EMPIRICS OF TORT*

Ben C.J. van Velthoven

1. Introduction

People can incur damage in many ways. A person’s feelings and property may get hurt by slander, deceit, assault or battery; a consumer may be injured by a defective product; a motorist may see his car being driven into from behind; an employee may get sick from the working conditions in his job; a patient may be harmed by a medical error. In each of these instances, the victim can turn to the tort liability system and try to obtain a court order to make the injurer pay for his losses. In this way, the tort liability system serves three purposes. It provides a forum for the victims to be heard and to oblige the injurers to make up for morally culpable and egregious behavior (corrective justice). It provides compensation to those who are harmed (distributive justice). And it provides incentives for individuals and firms to take appropriate care and to reduce the number of injuries (prevention or deterrence).

The economic approach, drawing on concepts of efficiency, tends to emphasize the deterrent objective of the tort system. The efficiency question is about minimizing the total costs associated with injuries, which includes:

- the costs of prevention, when efforts are put in place to take a certain level of care to avoid injuries;
- the costs of the injuries that nonetheless occur (both economic losses such as material damage, medical care and decreased worker output, and non-economic losses such as pain and suffering);
- the costs of administrative resources (such as attorneys’ fees) to obtain compensation through a court order or a settlement in the shadow of the law;
- and the costs of uncertainty, when potential victims and injurers are risk-averse and may try to reduce that burden through insurance or other risk-spreading mechanisms (with associated transaction costs).

From this perspective, law and economics scholarship over the past 30 or so years has greatly enhanced our understanding of tort rules (Calabresi, 1970; Landes and Posner, 1987; Shavell, 1987 and 2004). It has analyzed the pros and cons of negligence versus strict liability, of contributory versus comparative negligence, of punitive damages, and so on. And it has done so from various angles and for various cases: level of care and level of activity, unilateral and bilateral accidents, injurer and victim being strangers and involved in a contractual relationship, and so on. But much of the analysis is done in a more or less partial framework. That is, the elements of the tort system are studied one by one and in isolation from other social institutions that affect the frequency and severity of

injuries occurring and the handling of claims when losses have been sustained. When it comes to abstract theorizing, such a partial approach is the only fruitful way to proceed, taking full advantage of the *ceteris paribus* setting.

But in the end, theoretical analysis cannot give us a definitive and clear-cut answer as to how the tort system can best be organized in the real world. For, as Schwartz (1994, pp. 382 ff) points out, tort law need not be a *necessary* factor in achieving deterrence. It may be rendered superfluous by other incentives operating on the parties to avoid unduly risky behavior and accidents. These other incentives include: moral principles discouraging people from needlessly inflicting risk and harm on others, the risk of hazardous behavior for the acting party’s own safety, market forces driving unsafe products out or internalizing job hazards in wage differentials, and the regulatory programs put in place by the government for the purpose of achieving specific safety goals in society. At the same time, tort law also might not be a *sufficient* factor to achieve deterrence. It may turn out to be futile in its efforts, for various reasons. Because individuals operate under cognitive and psychological limitations that stand in the way of fully rational behavior towards accident risks, they may be ignorant of the legal rules and the due standard of care, or they may discount a small chance of a major future liability. Negligent conduct can be inadvertent and result from lapses by parties that are genuinely accidental. Liability insurance can intervene and reduce or eliminate the incentive effects of the threat of liability. The same holds for tort litigation, where the prospect of substantial legal costs and an imperfectly predictable verdict may withhold victims from filing a claim or induce them to accept a settlement that does not cover all losses.

For all these reasons, exactly how much deterrence tort law provides is ultimately an empirical question. However, it is not so easy to find the answer. For one thing, we need reliable data, on the level of care, the number of accidents, the frequency and severity of injuries, and the costs of prevention, which in general are not readily available. We also need variation in the tort rules, across time or space, for it is the differential impact that can truly inform us of the effects of the rules. But as soon as we start to compare outcomes across time or space, we have to control for other social, economic and technological developments that might be responsible for the differences. Furthermore, we have to ascertain that the substantive effects of tort law can be distinguished from the effects which may be due to the process of litigation, the organization and regulation of the insurance industry, and governmental safety policies. For these other social institutions condition the working of tort rules, and thus may dampen and mask the actual effects of (variations in) these rules. And they exhibit variations of their own across time and space, which perhaps may give a better explanation of observed differences in the frequency and severity of accidents than (the variation in) tort law. A final remark relates to the fact that decisions on tort reform, as well as on governmental safety policies, are not made within a political void (Epstein, 1988; Rubin, 2005). They are influenced by lobbying efforts from special interest groups (business, the insurance industry, the medical profession, the consumer movement, trial lawyers). As a consequence, the relationship between tort law and the frequency and severity of accidents in society may well be bi-directional, and difficult to disentangle.

Of course, the above problems are not unique for tort law. Therefore it did not come as a real surprise when Landes (2003) showed that empirical analysis plays a much smaller role in the economic analysis of law than in economics in general. He argues that
law and economics scholars are more likely to choose theoretical projects because they hold out the prospect of lower costs and greater rewards than empirical projects. But that is not to say that there is no empirical work at all, as this survey will show. Posner (1972) used a sample of over 1,500 US appellate court decisions in accident cases over the period 1875-1905 to illustrate the ‘classical’ application of the negligence concept. In his interpretation of the historical facts, the standards of due conduct were broadly designed to bring about the efficient (cost-justified) level of accidents and safety. Nevertheless, he also notes that the courts did not lead the way to major innovations in safety methods. In his sample, no enterprise was ever held negligent for having failed to introduce a safety method or appliance that was not already broadly used in the industry.

Since Posner, two main lines of inquiry can be distinguished within the empirical literature on tort law, by far the greater part of which concentrates on the situation and developments in the US. Many empirical studies relate to the operation of the tort litigation system (cf. survey articles by Saks, 1992; Galanter, 1996; Schwartz, 2002; CBO, 2004). A series of questions is addressed in this line of inquiry. Is the number of tort claims more or less stable, or is it getting out of control? Are all claims more or less valid, or is a substantial fraction inflated or even totally fake? Are awards in general and punitive damages in particular moderate and predictable? Are the results about the same if the trial is by judge rather than by jury? Which percentage of tort victims is able to recover for injury losses, and to what extent? And how large are the overhead costs of the tort system?

The second line of research is concerned with the safety effects of the tort system (cf. general surveys by Dewees and Trebilcock, 1992; Schwartz, 1994; Dewees, Duff and Trebilcock, 1996). What are the effects of tort law on the parties’ level of care, on the number of accidents, and on the frequency and severity of injuries? How do these effects compare to the safety effects from regulatory interventions? And what would be the verdict if we - cautiously - try to implement a cost-benefit analysis of the tort system?

