part of this section closely follows Shavell (1987) at 182-185, except for the introduction of positive monitoring costs.

¹¹ See Shavell (1987) at 182, Sykes (1981).

¹² Shavell (1987) at 185.

will choose his level of precaution by maximizing the left-hand side of Eq. (6). The first order condition yields:

$$(8) \quad p_x(x^a) = -\frac{1}{w_n - w_h},$$

where x^a is the level of the agent's precaution that the principal can attain without monitoring. The principal can at most offer a contract in which $w_h = -a$, that is the agent pays his total assets to the principal when an accident occurs¹³. However, we have assumed at the beginning that $w_n + a < h$; therefore, the level of precaution chosen by the agent is lower than the optimal level of precaution under personal liability, $x^a < x^p$.

The monitoring cost is, hence, zero for $x = [0, x^a]$, and starts rising to the right of the interval¹⁴. By investing in monitoring, the principal can attain a level of precaution, x^v , falling between x^a and x^p . The assumptions made assure the convexity of the principal's minimization problem between those two limits. Therefore, the conclusions reached supra do not change.

Proposition 3. If the principal bears the agent's precaution cost, strict liability on the principal and duty-based liability on the agent might yield overprecaution and over-monitoring if the standard of negligence for agent's behavior is set at x^p .

Let x^p be the standard of negligence for the agent's behavior. After adapting Exp. (7), the principal's minimization problem becomes:

$$(9) \quad \min_x \begin{cases} x + m(x) & \text{if } x \geq x^p \\ x + p(x)h + m(x) & \text{if } x < x^p \end{cases}.$$

The principal will choose x^p ¹⁵ if

$$(10) \quad x^p + m(x^p) < x + p(x)h + m(x) \quad \text{for } x < x^p,$$

and x^* otherwise¹⁶.

From Proposition 1, it follows that x^p is too high a level of precaution and, hence, $m(x^p)$ is too high a level of monitoring. If the standard of agent's negligence under vicarious liability is set at the same level as the standard of agent's negligence under personal liability, principals might be too strict in order to escape liability, and enforce too high a level of precaution.

¹³ Such a contract would be illegal under many legal systems, as it makes the agent completely judgment proof. The limit should therefore be $w_h \geq 0$; however, that does not change the result.

¹⁴ Hence, $m_x(x) = 0$, $m_{xx}(x) = 0$ for $x \leq x^a$; $m_x(x) > 0$, $m_{xx}(x) > 0$, otherwise.

¹⁵ The principal will not choose $x > x^p$ as the first expression in (9) is increasing in x .

¹⁶ This follows from Proposition 2.

Nevertheless, if the standard of agent's negligence is correctly set at x^* , as determined by Proposition 1, the outcome will be efficient.

Proposition 4. If the principal bears the agent's precaution cost, duty-based liability on the principal achieves the social optimum even if the standard of agent's negligence is set at x^p .

Let us assume that the standard of principal's negligence is correctly set at $m^*=m(x^*)$. The principal's minimization problem is:

$$(11) \quad \min_x \begin{cases} x + m(x) & \text{if } m \geq m^* \\ x + p(x)h + m(x) & \text{if } m < m^* \end{cases}$$

As $m(x^p) > m(x^*)$, the principal will choose $m(x^*)$: any higher level of monitoring would only increase the monitoring cost and thus the precaution cost without decreasing principal's expected liability, which is already equal to zero.

8.2.A.III. Non-contractual settings

The concern here is with the monitoring decision of a principal who is not in a contractual relationship with the tortfeasor (we shall continue calling him the agent). Parents, supervisors, and *patresfamilias* are in this situation.

Proposition 5. If the principal only partially bears the agent's precaution cost (or underestimates it), strict liability on both the principal and the agent yields overprecaution and over-monitoring. If the principal overestimates the agent's precaution costs, strict liability on both the principal and the agent yields underprecaution and under-monitoring.

When the parties are not in a contractual relation, the principal might not internalize the agent's precaution cost perfectly, for he does not have to compensate the agent for his effort. On the one hand, he can use authority over the agent and require a certain level of precaution, without having to pay for it. On the other hand, while trying to take the agent's precaution cost into account, the principal might over or underestimate that. Let b denote the portion of x that the principal internalizes. $b < 1$ if the principal does not fully internalize the agent's cost or underestimates them. $b > 1$ if the principal overestimates the agent's cost.

$$(12) \quad \min_x [bx + p(x)h + m(x)].$$

As Exp. (12) is convex¹⁷, from the first order condition we have:

$$(13) \quad p_x(x^r) = -\frac{b+m_x}{h},$$

where x^r is the level of monitoring that minimizes Exp. (12). If $b < I$, $p_x(x^r)$ in Eq. (13) is clearly greater than $p_x(x^*)$ in Eq. (4). As $p_{xx} > 0$, it follows that $x^r > x^*$, and hence $m(x^r) > m(x^*)$. If the principal does not bear the agent's precaution costs fully, he will spend too much on monitoring and induce too high a level of the agent's precaution. On the contrary, if $b > I$, $x^r < x^*$, and, hence $m(x^r) < m(x^*)$. If the principal overestimates the agent's precaution cost, he will spend too little on monitoring and induce too low a level of the agent's precaution.

Proposition 6. If the principal only partially bears the agent's precaution cost (or underestimates them), duty-based vicarious liability or an appropriate maximum upper threshold on damages achieves the social optimum. If the principal overestimates the agent's precaution cost, duty-based vicarious liability achieves the social optimum under certain restrictive conditions.

The inefficient-monitoring / inefficient-precaution problem described in Proposition 5 can be dealt with by introducing a negligence standard concerning the principal's monitoring level or a maximum upper threshold on principal's liability.

Under duty-based liability, if the standard of negligence is correctly set at $m^* = m(x^*)$ ¹⁸, the principal's minimization problem is:

$$(14) \quad \min_x \begin{cases} bx + m(x) & \text{if } x \geq x^* \\ bx + p(x)h + m(x) & \text{if } x < x^* \end{cases}$$

If $b < I$, then $x^r > x^*$, and the principal will choose x^* : any higher level of precaution would only increase the cost of monitoring without decreasing the expected liability, which is already equal to zero. Therefore, duty-based vicarious liability solves the overprecaution / over-monitoring problem completely.

If $b > I$, then $x^r < x^*$, and the principal will choose x^* only if $bx^* + x^* \leq bx^r + p(x^r)h + m(x^r)$. Consequently, duty-based vicarious liability does not always solve the underprecaution / under-monitoring problem.

Let us now turn to the second proposed solution. If a threshold $t < h$ is introduced on

¹⁷ First order condition: $b + p_x h + m_x = 0$. Second order condition: $p_{xx} h + m_{xx} > 0$.

¹⁸ Note that the standard of principal's negligence can be expressed indifferently in terms of m or x . We shall use the latter as it simplifies the analysis.

principal's liability, the principal's minimization problem becomes¹⁹:

$$(15) \quad \min_x [bx + p(x)t + m(x)].$$

The level of t should be set so that x^* minimizes Exp. (15). As Exp. (15) is convex²⁰, from the first order condition we have:

$$(16) \quad p_x(x) = -\frac{b + m_x}{t}.$$

By substituting the right-hand side of Eq. (16) in the left-hand side of Eq. (4), we obtain:

$$(17) \quad t^v = \left[\frac{b + m_x(x^*)}{1 + m_x(x^*)} \right] h.$$

It is clear that $t < h$ can yield x^* only if $b < l$. Hence a maximum upper threshold can only solve the overprecaution / over-monitoring problem, but would be an inappropriate solution for the case of underprecaution / under-monitoring.

8.2.B. Bilateral accidents

Let us now consider the case of victims being able to take some precautions in order to curtail the probability of accidents. In addition to what has already been specified for the case of unilateral accidents, let:

y = victim's level of precaution, $y \in [0, \infty)$;

$p(x, y)$ = probability of an accident occurring, $p \in [0, 1]$; $p_x < 0$; $p_y < 0$; $p_{xx} > 0$; $p_{yy} > 0$.

8.2.B.I. Social optimum

If liability is personal, the socially optimal level of precaution, (x^p, y^p) , minimizes the sum of the precaution cost and the expected accident loss.

$$(18) \quad \min_{x, y} [x + y + p(x, y)h].$$

As Exp. (18) is convex²¹, from the first order condition we have:

¹⁹ The result would be the same even if we considered a stochastically variable and exogenous harm. The formal analysis would only slightly change.

²⁰ First order condition: $b + p_x t + m_x = 0$. Second order condition: $p_{xx} t + m_{xx} > 0$.

²¹ First order conditions: $1 + p_x h = 0$ and $1 + p_y h = 0$. Second order conditions $p_{xx} h > 0$ and $p_{yy} h > 0$.

$$(19) \quad \begin{cases} p_x(x^p, y^p) = -\frac{1}{h} \\ p_y(x^p, y^p) = -\frac{1}{h} \end{cases}$$

If liability is vicarious, the socially optimal level of precaution, (x^*, y^*) , minimizes the sum of the precaution cost, the expected accident loss and the monitoring cost.

$$(20) \quad \min_{x,y} [x + y + p(x, y)h + m(x)].$$

As Exp. (20) is convex²², from the first order condition we have:

$$(21) \quad \begin{cases} p_x(x^*, y^*) = -\frac{1+m_x}{h} \\ p_y(x^*, y^*) = -\frac{1}{h} \end{cases},$$

where (x^*, y^*) is the level of precaution which minimizes Exp. (20) and is clearly different from (x^p, y^p) . Two possible solutions are compatible with Eq. (19) and (21).

$$(i) \quad (x^* < x^p, y^* \leq y^p)$$

$$(ii) \quad (x^* \leq x^p, y^* > y^p)$$

Confronting the first Eq. in (21) and in (19), we have $x^*(y) < x^p(y)$ for any y . Hence, if $y^* \leq y^p$ also $x^* < x^p$. If $y^* > y^p$, then again $x^* \leq x^p$, as $(y^* > y^p, x^* > x^p)$ cannot be a solution: under vicarious liability the marginal benefit of precaution is the same as under personal liability but the marginal costs are higher, hence at least precaution level of either party must be lower under vicarious liability than under personal liability²³.

8.2.B.II. Contractual settings

In a contractual setting, a negligence rule can induce both parties' to take the optimal level of precaution. From the analysis under the former section it is clear that if the standard of negligence is set at the personal-liability level for either or both the agent and the victim, the result will be inefficient. Hence, any negligence rule ought to be implemented by setting a due level (for the victim, y^* , or the agent, x^*) that takes the monitoring cost into account.

8.2.B.III. Non-contractual settings

In a non-contractual setting the principal only partially bears x , the agent's cost of precaution. Therefore, the minimization problem is:

²² First order conditions: $1+p_x h+m_x=0$ and $1+p_y h=0$. Second order conditions $p_{xx} h+m_{xx}>0$ and $p_{yy} h>0$.

²³ If $(y^*, x^* \geq x^p)$ were a solution, then $x^p+y^p+p(x^p, y^p)h+m(x^p) > x^*+y^*+p(x^*, y^*)h+m(x^*)$, which can be rewritten as $x^p+y^p+p(x^p, y^p)h > x^*+y^*+p(x^*, y^*)h+m(x^*)-m(x^p)$, which is not possible as $m(x^*)-m(x^p) \geq 0$ and

$$(22) \quad \min_{x,y} [bx + y + p(x,y)h + m(x)].$$

As Exp. (22) is convex²⁴, from the first order condition we have:

$$(23) \quad \begin{cases} p_x(x^r, y^r) = -\frac{b + m_x}{h} \\ p_y(x^r, y^r) = -\frac{1}{h} \end{cases},$$

where (x^r, y^r) is the level of precaution which minimizes Exp. (22), and it is clearly different from (x^*, y^*) . If $b < 1$, two possible solutions are compatible with Eq. (21) and (23):

- (i) $(x^r > x^*, y^r \geq y^*)$
- (ii) $(x^r \geq x^*, y^r < y^*)$

Confronting the first Eq. in (21) and in (19), we have $x^r(y) > x^*(y)$ for any y . Hence, if $y^* \geq y^r$ also $x^* < x^r$. If $y^* < y^r$, then again $x^r \leq x^*$, as $(x^r < x^*, y^r < y^*)$ cannot be a solution: in a non-contractual setting the marginal benefit of precaution is the same as in the social optimum, but marginal costs are lower, hence at least the precaution level of either party must be higher in a non-contractual setting than in the social optimum²⁵.

If $b > 1$, the feasible solutions are:

- (i) $(x^r < x^*, y^r \leq y^*)$
- (ii) $(x^r \leq x^*, y^r > y^*)$

The reason is opposite to the one given above.

As in the unilateral precaution case, negligence or a maximum upper threshold can be implemented as solutions to the problem under the conditions already discussed.

8.3. Different combinations of the negligence criterion in employers' liability

The foregoing analysis suggests that if the principal is held liable for negligent agents' torts only, he will enforce too high a level of the agent's precaution unless:

- (i) While setting the due level of the agent's precaution, monitoring costs are also taken into account (the due level is set at x^*);
- (ii) Monitoring is inexpensive ($x^* = x^p$);
- (iii) A duty-based vicarious liability rule is implemented and an appropriate due level of monitoring is set [$m^* = m(x^*)$].

$x^p + y^p + p(x^p, y^p)h < x^* + y^* + p(x^*, y^*)h$. Hence, $x^* < x^p$. However, both $y^* > y^p$ and $y^* \leq y^p$ are possible.

²⁴ First order conditions: $b + p_x h + m_x = 0$ and $1 + p_y h = 0$. Second order conditions $p_{xx} h + m_{xx} > 0$ and $p_{yy} h > 0$.

²⁵ If $(y^r, x^r \leq x^*)$ were a solution, then $b x^r + y^r + p(x^r, y^r)h + m(x^r) < b x^* + y^* + p(x^*, y^*)h + m(x^*)$; it follows that $b x^r + y^r + p(x^r, y^r)h + m(x^r) < b x^* + y^* + p(x^*, y^*)h + m(x^*) + (1-b)(x^* - x^r)$, as $(1-b)(x^* - x^r) \geq 0$; the former can be rewritten as $x^r + y^r + p(x^r, y^r)h + m(x^r) < x^* + y^* + p(x^*, y^*)h + m(x^*)$. This is not possible as (x^*, y^*) minimizes the expression by

- (iv) The situation can be modeled as suggested by Grady (1983) and Kahan (1990); see my concluding remarks on this point.

If none of these conditions is fulfilled, a due level of the agent's precaution x^p that would be optimal under personal liability is excessive under vicarious liability. Unless monitoring is inexpensive, vicarious liability corrects the inefficiency due to personal liability of judgment-proof agents at the price of lower accident prevention.

This framework can be applied to the study of the different legal solutions adopted by various legal systems and to their efficiency evaluation. German law implements a duty-based liability for employers, while French and Anglo-American law generally opt for strict employers' liability (see Eörsi, 1975). A thorough comparative analysis should also take into account judicial decisions and the characteristics of the contexts in which employers' liability takes place (as indicated by points (ii) and (iv) above), but it would exceed the scope of this chapter. Here I simply suggest a theoretical framework which can be implemented in such an analysis.

8.4. Vicarious liability of parents and supervisors: negligence as a solution for over-demanding principals

If the principal only partially bears the agent's cost of precaution, he will enforce too high a level of precaution. In a non-contractual setting, the principal does not have to reward the agent for his effort. The relationships between parents and their children and between supervisors and the persons under their supervision resemble this situation. Parent will only partially take into account the cost of precaution borne by their children. Similarly, supervisors will not internalize the full cost of precaution borne by the supervised persons. Both of them will enforce too high a level of precaution.

It is also possible that parents and supervisors attempt to consider the costs of precaution that they impose on their children / supervisees, while taking their decisions concerning monitoring. If they overestimate those costs (they are too indulgent), then underprecaution will result.

An appropriate negligence standard can always solve the first problem and will sometimes solve the second as well. The liability of parents and supervisors is negligence-based in most legal systems (see le Gall, 1976). In the light of my analysis, this seems to be an efficient solution.

definition of the social optimum level of precaution. Hence, $x^p > x^*$. However, both $y^p \geq y^*$ and $y^p < y^*$ are possible.

8.5. *Noxae deditio* as a Roman law solution for over-demanding vicariously liable *paterfamilias*

In Roman families, the *paterfamilias* (the principal) did not have to reward his agents (the slaves and the members of his family, *filiifamilias*) for their efforts. If he had been obliged to pay any harm caused by any of them to another family clan, he would have enforced too high a level of precaution (at too high a monitoring cost). *Noxae deditio* limited his responsibility to the monetary value of the wrongdoer, who could be handed over to the victim as a form of compensation. This solution roughly corrects the overprecaution/over-monitoring problem.

The historical evolution of *noxae deditio* confirms these theoretical predictions. Originally, it was a form of vengeance on the injurer's body. The tortfeasor was handed over to the victim's family in order to be punished. The tortfeasor was often killed by his family and his corpse was handed over to the victim's family, for the only purpose of redressing the wrong committed. The same can be found in the *actio de pauperie*, which provided a similar solution for damages caused by animals: vengeance was directed against the body of the animal.

At a later stage, when tort law evolved from its early criminal law matrix, *noxae deditio* remained in force but changed its function. It simply meant the transfer of the property right on the slave or the member of the family to the victim's family. Our framework applies at this stage. The possibility of escaping liability by handing over the tortfeasor to the victim generated a sort of limited liability: the *paterfamilias* bore the lower cost between damage compensation and the market value of the slave / member of the family.

I have explained this peculiar system by noticing that the *paterfamilias* was likely to internalize only a portion of the slaves' efforts, and therefore under full liability he would have been too severe on them. The limitation on *paterfamilias*' liability mitigated this effect.

In a later period, when *filiifamilias*' liability could be personal for the existence of a personal patrimony (*peculium*), *noxae deditio* lost its function and tended to disappear, while it was retained for slaves. This is consistent with my explanation, as, when liability can be personal, the need for vicarious liability fades and the function of *noxae deditio* as a corrective device for the principal's enforcement level also disappears.

In modern societies, *noxae deditio* does not occur for two reasons. First, when the principal-agent relationship is based upon a contract, principals internalize the agents' precaution cost. I have demonstrated that, in such cases, since the principal has to reward the agent for his effort, he will consider the cost of precaution while setting the level thereof. Therefore, there is no need for a corrective device.