Both lines of inquiry will be addressed here, as the results from the second line of research cannot be viewed in the proper perspective without the findings from the first line. I will start, however, with a short detour among some more technical issues.

2. Prerequisites for empirical analysis

2.1 Variation

Most empirical studies on tort law focus on the US. The reason for this is easily understood, even apart from the presumably national bias of the leading law and economics scholars and journals. Although Congress has broad constitutional authority to change tort rules under its power to regulate interstate commerce, tort cases in the US until now have been primarily governed by state law (CBO, 2003). Hence, tort rules may vary across states, and indeed do so to a considerable extent. Tort rules have, moreover, been changing over time.

In the 19th century, US common law generally established negligence as the basis for tort liability. Plaintiffs had to provide sufficient evidence that defendants had been negligent, chiefly defined in terms of the extent to which injury-causing behavior deviated from the normal. In practice, the requirements turned out to be rather restrictive and hence severely limited the scope of the tort system (Priest, 1991). In the course of the
20th century, public debate increasingly emphasized victim compensation and accident reduction. This led to several ‘waves’ of changes in legislation and in court-imposed standards and rules.

The first wave had to do with workers’ injuries (Chelius, 1976; Schwartz, 1994). Around 1900, a number of states passed legislation, expanding employers’ liability by modifying various defenses. Congress followed in 1908 with the Federal Employers’ Liability Act, which broadened the liability of railroads for workers’ injuries by abrogating the fellow-servant rule and replacing contributory negligence as a full defense with comparative negligence as a partial defense. Then, beginning in 1911, the states started to switch from a negligence system to the by now general program known as workers’ compensation. Under this program, an employer is required to pay employees compensation for work-related accidents, regardless of the cause. On the other hand, the employee who is covered by workers’ compensation is barred from any negligence proceeding against his employer. The amount of compensation, which is established individually according to statutes that vary by state, is substantially less than the employee’s full accident costs. So, the program operates as a system of shared strict liability.

A second major wave of changes affected product liability (Higgins, 1978; Priest, 1991). Until the 1950s, manufacturers’ liability for defective products had generally been subject to a regime of negligence under privity of contract. The plaintiff could not sue the manufacturer unless the commodity was purchased directly from the producer. And he had to show that the manufacturer had not exercised reasonable care in the design or construction of the defective product. In the course of the 1960s, most states switched to a standard of strict liability. Strict liability is not absolute, however, as the carelessness of the plaintiff may be brought in defense. Furthermore, it became accepted that defects not only could relate to design or manufacturing, but also to warnings for a proper use.

Liability for automobile accidents underwent a significant change between 1971 and 1975, when a number of states switched from negligence to a system of no-fault for bodily injuries (Schwartz, 2000; Cohen and Dehejia, 2004). No-fault systems generally require drivers to purchase insurance that provides first-party coverage for economic losses, regardless of who was at fault. And they limit the extent to which drivers can initiate tort suits. In a pure no-fault system, victims do not have any recourse to negligence-based suits. However, all US states provide for a threshold, beyond which parties to an accident may sue for negligence. The no-fault systems in the US thus have a hybrid character.

In the context of automobile accidents, there have also been two other relevant changes in liability law. The first one is that over time most states, many of them in the 1970s, replaced the rule of contributory negligence by comparative negligence (White, 1989). Secondly, since the 1980s, many states imposed tort liability on commercial servers of alcoholic beverages, either by statute or by case law (Sloan, Reilly and Schenzler, 1994). Through these so-called dram shop laws, servers can be found liable for accident costs from injuries caused by their intoxicated customers.

The changes in tort law have not happened without fierce discussion. Legal scholars and interest groups argued that the expansion of the tort system went too far. Lawsuits were started frivolously; damages were awarded arbitrarily; the administrative costs of the system, particularly attorneys’ fees, had gone too high; the costs of liability
insurance were rising, to the point that availability became an issue; business was saddled with excessive costs, driving up consumer prices; doctors were restricting their practices or submitting to defensive medicine. In response to these criticisms, a large majority of states has since the mid-1980s enacted statutes to restrict the number of tort lawsuits filed and/or the damages awarded (CBO, 2004; Rubin and Shepherd, 2007). The list of tort reform measures includes: caps on non-economic damage awards; caps on punitive damage awards; higher evidence requirements for punitive damages; allowing the admissibility of evidence of collateral source payments; allowing courts to offset awards by the amount of collateral source payments; restrictions on prejudgment interest; reductions in the statute of limitations; restrictions on contingent fee arrangements; restrictions on joint-and-several liability rules; and certain limitations in manufacturers’ product liability. Interestingly enough, while tort reform has been a national trend, the extent and specifics of that reform vary from state to state.

And that brings us back to the main point of this section. The fact that liability rules have varied so much across time and states makes the US the obvious choice for a comparative analysis of the empirical effects of tort law. That is not to say that there are no alternative objects of study at all. In the field of automobile accidents, for instance, no-fault systems have also been introduced in (parts of) Canada, Australia and New Zealand.

2.2 Data

For an empirical analysis of the operation and efficiency of the tort system, we need data. Alas, getting a more or less complete picture turns out to be an almost impossible job.

To start with, no data have been available until recently that cover all of the tort cases brought in the various jurisdictions across the US (Saks, 1992; CBO, 2003). Data about cases disposed of in the federal courts are available from the Administrative Office (AO) of the US Courts. The National Center for State Courts (NCSC) has been working over the years to bring about more uniform data collection and reporting on trends in civil filings in general jurisdiction state courts. It also conducts periodic surveys of civil trials in the nation’s 75 largest counties for the Bureau of Justice Statistics (BJS).

Even with full court data now, our information would only be partial, for the majority of tort disputes never reaches a trial verdict. Generally, settlement details are not reported to the courts. However, since a large percentage of tort awards is actually paid by defendants’ insurers, insurance company records may be of help here. But note that these companies do not regularly make their records available.

To put the number of tort claims in proper perspective, we should have data on accidents and injuries. The National Safety Council reports on unintentional injury deaths and on medical visits and hospitalization as a result of accidents. There is no institute that keeps track of developments in the frequency and severity of accidents, let alone of the degree to which accidents were due to negligence or pure coincidence. Information can only be obtained through time- and cost-consuming surveys, be it nationally (Hensler et al., 1991) or by sector (see for medical injuries: Mills, 1978; Harvard Medical Practice Study, 1990; Studdert, Brennan and Thomas, 2000).