Second, when the principal-agent relationship is non-contractual, in modern societies the

principal's enforcement might also lead to too high a level of precaution. However, a better solution is available: negligence rules. Negligence rules, if appropriately set, are more precise a mechanism to deal with the over-monitoring / overprecaution problem than *noxae deditio*, which only provided a rough correction of the problem, for it limited principal's liability to the market value of the slave, which did not necessarily correspond to the optimal limit.

On the contrary, negligence rules can be set exactly at the optimal level and correct the vicarious liability rule perfectly. Nevertheless, the application of a negligence rule triggers higher administrative costs than *noxae deditio*, as it requires an analysis of the level of precaution actually taken by the parties and of the precaution and monitoring technology. Those costs are likely to be lower in modern societies than in ancient societies, for a much more developed judicial system is available, the cost of acquiring information is lower, the understanding of the laws of nature is more precise and systems of records are more common²⁶.

8.6. Concluding remarks

Some of the results of this analysis crucially depend on the discontinuity created by the application of a negligence rule to the injurer's cost function. Such discontinuity is due to the fact that negligent injurers also pay damages for the accidents that would have occurred anyway, even if they had been non-negligent. Grady (1983) and Kahan (1990) have demonstrated that, if such discontinuity is eliminated, a due level of precaution that is higher than the optimal level does not induce overprecaution. However, a due level of precaution that is lower than the optimal level of precaution yields underprecaution just as in the standard model.

It follows that the first part of this analysis does not hold true in the Grady-Kahan model: principals will choose the optimal level of precaution even if the agents' standard of negligence is too high. Nevertheless, the second part of the analysis still holds true: those principals, who do not bear the full cost of the agent's precaution, will monitor excessively and a negligence rule can lower their monitoring levels.

8.7. References

Arlen, Jennifer H. (1994), 'The Potentially Perverse Effects of Corporate Criminal Liability', *23 Journal of Legal Studies*, 832-867.

²⁶ See Posner (1980).

- Chapman, Bruce (1992), 'Limited Auditors' Liability: Economic Analysis and the Theory of Tort Law', **20** *Canadian Business Law Journal*, 180-214.
- Chapman, Bruce (1996), 'Corporate Liability and the Problem of Overcompliance', **69** *Southern California Law Review*, 1679-1703.
- Choi, Stephen (1998), 'Market Lessons for Gatekeepers', **92** *Northwestern University Law Review*, 916 ff.
- Chu, C.Y. Cyrus and Qian, Yingy (1995), 'Vicarious Liability under a Negligence Rule', **15** *International Review of Law and Economics*, 305-322.
- Croley, Steven P. (1996), 'Vicarious Liability in Tort: On the Sources and Limits of Employee Reasonableness', **69** *Southern California Law Review*, 1705-1738.
- Eörsi, Gyula (1975), 'Private and Governmental Liability for the Torts of Employees and Organs', in André Tunc (ed.), *Torts, International Encyclopedia of Comparative Law*, Tübingen, Mohr.
- Fischel, Daniel R. and Sykes, Alan O. (1996), 'Corporate Crime', **25** *Journal of Legal Studies*, 319-349.
- Grady, Mark F. (1983), 'A New Positive Economic Theory of Negligence', **92** *Yale Law Journal*, 799-829.
- Kahan, Marcel (1989), 'Causation and Incentives to Take Precaution under the Negligence Rule', **18** *Journal of Legal Studies*, 427-447.
- Khanna, V.S. (1996), 'Corporate Criminal Liability: What Purpose does it Serve?', **109** *Harvard Law Review*, 1477-1534.
- Kornhauser, Lewis A. (1982), 'An Economic Analysis of the Choice between Enterprise and Personal Liability for Accidents', **70** *California Law Review*, 1345-1392.
- Kraakman, Reinier H. (1984a), 'The Economic Functions of Corporate Liability', in Hopt, Klaus J. and Teubner, Gunther (eds), *Corporate Governance and Directors' Liabilities*, New York, de Gruyter, 178-207.
- Kraakman, Reinier H. (1984b), 'Corporate Liability Strategies and the Costs of Legal Controls', **93** *Yale Law Journal*, 857-898.
- Kraakman, Reinier H. (1986), 'Gatekeepers: The Anatomy of a Third-Party Enforcement Strategy', **2** *Journal of Law, Economics, and Organization*, 53-104.
- Landes, William M. and Posner, Richard A. (1987), *The Economic Structure of Tort Law*, Cambridge (MA), Harvard University Press.
- Le Gall, Jean-Pierre (1976), 'Liability for Persons under Supervision', in André Tunc (ed.), *Torts, International Encyclopedia of Comparative Law*, Tübingen, Mohr.
- Levmore, Saul (1995), 'Gomorrhah to Ybarra and More: Overextraction and the Puzzle of Immoderate Group Liability', **81** *Virginia Law Review*, 1561-1604.
- Newman, Harry A. and Wright, David W. (1990), 'Strict Liability in a Principal-Agent Model', **10** *International Review of Law and Economics*, 219-231.
- Parker, Jeffrey (1996), 'Doctrine for Destruction: The Case of Corporate Criminal Liability', **17** *Managerial and Decision Economics*, 381-398.
- Posner, Richard A. (1980), 'A Theory of Primitive Society with Special Reference to Law', **23** *Journal of Law and Economics*, 1-53.
- Schwartz, Gary T. (1996a), 'The Hidden and Fundamental Issue of Employer Vicarious Liability', **69** *Southern California Law Review*, 1739-1767.

- Shavell, Steven (1986), 'The Judgment Proof Problem', **6** *International Review of Law and Economics*, 45-58.
- Shavell, Steven (1987), *Economic Analysis of Accident Law*, Cambridge (MA), Harvard University Press.
- Shavell, Steven (1997), 'The Optimal Level of Corporate Liability Given the Limited Ability of Corporations to Penalize Their Employees', **17** *International Review of Law and Economics*, 203-213.
- Sykes, Alan O. (1981), 'An Efficiency Analysis of Vicarious Liability Under the Law of Agency', **91** *Yale Law Journal*, 168-206.
- Sykes, Alan O. (1984), 'The Economics of Vicarious Liability', **93** *Yale Law Journal*, 1231-1280.
- Sykes, Alan O. (1988), 'The Boundaries of Vicarious Liability: An Economic Analysis of the Scope of Employment Rule and Related Legal Doctrines', **101** *Harvard Law Review*, 563-609.

EFFICIENCY WAGES, CONDITIONAL BONUSES, AND PUNITIVE DAMAGES: MONITORING LEVELS WHEN SANCTIONS ARE FINANCED BY VICTIMS*

Gerrit De Geest
Utrecht University and Ghent University

Giuseppe Dari Mattiacci
Utrecht University

Jacques J. Siegers
Utrecht University

ABSTRACT

According to the efficiency wage theory by Shapiro and Stiglitz (1984), it can be rational for an employer to pay higher wages to his employees in order to make the sanction of dismissal more severe. This increased sanction allows the employer to lower monitoring levels and hence to save on monitoring costs. However, if all employers do so, the effect of such a relative overpayment disappears and the only sanction to shirking employees will be provided by the risk of unemployment. In this chapter, we argue that efficiency wages are almost never an efficient sanction system. Rational parties will nearly always prefer negative sanctions to positive sanctions. In the few cases where positive sanctions will be employed, parties will choose conditional bonuses (an alternative positive sanction overlooked in the labor economics literature) rather than efficiency wages. Therefore, it seems extremely implausible that employers will ever opt for efficiency wages on such a massive scale that structural unemployment results. In addition, employers determine the monitoring level and the sanction in order to minimize the sum of the monitoring costs and the expected bonus payment expenses; we find that efficiency wages lead to higher than optimal monitoring levels while monitoring levels are optimal under punitive damages and conditional bonuses. Our analysis can also be applied to principal-agent relationships different from labor contracts. It suggests, for instance, that overpaying judges or notaries is not the optimal method to conquer corruption as long as criminal sanctions can be increased.

JEL classification: J41, K12, K14, K31

Keywords: negative sanctions, positive sanctions, corruption, penalty clauses, unemployment.

* Gerrit De Geest, Economic Institute/CIAV, Utrecht University, and School of Law, Ghent University. Email: gerrit.degeest@rug.ac.be. Giuseppe Dari Mattiacci, Economic Institute/CIAV, Utrecht University, Kromme Nieuwegracht 22, 3512 HH Utrecht, the Netherlands. Email: g.darimattiacci@econ.uu.nl. Jacques J. Siegers, Economic Institute/CIAV, Utrecht University, Kromme Nieuwegracht 22, 3512 HH Utrecht, the Netherlands. Email: J.J.Siegers@econ.uu.nl. The authors would like to thank Mireia Artigot Golobardez, Roger Bowles, Jef De Mot, Robert Dur, Ehrling Eide, Winand Emons, Carlo Ferrigno, Gianluigi Galeotti, Hans Peter Grüner, Tina Heubeck, Manfred Holler, Mitch Polinsky, Hein Roelfsema, Caspar Rose, Alessandra Ruggeri and participants in the Annual Conference of the American Law and Economics Association, Washington, May 11th-12th, 2001 for helpful suggestions and comments on early drafts of this chapter. Updated versions of this chapter will possibly be available on my web page: www.law.uu.nl/eco/gdm.htm.

9.1. Introduction

According to the efficiency wage theory by Shapiro and Stiglitz (1984), it can be rational for an employer to overpay his employees since above-market wages make the sanction of dismissal more severe. This increased sanction allows the employer to lower monitoring levels and hence to save on monitoring costs. But if many employers try to save on monitoring costs by increasing wages, the effect disappears since a worker who loses an ‘overpaid’ job can easily find another ‘overpaid’ job. In this case, structural unemployment is the only way for the economy to create incentives for employees not to be dismissed.

Shapiro and Stiglitz’ article (1984) was written in the 1980s, at a time when many economists started to believe that unemployment was structural, rather than cyclical. Their theory was and is still considered as a possible explanation of unemployment and is mentioned in textbooks on labor economics.

Other authors have suggested different justifications for efficiency wages, such as the income-satiation (Rasmusen, 1992) and the gift-reciprocity (Akerlof, 1982) arguments. In this chapter we look at efficiency wages only as an enforcement device, and argue that they are less rational a choice than generally accepted and that it is extremely implausible that employers will ever opt for efficiency wages on such a massive scale that structural unemployment will result.

A second reason to reconsider the efficiency-wage theory is that law and economics scholarship has not realized the potential relevance of labor economics insights yet¹. Efficiency wages are just another technique to avoid breach of contract: in order for the employee to have an incentive to perform (working hard) he is overpaid. If he breaches the contract he does not have to pay damages, but loses a bonus. The economic literature on contract remedies has focused only on negative sanctions: specific performance, expectation damages, reliance damages, penalty clauses, and so forth. It has not compared the advantages of positive sanctions to those of negative sanctions yet. The overpayment technique is also relatively unexplored in the economic literature on public law enforcement, and corruption in particular².

In this chapter, we analyze under what conditions rational employers to a contract would prefer efficiency wages (a specific positive sanction) to negative sanctions. We will consider

¹ One exception is Shavell (1997), who took into account the possibility for firms to pay efficiency wages to employees, in a discussion on the optimal level of corporate liability.

² Most models on corruption do not take bonus payments into account as an alternative solution to corruption. Exceptions are Polinsky and Shavell (2000b), and Cooter and Garoupa (2000).

only one type of negative sanction: ‘*punitive damages*’³. Punitive damages are monetary sanctions with an expected value that equals the true harm⁴. For instance, if the harm of contract breach is \$2 and the chance that the breach will be discovered by the promisee is 0.1, then punitive damages are set at \$20.

The choice of punitive damages may at first sight puzzle some readers, since the economic literature on contract breach normally focuses on other sanctions, like the expectation measure or specific performance. However, that literature is mainly concerned with situations in which there is a 100% chance that the breach will be discovered. This is the case in many, if not most, contracts. For instance, if a builder promises to build me a house and he does not, it is extremely likely that I will notice it.

The efficiency-wage model focuses on those contracts in which monitoring is quite expensive, so that the apprehension rate of contract breach is less than 100%. This is a typical feature of many labor contracts. If a researcher reads a newspaper instead of doing research for which he is paid, the university might not notice it. Similarly, if a manager of a firm enriches himself in a fraudulent way, it is not sure that the shareholders will notice it in time. If the apprehension rate is lower than 100%, it is clear that compensatory sanctions do not provide high enough incentives. Since under punitive damages the sanction equals the harm in expected terms, punitive damages have the same incentive effects as normal damages in a 100% apprehension rate context.

Besides efficiency wages and punitive damages, we will consider a third alternative, which seems to have been completely overlooked in the literature⁵. We will call this alternative *conditional bonuses*. It is a positive sanction, just like efficiency wages, but unlike efficiency wages the overpayment is awarded only if monitoring takes place (and if the employee is not found shirking).

³ One remark should be made to avoid confusion. From a legal-technical viewpoint, it is possible that what we (and other economic theorists) call ‘punitive damages’ gets another name in real-world legal systems. For instance, many contract law systems will consider them as ‘penalty clauses’ (see De Geest and Wuyts, 2000). If the employer is the government and the employee a citizen or civil servant, the sanction may be called a ‘fine’ (and fall under criminal law rather than contract law). As a matter of fact, ‘punitive damages’ is a concept derived from American tort law, which is unknown (as such) in continental legal systems.

⁴ For an overview on the economic literature on punitive damages, see Polinsky and Shavell (2000a).

⁵ Shapiro and Stiglitz (1984) also briefly discuss *bonds* as an alternative sanctions mechanism. A bond is an amount of money from the employee that is posted on a blocked bank account before the contract is performed (ex ante). If the employee breaches the contract, this amount of money is paid to the employer. Damages are an amount of money that the employee pays to the employer after breaching the contract (ex post). Analytically both are equivalent, at least for our purposes. The maximum bond equals the total wealth of the employee, just as the maximum damages. Making a distinction between both is useful only when the changed wealth over time is a crucial variable, like in models on bankruptcy.

Suppose, for instance, that the employee chooses between shirking and non-shirking and that shirking gives him a benefit of +2. The market wage is 100. If the employer wants to employ efficiency wages and to adopt a 10% monitoring level (i.e. a 10% chance that shirking is discovered), then the overpayment (that the employee loses if shirking is discovered) should be at least 20 (which brings the employee's wage to at least 120). Now, let us consider the following alternative. The employer pays a wage of 100, but announces that there will be a 10% chance that the employee will be monitored. If monitoring takes place, and the employee is not found shirking, then he will receive a bonus of 20. This creates exactly the same incentives for the employee: shirking is sanctioned by a 10% chance to lose (the opportunity to receive) 20. But the expected cost for the employer is 10 times lower (if employees do not shirk): a conditional bonus has an expected cost of only 2.

This chapter is organized as follows. In section 9.2 we will elaborate our general framework (one-period model). The employer first chooses the monitoring level which minimizes his expected total enforcement expenses (i.e. the sum of monitoring costs and the expected overpayment expenses) for each sanction system. We will demonstrate that monitoring levels are set at an optimal level under punitive damages and conditional bonuses, but at a higher than optimal level under efficiency wages (section 9.3).

Next, the employer chooses the sanction type that costs least to him while satisfying three constraints: the non-shirking constraint, the maximum sanction constraint and the employee participation constraint. We will demonstrate that rational employers will always prefer punitive damages or conditional bonuses to efficiency wages (section 9.4). In the same section we will show that rational employers will in principle prefer to employ punitive damages; only when the employee's wealth (i.e. the maximum sanction available under punitive damages) is lower than a certain value, will it be profitable to employ conditional bonuses. We will determine the conditions for that shift.

In section 9.5 we will analyze combined sanctions and find that only under certain conditions will employers use combined sanctions (punitive damages and conditional bonuses) rather than punitive damages. In section 9.6 we will replace the one-period model by an n -period model. We will demonstrate that the conclusions reached under the one-period model also hold in the n -period model and show that it is rational for the employer to divide the contract in more than one period. In section 9.7 we will introduce turnover costs into the model and show that they do not bear on our findings. Finally, we will provide the reader with a note on the appropriation problem (section 9.8) and some concluding remarks (section 9.9).

9.2. The basic framework (one-period model)

An employee is hired to work for an employer during a specified period of time. The employee has to choose between shirking and non-shirking. For reasons of simplicity, we will assume that this choice is discrete: the employee fully shirks or does not shirk at all. If he shirks, he obtains a *shirking benefit* e (an avoided effort cost). e is always positive: in the absence of a sanction, the employees prefers shirking. For mathematical simplicity, we assume that if the expected sanction on shirking equals the shirking benefit, the employee is not indifferent but prefers non-shirking.

We further assume that the harm caused by shirking is so high that the employer always prefers non-shirking⁶ and always chooses a sanction-apprehension rate combination that guarantees non-shirking behavior (non-shirking constraint).

At the end of the period, the employer has the opportunity to monitor the employee. If monitoring takes place, shirking will always be discovered. The contract states the *monitoring level*, x , i.e. the chance that monitoring will take place at the end of the period (the contract is of course binding, so the employer cannot change the monitoring level ex post)⁷. The choice of the monitoring level is not discrete but continuous: x can be any value between 0 and 1. To illustrate, if $x = 1$, this means that monitoring will always take place and that the apprehension rate for shirking is 100%. If $x = 0.2$, there is a 20% chance that monitoring will take place and consequently an apprehension rate of 20%.

The *cost of a monitoring session* is constant and is denoted by m . The expected monitoring costs are mx . We also assume that the cost of a monitoring session is lower than the employer's wealth, $m < w^p$.

The contract also states the magnitude of the *sanction* s that is applied to shirking. The sanction can be the payment of damages (negative sanction), or the loss of an overpayment (positive sanction). For mathematical simplicity, we will always give a positive value to s , which will be deducted from the employee's payoffs in the case of a negative sanction, and added to the employee's payoffs in the case of a positive sanction. Except for these damages or lost bonuses there are no other sanctions for breach. More specifically, we assume that there are *no turnover costs* t , and no transaction costs caused by the payment of bonuses or damages.