When it comes to efficiency, we have to balance the costs and benefits of the tort system, both in its current state and with alternative standards and/or rules. Estimates of the administrative costs of the system are published regularly now by Tillinghast, a management consulting firm that relies on data from the insurance industry (see, for
example, Tillinghast-Towers Perrin, 2003). Data on damages avoided because of (additional) precaution are, as a matter of fact, not readily available, but can perhaps be inferred from econometric analyses of the relationships between tort rules and the frequency and severity of injuries (see further on). What we certainly don’t have is information on most of the indirect costs of the tort system, such as the costs of precautions taken by potential injurers.

2.3 Methods
Suppose we wish to compare liability rules across time and/or across states, for example, tort versus no-fault in case of automobile accidents. Our working hypothesis reads that the liability rule affects the average level of care by drivers. For that reason, we may want a statistical estimate and test of the relationship between the liability rule and the number of fatalities in road traffic.

If so, we should take account of other factors that may influence the relationship. For the number of fatalities may also depend on the number of cars and the amount of miles driven, on the presence of speed limits and the intensity of their enforcement, on the degree of experience rating in drivers’ insurance contracts, and so on. One reason for introducing all possibly relevant variables in the relationship to be estimated is to find out which are the real explanatory factors. But there is an even more important reason: failure to control for some explanations by leaving out critical variables may bias the estimates of the effects of the remaining variables. In general, omitting a variable that is not correlated with the variables that are included in the equation does not affect the results. If the correlation is positive (negative), however, the coefficients of the included variables will to some degree take on (be dampened by) the explanatory power of the omitted variable. The included variables will appear more (less) important than they really are. Moreover, standard errors will also be biased, so confidence intervals and hypothesis tests are inaccurate. In the case of automobile accidents, for example, the researcher should be aware that the transition to no-fault may have coincided with certain changes in the insurance regime and in traffic safety measures.

While the obvious solution to the omitted variable problem is to include all potentially explanatory factors, there may be good reasons for not doing so. Available data put more or less stringent limits on our ability to add new variables. If the number of degrees of freedom becomes too small, the standard errors and confidence intervals associated with our estimates will be too large to tell us much. And when the explanatory variables are indeed highly correlated among each another, it may prove difficult to pull the individual effects apart, the so-called multicollinearity problem.

A frequently used method in the empirics of tort is the panel data approach, which yields more observations than pure cross-sectional or pure time-series data. Mostly, the analysis then controls for fixed effects of each state and each year, through the inclusion of dummy variables. The state dummy variables will capture the effects of those variables that differ more or less permanently among the states, but have not been included in the equation (weather conditions, maybe). The year dummy variables will capture the effects of those omitted variables that affect all states over time.

Another issue that deserves attention is the problem of simultaneity or endogeneity. For instance, when a state has decided to make the transition from tort to no-fault, that shift may affect the average level of care by drivers and hence the number
of fatal accidents. But note that the decision in itself may have been taken as a result of a large number of automobile accidents within the state that produced (too) high pressure on the tort system. What then, if we find that no-fault states have more fatal accidents than tort states? Is it because drivers tend to be less careful under no-fault? Or is it because only those states have switched over that for some reason or other (for example, bad weather conditions) are accident-prone, while drivers’ care is not affected at all, or maybe even in the opposite direction but to a relatively lesser degree? Unless the two effects can be separated, any attempt to estimate either relationship alone is bound to bring in the effects of the other.

The appropriate method to deal with this problem is to find exogenous variables that may affect one relationship (for example, the decision to switch from tort to no-fault), but not the other (for example, the number of fatal accidents). The usefulness of these so-called instrumental variables, or, in short, instruments, depends on two factors. Firstly, they must be good predictors for the relationship in which they are included. Secondly, it must be correct to exclude them from the other relationship. In practice, it often turns out to be quite difficult to find instrumental variables that satisfy these two conditions.

A final remark pertains to the functional form of the relationship to be estimated. As it turns out, all kinds of specifications can be found in the empirical literature. For example, the dependent variable sometimes reads in terms of fatalities per (1, 10 or 100) million miles traveled, sometimes in fatalities per 100,000 inhabitants. Sometimes, the dependent variable is in level form, sometimes it is converted by taking logarithms. Alas, theory generally does not give us much to hold on to with respect to the functional form. It can be noted though that a logarithmic transformation may yield certain technical benefits. If the dependent variable varies considerably across states and/or time, as the accident rate certainly does, its logarithm presumably has a distribution that is less skewed. This will tend to make the coefficients and their confidence intervals more reliable. Moreover, coefficients in a log-log equation can be read as elasticities.

This is not the place to delve any further into technical details on the appropriate statistical methods to estimate coefficients and test hypotheses in relationships. For that, the reader is referred to the econometric literature; see, for example, Greene (2008). Spelman (2000) provides a highly readable guide to the various methodological problems and the ways to solve them, for the admittedly different, but as to the essentials very much comparable, context of crime and punishment.

3. The tort litigation system

This section surveys the empirical evidence concerning the operational behavior of the tort litigation system. The evidence will be organized within a general model of the flow of disputes into and through the system (Saks, 1992). This model may be visualized in the standard approach of legal scholars as a ‘dispute pyramid’ made up of successive layers (Felstiner, Abel and Sarat, 1981; Galanter, 1996). The model tries to unravel the path from adverse events to final dispositions. After an adverse event occurs, a victim must decide whether or not to complain; lawyers must decide whether or not to accept and file the cases offered to them; a process of negotiation resolves most cases short of a trial; an important minority of cases will be resolved by trial; and various post-verdict remedies are available. A parallel course is followed by plaintiffs who have suffered non-
compensable injuries which are nonetheless brought to the system; these claims are either correctly put aside or mistakenly granted compensation at one stage or another.

Examining the system in this way brings to light how decisions at one stage affect what may come to the surface at later stages. It can tell us to what degree victims are compensated for the tortious conduct of their injurers. And it informs us to what extent the system makes (potential) injurers internalize the damages they bring about by their conduct, and under what margin of error. Thus, it provides a frame of reference for studying the effects of changes in the system and in the external environment.

3.1 Base rate
To have a full view of the dispute pyramid we should start at the base, consisting of all the events in which, for example, a particular product is used. In a - small - fraction of these events someone gets hurt. This yields the first layer of adverse events.

Hensler et al. (1991) report on a large-scale national survey by the RAND Institute for Civil Justice in 1988/89. They find that a person has a 15.8 percent probability of suffering some economic loss from a non-fatal injury in a year. Some 38 percent of the person-incidents are caused by slips and falls; 30 percent involve a product, such as a tool, a household appliance or sports equipment; 18 percent involve motor vehicles; and 1 percent results from medical treatment. Motor vehicle accidents are much more likely than product-associated incidents or slips and falls to result in very serious injuries and to require hospitalization or surgery.