⁶ Assuming that the optimal shirking level is zero is plausible for many (though not all) contracts. For instance, banks want to reduce the employee's fraud to 0. Similarly, states want judges never to be corrupt. The assumption that the optimal shirking level is 0 can be restated as the equivalent assumption that the marginal enforcement costs are always lower than the marginal net social costs of shirking – an assumption which is common in many economic models on law enforcement.

⁷ By the same token, we also assume that renegotiation between the employee and employer is not possible.

The employer sets the monitoring level x in such a way that his total enforcement costs (i.e. the sum of expected monitoring costs and expected bonus payments) are minimized. His choice, however, is limited by three constraints. First, the *non-shirking constraint*: the expected sanction must be high enough to prevent employees from shirking. Second, the *employee participation constraint*: the expected wage (taking bonuses and damages into account) of a non-shirking employee must at least equal the market wage (or price) p^m . Third, the *maximum sanction constraint*: the maximum sanction or bonus is not infinite. In the case of damages (negative sanction), we assume that the maximum feasible sanction is equal to the employee's wealth w^a . In the case of bonuses, we assume that the maximum feasible bonus is equal to the employer's wealth w^p ⁸.

All parties are assumed to be rational, utility maximizing, and risk neutral. All parties have perfect information on e , x , s , m and the market wage for labor (given the required effort level) p^m .

Notation:

- e = the employee's shirking benefit (avoided effort), $e > 0$;
- x = the probability of employer's monitoring (monitoring level); in theory $x = [0, 1]$; however, in the analysis we will determine a feasibility range $x = [x^{min}, 1]$;
- s = the sanction; $s = [0, s^{max}]$; $s^{max} = w^a$ (the employee's wealth) for punitive damages, while $s^{max} = w^p$ (the employer's wealth) for conditional bonuses and efficiency wages;
- m = the cost of a monitoring session; $m < w^p$;
- p^m = the market wage.

9.2.A. *Non-shirking constraint under efficiency wages*

Under efficiency wages, the employee's wage is the market-clearing wage p^m plus a bonus s . If monitoring takes place and he is found shirking, he loses the bonus⁹. The employee's expected net returns in the case of non-shirking are: $p^m + s$. The employee's expected net

⁸ The assumption that the maximum sanction corresponds to the maximum wealth of the party who has to pay or finance the sanction may seem to be extreme and unrealistic, but it is frequently made in the economic literature on law enforcement (Becker, 1968, at 193, Polinsky and Shavell, 2000c, at 310) and it is useful for analytical purposes. Alternatively, w^a and w^p could be defined as the maximum budget of the agent and the principal respectively that is available for sanctions. This would not change the analysis and the conclusions with respect to the superiority of punitive damages over conditional bonuses and conditional bonuses over efficiency wages.

⁹ We assume in this one-period model that when the employee is caught shirking at the end of the period he will retroactively lose the overpayment of that period. Without such retroactivity, efficiency wages would not make sense, as they would provide no incentive (whether or not the employee is

returns in the case of shirking are: $p^m+e+(1-x)s$. The employee will prefer non-shirking to shirking if: $p^m+s \geq p^m+e+(1-x)s$, which can be re-written as: $xs \geq e$ or $s \geq e/x$.

The lowest sanction that satisfies the non-shirking constraint at a given monitoring level is: $s=e/x$.

9.2.B. *Non-shirking constraint under conditional bonuses*

Under conditional bonuses, the employee's wage is the market clearing wage p^m plus a bonus s that is only paid if monitoring effectively takes place and he is found non-shirking. The employee's expected net returns in the case of non-shirking are: p^m+xs . The employee's expected net returns in the case of shirking are: p^m+e . The employee will prefer non-shirking to shirking if: $p^m+xs \geq p^m+e$, which can be re-written as: $xs \geq e$ or $s \geq e/x$.

The lowest sanction that satisfies the non-shirking constraint at a given monitoring level is: $s=e/x$.

9.2.C. *Non-shirking constraint under punitive damages*

Under punitive damages, the employee's wage is the market clearing wage p^m . He never receives a bonus. If monitoring takes place and he is found shirking, he has to pay a sanction s . The employee's expected net returns in the case of non-shirking are: p^m . The employee's expected net returns in the case of shirking are: p^m+e-xs . The employee will prefer non-shirking to shirking if: $p^m \geq p^m+e-xs$, which can be re-written as: $xs \geq e$ or $s \geq e/x$.

The lowest sanctions that satisfy the non-shirking constraint at a given monitoring level is: $s = e/x$.

9.2.D. *Participation constraint*

The employer may determine the terms of the contract (wage, monitoring level, sanction) but these terms must not be so uninteresting that no employee is willing to participate. Since the employer wants to attract non-shirking employees only, we have to take their payoffs into account when analyzing whether the participation constraint is satisfied. If the employer pays the market wage, the non-shirking employee's expected payoffs are the following:

- (i) Under efficiency wages: p^m+s .
- (ii) Under conditional bonuses: p^m+xs .
- (iii) Under punitive damages: p^m (non-shirking employees never pay damages).

caught shirking, he can keep the overpayment). In the n -period model, partially retroactive or non-retroactive efficiency wages still provide with some incentives. See section 9.6.C for a further analysis.

Since the participation constraint is satisfied if employees receive at least p^m (the market wage given the required effort level) and since x and s are always positive, the participation constraint is always fulfilled under any of these three contracts.

9.2.E. Maximum sanction constraint (creating a feasibility constraint for x)

The fact that sanctions (damages or bonuses) cannot be infinitely high restricts the range of x in the non-shirking constraint $s=e/x$. x can no longer be any value between 0 and 1, but has to be a value between the *minimum monitoring level* x^{min} and 1. The minimum monitoring level that satisfies the non-shirking constraint is found by the following formula:

$$(1) \quad x^{min} = e/s^{max}$$

Since s^{max} corresponds to the employer's wealth w^p in the case of bonuses and to the employee's wealth w^a in the case of damages, x^{min} may have a different value for positive sanctions (efficiency wages and conditional bonus) than for negative sanctions (punitive damages).

9.2.F. Normalized efficiency wages

We now introduce a fourth type of sanction, which is a variant of efficiency wages: normalized efficiency wages. Here, the employer sets the normal wage at (p^m-s) and adds a bonus s , which the employee can keep if he is not found shirking. Hence, a non-shirking employee always receives $(p^m-s)+s=p^m$, while a shirking employee receives the same if no monitoring takes place and p^m-s if monitoring takes place. Since a non-shirking employee receives p^m , the participation constraint is satisfied. The non-shirking constraint is the same as under efficiency wages, since s is the same.

At first sight, 'normalized efficiency wages' seems to be just another name for punitive damages: a non-shirking employee is given the market wage p^m while a shirking employee risks to receive less than the market wage, i.e. less than he could earn elsewhere. This means that a shirking employee risks a negative sanction. In theory, the maximum sanction under normalized efficiency wages can be the same as under punitive damages: the wealth of the employee w^a . If w^a is higher than p^m (which is extremely plausible if the period is shorter than the total career of the employee), the starting wage p^m-s would then become negative. For instance, if the wealth of the employee is 1000 and the market wage is 100, the employee would have to pay 900 to obtain the right to work for the employer, but if he is not fond shirking he would receive a bonus of 1000.

Although normalized efficiency wages are the same as punitive damages if maximum sanctions are unrestricted by the legal system, there can be a difference when regulation sets a

different maximum. For instance, a legal system could forbid punitive damages, while allowing normalized efficiency wages as long as the bottom wage (for shirking employees) is not below the legal minimum wage. If the legal minimum wage p^{\min} is lower than the market wage p^m , the maximum sanction under normalized efficiency wages is $p^m - p^{\min}$.

9.2.G. Normalized conditional bonuses

Under normalized conditional bonuses, the employer sets the normal wage at $(p^m - xs)$ and adds a bonus s , which the employee receives if monitoring takes place (probability x) and he is not caught shirking. So a non-shirking employee always receives $(p^m - xs) + xs = p^m$, while a shirking employee receives $p^m - xs$ (since he will never receive the bonus if monitoring takes place).

Since a non-shirking employee is paid p^m , the participation constraint is satisfied. The non-shirking constraint is given by $p^m = p^m + e - xs$, so that $s = e/x$. As already shown for normalized efficiency wages the maximum sanction can be negative and is given by the wealth of the employee, unless the legal system sets a minimum wage. Indeed, normalized conditional bonuses also have the characteristics of punitive damages.

9.3. Chosen monitoring levels (one-period model)

In this section, we will analyze what monitoring levels the employer chooses under each sanction, and compare them to the optimal monitoring level under that sanction.

Proposition 1. For each type of sanction, the optimal monitoring level is the lowest possible monitoring level x^{\min} i.e. the level that corresponds with the highest possible sanction.

Under punitive damages and conditional bonuses the employer chooses x^{\min} , while under efficiency wages the monitoring level will generally be higher than optimal.

From a social point of view, damages are purely redistributive: what the employee loses corresponds to what the employer gains¹⁰. This also applies to bonuses: what the employer pays corresponds to what the employee receives. Consequently, the only costs that remain to

¹⁰ Our analysis does not take activity levels into account. For instance, if employers have to pay bonuses, they buy labor at a price that is above the real costs (if we assume that market prices and market wages reflect the real costs in competitive markets). Consequently, employers will buy less labor than optimal. As a result, the socially optimal monitoring level may be higher than the minimum: the lower the monitoring level, the lower the monitoring cost (which is a good thing for society) but also the higher the substitution effects (which is a bad thing). This trade-off between monitoring costs and substitution costs does not exist for punitive damages. Under punitive damages, the employer pays the market wage to non-shirking employees. So introducing activity levels into the model would make the case for punitive damages even stronger.

be minimized are the expected monitoring cost mx . Therefore, the optimal monitoring level is the lowest possible monitoring level x^{min} .

As argued in the former section, the lowest possible monitoring level is not equal to 0, due to the maximum sanction constraint. x^{min} is the monitoring level that corresponds with the highest possible sanction¹¹.

Under conditional bonuses, the total enforcement costs, that the employer seeks to minimize, are the sum of the expected monitoring costs and the expected bonus payment costs, xs . Since as a consequence of the non-shirking constraint, $xs=e$, the expected bonus payment costs are always equal to e , which is a constant. Therefore, the employer's minimization problem is:

$$(2) \quad \min_x [mx + e]$$

The lowest value is given by x^{min} , which is the optimal monitoring level.

Under punitive damages, the employer does not have to pay any bonuses. The damages are entirely paid by the employee and are a bonus the employer receives. However, since the non-shirking constraint has to be fulfilled, the employee will never shirk and consequently never pay damages. Therefore, the costs that the employer seeks to minimize only consist of the expected monitoring costs.

$$(3) \quad \min_x [mx].$$

The chosen monitoring level will clearly be the lowest possible monitoring level x^{min} .

Under efficiency wages, the employer will minimize his total enforcement costs, which is the sum of the expected monitoring costs and the expected bonus payments:

$$(4) \quad \min_x [mx + s] = \min_x \left[mx + \frac{e}{x} \right].$$

The cost function is no longer linear, as under punitive damages and conditional bonuses, but, to the contrary, is convex¹². Let x^\wedge denote the solution. The first order condition yields

$x^\wedge = \sqrt{\frac{e}{m}}$ or $x^\wedge = \sqrt{\frac{x^{min} w^p}{m}}$, as $x^{min} = e/w^p$. It can be easily shown that $x^\wedge \leq x^{min}$ only if $x^{min} \geq w^p/m$, which can occur only if the minimum monitoring level, x^{min} , is equal to 1, as we

¹¹ The lowest possible monitoring level x^{min} , may be different for negative and positive sanctions since $x^{min} = e / s^{max}$ and s^{max} equals the wealth of the employee w^a for negative sanctions and the wealth of the employer w^p for positive sanctions. However, in this section we compare the chosen monitoring level under each sanction with the optimal monitoring for that particular sanction. Therefore, it makes sense to use the general symbol x^{min} for all sanctions in this section.

¹² The first order condition is $m - ex^{-2} = 0$. The second order condition is $2ex^{-3} > 0$.

have assumed $m \leq w^p$. Only under these extreme conditions would efficiency wages lead to efficient monitoring, while generally they will lead to excessive monitoring levels.

9.4. Chosen sanction (one-period model)

In relation to the choice of the sanction system, employer will opt for the system which yields the lowest total costs.

- (i) Under efficiency wages the employer's total enforcement costs are $mx + e/x$.
- (ii) Under conditional bonuses the total enforcement costs are $mx + e$.
- (iii) Under punitive damages the total enforcement costs are mx .

Proposition 2. Rational employers always prefer conditional bonuses to efficiency wages, unless the optimal monitoring level is equal to 1

The employer will always prefer conditional bonuses to efficiency wages for a simple reason. In spite of the fact that both mechanisms are subject to the same maximum sanction constraint, the employer's wealth, and hence to the same lower limit to the monitoring level, conditional bonuses have two advantages: first, the expected payment to the employer is always lower, as it takes place only if monitoring occurs ($e \leq e/x$, since $x \leq 1$); second, the expected monitoring costs mx are lower too, since supra we have demonstrated that monitoring levels under conditional bonuses will be lower than or equal to the monitoring level under efficiency wages. The total enforcement costs would be the same only if $x^{\min} = 1$ ¹³.

Proposition 3. Rational employers always prefer punitive damages to conditional bonuses, if the employee's wealth w^a is greater than the cost of a monitoring session m .

We have already shown that under both sanctions, the employer will choose the minimal monitoring level x^{\min} . However, the lowest possible monitoring level x^{\min} , may be different for the two cases, since $x^{\min} = e/s^{\max}$ and s^{\max} equals the wealth of the employee for punitive damages and the wealth of the employer for conditional bonuses. In the former sections we used the same symbol x^{\min} to denote both (which made sense since we compared the chosen monitoring level under each sanction with the optimal monitoring level for that particular sanction). In this section we will use two symbols to avoid confusion. $x^{\min_{dam}}$ will denote the minimum monitoring level under punitive damages and $x^{\min_{bonus}}$ will denote the minimum monitoring level under conditional bonuses. Under punitive damages, the employer's total

¹³ See proposition 1.

enforcement costs are: $mx^{\text{mindam}}=me/w^a$. Under conditional bonuses, the employer's total enforcement costs are: $mx^{\text{minbonus}}+e=me/w^p+e$.

The employer will prefer punitive damages to conditional bonuses (and this will be accepted by the other party since the participation constraint is always satisfied) if $me/w^a < me/w^p+e$. This expression can be written as:

$$(5) \quad w^a > m \frac{w^p}{m + w^p}.$$

Since we have assumed that $w^p > m$, whenever the employee's wealth is greater than the cost of a monitoring session, a rational employer will choose punitive damages. However, this is just a sufficient condition, as when the employer's wealth is very low, even an employee's wealth lower than the cost of a monitoring session justifies the use of punitive damages.

9.5. Chosen monitoring levels and sanctions under combined sanctions (one-period model)

Let us now examine the chosen monitoring level under a combined sanction. The advantage of combining a positive with a negative sanction lies in the possibility to increase the maximum available sanction and hence to reduce the monitoring level. If combined sanctions are employed, the maximum feasible sanction is given by the sum of the employee's wealth (the maximum feasible sanction in the case of punitive damages) and the employer's wealth (the maximum feasible sanction in the case of conditional bonuses and efficiency wages).

We have already demonstrated that conditional bonuses will always be preferred over efficiency wages; therefore, also in the case of combined sanctions, the employer will prefer to combine punitive damages with conditional bonuses rather than with efficiency wages.

Proposition 4. Rational employers will prefer a combined punitive-damages-plus-conditional-bonuses sanction to punitive damages only if the employee's wealth is lower than the cost of a monitoring session. They will employ punitive damages otherwise.

Corollary. Rational employers will always prefer a combined punitive-damages-plus-conditional-bonuses sanction to pure conditional bonuses.

Let s^b denote the positive sanction (bonus) and s^d the negative sanction (punitive damages). Under a combined punitive-damages-plus-conditional-bonuses sanction, the employer will minimize the sum of expected monitoring costs and expected bonus payments:

$$(6) \quad \min_x [mx + xs^b],$$

where x is the monitoring level given by $x=e/(s^d+s^b)$. Therefore, Exp. (6) is equal to

$$(7) \quad \min_{s^d, s^b} \left[\frac{e}{s^d + s^b} m + \frac{e}{s^d + s^b} s^b \right] = \min_{s^d, s^b} \left[e + e \frac{m - s^d}{s^d + s^b} \right].$$

With respect to punitive damages, the function is clearly decreasing in s^d , so that the employer will choose the maximum value of $s^d=w^a$ for punitive damages.

With respect to conditional bonuses, on the one hand, the function is decreasing also in s^b if $s^d < m$, that is if $m - s^d > 0$. Since the employer will choose always $s^d=w^a$, the condition becomes $w^a < m$. In this case the value which minimizes the total enforcement cost is the maximum sanction also under conditional bonuses $s^b=w^p$. On the other hand, if $w^a > m$, the function is increasing in s^b , so that the value of s^d which minimizes the total enforcement costs is $s^d=0$. Therefore, rational employers will choose a combined punitive-damages-plus-conditional-bonuses sanction instead of punitive damages only if the employee's wealth is lower than the cost of a monitoring session.

From this result it is possible to conclude that, since the total enforcement cost is decreasing in s^d , the total enforcement cost if $s^d=0$ will be higher than the enforcement cost for any positive value of s^d . That means that a rational employer will always prefer a combined punitive-damages-plus-conditional-bonuses sanction to conditional bonus, which proves the corollary¹⁴. In synthesis, if combined sanctions are feasible, the employer will opt either for punitive damages plus conditional bonuses or for punitive damages.

9.6. *N*-period model

In the former sections, we have assumed that the contract is to be performed in only one period. Here we will analyze the case of contracts subdivided into two or more periods and demonstrate that the conclusions reached in the one-period model are confirmed.

By more than one period, we do not mean that the contract lasts longer (for instance 30 years instead of 1 year). Since we intend to compare the characteristics of sanction systems, we shall keep the total duration of the relationship constant. Therefore, n -period will mean that the original period is divided in n shorter periods. For instance, a 30-years contract is divided in 360 periods of 1 month, and at the end of each month there is a chance that monitoring will take place (instead of at the end of the total period, i.e. at the end of the

¹⁴ Monitoring level will obviously be efficient.

career). Similarly, a one-month contract (with an opportunity to monitor at the end of the month) could be divided in 30 days with an opportunity to monitor at the end of each day.

9.6.A. The n -period model

Let e still denote the gain of a shirking employee in the whole contract and as e/n the gain in each of the n periods. Let m still denote the cost per monitoring session if the employer controls the activity of the employee over the period only at the end the contract. Let m_i denote the cost per monitoring session if the employer controls the activity of the employee at the end of each period and with respect to only that period. The cost of a monitoring session is likely to be lower if monitoring takes place with respect to the employee's performance in only one period than if it concerns the entire duration of the contract. In the best scenario $m_i = m/n$; in other cases the cost of a monitoring session in each period might be $m_i = (m/n) + c$, where c is a constant cost of each monitoring session.

The probability of control, x , will be by definition the same both if the employer monitors the whole activity only at the end of the contract and if he does it at the end of each single period¹⁵.

Let s denote the bonus that the non-shirking employee could receive in a one-period contract. Therefore, if the employer wants to keep the total bonus payment costs constant when switching from a one-period contract to an n -period contract, the bonus that the employee can receive in each of the n periods is s/n .

9.6.B. Constraints and discount factor in a n -period model

The non-shirking constraint for one-period contract was $s = e/x$. In an n -period model, the

non-shirking constraint per period i is $s_i = \frac{s}{n} = \frac{e}{n} \frac{1}{x}$, $\sum_{i=1}^n s_i = s$.

The participation constraint for a one-period contract was p^m . In an n -period model, the

participation constraint per period is $p_i = \frac{p^m}{n}$, $\sum_{i=1}^n p_i = p^m$.

The maximum sanction constraint for a one-period contract was $s^{\max} = w^a$ for punitive damages and $s^{\max} = w^b$ for conditional bonuses and efficiency wages. In an n -period model,

the maximum-sanction constraint per period is s_i such that $\sum_{i=1}^n s_i^{\max} = w^p$ for conditional

¹⁵ Imagine the employer uses a lottery system, which randomly selects 5 out of the 20 employees to be controlled. If the monitoring session takes place at the end of a one-year contract each employee has a 25% probability to be controlled. Similarly, if the monitoring takes place every month, each employee has a 25% chance to be checked each month.

bonuses and efficiency wages, while it is w^a for each period in the case of punitive damages since a non-shirking employee will not pay damages at any period.

Furthermore, let θ denote the interest rate for each period, so that $\delta = 1/(1+\theta)$ is the net present value in period i of a unitary gain in period $i+1$ (discount factor). $\delta, \theta \in (0, 1)$.

9.6.C. *Efficiency wages and retroactivity*

A little clarification is worthwhile with respect to the implementation of efficiency wages in an n -period model. In particular we should clarify what is the magnitude of the sanction (foregone bonus) that the employee pays if found shirking. There are three possibilities.

(i) *Full retroactivity*. If the employee is found shirking he loses the overpayment with respect to all the previous periods of the contract. For example, if he is found shirking in the 12th period, he will have to return the overpayment received at the end of all previous 11 periods. This is, however, the equivalent of saying that there is no overpayment in the first 11 periods and one big bonus at the end. In this case conditional bonuses are superior to efficiency wages for the same reason as in the one period model. The only difference between a conditional bonus applied at the end of the contract and fully retroactive efficiency wages is that under the latter the employee also receives the bonus in the case that no monitoring takes place.

(ii) *Partial retroactivity*. If the employee is found shirking, he does not receive the overpayment for the last period but can keep the overpayments relative to the earlier periods. For example, if he is found shirking in the 12th period, he will not have to return the overpayment received at the end of all previous 11 periods, but will not receive any overpayment for period 12. Also this solution is inferior to conditional bonuses, for the analysis of the one-period model proves that conditional bonuses would ensure the same incentive effect at a lower cost for the employer (non monitored employee would not be rewarded) in any period.

(iii) *No retroactivity*. If the employee is found shirking he only loses the opportunity of receiving an overpayment in the next period, but can keep the overpayment relative to the current period. For example, if he is found shirking in the 12th period, he will not have to return any overpayment, will still receive the overpayment for period 12, and will only lose the opportunity to receive the overpayment in period 13 and in the following periods. This solution is inferior to conditional bonuses for a stronger reason than before: in the last period the employee has no incentive to work.

9.6.D. Sanctions and monitoring level in an n -period model

We have already noticed in section 9.6.C that conditional bonuses are superior to efficiency wages in an n -period contract because they are superior in any of the single periods, given any retroactivity rule for efficiency wages. Therefore, our contention is that conditional bonuses will still be preferred in a n -period contract.

Here we will compare punitive damages and conditional bonuses and show that parties will prefer punitive damages, or a combination of the two also in an n -period contract. Several elements bear on our analysis, as the division of the contract in periods might increase the efficacy of both punitive damages and conditional bonuses in relation to the possibility to lower monitoring levels.

Under punitive damages, since a non-shirking employee does not pay damages, the maximum sanction for each period is the total wealth of the employee. However, the shirking gain is equal to e/n . From the non-shirking constraint we have $x=e/s$. In a one-period contract the employer will minimize $\min[mx]$, $x^{\min}=e/s^{\max}$, while in an n -period he will minimize at each period $\min[x_i m_i]$, and $x_i^{\min}=(e/n)(1/s)$. So that $x_i^{\min} = \frac{e}{s^{\max} n} = \frac{1}{n} x^{\min}$. Since the non-shirking employee does not pay damages, this lower monitoring level can be hold constant over the n periods. The division of the contract in periods works as a multiplier for the maximum sanction¹⁶.

Under conditional bonuses (and the inferior alternative of efficiency wages), if the sanction is constant over the contract, the monitoring level in an n -period contract is the same as in a one-period contract. In fact, in the last period of an n -period contract, the non-shirking constraint requires $s_n=s/n=(e/n)x$. Since x is constant over the contract, if s is assumed to be constant, then also in any other period $s_i=s/n=e/n.x$, for any $i=1,2,\dots,n$. The total sanction to

be applied is $\sum_{i=1}^n s_i = \sum_{i=1}^n \frac{e}{nx} = \frac{e}{x} = s$.

It follows that the monitoring level, $x=e/s$, will be the same under an n -period contract as under a one-period contract¹⁷.

¹⁶ We have assumed that the benefit that the employee receives from shirking in one period is n times smaller than the benefit of shirking during the entire contract. It is worthwhile to notice that our argument does not crucially depend on this assumption. It is sufficient that the shirking gain in one period is lower than the shirking benefit concerning the entire contract for the monitoring level to be lower if the contract is divided in periods. By dividing the contract in periods, the employers can reduce the monitoring level, but might bear a higher total monitoring cost, if the sum of the monitoring costs for single periods is larger than the cost of only monitoring at the end of the contract. This factors bear on the decision of whether or not to divide the contract in periods, which is not our concern here and does not affect our discussion with respect to monitoring levels.

¹⁷ Note that if the shirking benefit in one period were higher than e/n (see footnote 16) the monitoring level in an n -period contract with constant sanctions would be higher than in a one-period contract.

If the sanction can vary, in the last period the non-shirking constraint is $s_n = s/n = (e/nx)$, while in the second but last period the employee will face a potential sanction at the end of the period and, since he will be dismissed, a potential sanction also for the last period, that he discounts by δ . Therefore the shirking constraint in the second but last period is $s_{n-1} = (e/nx) - \delta s_n$. Since $(e/nx) = s_n$, then $s_{n-1} = s_n(1 - \delta)$. In general we can define the shirking constraint in the case of variable sanctions as $s_i = s_n(1 - \delta)^{n-i}$, for any $i = 1, 2, \dots, n$.

Therefore, sanctions for earlier periods of the contract can be lower than s_n , although the level of employee's effort will be the same in any period. Consequently, the same total sanction can be redistributed over the contract, so that the sanction increases for later periods (as they also have incentive effects on the preceding periods), and the monitoring level decreases.

If $\delta = 1$, the sanction will be zero in all the periods but the last, in which the maximum sanction will be applied. By applying the maximum sanction only at the last period so that $s_n = s^{max}$, rational employers will lower x by n times. Nevertheless, incentives will be produced for all the preceding periods as well, as a dismissed employee loses the opportunity to obtain a future benefit. In this case conditional bonuses and efficiency wages improve as punitive damages due to the division in n periods. If $\delta = 0$, the total sanction under the n -period model will be equal to the sanction in the one-period contract, and so will the monitoring level.

From the former reasoning, it is clear that the division in n periods multiplies the maximum sanction by n under punitive damages, and, if the discount factor is equal to 1, also under conditional bonuses and efficiency wages¹⁸. That means that they can be compared as in the case of one-period contracts and the result will be the same. If the discount factor gets smaller than 1, the gain of the division in n periods decreases for conditional bonuses and efficiency wages, while it remains constant for punitive damages. Therefore, for low values of the discount factor, the argument for punitive damages becomes stronger.

An additional argument which might be made against positive sanctions is that the incentive effect of future payments disappears if the same future benefits are promised by other employers in the market. In this case only the positive sanction applied at the end of the current period would have effects. Therefore, the conclusions reached for the one-period apply also in the n -period model.

¹⁸ Both conclusions would be weakened if the shirking benefit per period were higher than e/n , but their relative balance would be unchanged.

9.7. Turnover costs

In this section we will introduce turnover costs into the model and demonstrate that they do not alter our conclusions.

Let t denote the turnover costs, the employee's costs due to dismissal. Turnover costs can be seen as an additional negative sanction. The legal system may change the allocation of turnover costs, for instance by requiring the employer to pay damages to the employee (that compensate turnover costs). In this section we assume that turnover costs are transaction costs borne by the employee¹⁹.

Turnover costs increase the maximum sanction available under all types of sanctions by the same amount and do not change the conclusions already reached. If turnover costs are greater than zero, the minimum monitoring level becomes $x^{min}=e/(s^{max}+t)$. Under punitive damages the enforcement costs $mx=e.m/(s^d+t)$, are still minimized by the maximal sanction w^a ²⁰. Therefore, the employer will choose the minimum monitoring level also in the case of positive turnover costs.

Under conditional bonuses the enforcement costs $mx+xs^b=em/(s^b+t)+e$, are still minimized by the maximal sanction w^p . Therefore, the employer will choose the minimum monitoring level also in the case of positive turnover costs.

Under efficiency wages the enforcement costs are $mx+s^w=em/(s^w+t)+s^w$, and it can be shown that the employer will always choose a lower sanction than maximum and, therefore the monitoring level will always be higher than optimal, as in the case of $t=0$.

As already shown, conditional bonuses will always be preferred over efficiency wages, while punitive damages will only be preferred to conditional bonuses at the conditions mentioned in proposition 3.

Suppose now that punitive damages are not allowed by the legal system, but the legal system allows turnover costs to be borne by the employee. In that case the maximum

¹⁹ Under this interpretation turnover costs are a waste of resources (a destructive sanction), as the cost to the employee does not correspond to a benefit for anyone else. They might also take the form of a transfer from the employee to the employer (a redistributive sanction); however, in this case they would simply be a different technique to employ punitive damages, hence our former analysis would apply.

²⁰ Strictly speaking, since both turnover costs and punitive damages are borne by the employee and the sanction is already at its maximum, monetary turnover costs do not increase the maximal feasible sanction but decrease the maximum damages applicable to the employee, as a portion of his wealth will be eroded by turnover costs. Hence the maximum sanction remains constant. Non-monetary turnover costs, however, might be borne over the limit of the maximum (monetary) negative sanction. Nevertheless, our conclusions do not change even if monetary turnover costs are considered. In particular, since we have shown that parties will have an incentive to employ either punitive damages or punitive damages combined with conditional bonuses, their decisions concerning the optimal combination would either be unchanged (as monetary turnover costs leave both the maximum positive

punitive-damages sanction is not w^a but t . Our conclusions are quantitatively affected but not qualitatively: if only positive sanctions are allowed, rational employers will prefer a combined punitive-damages-plus-conditional-bonuses sanction to punitive damages only if the turnover costs are lower than the cost of a monitoring session. Efficiency wages will never be chosen.

9.8. A Note on the Appropriation Problem

Shapiro and Stiglitz (1984) briefly discussed ‘bonds’²¹ as an alternative to efficiency wages and remarked that bonds may not be feasible because of the appropriation problem: the employer can falsely pretend that the employee shirked and might consequently refuse to pay (appropriate) the bond. The questions that we pose here are whether efficiency wages are less vulnerable to the appropriation problem than the two alternatives we have considered, and whether it is possible to find a solution for this problem.

All alternatives face the same appropriation problem. In the case of efficiency wages, at the end of the period the employer might increase the monitoring level x (e.g. he sets it to 100%) and might falsely declare that the employee shirked. In the case of conditional bonuses, at the end of the period the employer might reduce the monitoring level to zero or falsely declare that the employee shirked. In the case of punitive damages, at the end of the period the employer might increase the monitoring level (e.g. he sets it to 100%) and might falsely declare that the employee shirked. Thus, the appropriation problem seems to affect all the sanction systems we have analyzed to the same extent.

In any case, the appropriation problem can be solved by ‘decoupled liability’ (Polinsky and Che, 1991). Decoupled liability means that what the injurer (employee) pays differs from what the victim (employer) receives. In the case of negative sanctions, the damages do not go to the employer but to a third party (for instance, a charity fund, or even better for the parties involved, the non-shirking employees of the same firm). An example can be found in professional soccer teams: if a player arrives late at the training, he has to pay a fine that is used to offer all players things like a Christmas dinner.

In the case of positive sanctions, the bonuses are not financed by the employer but by all other employees of the same firm. Since the expected net wage for an employee (the market wage p^m plus the expected bonuses minus the expected bonus payment expenses) still equals the market wage p^m (the expected bonuses equal the expected bonus payment expenses), this bonus financing technique has no influence on the employee participation constraint.

and negative sanctions unchanged) or be biased towards punitive damages (as non-monetary turnover costs increase the maximal negative sanction and limit the need to combine it with a positive sanction).

²¹ See Shapiro and Stiglitz (1985) for further discussion on these issues.

Alternatively, the employer could create a ‘bonus fund’, and divide in each period a fixed amount among all non-shirking employees, so that declaring one more employee ‘shirking’ does not reduce the bonus payment expenses of the employer.

Ex post fraud by the employer as to the effective monitoring level x can also be solved by letting a third party organize the ‘lottery’ that decides whether monitoring will effectively take place. The contract with that third party can state that the chances of the ‘lottery’ are the one defined in the contract (x). It can also be solved by letting a third party monitor and ex ante determining the budget of that third party (which determines the monitoring level x).

9.9. Concluding Remarks

9.9.A. *Implications for contract law*

In relation to *contract law* issues, we argue that negative sanctions for breach are generally superior to (non-normalized) positive sanctions.

9.9.B. *Implications for labor law*

With respect to *labor economics*, we argue that efficiency wages are less rational a choice than generally accepted and that it is rather unlikely that they can explain unemployment. In this chapter, we argue that only under very implausible assumptions do efficiency wages make sense. Rational parties will nearly always prefer negative sanctions to positive sanctions. In the few cases where positive sanctions will be preferred, parties will choose conditional bonuses (an alternative positive sanction overlooked in the labor economics literature) rather than efficiency wages. And even if efficiency wages were chosen, parties would try to work with normalized efficiency wages, which do not result in overpayment. Therefore, it seems extremely implausible that employers will ever opt for non-normalized efficiency wages on such a massive scale that structural unemployment would result.

9.9.C. *Implications for corruption issues*

Finally, our analysis suggests that the overpayment of judges and other public officials is not a rational method to conquer *corruption* as long as criminal sanctions are less than the maximum. The first best solution is provided by criminal sanctions on corruption (equivalent of punitive damages in our model).

9.10. References

- Akerlof, George (1982), 'Labor Contracts as Partial Gift Exchange', **97** *Quarterly Journal of Economics*, 543-569.
- Becker, Gary S. (1968), 'Crime and Punishment: An Economic Approach', **76** *Journal of Political Economy*, 169-217.
- Cooter, Robert D. and Garoupa, Nuno (2000), 'The Virtuous Circle of Distrust: A Mechanism to Deter Bribes', working paper.
- De Geest, Gerrit and Wuyts, Filip (2000), 'Penalty Clauses and Liquidated Damages', in Bouckaert, Boudewijn and De Geest, Gerrit (eds.), *Encyclopedia of Law and Economics*, Cheltenham, Edward Elgar, vol. III, 141-161.
- Fella, Giulio (2000), 'Efficiency Wage and Efficient Redundancy Pay', **44** *European Economic Review*, 1473-1490.
- Polinsky, A. Mitchell and Che Yeon-Koo (1991), 'Decoupling Liability: Optimal Incentives for Care and Litigation', **22** *Rand Journal of Economics*, 562-570.
- Polinsky, A. Mitchell and Shavell, Steven (2000a), 'Punitive Damages', in Bouckaert, Boudewijn and de Geest, Gerrit (eds.), *Encyclopedia of Law and Economics*, Cheltenham, Edward Elgar, vol. II, 764-781.
- Polinsky, A. Mitchell and Shavell, Steven (2000b), 'Corruption and Optimal Law Enforcement', *Journal of Public Economics*, forthcoming.
- Polinsky, A. Mitchell, and Steven Shavell (2000c), 'Public Enforcement of Law', in Bouckaert, Boudewijn and de Geest, Gerrit (eds.), *Encyclopedia of Law and Economics*, Cheltenham, Edward Elgar, vol. V, 307-344.
- Rasmusen, Eric (1992), 'An Income-Satiation Model of Efficiency Wages', **30** *Economic Inquiry*, 467-78.
- Shapiro, Carl and Stiglitz, Joseph E. (1984), 'Equilibrium Unemployment as a Worker-Discipline Device', **74** *American Economic Review*, 433-444.
- Shapiro, Carl and Stiglitz, Joseph E. (1985), 'Can Unemployment be Involuntary: Reply', **75** *American Economic Review*, 1215-1217.
- Shavell, Steven (1997), 'The Optimal Level of Corporate Liability Given the Limited Ability of Corporations to Penalize Their Employees', **17** *International Review of Law and Economics*, 203-213.