More recent statistics from the National Safety Council are on rather similar lines. It is estimated that in 2004 11.5 percent of the population sought medical attention for an injury. From the unintentional-injury deaths estimated to total 112,000, that is 38.1 per 100,000 population, some 45,300 were killed in motor vehicle crashes and 5,000 at work.

Ideally, we should be able to separate the accidental injuries into those which are tortious and those whose costs must remain with the victim. When the RAND study asked respondents about what they considered to be the main cause of their injury, 10 percent attributed the accident to chance only, 50 percent mostly to one’s own behavior, 29 percent mostly to someone else, and 11 percent equally to oneself and another person. But these subjective judgments do not constitute solid ground for research into negligence questions.

The most useful studies in that regard relate to medical malpractice. Three large-scale surveys of the medical records of hospitalized patients have investigated the incidence of injury due to medical care and the subset caused by negligence. The first study was done in California in 1974 (CMA and CHA, 1977; Mills, 1978), the second in New York in 1984 (Harvard Medical Practice Study, 1990; Brennan et al., 1991; Leape et al., 1991; Weiler et al., 1993) and the third in Utah and Colorado in 1992 (Studdert, Brennan and Thomas, 2000; Studdert et al., 2000; Thomas et al., 2000). The California study concluded that 4.65 percent of hospitalized patients suffered an injury due to medical care. Of these, 17 percent involved a negligent act or omission. In the New York study 3.7 percent of patients suffered an injury due to medical care, 28 percent of which were attributed to negligence. And in the Utah and Colorado study, adverse events occurred in 2.9 percent of hospitalizations, 29 percent of which were due to negligence.

Thus, the results are remarkably similar as they point out that between 0.8 and 1.0 percent of all hospitalized patients suffer a negligent injury. The results differ, however, with respect to mortality rates. Extrapolating the Utah and Colorado results to all hospital admissions in the US in 1992 reveals that 980,000 adverse events might have led to 65,000 deaths, of which 25,000 can be attributed to negligence. That burden of mortality is considerably less than the estimate from the New York study: nearly 200,000 deaths a year due to adverse events, of which 120,000 are negligent. But the mortality rate remains a startling figure, as such, and in comparison to the fatality rate in motor vehicle and workplace accidents.

The figures on iatrogenic injury rates, instructive as they are, should nevertheless be viewed with some caution (Danzon, 2000). A substantial - but undetermined - proportion of the patients were seriously ill and many would have died from their underlying illness anyway, whereas most victims of automobile and workplace injuries were healthy. Secondly, the findings reflect the broad definitions used in the studies. A negligent adverse event was defined as the consequence of treatment that failed to meet the standard of the average medical practitioner. None of the studies attempted to define negligence by weighing marginal costs and benefits of additional precautions. So, the resulting count of negligent injuries does not necessarily correspond to economically inappropriate injuries.

3.2 Claiming rate

The second layer of the dispute pyramid discloses how many of the injury victims take measures to obtain compensation from those who injured them.

The RAND study cited above found that 90 out of 100 accidentally injured people did not take any action at all. The remaining ten that took some action pursued multiple paths; two tried to negotiate directly with the injurer and four with his insurer, while seven consulted an attorney at one time or other. Eventually, four out of 100 hired a lawyer and only two filed a lawsuit.

It is tempting to compare the 10 percent of injured people that took some action and the 2 percent filing a lawsuit with the 29 percent that attributed the cause of the accident mostly to someone else’s behavior. This suggests that a large number of potential plaintiffs with a valid claim never initiate one and thereby become instant ‘false negatives’.

The RAND study further points out that claiming patterns may differ sharply between subclasses of accidental injuries. While 39 percent of the persons who were injured in a motor vehicle accident took some action to obtain compensation, this only holds for 11 percent in the case of work accidents and 4 percent with other injuries. Note that this pattern correlates well enough with the percentage of accidents being attributed mostly to others, which ranges from 71 percent in case of motor vehicle accidents, 38 percent for work-related accidents, and 13 percent for other injuries.

More specific information is available for medical injuries. The California study did not take a direct look at claims filed by patients within the sample, but the other two studies did. In the New York study (Localio et al., 1991) only 1.5 percent of the patients who were identified as having sustained an injury due to negligence filed a malpractice claim, and in the Utah/Colorado study the figure was 2.5 percent. These results affirm that claims lag well behind the incidence of negligent injury. Moreover, factors other than
individual merit appear to play a role in determining who uses the malpractice system. Compared with patients who did sue for negligence, non-claimants were found more likely to have suffered minor injury and to be Medicare/Medicaid recipients, elderly and low income earners (Danzon, 1985; Studdert et al., 2000).2

There is also a significant number of ‘false positives’ among medical malpractice claims. Starting from the California study, Danzon (1985) made a comparison between the extrapolated total number of negligent injuries and aggregate claim data from insurers’ records. It appeared that, overall, 1 in 10 negligent injuries resulted in a claim. Later studies found this ratio to be 1 in 7.6 (New York), 1 in 5.1 (Utah) and 1 in 6.7 (Colorado). Once again, this claim ratio of between 10 and 20 percent shows how negligent behavior frequently escapes redress. But note that this overall claim ratio is definitely higher than the 1.5 to 2.5 percent figure that followed from a direct look at claims filed by patients who were known to be treated negligently. Indeed, it was estimated from the New York results that only 1 in 6 malpractice claims responded to an identifiable injury due to negligent medical behavior (Localio et al., 1991).

Upon closer scrutiny, the picture is somewhat less alarming. Several studies have investigated sets of closed malpractice claims, focusing on a single specialty (Cheney et al., 1989; Sloan et al., 1997), insurer (Taragin et al., 1992), hospital (Farber and White, 1994) and state (Harris et al., 2006), or covering the whole terrain (Studdert et al., 2006). Examination of medical records made it clear that 40 to 60 percent of claims did not involve any negligence and hence had no merit. That fraction is still considerable, but it is, notably, an ex post and expert judgment. Plaintiffs may have filed the claims in good faith, from a state of imperfect information. It is up to the tort litigation system, then, to separate the rightful claims from the non-deserving ones.

3.3 Disposition of claims
The third layer of the dispute pyramid discloses how filed claims fare in the system.

Although there are no all-encompassing statistics of tort filings in the US, combining data from various sources provides a useful overview. BJS data for the fiscal year 2002-03 (Cohen, 2005) show that tort filings in federal courts amount to some 34 per 100,000 population. Tort filings in state courts are far more numerous, at an average rate of some 280 claims per 100,000 persons, according to NCSC data for 2004 (Schauffler et al., 2006). Around this average there is substantial variation in the rate at which tort cases are filed in state courts, from 90 per 100,000 in North Dakota to 786 in New Jersey. Densely populated states in the North East generate many of the highest figures.

Available data also give some insight into the composition of the incoming tort cases. Schauffler et al. (2006) report how automobile accidents clearly dominate in state courts with a share in torts of 51 to 67 percent. Medical malpractice and product liability cases amount to no more than 3 and 4 percent, respectively. These proportions are corroborated by somewhat more detailed findings from a sampling of tort cases in the nation’s 75 largest counties (Smith et al., 1995) and in Georgia (Eaton and Talarico, 1996; 2 Posner (1997) shows that interstate differences in factors like these (income, education, age distribution) may help to explain the variance in overall per capita tort filing rates across the US, and between the US and England. He suggests that, having corrected for these factors, England might be more litigious than the US.
When it comes to the disposition of the claims, we must first of all face up to the fact that a large majority is settled in the ‘shadow of the law’. Cohen (2005) notes that only 2 percent of tort cases in federal courts in 2002-03 were decided by trial. Smith et al. (1995) give more details for tort case disposition in state courts. The most common method is an agreed settlement (73 percent), followed by dismissal (10 percent), transfer (5 percent), arbitration award (4 percent) and default judgment (3 percent). Alas, no general and systematic data are available regarding the substantive outcome of the underlying dispute under these various headings. But this much is clear that a trial verdict resolves relatively few cases (3 percent).  

The composition of trial cases differs between courts (Cohen, 2004 and 2005). Automobile accident, product liability and medical malpractice cases make up 20, 15 and 10 percent, respectively, of federal court trials, and 53, 2 and 15 percent, respectively, of state court trials. When the latter distribution is compared to the incoming case load, it becomes apparent that medical malpractice cases are relatively more often brought to trial than other tort cases.

The average plaintiff win rates in federal and state court trials, 48 and 52 percent, are almost perfectly in conformity with the 50 percent rule derived by Priest and Klein (1984). If there is no real asymmetry in the stakes of the parties, and if significant legal costs can be avoided by settling, they predict that only close call cases will proceed to trial. But going into more detail, plaintiff win rates differ markedly between automobile accidents on the one hand (around 60 percent) and product liability and medical malpractice cases on the other hand (some 40 and 30 percent, respectively). Differential stakes and information clearly play an important role, as automobile accident cases are mostly between individuals, while defendants in product liability and medical malpractice cases are mostly ‘repeat players’ (business, hospitals, insurers). The hospital or producer whose reputation may be harmed has more at stake than the damages sought by the plaintiff. Therefore, defendants may be more willing to settle strong plaintiff suits.  

The final amount awarded in federal and state court trials has a median of $210,000 and $27,000 respectively. Here too, major differences can be observed between case types. The trial awards are on average substantially lower in automobile accident cases, especially in state courts (median $164,000 and $16,000), than they are in product

3 See also Eaton, Mustard and Talarico (2005) with rather more recent, but somewhat comparable results for Georgia.

4 Viscusi (1986, 1988, 1989) provides detailed analyses of the disposition of product liability claims, using insurance data from all US states on claims closed between 1976 and 1977. See Viscusi and Scharff (1996) for a summary. Of all claims, 19 percent are dropped. Of the claims not dropped, 95 percent are settled. The plaintiff win rate at verdict is 37 percent. To assess the measure of compensation, Viscusi considers the replacement ratio, that is the ratio of the payment (through settlement or court verdict) to financial losses due to bodily injury (thus, excluding compensation for pain and suffering). For the full sample, on average, this replacement ratio is 1.05, so there is slight overcompensation of economic losses. The replacement ratio declines quite steadily with the size of losses, and drops below unity for losses in excess of $100,000. For losses over a million dollars, the replacement ratio is only 0.25.

Overall, the performance of the tort litigation system closely reflects the predictions of economic models. Decisions to drop and settle claims are dictated by economic stakes, not by unconstrained commitments to justice. Similarly, the magnitude of out-of-court settlements reflects the character of the legal bargaining game.
liability and medical malpractice cases (where the median ranges from $350,000 to $600,000).

An additional source of variation stems from the fact that trial can be by jury or by judge. The plaintiff win rate appears to be somewhat higher in a bench trial, while the awarded amount tends to be higher in a jury trial.5

Finally, it should be noted that a trial award needs not be the final result of the litigation process. A party that is unhappy with the size of a jury award can request for a review from the trial judge. And parties can, of course, enter an appeal, the threat of which in itself can lead to further negotiations. In a study by Shanley and Peterson (1987) 20 percent of jury awards were found to be adjusted in the post-trial stage. On average, defendants paid 71 percent of what juries had awarded.

Instructive as they are, the preceding figures on plaintiff win rates and awarded amounts do not tell us whether the tort litigation system reached the correct decisions. Were false positive claims at one moment or other put aside? And were rightful claims remunerated in such a way as to adequately compensate for false negatives in creating incentives for the (potential) injurers?

Once again, analysis of medical malpractice cases has raised a corner of the veil. Several studies have modeled the handling of claims (drop, settle, proceed to trial verdict) as a process in which the parties form expectations about the probability and size of award at verdict and about the litigation costs. Empirical tests lend support to the non-random selection of cases to trial verdict. The data are more or less consistent with a model in which plaintiffs are poorly informed ex ante about whether there has been negligence, file suit to gather information, and either drop the case if they find that negligence is unlikely or settle if negligence is likely (Farber and White, 1991 and 1994). Thus, legal standards (that is, prevailing practice) influence court verdicts directly and settlements indirectly. And plaintiff win rates at settlement (around 50 percent) are higher than at verdict (30 percent) (Danzon and Lillard, 1983). Overall, compensation in most cases falls short of plaintiff’s losses, especially for more serious injuries (Sloan and Hoerger, 1991). Harris et al. (2006) also find that cases are generally settled when negligence is rated as probable by the defendant’s insurer. But when negligence is rated as uncertain or unlikely, ‘strategic variables’, such as the witness potential of the plaintiff versus the defendant physician and the reputation of the plaintiff’s attorney, play a supplementary role. It appears that cases in which the defendant (plaintiff) has a strategic advantage are much less (more) likely to settle.