CONCLUSIONS

This work is a search for general principles in the field of the economic analysis of law. The application of microeconomic theory to the study of law has presented legal scholars with helpful insights into the way in which the law shapes people's behavior. This has induced law and economics scholars to "go wider". As a result, law and economics analysis has pervaded all fields of law.

My aim has been to "go deeper". I have tried to delve into the foundations of law and economics and, for this reason, the main grounds of my analysis have been the most general and those that have been studied the most: torts and contracts. The analysis has focused on the fundamental characteristics of incentive streams created directly by the legal system by means of tort law and contract law, and by private parties within the legal system by means of contracts.

A stream of water is something physical and tangible, which runs and carries things along with it, which, on the one hand, can be destructive or, on the other hand, can be used to make a watermill work, depending on many factors and not only on how much water we allow to flow. The same waterfall can demolish a dam or produce electricity.

Similarly, the same incentive stream may produce different effects on people's behavior depending on the way it is handled. Neither does a general theory that explains how incentive streams work exist, nor I have attempted to provide one here. My task in this research has rather been a limited one. I have asked some still-unanswered (and sometimes not even asked as yet) questions about the general characteristics that incentive streams happen to have in some legal situations, in order to achieve two goals: first, to demonstrate that a more in-depth look at law and economics theory is possible and worthwhile and, second, to contribute to the establishment of such a deeper level of law and economics analysis¹.

The introduction to this thesis has already provided the reader with a chapter-by-chapter overview of the contents of my research. I will provide here a cross-sectional comment on the results, their relation to the literature on law and economics and their implications for policy makers. Therefore, in section 10.1 I will present the logic governing the results of my research; in section 10.2 I will discuss their relation to the literature and underline the original aspects of my approach; in section 10.3 I will account for the policy implication of my findings.

¹ See the introduction for a broader discussion on the aim and the scope of this research.

10.1. Six general characteristics of incentive streams

In this section, I will focus on six² general characteristics of incentive streams created by legal systems, such as they have emerged from the analysis of the previous chapters, which have been centered on torts and contracts. Here, I will present them in a very general and somewhat metaphorical way, in order to demonstrate that those findings may have a very broad application in law and economics and draw the reader's attention in particular to their logic.

I will first focus on two major problems: How can a single incentive stream be designed in order to produce effects on two people? How does the effect of the incentive stream change if parties are partially shielded against it? These questions attempt to reveal the ultimate mechanisms of legal incentives. In tort law, they refer to the setting of the appropriate negligence criterion and to the effect of insolvency on incentives. In order to answer those questions I have inquired into the historical evolution of liability systems, the choice of liability rules in different areas of accident law and the use of tort law and regulation as means of dealing with insolvent injurers.

Later, I will take into consideration the problem of how the legal system can delegate the administration of incentive streams to private parties by means of vicarious liability. I will ask which forms vicarious liability should take in different situations. In order to answer this question I have inquired into modern and Roman law.

At last, I will explore the field of privately generated incentive streams and ask the question of which incentive devices parties to a contract would rationally prefer. In order to answer this question I have inquired into the fields of labor law and labor economics.

10.1.A. Characteristic I: "One bullet for two gangsters"³

A single incentive stream can be fully directed towards one party (the residual bearer) and to a variable extent towards a second party at the price of controlling that party directly. The extent of such a control depends on its costs and benefits.

The first problem has long been debated in law and economics and seems to have been solved by finding that a single (tort law) incentive stream gives both parties incentives to take optimal care⁴ but only one of them incentives to choose the optimal activity level⁵, if a

² The reader might be puzzled, for in the introduction I listed seven questions while here I account only for six characteristics. The reason is that my fourth question deals with the problem of how to reduce the effect of insolvency on incentives. Therefore, the analysis thereof is rather relevant for what concerns policy making (an aspect dealt with in section 10.3 of this chapter) than for the theoretical discussion on the general characteristics of incentive streams, which we are embarking on in this section.

³ See chapter 3.

⁴ Landes and Posner (1980).

⁵ Shavell (1980).

negligence criterion is applied. Hence, the inability to control the activity level has been traditionally seen as a failure of the tort law system.

I have demonstrated that this is not true. In theory, the legal system could control both parties' care and activity level, but deliberately does not do so, for that would frequently be too expensive a solution. One party is the residual bearer of the accident loss, the one who bears the loss if both parties behave non-negligently: this party has an incentive to take optimal care and activity level. The other party is controlled through the negligence criterion: he pays the accident loss only if found negligent. This party will only take the precautionary measures required by the negligence criterion (namely, care) and no more. Those precautionary measures which are excluded from the negligence criterion can be defined as the activity level.

The legal system decides how many precautionary measures should be omitted from the negligence criterion, i.e. it decides how large the activity level should be. In a world without transaction costs, the legal system could opt for a fully comprehensive negligence criterion, where the non-residual bearer is fully controlled and no activity level exists. In this situation both parties would take optimal precaution.

Unfortunately, on the one hand, observing parties' precaution is an expensive exercise, and the more precautionary measures are included in the negligence criterion, the more expensive that becomes. On the other hand, if some precautionary measures are not controlled, accident prevention will be lower than optimal. Therefore, the courts will control the behavior of the non-residual bearer up to a certain point only, where the marginal cost of observing one more precautionary measure is equal to the marginal benefit (accident prevention closer to the optimal level). Therefore, the fact that some precautionary measures (those measure referred to as activity level) are not controlled under the negligence inquiry does not represent a failure of the liability system; on the contrary, it is the result of an efficient balance.

The theoretical consequence of my model is that care and activity level should not be seen as two different types of precaution, as they have always been described in the literature. On the contrary, they are two subsets of the set of parties' precautionary measures: care is the subset comprising the precautionary measures included in the negligence criterion, while the activity level is the subset comprising the precautionary measures omitted from it.

In addition, those two concepts are not determined by nature. The legal system defines them by means of the negligence criterion; therefore, they are endogenous. We do not have to assume a distinction between activity level and care but that distinction is produced by the application of the negligence criterion itself, it is indeed produced by the legal system. That

means that the distinction is variable and can be manipulated by the legal system by enlarging or restricting the negligence criterion.

It is interesting to look at the problem of defining care and activity level from a different perspective, by using a metaphor.

The incentive stream created by a tort law rule is fully directed towards one party only (e.g. the victim of an accident): like a gun pointing at a gangster. The legal system does not do much about it; it simply points the gun (i.e. lets that party pay the accident loss). If the legal system wants to create an additional incentive for a second party (e.g. the injurer) by using the same stream, it does so by creating a threat: the second party is sanctioned if found negligent. That party will react by being non-negligent, and therefore the incentive stream will never have to be deviated towards him (the state never shoots). However, the threat is a costly device, as it presupposes that the legal system controls what the party does, in order to decide on the issue of negligence. The literature has always considered the size of such a threat (i.e. the extent of the negligence criterion) as being fixed. I have shown that this is not the case, and the legal system itself determines how much control the second party should be subjected to.

*10.1.B. Characteristic II: "Two gangsters in the dark: shoot either one or point in the middle and wait?"*⁶

In some situations, dividing a single incentive stream into two half streams directed towards two different parties might be more desirable than directing the full original stream towards either party, for the former strategy can have a filtering effect on violations, which prevents the most harmful of them and allows the least harmful to occur.

A second well established result is that any division of the accident loss among the parties to an accident produces the same effect with respect to incentives⁷. I have analyzed the properties of different rules for sharing the accident loss among the parties in an imperfect information environment.

The situation is similar to the one experienced by a one-bullet policeman who notices two potential gangsters approaching in the dark: one of them has a weapon, while the other does not. What is the best strategy for the policeman? He might fire at either of the gangsters in the darkness: he has a 50% chance to stop the armed one and a 50% chance to stop the unarmed and be killed by the other. The second possibility is to wait until they come closer, so that he can see who has the weapon. In this case he can fire at the right one. Nevertheless, there is still a possibility that he will wait too long and will have no time to fire at either one.

⁶ See chapter 4.

My analysis has explored virtues and shortcomings of the two strategies and has shown that a wait-and-see policy (pointing the gun in between the two gangsters) can be a winning strategy in most cases because it has a filtering effect: the greater the danger, the sooner the policeman will realize it, and the sooner he will shoot. Therefore, under a wait-and-see strategy our policeman is likely to miss the target only if the danger is not so great and evident, while under a shoot-either strategy he will miss the target in 50% of the cases, without any selection being made between less-dangerous gangsters and more-dangerous ones.

Such a finding provides a theoretical support for those legal rules which allow the court to share the accident loss in unclear cases, such as contributory negligence and the rule of contribution in the case of multiple tortfeasors.

*10.1.C. Characteristic III: "Shooting bulletproof gangsters"*⁸

When parties are partially shielded against sanctions, incentive streams have a limited effect. If parties can control the probability of the sanction, the incentive stream is undermined. If parties can control the magnitude of the sanction, the outcome is different: the incentive stream either remains intact or is completely neutralized.

The problem of insolvency has found an almost unquestioned formulation in the judgment-proof effect: if injurer's liability is limited, he will take less precaution⁹. I have demonstrated that such a result is strictly dependent on the mathematical form¹⁰, which has been given to the problem. By using a more general model, I have shown that this conclusion is not always true: optimal precaution will be the outcome in some situations, while no precaution at all will result in other situations.

The problem is similar to questioning whether a gangster who is wearing a bulletproof jacket will be frightened by the policeman's gun. The inability to pay damages is indeed a sort of judgment-proof garment, as it shields injurers against liability. The literature has always answered: he will be less frightened than a gangster without a bulletproof jacket. This result looks intuitive but it is slightly misleading. Let us look at the jacket.

If the jacket is such that it reduces the harm caused to the gangster by any bullet, the gangster will be less frightened indeed, as the effect of the police's fire is reduced: instead of being killed he will be wounded. This type of jacket corresponds to the probability models,

⁷ Landes and Posner (1980), Cooter and Ulen (2000).

⁸ See chapters 5 and 6.

⁹ Shavell (1986).

¹⁰ The literature has until now employed a probability model. I have analyzed the employed also a magnitude model and two mixed probability-magnitude models.

which have been employed in the literature until now in order to describe judgment-proof situations: judgment-proof injurers take less than optimal precaution.

Nevertheless, we can consider a different bulletproof jacket (i.e. a different model of accident prevention, namely a magnitude model), a jacket which stops all bullets under a certain caliber, but is completely ineffective for bullets above that critical caliber. If the policeman shoots, the gangster will either be killed or will remain uninjured. Everything depends on the resistance of the jacket. If it is sufficiently low, the gangster will be just as frightened as if he had no jacket. If it is very high, the gangster will not be frightened at all. In this case we observe a binary outcome.

Similarly, a magnitude model predicts that judgment-proof injurers will take either optimal precaution or no precaution at all, a very dramatic outcome, and not always simply underprecaution as in probability models. The results have also been extended to mixed probability-magnitude models and prove that incentive-proof injurers do not all react to incentives in the same way: their reaction depends on whether they can reduce the probability of an accident or the magnitude of the harm. Therefore, the study of incentive-proof situations has to take the characteristics of the “shield” into account, as indicated in my analysis.

10.1.D. Characteristic IV: “Are gangsters easier to capture if the police have orders to kill?”¹¹

If the investment in compliance reduces the sanction, even judgment-proof parties have an additional incentive to comply. However, this increases parties’ compliance only when parties control the probability of the sanction, while it has no effect when parties control the magnitude of the sanction

Imagine a gangster trying to escape from the police. The police have orders to arrest him alive, and he will receive a shorter imprisonment term if he turns himself in spontaneously. He faces a dilemma: the more policemen he kills the greater the probability to escape, but the longer the imprisonment term. The possibility to have a shorter sentence might make him less brutal and the police’s task easier.

Judgment-proofness might create a similar dilemma for injurers, because in some cases the expenditure on precaution reduces the assets available to pay damages.

The judgment-proof effect creates two contrasting implicit subsidies: first, it subsidizes a part of the harm, that part which exceeds injurer’s assets, and hence provides the injurer with an incentive towards less precaution; second, it subsidizes some precaution costs, as the expenditure on precaution reduces the assets available for damage compensation, and hence provides the injurer with an incentive towards more precaution. The stronger subsidy (in

¹¹ See chapter 7.

marginal terms) will prevail. This means that judgment-proof injurers might even take overprecaution¹², if the precaution subsidy prevails over the harm subsidy.

The former result holds in probability models, where the injurer can reduce the probability of an accident by taking more precaution (and gangsters can reduce the probability of being arrested), while the magnitude of the harm is exogenous.

Let us consider again the same situation as before, but now the police have orders to kill. It is obvious that the possibility of obtaining a shorter imprisonment term, does not produce any benefit for a dead gangster...hence, he will not be less brutal because of this.

This situation is similar to the one in which the injurer can only reduce the magnitude of an accident. As he cannot reduce the probability of the accident, once he is bankrupt (dead) any investment in precaution is rendered worthless. Hence, there is no extra incentive towards precaution due to bankruptcy itself. In magnitude models the harm subsidy always prevails over the precaution subsidy: the incentive towards less precaution is stronger than the incentive towards more precaution.

10.1.E. Characteristic V: “Are parents too strict with badly behaved children?”¹³

When a party is shielded against sanctions, the incentive stream can be directed towards a second party, which has control over the first party’s compliance. If the second party bears both the sanction and the cost of compliance, the outcome is efficient. However, if the second party bears only the sanction, the outcome will be over-compliance. Such inefficiency can be corrected by controlling first party’s compliance, or second party’s enforcement through negligence, or by limiting the sanction to the second party.

Parents bear the cost of accidents caused by their children, as they have to pay for them. This is the same as saying that parents receive the benefit of their children’s precaution. However, parents do not bear the full cost of such precaution, as they do not pay their children for being careful; hence, they might be too strict.

Paradoxically, it would be a sound economic proposal to oblige parents to reward their children for being careful, in order to induce a perfect internalization of costs and benefits and obtain optimal precaution. It sounds a less anomalous solution, if we think that employers actually do so: they also pay for their employees’ efforts and not only for the accidents caused by them.

My analysis can explain why parents and supervisors are only usually liable if negligent for the damages caused by their children and persons under their supervision, while

¹² Beard (1990).

¹³ See chapter 8.

employers are generally strictly liable. Negligence has the effect of reducing the incentives towards excessive precaution, because a non-negligent injurer never pays damages.

I have shown that a limit on injurer's liability can also be profitably implemented in order to solve overprecaution problems. Roman law gives an example of such technique with the rule of *noxae deditio*, which reduced *paterfamilias*' liability to the monetary value of the tortfeasor. A cap on *paterfamilias*' liability reduces the incentive towards overprecaution, and it is simpler to implement than a negligence rule, even if less accurate.

*10.1.F. Characteristic VI: "Should broken promises be sanctioned or should kept promises be rewarded?"*¹⁴

In a contract, both parties will prefer an incentive stream based on sanctions rather than on rewards, if the victim finances the reward.

Parties to an agreement will prefer sanctions to rewards. The reason is that both sanctions and rewards will ensure that the promisor adheres to the promise; however, in such a case sanctions will never be paid, while rewards would be.

I have analyzed incentives between parties to a labor contract and have discussed the well-known problem of incentives to provide effort when effort is expensive to monitor. The analysis provides a counter-argument to the efficiency wage theory, which predicts that parties will adopt rewards rather than sanctions, and namely a specific kind of reward (higher wage) that will also be paid in the case of shirking. I have shown that, if other incentive streams are available, parties will never choose efficiency wages.

10.2. Relation to the literature

This work is an attempt to analyze the way incentive streams channel people's behavior in various circumstances. In this respect, it responds to a strongly felt need to go beyond the traditional boundaries of law and to provide a general theory. My contribution highlights some characteristics of incentive streams that until now have been overlooked or misinterpreted by the literature. Although my aim is to generalize the analysis, my results have an impact on several specific areas of law and economics, which have long been debated in the literature.

In this section, I relate my findings to the literature and try to draw the attention of the reader towards the differences between my approach and the traditional one in several areas of research.

¹⁴ See chapter 9.

10.2.A. Activity level vs. care¹⁵

The definition of the concepts of care and activity level is central in tort law and economics. Surprisingly, after 40 years of economic analysis of tort law¹⁶ and 20 years after Shavell wrote his seminal article on strict liability vs. negligence¹⁷, we still do not know what activity level and care mean.

The distinction was introduced by Shavell, who defined the two concepts as precaution that reduces the expected accident loss (care), and as the number or frequency of a repeated risky action (activity level). These definitions have been highly successful and have become the basis of many subsequent analyses.

In chapter 3 I have demonstrated that those concepts are tautological and have proposed a more precise definition thereof: care as precautionary measures included in the negligence inquiry, activity level as precautionary measures omitted from the negligence inquiry. My approach differs from the traditional one in three respects. First, the literature considers precaution as a single homogeneous action. I consider precaution as being composed of many different precautionary measures. For example a car driver can reduce his speed, drink less alcohol, control the quality and the condition of the tires, have a long rest before driving, and so forth. The legal system observes ex post these precautionary measures at different costs. Checking the speed of a car is, for example, easier than ascertaining whether the driver had had a sufficient sleep the night before. In this respect, my model captures more of reality than the traditional one.

Second, my definitions are endogenous to the model, while the traditional definition is an assumption of the model itself, and is therefore exogenous to it. Endogeneity brings along two theoretical advantages: first, the model is based on fewer assumptions, and hence it is more flexible and theoretically more powerful, as it requires less unproved points to begin with; second, the definitions I propose are mutually exclusive and non-tautological.

Third, my definitions are variable, merely because they are endogenous. They vary with the characteristics of the model and, therefore, they account for changes overtime and over different legislations. An historical analysis of tort liability systems can be brought into the model by no need of external assumptions, as I will explain in a coming section.

Apart from these three differences, my model adheres to the lines already traced by the traditional approach and confirms the results thereby obtained. However, it proves to be a better tool to explain some previously unsolved puzzles concerning the choice of the most suitable rule in practical cases, by providing a precise economic criterion to compare

¹⁵ See chapter 3.

¹⁶ See Calabresi (1961) and Trimarchi (1961).

¹⁷ See Shavell (1980).

(information) costs and (allocative) benefits of tort law rules, as I will show in the next section.