Other studies provide illustrative descriptive data. White (1994) summarizes existing evidence and concludes that the probability of a claim is 0.026 per negligent injury, 0.01 per non-negligent injury and 0.001 per non-injury. Quality of care also has a strong impact on the probability and size of a compensatory award. Valid claims are

5 It has to be acknowledged that most of these differences in mean awards and win rates may be due to selection effects. Clermont and Eisenberg (1992) suggest that parties, through their lawyers, send quite different mixes of cases to the two modes of trial as a result of persistent misperceptions of judge/jury differences for various tort categories. But judges and juries also exhibit some ‘real’ differences as adjudicators. Helland and Tabarrok (2000) note that juries, in particular if they are drawn from pools with high poverty rates, appear to be more receptive to ‘redistribute wealth’ arguments than judges. Tabarrok and Helland (1999) pursue a similar line, when they argue that elected judges may grant higher awards in cases with out-of-state defendants, ceteris paribus, than appointed judges.
much more likely to receive payment than invalid claims, and at a higher average amount. The percentage of patient claims that result in actual compensation varies between 73 and 91 percent in case of negligent injury, and between 16 and 47 percent in case of non-negligent care (Cheney et al., 1989; Taragin et al., 1992; Farber and White, 1994; Studdert et al., 2006).

The tort litigation system is not perfect, then. It sometimes makes doctors - or their insurers - pay damages for non-negligent care. But the system is clearly not a random lottery. As a result of a selection process, negligent injuries are at least ten times as likely to end up in compensatory payments as non-negligent injuries. More disturbing for the proper working of the system seems to be the high rate of false negatives, mainly because a large fraction of valid claims is not filed, but to a lesser degree also because not all valid claims that are filed get honored. If that result is combined with the finding that compensation generally falls short of victims’ losses, the deterrent function of the system might be insufficient after all.

3.4 Punitive damages
Punitive damages have attracted special attention over the past decades, both in public policy debates and in empirical research. From a deterrence perspective, the rationale of punitive damages lies in the fact that an injurer sometimes may escape liability. The level of damages imposed on him when he is found liable then needs to exceed a compensatory award, so that, on average, he can expect to pay for the full harm he inflicts upon his victims. To play this incentive role properly, punitive damages should be set equal to compensatory damages multiplied by \( \frac{1-p}{p} \), where \( p \) denotes the probability that the injurer is detected and found liable. However, the incentive role loses weight, if the frequency with which punitive damages are awarded is basically unpredictable and the size arbitrary.

Eisenberg et al. (1997) studied the predictability of punitive damage awards in jury trials. They note that punitive damages are awarded in only 6 percent of jury trials where plaintiffs prevail. Most of the punitive damage awards occur in business/contract and intentional tort cases, and only a small minority of 4 percent in product liability and medical malpractice cases. The median punitive damage award is $50,000, which is not strikingly high. However, the mean value of $534,000 shows that the distribution is rather skewed, with some high-end awards. Most importantly, regression analysis shows that the size of the punitive damage award is strongly and significantly correlated with the compensatory award.\(^6\)

The implications which seem to follow from these findings, that punitive damages are (1) rare and insignificant and (2) predictable, have not gone unchallenged (Polinsky, 1997). Firstly, even if punitive damages are on average rather insignificant at trial, it does

---

\(^6\) Several studies compare judge and jury performance in awarding punitive damages, with mixed results. Eisenberg et al. (2002) find no substantial evidence that judges and juries differ in the rate at which they award punitive damages or in the relationship between the size of the punitive and compensatory awards. Hersch and Viscusi (2004), on the contrary, report that juries are significantly more likely to award punitive damages and award higher levels of punitive damages than judges. The results in Eisenberg et al. (2006) affirm that juries and judges award punitive damages in approximately the same ratio to compensatory awards, conditional on the existence of a punitive award. But they also reveal some differences in juries’ and judges’ tendencies to award punitive damages between bodily and no-bodily injury cases.
not follow that the effect on settlements is minimal as well. However, empirical findings with respect to the settlement process suggest that there is no need for alarm. Eaton, Mustard and Talarico (2005) study whether and how the major decisions in the tort litigation process are affected by a request for punitive damages. In most phases of the litigation process they can find no significant impact. Koenig (1998) analyses liability insurance data from Texas, where claims adjusters must report on the components of all closed claims settled without a verdict. In only 5.5 percent of the cases which settled for $10,000 to $25,000 did the threat of punitive damages play any role. If so, the payout was inflated by an average of only 12.5 percent. In larger settlements, 43 percent of claimants received more than economic losses. If so, punitive damages made up 11 percent of the settlement amount, against 56 percent for non-economic damages.

The second target of criticism is the predictability of punitive damages. It may be true that a substantial part - some 50 percent - of the variation in the size of punitive damages can be explained, conditional on knowledge of the compensatory award and of the assignment of a punitive award. But that still leaves the issue whether punitive damages will be awarded at all. Karpoff and Lott (1999) show that without prior information on the allotment of a punitive award less than 2 percent of the variation in punitive damages can be explained. An additional problem is that the compensatory award is not known until after the verdict, and that it is unlikely that parties might forecast it without error. Viscusi (1998b) concludes that there is no real basis for decision makers to predict the punitive awards arising from different safety choices.7

3.5 Administrative costs

A major source of concern is the considerable cost of administering the tort system. Kakalik and Pace (1986) estimated total expenditure for tort litigation in state and federal courts in 1985 to be between $29 and $36 billion. From this total 46 percent was paid to plaintiffs as net compensation, while the rest was taken up in legal fees, insurance company costs of claim processing, the value of litigants’ time, and the costs of operating the court system. Defendants’ costs made up 28 percent of total expenditure, plaintiffs’ costs 24 percent, and court expenditure 2 percent. Apparently, it takes society on average more than $2.10 to deliver $1 of net compensation to a tort victim.

Automobile tort cases, which account for 53 percent of all tort filings, only cover 41 percent of total expenditure. The average net compensation per claim is lower, but more importantly, so are litigation costs, as these cases are usually relatively straightforward to handle. For that reason, delivering $1 of net compensation in an automobile tort case costs society $1.90, and over $2.30 in other tort cases.

More recent estimates corroborate the above findings. Studdert et al. (2006) provide some detail on medical malpractice claims, suggesting that the average administrative cost of handling such a claim at present amounts to more than $140,000. Tillinghast-Towers Perrin (2003) estimate total expenditure on tort claims nationwide at $290.0 billion in 2005, that is 2.33 percent of GDP. Plaintiffs receive 46 percent in net compensation, to wit: 22 percent for economic losses and 24 percent for non-economic loss. Defendants’ costs (insurance companies’ administration and defense costs) amount to 35 percent, and plaintiffs’ attorney fees to 19 percent. Once again, it appears to cost

---

7 Indeed, Viscusi (1998a) cannot find any systematic differences in the safety and environmental performance between states with and without punitive damages.
society overall more than $2.10 to deliver $1 of net compensation.  

4. *Effects of tort reform on litigation and insurance*

The previous section gave an overview of the main characteristics of the operational behavior of the tort litigation system. Historical developments were left aside. However, some of the trends over time are highly relevant, as they provided the trigger to tort reform initiatives.

A brief look at the scarce figures that sketch long-term trends reveals that the total number of tort filings rose steadily between 1970 and the mid-1980s, both in state and federal courts (Ostrom et al., 2004; Schauffler et al., 2006; Cohen, 2005). Observing this trend in litigation, Hensler (1987) noted that there was no single tort system. At least three different types of tort litigation should be distinguished: (1) ordinary accident litigation, best illustrated by cases that arise out of automobile accidents; (2) ‘high stakes’ litigation, illustrated by product liability and medical malpractice lawsuits; and (3) mass latent injury cases, such as asbestos. The observed steep growth in (per capita) numbers of lawsuits and (inflation-corrected) median verdicts could mainly be attributed to product liability and medical malpractice cases, which showed much more volatility over time than the rather stable development of automobile accident cases. The ensuing escalation in liability insurance costs caused problems with insurance premiums, and with insurance availability (Priest, 1987 and 1991; Viscusi, 1991a, 1991b). A first ‘crisis’ in the mid-1970s was more or less concentrated in the medical malpractice sphere and remained confined to a number of states, the next in the mid-1980s was much more general. Following these ‘crises’ most US states have adopted tort reform measures. The extent and specifics of that reform, however, vary from state to state.  

Some reforms make it more costly or difficult to file tort cases, such as by imposing limits to the application of the joint-and-several liability rule, reductions in the statutes of limitation, and caps on legal contingent fees. Other reforms aim at a reduction of damage awards, such as by limits on non-economic damages, limits on punitive damages, and reforms of the collateral source rule which prohibits the introduction at trial of evidence about the plaintiff receiving payments from other sources.

The figures for long-term trends suggest that tort reform has had some success. The steadily upward trend in tort filings was curbed at the end of the 1980s, with some serious up- and downswings until the end of the 1990s, and an apparent downward movement since then. Notwithstanding, tort reform is a continuing issue in the U.S. (see, for example, Black et al., 2005). That makes it interesting to know which of the tort reform measures have been the more effective.

Several empirical studies have analyzed the effects of tort reform on the frequency and severity of tort claims and on liability insurance losses and premiums. In general, they did not address the social desirability of the reforms.  

---

8 Hersch and Viscusi (2007) analyze tort litigation costs in Texas over the years 1988-2004 and report total transaction costs for each dollar received by claimants at $0.75.

9 A listing of tort reforms by state up till now can be obtained from the American Tort Reform Association (ATRA) at www.atra.org. ATRA was co-founded in 1986 by the American Medical Association and the American Council of Engineering Companies.

10 The one or two studies that analyzed the effects of tort reform on the number of accidental injuries and on defensive medicine will be addressed in section 5.
know whether the use of the tort system was indeed too intense, leading to overinvestment in care and defensive medicine. And one should have more precise information on whether the problems in the insurance industry were indeed due to overclaiming, or resulted from more or less ‘natural’ cycles in building up loss reserves and in investment returns on those reserves.

Most attention has been given to medical malpractice (cf. partial surveys in Studdert, Mello and Brennan, 2004 and Rapp, 2006). First, Danzon (1984) showed that pro-plaintiff common law doctrines adopted prior to 1970 contributed significantly to claim frequency. Subsequently, Danzon (1986) analyzed the tort reforms that followed in the mid-1970s. States that enacted shorter statutes of limitations and set outer limits on discovery rules had less growth in claim frequency than states that were more lenient to patients. Statutes permitting or mandating the offset of collateral benefits reduced both claim frequency and awards. And damage caps were effective in reducing plaintiffs’ recovery. The latter result is also found by Yoon (2001), who reports on the implementation and nullification of damage-cap laws in Alabama between 1987 and 1995.11 Other studies examined liability insurance company losses and premiums. The only reforms found by Zuckerman, Bovbjerg and Sloan (1990) to significantly lower premiums were those that impose a cap on the provider’s liability or reduce the statute of limitations. That damage caps may indeed be an effective instrument (and perhaps the only one) to reduce losses, lower premiums and enhance insurer profitability is corroborated by the empirical results in Born and Viscusi (1998), Thorpe (2004) and Viscusi and Born (2005).

Some studies discuss the consequences of reform in other fields of tort. Viscusi et al. (1993) report that limits on non-economic damages and modifications of the joint-and-several liability rule have been able to restrain the costs of general liability insurance (which includes product liability). On the other hand, Lee, Browne and Schmitt (1994) find no real evidence that joint-and-several liability reform reduced the total number of tort filings, apart from a surge in the period before the reform took effect. And Browne and Puelz (1999) examine how tort reform has affected automobile accident claims. Caps on non-economic damages and collateral source reforms result in lower awards, while joint-and-several reforms are associated with an increase in non-economic damages. Caps on non-economic damages also reduce the number of filings, while reforms of the joint-and-several and collateral source rules have no significant effect.

On the whole, the most consistent finding is that caps on damage awards reduced the number of lawsuits filed, the value of awards, and insurance costs. There is little systematic evidence that any other type of reform had a significant impact on any of the various outcome measures studied. But that result may, after all, be more reflective of the lack of data than of a failure of the reforms (CBO, 2004).

5. Safety effects

Let us now address the incentive and safety effects. What is the impact of liability rules on the parties’ level of care, on the number of accidents, on the frequency and severity of

---

11 In another article, Yoon (2004) examines the introduction of pretrial mandatory arbitration by screening panels in Nevada in 1986. The reform led to a decrease in the percentage of claims being resolved by courts, but did not affect plaintiffs’ recovery.
injuries? To survey that literature in a meaningful way, it is an absolute necessity to disaggregate and categorize. Between categories of accidents, there are large differences in the factors that bear on incentives. Some categories of accident are generally of a bilateral nature, where the level of care by the victim plays its role; medical malpractice on the other hand mostly has a unilateral character. Some categories of accident occur in a contractual setting, where parties can bargain on safety elements and where prices can adjust; automobile accidents, on the other hand, generally occur between strangers. There are also important differences between accident categories in the price structure of insurance, and in the extent of public safety regulation. The evidence is therefore presented on a sector-by-sector basis.