10.2.B. *Strict liability vs. negligence*¹⁸

The literature¹⁹ has traditionally approached the problem of how to choose between strict-liability-based rules (traditionally referred to simply as “strict liability”, the injurer is the residual bearer) and no-liability-based rules (traditionally referred to simply as “negligence”, the victim is the residual bearer) by applying the dangerousness criterion.

“If it is more important to control injurers’ level of activity than victims’” the rules that result in greater social welfare are strict-liability-based rules²⁰. The legal system should otherwise adopt no-liability-based rules. In turn, the “importance” of controlling one party’s activity can be seen in terms of riskiness or dangerousness of that party’s activity²¹.

This criterion has provided a classical and broadly accepted explanation for the choice between strict liability and negligence. However, not all strict liability regimes can be justified in terms of dangerousness. “Is the chance of a wild animal escaping from a zoo and doing harm, for which strict liability would probably result in the United States, greater than that of an automobile running down a pedestrian, for which the negligence rule would govern?”²². The criterion is indeed “somewhat rough”²³. Many actual legal solutions are unclear, as “the choices made between strict liability and negligence rules are not always easy to explain on the basis of differences in riskiness”²⁴.

I have provided an additional explanation: the choice of the liability rule can be justified in terms of verifiability of the parties’ precautionary measures. The reason is that if the precautionary measures of one party are particularly easy to verify, the application of the negligence inquiry to that party will be particularly effective, as the negligence criterion can cover almost all his precautionary measures and the measures omitted from it (activity level) will only be few. Therefore, the verifiability criterion urges us to choose as the residual bearer the other (less verifiable) party.

In chapter 3, I have demonstrated that when strict liability is applied to the regulation of non-dangerous activities, the reason for its implementation resides in the scarce verifiability of the injurer’s precautionary measures.

¹⁸ See chapter 3.

¹⁹ See Shavell (1980) and Schäfer and Schönenberger (2000).

²⁰ Shavell (1987) at 29. See also Landes and Posner (1987) at 70.

²¹ Shavell (1987) at 31-32.

²² Shavell (1987) at 31.

²³ Shavell (1987) at 32.

²⁴ Shavell (1987), at 31-32. See also Landes and Posner (1980) and Schäfer and Schönenberger (2000).

10.2.C. Sharing damages among multiple tortfeasors²⁵

Landes and Posner (1980) proved that in a full-information environment a 50/50 sharing rule for the allocation of accident losses between victims and injurers, when both parties are found non-negligent, is as efficient as any other sharing rule.

This argument has been used in the comparison between contributory and comparative negligence. As they are equally efficient, it has been said, then the legal system should opt for the rule that is easier to implement and triggers less administrative costs. Contributory negligence meets these requirements, but in reality comparative negligence is the most common rule (Curran, 1992). This evidence calls for an explanation.

The opinions of early writers concerning the efficiency of comparative negligence was largely negative (Brown, 1973, and Green, 1976, at 558) or was that the incentive effect of that rule strictly depends on the liability fraction (Diamond, 1974a, at 140-145). Nevertheless, more recently and after the emergence of the efficiency equivalence theorem, some authors have developed economic arguments for comparative negligence, in the case of some of the assumptions of the standard model being not satisfied²⁶.

Cooter and Ulen (1986) argue that in the case of the court making random errors in comparing the due level of care with the level of care actually taken by the parties (evidentiary uncertainty), parties are induced to take an inefficient level of precaution in order to minimize their expected liability; comparative negligence, by distributing the loss between the parties, reduces the effect of errors and leads to less distortion. Edlin (1994) argues instead that in these situations comparative and contributory negligence only differ with respect to the standard of negligence which is to be implemented in order to attain an efficient outcome.

It has been said that comparative negligence is more efficient than contributory negligence when the standard of care is uniform for all parties, but the individual costs of care differ (Rubinfeld, 1987), and when the courts make errors regarding the level of care cost actually borne by the parties (Haddock and Curran, 1985). Shavell (1987) showed that comparative negligence can be superior in non-equilibrium situations, when (for some reasons) victims or injurers are found negligent. Nevertheless, a counter-argument against comparative has been recently proposed by Bar-Gill (2001).

Under a different perspective, comparative negligence can enhance a more efficient outcome when individuals are not identical (Emons, 1990a, 1990b, Emons and Sobel, 1991) and when parties take precaution decisions sequentially, so that one party can observe the other party's level of precaution before choosing his (Rea, 1987, Grady, 1990). A completely different approach eventually leads to fairness and equity arguments which favor comparative

²⁵ See chapter 4.

negligence, insofar as sharing losses when parties are similarly at fault is socially considered to be just.

All the studies mentioned above focused on the use of comparative negligence as a criterion to share the accident loss when both parties are at fault in order to improve the parties' incentives to take care. Parisi and Fon (2001) highlight instead a different point, as they argue that loss sharing can be profitably implemented even in those situations in which both parties have been found non-negligent, in order to provide them with improved incentive to take optimal activity level. In those cases the accident loss is apportioned on the basis of the parties' relative causal contribution to the loss rather than on the basis of their relative degree of negligence.

In chapter 4, I have developed a model which describes the advantages of a comparative negligence rule in an incomplete information environment. Those situations are considered in which the judge does not know the precise cost of taking precaution, and therefore cannot appropriately apply the negligence criterion. Comparative negligence allows the judge to share the loss between the parties and can work as a filter for accidents: it let occur the least harmful accidents and prevents the most harmful ones. The logic of the filtering effect has been already explained in section 10.2 of this chapter.

The same argument applies to the allocation of the accident loss among multiple injurers in unclear cases. This solution is explicitly implemented in the Italian civil code, and it has been adopted more or less openly by the courts in many other countries.

My argument directly relates to the efficiency literature on comparative negligence, by providing an alternative explanation of the success of more or less equal sharing in unclear cases. They might be defended as being just; I argue that they are also efficient.

10.2.D. The judgment-proof problem²⁷

An injurer is considered to be judgment-proof when his assets are not sufficiently large to pay all the damages arising from an accident.

The literature on the judgment-proof effect is almost entirely based on an article by Shavell (1986), who elaborated on Summers (1983). Shavell showed that a judgment-proof injurer takes systematically less than optimal care. His model has two limitations. First, it assumes that the injurer can only influence the probability of the accident, while the harm is exogenous. Second, it does not consider the effect of precaution expenditures on the injurer's assets. When the injurer spends on precaution, the assets available for damage compensation decrease, while Shavell assumed that the assets remain at a constant level.

²⁶ See on this point Chung's (1993) comment on Orr (1991).

This second shortcoming has been thoroughly analyzed by Beard (1990), who modified Shavell's model in order to include the assets-reduction effect of precaution expenditures. Contrary to Shavell's conclusions, he found that underprecaution does not always result. Sometimes, injurers will be led to overprecaution, as by investing in precaution they reduce their exposure to liability. However, Beard's model was also based on the assumption that the injurer reduces only the probability of the accident.

The subsequent literature on the problem has largely ignored Beard's results, and has based its contributions on Shavell's model. Boyd and Ingberman (1994) considered a simple case in which the injurer has two different precautionary measures to take, which trigger different damage magnitudes. This has been the only attempt to include magnitude-reducing precaution in a model analyzing the judgment-proof problem. Their focus, however, is on the effectiveness of punitive damages as a corrective measure.

My contribution tries to shed some light on three aspects of the problem, which until now have been almost totally ignored.

First, I use a formal model, which takes four possible accident prevention situations separately into account. The first is the case in which the injurer only reduces the probability of the accident by taking precaution (the traditional case analyzed in the literature). The second is a pure magnitude case: the injurer cannot do anything about the probability of the accident, but can reduce its magnitude. The other two cases concern mixed probability-magnitude models. An explicit analysis of magnitude models and of mixed models is completely new in the literature and reveals interesting mechanisms of injurers' behavior, in particular, the injurers' decision on the level of precaution to take is sometimes binary: either optimal precaution (judgment-proofness does not undermine the incentives towards precaution) or no precaution at all (the result is more dramatic than it would seem under the traditional approach).

Secondly, I consider two situations separately. In chapter 5 I analyze the case of the injurer's assets not being reduced by precaution expenditures (hence, I build on Shavell); in chapter 7 I analyze the mirror image of the injurer's assets being reduced by precaution expenditures (hence, I consider the situation analyzed by Beard). I show that Beard's predictions concerning overprecaution do not hold in magnitude models and in some mixed models.

Third, I consider and analyze a broad range of possible ways to solve the inefficient precaution problems arising from judgment-proofness. In chapter 5, I consider solutions which can be implemented within the domain of tort law. Until now, the literature has focused

²⁷ See chapters 5 and 7.

on only few solutions: negligence (Shavell, 1986), punitive damages (Boyd and Ingberman, 1994), and regulation in general (Shavell, 1984a). I consider a broader picture and add an analysis of under-compensatory damages and average compensation of the harm. The next section also explores regulatory solutions to the judgment-proof problem.

10.2.E. The combined use of regulation and tort law²⁸

The literature has mainly analyzed regulation and tort law as alternative ways of controlling risk. However, in reality we can often observe that they are combined. The only formal analyses of tort law and regulation as complements have been provided by Shavell (1984b), Kolstad, Ulen and Johnson (1990) and Burrows (1999)²⁹.

Shavell (1984b) discussed the joint use of regulation and tort law in the presence of judgment-proofness (as a limit for tort law) and limited information on risk (as a limit for regulation). In his model, regulation assists tort law in controlling judgment-proof injurers: those injurers, who do not react to tort law incentives for they are insolvent, will be induced to take precaution by regulation. This solution leads to a second-best outcome, as regulatory bodies lack perfect information concerning risk and, therefore, the regulatory standard will not correspond to the injurers' optimal level of precaution.

Such an approach only considers probability models (where precaution reduces the probability of an accident); to the contrary, I have also accounted for three other models in which the injurer can influence the magnitude of the accident.

When the injurer can influence only or also the magnitude of the accident, he has actual control on whether or not he will be potentially bankrupt. If he takes low levels of precaution, the harm will be great, and it might be greater than his assets. If he takes more precaution the harm decreases and he will eventually be able to pay for it. Therefore, the injurer's total cost in magnitude models is discontinuous, for there are two zones: in one he is judgment-proof (little precaution), while in the other he is solvent (high level of precaution). In probability models the injurer finds himself in either one, as the magnitude of the harm is exogenous. On the contrary, in magnitude models the injurer can decide in which zone to be by setting his level of precaution.

Consequently, in magnitude models regulation can be employed in order to influence such an injurer's decision, by increasing the cost associated with the judgment-proof zone. I

²⁸ See chapter 6.

²⁹ Kolstad, Ulen and Johnson (1990) discussed the combined use of regulation and tort law in a different context. Tort law is said to generate incentive problems because of uncertainty in the application of the negligence rule. Regulation is used to repair such shortcoming. Burrows (1999) focused his analysis on instrumental uncertainty: a situation in which the injurer does not know in advance the severity of either or both the negligence rule and the regulatory standard. Shavell (1984b) is accounted for in the text.

have defined this type of regulation as anti-judgment-proof regulation, for it has the only purpose of removing the judgment-proof problem, i.e. the incentives for injurers to choose the judgment-proof zone.

Therefore, regulation can only target those levels of sub-optimal precaution that would be chosen by judgment-proof injurers. This is enough to prevent injurers' from deciding to expose themselves to insolvency. Moreover, in many cases injurers can take a precautionary measure to reduce the probability and another to reduce the magnitude of the accident. They will be either potentially judgment-proof or not depending on the level of their magnitude precaution. Therefore, a regulatory standard that only targets magnitude precaution might be enough to prevent injurers from deciding to expose themselves to insolvency.

It is worth noticing that regulatory bodies do not need perfect information on risk, for anti-judgment-proof regulation only targets serious underprecaution and does not have to be set at the optimal level of precaution. Once regulation has removed the incentives to be judgment-proof, the task of inducing optimal precaution is left to the normal functioning of tort law. The appropriate use of anti-judgment-proof regulation leads to a first-best outcome.

The same approach justifies the implementation and the enforcement of negligence standards independently from regulatory standards.

*10.2.F. Vicarious liability*³⁰

Injurers are often judgment-proof. The legal system opts, when this is feasible, for vicarious liability. A second party is held liable for the damage caused by the injurer. The condition in order for this mechanism of delegated control to work is that the vicariously liable party should be able to monitor the injurer and to induce him to take the optimal level of precaution.

In other situations, vicarious liability simply enhances a better risk allocation between two parties to a contract (for example an employer and his employee) and saves the transaction costs of negotiating such allocation.

After the early contributions by Sykes (1981,1984 and 1988) and Kornhauser (1982), who clarified the logic of vicarious liability and delimited its range of application, the literature has mainly been interested in corporate liability, and, in general, in the vicarious liability of firms or employers³¹, the main concern being the desirability of a vicarious liability rule. Little attention has been devoted to two other aspects of the problem: the cost borne by the vicariously liable party in order to monitor the potential injurer and the cost of taking precaution.

³⁰ See chapter 8.

³¹ See Arlen (1994), Chapman (1992 and 1996), Choi (1998), Croley (1996), Fischel and Sykes (1996), Khanna (1996), Kraakman (1984a, 1984b and 1986), Parker (1996), Schwartz (1996), Shavell (1997).

With respect to the former aspect, the literature has always considered the world in a discrete fashion: either monitoring is possible at a cost equal to zero, or it is not (i.e. the cost is extremely high). Sharp assumptions yield sharp conclusions: vicarious liability is desirable in the former case and it is an ineffective solution in the latter. I consider what lies in between: positive costs of monitoring, variable between zero and infinity and dependent on the level of precaution: I show that vicarious liability can always be implemented at the price of some allocative losses, due to the fact that the costs of monitoring makes precaution more expensive. Under direct liability, the benefit in accident loss reduction has to be balanced against the precaution costs. Under vicarious liability, the benefit in accident loss reduction has to be balanced against the precaution costs plus the monitoring costs. Hence, I have found that the optimal level of precaution under vicarious liability must be lower than under direct liability, and that the greater the monitoring costs, the greater the difference between the two.

With respect to the second aspect, it is important to distinguish between two situations. In contractual settings, such as for example labor contracts, the employer pays the employee for his effort; hence he internalizes the costs of taking precaution borne by the employee. For this reason, the employer will make the employee take optimal precaution. However, vicarious liability also applies in non-contractual relationships: parents are liable for their children and supervisors are liable for persons under their supervision. Parents and supervisors do not pay the precaution costs, and therefore they might be too concerned about preventing accidents and might require too high a level of precaution. An appropriate negligence rule corrects this inefficiency. In reality we observe, in most legal systems, strict employers' vicarious liability and negligence-based parents and supervisors' vicarious liability, which can be explained by my findings.

The same framework can be used for an analysis of Roman law, on which I will comment in one of the next sections.

10.2.G. The efficiency wage theory³²

According to the efficiency wage theory by Shapiro and Stiglitz (1984), it can be rational for an employer to overpay his employees since higher wages make the sanction of dismissal more severe. This increased sanction allows the employer to lower monitoring levels and hence to save on monitoring costs. But if many employers try to save on monitoring costs by increasing wages, the effect disappears since a worker who loses an 'overpaid' job can easily find another 'overpaid' job. In this case, structural unemployment is the only way for the economy to create incentives for employees not to be dismissed.

³² See chapter 9.

Shapiro and Stiglitz' article (1984) was written in the 1980s, at a time when many economists started to believe that unemployment was structural, rather than cyclical. Their theory was and is still considered as a possible explanation of unemployment and is mentioned in textbooks on labor economics.

Other authors have suggested different justifications for efficiency wages, such as the income-satiation (Rasmusen, 1992) and the gift-reciprocity (Akerlof, 1982) arguments. In chapter 9 I have examined efficiency wages only as an enforcement device, and argued that they are less rational a choice than generally accepted and that it is extremely implausible that employers will ever opt for efficiency wages on such a massive scale that structural unemployment will result.

A second reason to reconsider the efficiency-wage theory is that law and economics scholarship has not realized the potential relevance of labor economics insights yet. Efficiency wages are just another technique to avoid breach of contract: in order for the employee to have an incentive to perform (working hard) he is overpaid. If he breaches the contract he does not have to pay damages, but loses a bonus. The economic literature on contract remedies has focused only on negative sanctions: specific performance, expectation damages, reliance damages, penalty clauses, and so forth. It has not compared the advantages of positive sanctions to those of negative sanctions yet. The overpayment technique is also relatively unexplored in the economic literature on public law enforcement, and corruption in particular³³.

I have analyzed under what conditions rational principals to a contract will prefer efficiency wages (a specific positive sanction) to two other incentive devices: punitive damages³⁴ and conditional bonuses. Punitive damages are monetary sanctions with an expected value that equals the true harm. Conditional bonuses are a positive sanction, just like efficiency wages, but, unlike efficiency wages, the overpayment only occurs if monitoring takes place (and the agent is not found shirking). I have found that if these additional incentive devices are available, a rational employer will never adopt efficiency wages.

³³ Most models on corruption do not take bonus payments into account as an alternative solution to corruption. Exceptions are Polinsky and Shavell (2000b), and Cooter and Garoupa (2000).

³⁴ One remark should be made to avoid confusion. From a legal-technical point of view, it is possible that what we (and other economic theorists) call 'punitive damages' is given another name in real-world legal systems. For instance, many contract law systems will consider them as 'penalty clauses' (see De Geest and Wuyts, 2000). If the principal is the government and the agent is a citizen or civil servant, the sanction may be called a 'fine' (and fall under criminal law rather than contract law). As a matter of fact, 'punitive damages' is a concept from American tort law, which is unknown (as such) in continental systems. For an overview of the economic literature on punitive damages, see Polinsky and Shavell (2000a).

10.2.H. Towards an economic analysis of Roman law³⁵

Roman law offers a broad variety of legal solutions to many practical problems. Although it would be extremely interesting, an economic analysis of Roman law is still lacking. Few law and economics scholars have devoted their energies to the study of ancient law³⁶, and Roman law is often mentioned aside in economic analyses of modern law.

The Roman legal tradition is well documented and covers a long period of time, much longer than any modern legal system. It would therefore be a profitable field for a historical analysis of the development of legal systems and a diachronic comparison with modern legal solutions under an economic approach.

In chapter 8, I have analyzed a specific Roman norm, *noxae deditio*, which allowed the *paterfamilias* to choose between paying the damages caused by a member of his family and hand over the tortfeasor to the victim. The reason why this is of great interest to modern scholars is threefold. First, it is an example of a very old rule that survived over time by changing its function within the tort law system. Second, it almost completely disappeared in modern tort law. Third, at first sight it is a very strange mechanism, without any plausible explanation. *Noxae deditio* has never been analyzed in a law and economics perspective.

Paterfamilias were originally liable for their slaves or other members of the family, including their descendants and their wives; they had to punish the tortfeasor directly or hand him over to the victim's family. Often a dead body was given in order to redress the wrong. It was indeed a form of vengeance. The ultimate aim was to assure that the person responsible would pay the price.

At a later stage, the function of tort law evolved from regulated vengeance to damage compensation. The tortfeasor was no longer killed, but would be transferred from his original family to the victim's family. Although its function changed, *noxae deditio* was retained, and *paterfamilias* had the choice between monetary or in-kind compensation and the handing over of the tortfeasor. That norm survived for many centuries.

In an economic perspective this solution is rather peculiar, as it realizes a form of limited *paterfamilias*' liability. If a slave caused an accident, the *paterfamilias* would pay the losses up to the market value of the slave and opt for *noxae deditio* in the case of heavier losses. That rule generates a liability cap that undermines *paterfamilias*' incentives towards monitoring slaves' precaution. It certainly calls for an explanation.

In a slave-based economy, *patresfamilias* obviously do not pay for the slaves' effort. Hence, they do not internalize the costs of precaution, in a similar way as modern parents and supervisors do not internalize the children and supervisees' costs of precaution. Both Roman

³⁵ See chapter 8.

patresfamilias and modern parents and supervisors might be tempted to enforce too high a level of precaution, which would be inefficient.

In theory, at least two solutions are available: a negligence rule and a liability cap. Both have the effect of reducing precaution. A negligence rule is more accurate, but it requires a well-developed judicial system and quite a large amount of information on the costs and benefits of accident prevention; hence, it triggers high administrative costs.

A liability cap is simple to implement, especially in the way in which Roman law interpreted it in *noxae deditio*. It does not require information or specific judicial expertise and is quite clear a rule. Nevertheless, it is less accurate than a negligence rule and only roughly corrects the overprecaution problem.

Roman law opted for the latter, for the sake of saving administrative costs. Modern law can count on a more advanced judicial system, and hence it can exploit the advantages of a negligence rule at lower administrative costs and can obtain a more accurate correction of the inefficiency.

Besides the interest in the study of ancient legal solutions, my analysis has shown that ancient law can be analyzed by using the same approach as is used for modern law. This proves once more that law and economics is not a regional discipline which is suitable for the study of the legal system in which it has been developed (namely American law), but it can be profitably applied to any modern and ancient legal system.

In addition, the results show how a law and economics approach can help to understand differences and similarities among different legal traditions, and can provide an additional tool for the historical analysis of law.

10.2.1. The economic analysis of the evolution of legal systems³⁷

In the former section I have compared modern and Roman legal norms in a law and economics perspective. In this section I will discuss a more dynamic approach for a law and economics analysis of the evolution of legal systems over time. In chapter 3, I have discussed the evolution of tort law systems, and have drawn a path which seems to be able to explain the evolution of early Roman and English law.

My analysis has shown that legal systems tend to stay as close as possible to a no-liability rule, and slowly develop specific rules of strict liability in determined areas of the law. This is due to the fact that strict liability involves damage compensation, while no liability does not. Hence, unless there is a clear advantage in giving incentives to the injurer (strict liability) rather than to the victim (no-liability), the legal system stays on the latter's side, in order to

³⁶ Posner (1980) and Parisi (2001) are two examples.

save compensation costs. At such an early stage, the judiciary system is likely to be still underdeveloped, and hence a negligence criterion, which implies an analysis of the costs and benefits of parties' precaution, happens to be too expensive.

When the legal and judicial systems develop, the cost of ascertaining parties' precaution diminishes and two more rules become available: strict liability with defense of contributory negligence and simple negligence. Both Roman and English legal traditions evolved by initially developing a list of torts subject to strict liability. Later, both systems tended towards a greater generalization and a more or less explicit use of negligence defenses.

Once negligence defenses have been introduced, the choice of the best tort law rule becomes more complex. Two decisions should be taken concerning the choice of the residual bearer and the extent of the negligence criterion. In chapter 3 I have formalized this analysis into a model and obtained a criterion to choose of the best rule. My findings are original in two respects.

First, my model is a formalized analysis of the choice of the optimal tort law rule and discusses the results obtained by Shavell (1980a and 1987) and Landes and Posner (1987). In their work the choice between different rules was discussed in a non-formalized way and did not take account of the fact that the extent of the negligence inquiry endogenously depends on an explicit choice of the law and it is not given by nature.

Second, the results that I have obtained can be used for reinterpreting some highly controversial issues, such as the switch from simple negligence to strict liability in industrial accidents and product liability, for which reference is made to chapter 3.

10.3. Relevance for policy makers

My contention in this work has been that law and economics analysis can be deepened by enriching the theory behind it. As such, my approach has been mainly theoretical. Nevertheless, in this section I will enumerate and briefly discuss some policy implications thereof.

A preliminary caveat is worthwhile, as the translation of any theoretical conclusion into practical guidelines has to account for the limits of the theory in the first place. As law and economics is a way to model human behavior into a simplified framework, it is intrinsically bounded by its own nature. It is an "*if...then...*" approach and as any synthesis of reality into models has an analytical cost, which cannot be overlooked. However, this would only be a shortcoming if the results of the analysis were uncritically taken as valid in any context.

³⁷ See chapter 3.

This work is an economic analysis, and heavily relies on several assumptions that have been clarified in the chapters. The fact that people are rational, perfectly informed and risk-neutral and respond to incentives is indeed the main standpoint. In reality this is not always true.

A second concern relates to values. Equity, fairness and justice can be considered in two different ways. As preferences that can be included in the efficiency analysis, or as external values that ought to be taken into account even if they are in contrast with the efficiency results of the model³⁸. In both cases, the results might change.

In general, my analysis focuses on incentives. Therefore, it leaves any discussion on risk allocation open. This aspect is particularly relevant with respect to the problem of uncompensated victims of an accident. Even when precaution is at the optimal level, some accidents result, and victims might suffer losses. Efficiency or fairness considerations might support arguments in favor of insurance or state intervention systems in order to reduce or eliminate those losses. Insurance is in general preferable to tort law as a system for risk reallocation, as it realizes transfers at a lower cost. The administrative costs of different tort law solutions are generally taken into account in the single chapters.

With respect to incentives, I have sometimes compared positive (bonuses, subsidies, rewards) with negative (damages, fines) devices, but have not considered behavioral implications thereof. A behavioral economics³⁹ approach could either strengthen or limit my conclusions.

Nevertheless, I believe that, if we are aware of the limits of the economic analyses that we perform, their results may prove helpful and furnish with perspective for us to consider some policy issues.

10.3.A. The choice of the residual bearer

Chapter 3 has provided an analysis of different tort law rules. One of the aspects is the choice of the residual bearer: which party should bear the cost of the accident when both the injurer and the victim have been found non-negligent. I have shown that the choice should take into account two elements: the cost of observing parties precaution within the negligence inquiry and the effect of the unobserved precautions on the expected accident loss. The literature has until now overlooked the first element and has focused its policy conclusions on the dangerousness criterion only.

³⁸ See Kaplow and Shavell (2001). They strongly support the first approach.

³⁹ See Sunstein (1999).

10.3.B. The extent of the negligence criterion

When parties can take different precautions in order to reduce the expected accident loss, the legal system has to decide which of these precautions will be considered in the negligence inquiry. Checking all of them is often too costly and sometimes impossible. However, excluding some precautions from the determination of negligence yields less effective accident prevention and some social losses. The criterion should be based on a balance between these two elements, as shown in chapter 3.

10.3.C. Comparative negligence and the rule of contribution

When it is not clear which of the two parties should be held responsible for an accident, fairness suggests that the loss should be equally divided. I have shown that efficiency requires the same in many situations, as a loss sharing between the parties has a filtering property: it only allows the least harmful accidents to occur, and prevents the most harmful ones.

Comparative negligence and a rule of contribution, which allow loss sharing in unclear cases, are therefore an efficient policy solution to the problem in some cases.

10.3.D. Insolvency

The analysis in chapters 4 to 6 has shown that insolvent injurers do not always tend to take less than optimal precaution. A variety of different outcomes can result. Injurers can take optimal precaution even when they are potentially insolvent. They might be systematically induced towards sub-optimal precaution. They might even take too much precaution. The outcome depends on two variables: first, it depends on whether the injurer can control the probability or the magnitude or both; second, it depends on whether or not the precaution expenditure is monetary.

In the model I have demonstrated that, depending on the situation, a specific corrective device can be used in order to induce optimal precaution. Possible solutions are sometimes counterintuitive. My analysis has shown that in some situations, in order to correct underprecaution, it is appropriate to subsidize the injurer by granting him a sort of damage allowance by means of under-compensatory damages. In other cases, punitive damages might be applied, but they are in general an inferior solution to the average compensation for the harm.

The use of regulatory sanctions has also been explored. I have found that it is often sufficient to sanction only serious underprecaution. When the injurer can separately control the probability and the magnitude of damages, it is generally necessary to control by regulation only the magnitude-reducing precaution; this is particularly relevant in the case of

safety measures, which are likely to reduce the magnitude rather than the probability of the accident.

10.3.E. Liability caps

The analysis in chapters 4 and 5 also applies to financial caps, which are limits on the amount of damages the injurer has to pay in the case of an accident and are set by law; this is often the case in nuclear accident liability.

It is generally believed that a cap reduces the incentives to take precaution. I have shown that this is not always true. Again, it is important to consider what the injurer can do in order to reduce the expected accident loss. If he can only reduce the magnitude of the harm, a financial cap on his liability might have no incentive-reduction effect if the cap is sufficiently high. However, if the cap is too low, a dramatic outcome might result: the injurer might find it advantageous to take no precaution at all.

It is therefore of central importance for the policy maker to consider these issues, while determining the desirability of limits on liability. In chapter 8 I have shown how Roman law used such a mechanism in order to correct the overprecaution problem arising from vicarious liability.

10.3.F. Negligence-based vs. strict vicarious liability

In the case of vicarious liability, the rule can be structured as a strict liability rule or as a negligence rule. In the latter case, the negligence criterion can be applied to the tortfeasor and/or to the vicariously liable party.

My analysis has shown that the setting of the negligence criterion should in both cases take the cost of monitoring into account, and it is in general efficient to have lower precaution than under direct liability, because of the cost of monitoring associated with vicarious liability.

In contractual settings (for example in the relationship between employers and employees), the vicariously liable party (the employer) internalizes the cost of precaution borne by the tortfeasor (the employee). A strict vicarious liability rule is efficient.

In non-contractual settings (for example in the relationship between parents and their children), the vicariously liable party (the parent) does not directly internalize the cost of precaution borne by the tortfeasor (the child). Thus, a negligence-based vicarious liability rule is efficient.

10.4. A final remark

My analysis started from the impression that law and economics analysis could be made more powerful by deepening its theoretical support. The sensation that accompanies me at the end of such an attempt is still that more could be done, both in the direction of improving the models employed for predicting human behavior and in the direction of applying these findings to new areas of research.

At least two of my findings might be profitably exploited: the recognition that the parties to an accident can take many different precautionary measures in order to reduce the expected accident loss and that they might be able to reduce either the magnitude of the harm or the probability of the accident or even both.

In general, all my findings could be extended to different areas of law from the ones in which they have been analyzed. This might lead to further research along the same lines of analysis.

10.5. References

- Akerlof, George (1982), 'Labor Contracts as Partial Gift Exchange', *97 Quarterly Journal of Economics*, 543-569.
- Arlen, Jennifer H. (1994), 'The Potentially Perverse Effects of Corporate Criminal Liability', *23 Journal of Legal Studies*, 832-867.
- Bar-Gill Oren (2001), 'Does Uncertainty Call for Comparative Negligence?', *John M. Olin Center for Law, Economics and Business, Discussion Paper No. 346*, Harvard Law School.
- Beard, Rudolph T. (1990), 'Bankruptcy and Care Choice', *21 Rand Journal of Economics*, 626 ff.
- Boyd, James and Ingberman, Daniel (1994), 'Noncompensatory Damages and Potential Insolvency', *23 Journal of Legal Studies*, 895-910.
- Brown, John Prather (1973), 'Towards an Economic Theory of Liability', *2 Journal of Legal Studies*, 323-350.
- Burrows, Paul (1999), 'Combining Regulation and Legal Liability for the Control of External Costs', *19 International Review of Law and Economics*, 227-244.
- Calabresi, Guido (1961), 'Some Thoughts on Risk Distribution and the Law of Torts', *70 Yale Law Journal*, 35-46.
- Chapman, Bruce (1992), 'Limited Auditors' Liability: Economic Analysis and the Theory of Tort Law', *20 Canadian Business Law Journal*, 180-214.
- Chapman, Bruce (1996), 'Corporate Liability and the Problem of Overcompliance', *69 Southern California Law Review*, 1679-1703.
- Choi, Stephen (1998), 'Market Lessons for Gatekeepers', *92 Northwestern University Law Review*, 916 ff.

- Chung, Tai-Yeong (1993), 'Efficiency of Comparative Negligence: A Game Theoretic Analysis', **22** *Journal of Legal Studies*, 395-404.
- Cooter, Robert D. and Garoupa, Nuno (2000), 'The Virtuous Circle of Distrust: A Mechanism to Deter Bribes', working paper.
- Cooter, Robert D. and Ulen, Thomas S. (1986), 'An Economic Case for Comparative Negligence', **61** *New York University Law Review*, 1067-1110.
- Cooter, Robert D. and Ulen, Thomas S. (2000), *Law and economics*, 3rd ed., Reading, Massachusetts, [etc.]: Addison-Wesley.
- Croley, Steven P. (1996), 'Vicarious Liability in Tort: On the Sources and Limits of Employee Reasonableness', **69** *Southern California Law Review*, 1705-1738.
- Curran, Christopher (1992), 'The Spread of the Comparative Negligence Rule in the United States', **12** *International Review of Law and Economics*, 317-332.
- De Geest, Gerrit and Wuyts, Filip (2000), 'Penalty Clauses and Liquidated Damages', in Bouckaert, Boudewijn and De Geest, Gerrit (eds.), *Encyclopedia of Law and Economics*, Cheltenham, Edward Elgar, vol. III, 141-161.
- Diamond, Peter A. (1974a), 'Single Activity Accidents', **3** *Journal of Legal Studies*, 107-164.
- Diamond, Peter A. (1974b), 'Accident Law and Resource Allocation', **5** *Bell Journal of Economics*, 366-406.
- Diamond, Peter A. and Mirrlees, James A. (1975), 'ON the Assignment of Liability: The Uniform Case', **6** *Bell Journal of Economics*, 487-516.
- Edlin, Aaron S. (1994), 'Efficient Standards of Due Care: Should Courts Find More Parties Negligent under Comparative Negligence', **14** *International Review of Law and Economics*, 21-34.
- Emons, Winand (1990a), 'Efficient Liability Rules for an Economy with Non-identical Individuals', **42** *Journal of Public Economics*, 89-104.
- Emons, Winand (1990b), 'Some Recent Developments in the Economic Analysis of Liability Law: An Introduction', **146** *Journal of Institutional and Theoretical Economics*, 237-248.
- Emons, Winand and Sobel, Joel (1991), 'On the Effectiveness of Liability Rules when Agents Are Not Identical', **58** *Review of Economic Studies*, 375-390.
- Eörsi, Gyula (1975), 'Private and Governmental Liability for the Torts of Employees and Organs', in André Tunc (ed.), *Torts, International Encyclopedia of Comparative Law*, Tübingen, Mohr.
- Fella, Giulio (2000), 'Efficiency Wage and Efficient Redundancy Pay', **44** *European Economic Review*, 1473-1490.
- Fischel, Daniel R. and Sykes, Alan O. (1996), 'Corporate Crime', **25** *Journal of Legal Studies*, 319-349.
- Garoupa, Nuno and Gravelle, Hugh (2002), 'Optimal Deterrence with Legal Defense Expenditure', *Economic Inquiry*.
- Grady, Mark (1990), 'Multiple Tortfeasors and the Economy of Prevention', **19** *Journal of Legal Studies*, 653-678.
- Grady, Mark F. (1983), 'A New Positive Economic Theory of Negligence', **92** *Yale Law Journal*, 799-829.

- Green, Jerry (1976), 'On the Optimal Structure of Liability Laws', *7 Bell Journal of Economics*, 553-574.
- Haddock, David D. and Curran, Christopher (1985), 'An Economic Theory of Comparative Negligence', *14 Journal of Legal Studies*, 49-72.
- Kaplow, Louis and Shavell, Steven (2001), 'Fairness vs. Welfare', 114 *Harvard Law Review*.
- Khanna, V.S. (1996), 'Corporate Criminal Liability: What Purpose does it Serve?', *109 Harvard Law Review*, 1477-1534.
- Kolstad, Charles D., Ulen, Thomas S., and Johnson, Gary V. (1990), 'Ex Post Liability for Harm vs. Ex Ante Safety Regulation: Substitutes or Complements?', *80 American Economic Review*, 888-901.
- Kornhauser, Lewis A. (1982), 'An Economic Analysis of the Choice between Enterprise and Personal Liability for Accidents', *70 California Law Review*, 1345-1392.
- Kraakman, Reinier H. (1984a), 'The Economic Functions of Corporate Liability', in Hopt, Klaus J. and Teubner, Gunther (eds), *Corporate Governance and Directors' Liabilities*, New York, de Gruyter, 178-207.
- Kraakman, Reinier H. (1984b), 'Corporate Liability Strategies and the Costs of Legal Controls', *93 Yale Law Journal*, 857-898.
- Kraakman, Reinier H. (1986), 'Gatekeepers: The Anatomy of a Third-Party Enforcement Strategy', *2 Journal of Law, Economics, and Organization*, 53-104.
- Landes, William M. and Posner, Richard A. (1980), 'Joint and Multiple Tortfeasors: An Economic Analysis', *9 Journal of Legal Studies*, 517-555.
- Landes, William M. and Posner, Richard A. (1987), *The Economic Structure of Tort Law*, Cambridge (MA), Harvard University Press.
- Orr, Daniel (1991), 'The Superiority of Comparative Negligence: Another Vote', *20 Journal of Legal Studies*, 119-129.
- Parisi, Francesco (2001), 'Genesis of Liability in Ancient Law', *3 American Law and Economics Review*, 82-124.
- Parisi, Francesco and Fon, Vincy (2001), 'Comparative Causation', *Law and Economics Working Paper Series 01-30*, George Mason University, School of Law.
- Polinsky, A. Mitchell and Shavell, Steven (2000a), 'Punitive Damages', in Bouckaert, Boudewijn and de Geest, Gerrit (eds.), *Encyclopedia of Law and Economics*, Cheltenham, Edward Elgar, vol. II, 764-781.
- Polinsky, A. Mitchell and Shavell, Steven (2000b), 'Corruption and Optimal Law Enforcement', *Journal of Public Economics*, forthcoming.
- Posner, Richard A. (1980), 'A Theory of Primitive Society with Special Reference to Law', *23 Journal of Law and Economics*, 1-53.
- Rasmusen, Eric (1992), 'An Income-Satiation Model of Efficiency Wages', *30 Economic Inquiry*, 467-78.
- Rea, Samuel A., Jr. (1987), 'The Economics of Comparative Negligence', *7 International Review of Law and Economics*, 149-162.
- Rubinfeld, Daniel L. (1987), 'The Efficiency of Comparative Negligence', *16 Journal of Legal Studies*, 375-394.

- Schäfer, Hans-Bernd and Schönenberger, Andreas (2000), 'Strict Liability versus Negligence', in Bouckaert, Boudewijn and De Geest, Gerrit (eds.), *Encyclopedia of Law and Economics*, II, Aldershot, Edward Elgar, 597-624.
- Schwartz, Gary T. (1996), 'The Hidden and Fundamental Issue of Employer Vicarious Liability', **69** *Southern California Law Review*, 1739-1767.
- Shapiro, Carl and Stiglitz, Joseph E. (1984), 'Equilibrium Unemployment as a Worker-Discipline Device', **74** *American Economic Review*, 433-444.
- Shavell, Steven (1980), 'Strict Liability versus Negligence', **9** *Journal of Legal Studies*, 1-25.
- Shavell, Steven (1984a), 'Liability for Harm Versus Regulation for Safety', **13** *Journal of Legal Studies*, 357-374.
- Shavell, Steven (1984b), 'A Model for Optimal Use of Liability and Safety Regulation', **15** *Rand Journal of Economics*, 271-280.
- Shavell, Steven (1986), 'The Judgment Proof Problem', **6** *International Review of Law and Economics*, 45-58.
- Shavell, Steven (1987), *Economic Analysis of Accident Law*, Cambridge (MA), Harvard University Press.
- Summers, John (1983), 'The Case of the Disappearing Defendant: An Economic Analysis', **132** *Pennsylvania Law Review*, 145-185.
- Sunstein, Cass R. (1999), 'Behavioral law and economics: a progress report', **1** *American Law and Economics Review*, 115-157.
- Sykes, Alan O. (1981), 'An Efficiency Analysis of Vicarious Liability Under the Law of Agency', **91** *Yale Law Journal*, 168-206.
- Sykes, Alan O. (1984), 'The Economics of Vicarious Liability', **93** *Yale Law Journal*, 1231-1280.
- Sykes, Alan O. (1988), 'The Boundaries of Vicarious Liability: An Economic Analysis of the Scope of Employment Rule and Related Legal Doctrines', **101** *Harvard Law Review*, 563-609.
- Trimarchi, Pietro (1961), *Rischio e responsabilità oggettiva* (Risk and Strict Liability), Milano, Giuffrè.

SAMENVATTING

Gedurende de afgelopen veertig jaar is binnen de rechtseconomie een indrukwekkende en vernieuwende methode ontwikkeld om juridische regels te analyseren. Dit heeft plaatsgevonden door het toepassen van micro-economische theorieën op juridische vraagstukken en heeft geleid tot de opkomst van specifiek rechtseconomische commentaren op alle gebieden van het recht.

De rechtsanalyse is van oudsher een analyse *de iure condito*. Het is de analyse van het recht zoals het is. Juristen hebben de taak het recht te interpreteren, het uit te leggen en de lacunes erin op te vullen, over het algemeen door interpretatie naar analogie. Zelfs wanneer een nieuw probleem zich voordoet dat om een nieuwe oplossing vraagt, is de rechtsanalyse gebaseerd op de veronderstelling dat elke mogelijke oplossing reeds besloten ligt in het recht dat voorhanden is en dat het er simpelweg om gaat deze oplossing te vinden. De rechtsanalyse vertaalt het (abstracte) recht naar (concrete) regels, die geschikt zijn om bestaande problemen op te lossen. Het bedenken van nieuwe regels wordt beschouwd als revolutionair.

De rechtseconomie is inderdaad revolutionair. Ze houdt zich bezig met regels, ongeacht of deze nu wel of niet zijn geïmplementeerd in het bestaande rechtssysteem. Ze onderzoekt de sociale wenselijkheid van bepaalde regels, waarbij als criterium de allocatieve efficiëntie wordt gehanteerd. En deze eenvoudige aanpak levert zeer veel op.

Wanneer de regel die het onderwerp is van de analyse al geïmplementeerd is in het bestaande rechtssysteem, dan kan het oordeel over die regel gebruikt worden om kritiek te uiten op het recht zoals dat is gemaakt door de wetgever of is voortgekomen uit een rechterlijke uitspraak. De conclusie zou kunnen luiden dat het, vanuit het standpunt van allocatieve efficiëntie, goed zou zijn het recht te wijzigen. Op vergelijkbare wijze kan het oordeel over de regel gebruikt worden om een bepaalde (*de iure condito*) uitleg van het recht te bekritisieren. De ene uitleg kan efficiënter zijn dan de andere.

Wanneer de regel die geanalyseerd wordt er een is die niet voorkomt in het bestaande recht, dan doet de rechtseconomie een stap vooruit door nieuwe regels aan te dragen bij de wetgever.

De rechtseconomie is in feite een methode voor een *de iure condendo* analyse, omdat ze het bestaande recht voortdurend aan vragen onderwerpt, zowel in haar pogingen tot toetsing als in haar pogingen tot het vormen van nieuw recht. Een bewijs voor deze stelling is dat rechtseconomie een analytisch kader biedt voor zowel de wenselijkheid van rechtmatige contractsclausules, als voor (onrechtmatig) crimineel gedrag en voor internationale interactie die dikwijls plaatsvindt buiten enig rechtssysteem.

De rechtseconomie kan derhalve de ingeburgerde juridische analyse niet vervangen, aangezien ze niet een nieuwe vorm van juridische analyse is maar iets anders.

Aanvankelijk concentreerde de rechtseconomie zich op de analyse van specifieke juridische problemen, maar de laatste tijd zijn ook pogingen ondernomen om te komen tot een generalisatie van de analyse. Men tracht dan te komen tot algemene concepten en uitspraken. Voorbeelden van deze tendens zijn de theorieën over *anti-commons*, *anti-insurance* en *givings vs. takings*. Enkele daarvan hebben pas later specifieke en soms onverwachte toepassingen gevonden. Dit onderzoek beoogt de fundamentele van de rechtseconomische analyse bloot te leggen.

Het benadrukken van algemene concepten die de fundering vormen van de rechtseconomie, is een proces dat nog niet voltooid is. Ik beoog deze onvoltooidheid aan het licht te brengen en een bijdrage te leveren aan het onderzoek. In mijn analyse concentreer ik mij op incentives. Wettelijke (maar ook contractuele) regels zetten een incentive-stroom in beweging die het gedrag van individuen beïnvloedt teneinde een efficiënt resultaat te bereiken. Mijn onderzoek richt zich in het bijzonder op een aantal basiskenmerken van incentives die tot nu toe over het hoofd zijn gezien in de literatuur.

De potentiële slachtoffers van ongevallen willen dat de potentiële schadeveroorzakers voorzorgsmaatregelen nemen en werkgevers willen dat werknemers werken. Om dergelijke doelen te bereiken, moeten hulpbronnen worden ingezet. Zoals een stroom water die van een heuvel glijdt, kan een incentive-stroom ook verschillende effecten teweeg brengen afhankelijk van de manier waarop hij gericht is op het doel.

Een volledige theorie over hoe de inzet van schaarse middelen kan leiden tot naleving van de wet, is nog niet voorhanden. Dit komt omdat incentive-stromen tot nu toe slechts geanalyseerd zijn in specifieke omstandigheden en er maar weinig moeite gedaan is om de uitkomst van de analyse te generaliseren.

Mijn analyse is onderverdeeld in verschillende delen die elk een specifieke kwestie behandelen. Het onderzoek beweegt zich op een gebied (met name de onrechtmatige daad) dat uitgebreid bestudeerd is in de jaren tachtig en vroege jaren negentig van de vorige eeuw en dat gezien wordt als een niet lonende investering voor jonge academici omdat er nauwelijks zaken overgebleven zijn waarover interessante ontdekkingen kunnen worden gedaan. Ik hoop het tegendeel te bewijzen.

Mijn eerste vraag is van fundamentele aard: wat wordt bedoeld met zorg en met activiteitsniveau? Mijn conclusie luidt dat het onderscheid tussen zorg en activiteitsniveau volgt uit de aansprakelijkheidsregels zelf en niet onafhankelijk van de regels zelf kan worden bepaald.

Deze conclusie ontwricht de bestaande economische analyse van de onrechtmatige daad. Op het traditionele onderscheid tussen zorg en activiteitsniveau zijn twee theorema's gebaseerd, het efficiëntie-equivalentie-theorema en het activiteitsniveau-theorema, de twee fundamentele uitkomsten van de economische analyse van de onrechtmatige daad. Ook de keuze tussen risicoaansprakelijkheid en schuldaansprakelijkheid vindt haar grond in genoemd onderscheid. Het is verbazingwekkend dat dit onderscheid altijd als gegeven is aangenomen en nog nooit ook maar enigszins ter discussie is gesteld.

De vraag naar het onderscheid tussen zorg en activiteitsniveau komt tevens neer op de vragen hoe een eenvoudige incentive-stroom (de plicht om de schade van het ongeval te dragen) twee individuen er toe kan aanzetten voorzorgsmaatregelen te nemen en hoe de grenzen van deze incentives te bepalen.

Wanneer een incentive-stroom eenmaal is opgezet, moet hij correct geïmplementeerd worden door het rechterlijke systeem, omdat beslissingen achteraf effecten vooraf veroorzaken. Er zijn echter situaties waarin de rechter over onvoldoende informatie beschikt om de zaak op de juiste wijze te beslissen. **Mijn tweede vraag** is dan ook: hoe moet de rechter de schade uitsplitsen over de twee partijen wanneer hij geen informatie heeft over hun gedrag voorafgaand aan de zaak?

De vraag rijst hoe het effect van een eenvoudige incentive-stroom gemaximaliseerd kan worden in een onduidelijke situatie. Moet het rechtssysteem de incentive-stroom geheel op een van de partijen richten of voor de helft op elk van hen?

Zelfs wanneer een optimale incentive-stroom is gevonden en geïmplementeerd, kunnen beperkingen van de aansprakelijkheid van de partijen het effect van de incentives verminderen. Om de analyses zo algemeen mogelijk te maken, maak ik gebruik van het begrip *maximum upper threshold*, hetgeen verwijst naar iedere mogelijke beperking van de blootstelling van partijen aan aansprakelijkheid, hetzij voortvloeiend uit het rechtssysteem zelf hetzij uit een tekort aan baten bij de partijen. De uitkomst van de traditionele analyse is tot nu toe altijd geweest dat een *maximum upper threshold* in toenemende mate de incentive tot het treffen van voorzorgsmaatregelen vermindert.

Mijn derde vraag is nu of dit inderdaad het geval is. Ik ben tot de conclusie gekomen dat het niet altijd zo is. Dat mijn resultaat afwijkend is, kan worden verklaard doordat de traditionele uitkomst in cruciale mate afhangt van de manier waarop de incentives altijd vorm hebben gekregen, dus van de manier waarop de vraag is gesteld of een *maximum upper threshold* de incentive stroom ondermijnt. Ik stel dezelfde vraag opnieuw, maar formuleer haar anders.

Mijn vierde vraag is geheel verbonden met de derde. Zodra het effect van *maximum upper thresholds* op de incentive-stromen is verduidelijkt, stel ik de vraag welke oplossingen in de vorm van regulering kunnen worden ingevoerd om eventuele inefficiënties te corrigeren.

In de literatuur worden situaties beschreven waarin *maximum upper thresholds* ertoe kunnen leiden dat schadeveroorzakers zich eerder zorgvuldiger gedragen dan onzorgvuldiger. In **mijn vijfde vraag** ga ik na onder welke omstandigheden dit het geval is. Door gebruik te maken van een ander formeel model dan in de literatuur wordt gehanteerd kon ik de logica van en de voorwaarden voor een dergelijke uitkomst, die tegen de intuïtie indruist, achterhalen.

Bij **de zesde vraag** wordt de analyse van *maximum upper thresholds* toegepast op een geheel ander gebied: de gedelegeerde aansprakelijkheid in het Romeinse recht. De *paterfamilias* in het oude Rome droeg de verantwoordelijkheid voor schade die was toegebracht door leden van zijn familie, maar dit was niet onbeperkt. Hij had de mogelijkheid degene die de onrechtmatige daad had begaan over te dragen aan de familie van het slachtoffer in plaats van schadevergoeding te betalen. Deze figuur werd *noxae deditio* genoemd. Deze regeling is enigszins atypisch en leidt feitelijk tot een beperking van de aansprakelijkheid. Iets dergelijks bestaat vandaag de dag niet meer in het geval van kwalitatieve aansprakelijkheid van werkgevers of ouders. Ik onderzoek de economische rechtvaardiging van *noxae deditio* en stel de daarmee verbonden vraag waarom tegenwoordig de kwalitatieve aansprakelijkheid van werkgevers meestal een vorm van risicoaansprakelijkheid is en die van ouders een vorm van schuldaansprakelijkheid.

Met mijn **zevende vraag** neem ik het geval onder de loep waarin de incentive-stroom direct gefinancierd wordt door de slachtoffers. Ik stel de vraag welk mechanisme calculerende slachtoffers zullen gebruiken om de meest effectieve incentive-stroom te scheppen tegen voor hen minimale kosten.

De antwoorden op alle hiervoor gestelde vragen zouden het bewijs op moeten leveren dat het onderzoek naar de algemene kenmerken van incentive-stromen nog lang niet voltooid is en dat we derhalve te maken hebben met een vruchtbaar onderzoeksgebied.

Alle vragen worden behandeld op rechtseconomische wijze: eerst worden ze vertaald naar een formeel economisch model, vervolgens geanalyseerd en tenslotte worden de resultaten besproken. De uitkomsten van de formele analyse worden daarna vanuit tweeërlei perspectief onder de loep genomen: er wordt gekeken naar de theoretische bijdrage aan het onderzoeksgebied en naar de gevolgen voor het beleid.

Het proefschrift is onderverdeeld in tien hoofdstukken: het eerste en laatste worden benut voor inleiding en conclusies. Het tweede hoofdstuk behandelt de rechtseconomische analyse van de onrechtmatige daad en is bedoeld als toelichting op de belangrijkste resultaten die veertig jaar onderzoek door vooraanstaande rechtsgeleerden en economen heeft opgeleverd.

In de andere zeven hoofdstukken worden de zeven hierboven beschreven vragen gesteld. De genoemde problemen worden zowel geanalyseerd door middel van een formeel model als bediscussieerd in de tekst. De formele modellen die worden gebruikt zijn geënt op het standaardmodel dat in de jaren tachtig van de vorige eeuw is ontwikkeld. Desalniettemin hangt het onvermogen van de literatuur antwoord te geven op de gestelde vragen met name af van de wijze waarop problemen tot nu toe in een model zijn ondergebracht. Ik probeer daarom een nieuwe formulering te geven waarbij ik toch zo dicht mogelijk bij de traditionele zienswijze blijf.

Een aantal tekortkomingen van het standaardmodel van de onrechtmatige daad wordt belicht. Ten eerste is het standaardmodel gebaseerd op een eendimensionale voorzorgsmaatregel: partijen bij een ongeval kunnen maar in één soort voorzorgsmaatregel investeren. In hoofdstuk 3 introduceer ik een model voor multidimensionale voorzorgsmaatregelen zodat het mogelijk wordt een realistischere situatie te bespreken waarin partijen in staat zijn ongevallen op vele verschillende manieren te voorkomen. Dit vormt de basis voor het aanvechten van het traditionele onderscheid tussen activiteitsniveau en zorg.

In hoofdstuk 4 wordt de kwestie van de beperkte rechterlijke informatie over het voorafgaande gedrag van partijen besproken. Tevens wordt de vraag naar de verdeling van de kosten van zorgvuldigheid, een vraag waar tot nu toe nooit bij stil werd gestaan, nader onderzocht. Hiertoe wordt een wiskundig model gepresenteerd van de effecten van verschillende verdelingscurven.

Waar het gaat om het bespreken van de grenzen aan de blootstelling van partijen aan aansprakelijkheid, gaat het standaardmodel ervan uit dat partijen de waarschijnlijkheid waarmee ongevallen voorkomen kunnen reduceren, maar niet de omvang van de schade, die immers exogeen blijft. Ik elimineer deze veronderstelling en overweeg gevallen waarin ook de omvang kan worden verminderd, wat bijvoorbeeld vaak voorkomt bij auto-ongelukken. De uitkomsten hiervan worden behandeld in de hoofdstukken 5, 6 en 7.

Hoofdstuk 8 behandelt de kwalitatieve aansprakelijkheid, zowel in het moderne recht als in het Romeinse recht, en met betrekking tot zowel de onrechtmatige daad als de overeenkomst.

Hoofdstuk 9 gaat in op incentives in de contractuele sfeer en betwist de gangbare uitleg voor het bestaan van werkloosheid (efficiënte-loontheorie) door de veronderstelling te laten

vallen dat werkgevers zich niet op geloofwaardige wijze kunnen committeren aan een ander instrument dan efficiënte lonen voor het opwekken van incentives.