5.1 Automobile accidents
The natural category to start with is that of automobile accidents. For one thing, it is the source of by far the largest number of claims. The most important characteristic of traffic accidents is that they are generally between strangers. Hence, transaction costs stand in the way of ex ante bargaining about the level of care and activities. Another important characteristic is the interchangeability of injurers and victims. The injurer who causes an accident by being careless may well get personally injured himself. This stimulates drivers to take a reasonable level of care, even in the absence of liability rules and public safety regulation. Most drivers carry some kind of insurance, which has generally been subject to a considerable degree of experience rating, at least under negligence.

Note that it is taken for granted (1) that the causes of most accidents can be identified with some degree of certainty, (2) that drivers can alter their accident-causing behavior (3) and that financial incentives can play a role (Bruce, 1984; Grayston, 1973).

No-fault
Starting with Massachusetts in 1971 some 19 US states, New Zealand, one province in Canada and three provinces in Australia, have replaced tort law with no-fault systems for the compensation of personal injury losses in traffic accidents. Under a pure no-fault system, injurers are not liable for victims’ damages. Instead, victims bear their own damage or collect compensation from their own insurance company. Proponents of no-fault argue that it has lower administrative costs because there are fewer lawsuits, and that it is more equitable because victims can collect for their damage regardless of whether injurers are negligent. In practice, there are several variants of no-fault. Under ‘pure’ no-fault, victims must always collect damages from their own insurance company. No compensation is paid for non-economic damages (pain and suffering) and there are limits on recovery for economic damages (such as lost wages and medical costs). Under ‘mixed’ versions of no-fault, victims are allowed to opt out and to sue injurers under the tort system if their losses exceed a monetary or verbal threshold. Under the tort system, they may receive higher compensation, including non-economic damages. Mixed versions of no-fault are the rule in US states; pure no-fault has only been adopted in New Zealand, Quebec and the Northern Territory of Australia.

---

12 Some US states that adopted no-fault later repealed it. By now, the number of states with a no-fault system has fallen to 14, with six states switching status in between.
13 This effect is documented in, for example, Devlin (1992) and Grabowski, Viscusi and Evans (1989).
Swan (1984) and McEwin (1989) present empirical evidence on the switch to no-fault in New Zealand and the Northern Territory, compared to other provinces of Australia; see also Brown (1985). The switch, with its abolition of negligence, was found to be associated with a substantial increase of 16 to 20 percent in the number of road fatalities. However, as McEwin points out, the results should be interpreted with care as the switch to no-fault was really a combination of changes in liability and insurance. Pure no-fault combined with compulsory non-merit-rated first-party insurance replaced tort with compulsory non-merit-rated liability insurance. The switch should be decomposed into its liability effects (the loss of the incentive to meet the negligence standard, but also an increase in uncompensated non-economic losses which might induce more care) and insurance effects (an increased coverage without merit rating to counter moral hazard). From his finding that the switch to mixed no-fault in Victoria and Tasmania did not affect road accident fatalities, McEwin concludes that compulsory first-party insurance, by itself, had no impact on road safety.

The switch to pure no-fault in Quebec was also a combination of changes in liability and insurance, but with somewhat different characteristics. Here, the existing private insurance system with experience-rated individual premiums was replaced by a public insurance system with a flat-rate pooled premium across all drivers. The vehicles that were previously uninsured, some 15 percent of the total, were forced to carry insurance. Devlin (1992) reports that the switch increased fatal road accidents by 9.6 percent. In the same vein, Gaudry (1992) concludes from his results that the switch certainly decreased road safety, although his finding of a 3.3 percent increase in fatal accidents is not very significant.

Let us now turn to the US states that adopted a mixed system of no-fault. A victim can bring a tort claim for his losses, but only if his injury is serious enough to exceed a ‘threshold’. Such a threshold can either take a dollar form for the victim’s costs of medical treatment or it may give a verbal description of the injury (‘serious impairment of body function’, ‘permanent serious disfigurement’). The real value of these thresholds varies widely, not only between states, but also over time (much of which is due to inflation). For example, in 1987, the proportion of personal insurance claims that was ineligible for tort recovery under a no-fault threshold varied between 37 percent for New Jersey and 88 percent for Michigan. There is also considerable variation between states and over time in the organization of insurance. Since 1970, when the purchase of automobile liability insurance was compulsory in only three states, many states have introduced that obligation. All states that introduced no-fault limitations on liability also adopted compulsory insurance requirements at the same time. Alas, the literature remains rather vague on the degree to which premiums depend on drivers’ behavior.

Landes (1982) delivered the first empirical study of the relationship between no-fault restrictions on tort and road fatalities in the US. She found that the adoption of no-fault resulted in 4 percent additional fatal accidents when the threshold barring claims from tort recovery was low ($500), and over 10 percent when the threshold was high.

---

14 Empirical studies in this field generally focus on fatalities, as it is the indicator which is most consistently measured over time and between jurisdictions. Derrig, Weisberg and Cheng (1994) and Cummins and Tennyson (1996), for example, provide evidence of significant ex post moral hazard effects in bodily injury claims.

15 Data from the Insurance Research Council, as reported in Cummins, Phillips and Weiss (2001).
($1,500). This decrease in road safety was subsequently confirmed by Medoff and Magaddino (1982), but contradicted by Kochanowski and Young (1985), DOT (1985) and Zador and Lund (1986), while Kabler’s (1999) results are mixed. However, there are doubts about the econometric methodology of these studies. They do not adequately control for differences in other state characteristics that may affect accident rates. Moreover, they do not take account of the potential endogeneity of no-fault (cf. Harrington, 1994). More recent studies try to avoid these problems. Cummins, Phillips and Weiss (2001) estimate that the switch to no-fault increased fatal accident rates by 7-13 percent, dependent on whether they use a 0-1 indicator or the tort claim ineligibility ratio. Sloan, Reilly and Schenzler (1994) and Derrig et al. (2002) also report an increase in road fatalities. Maybe the most careful analysis until now is by Cohen and Dehejia (2004), who try to unravel liability and insurance effects. Apart from the unequivocal moral hazard costs as a result of compulsory automobile insurance, the reductions in accident liability produced by no-fault laws appear to have led to an increase in traffic fatalities of about 10 percent.

A somewhat different perspective is followed by Cummins and Weiss (1992). They note that most US no-fault laws make no change in the legal (tort) rules involving property damage claims. As no-fault nevertheless appears to be positively associated with the frequency of total property damage claims, this yields indirect evidence that no-fault weakens driver incentives. Still another approach is taken by Devlin (1999), who uses micro-data on the severity of bodily injuries. She finds that the probability of sustaining a more serious accident is higher in no-fault states, which is taken as yet another sign that drivers take less care under no-fault. But it seems perfectly possible that her finding is after all due to a reporting effect.

To be fair, it must be observed that the more recent literature is not entirely unanimous. Loughran (2001) argues that no-fault in the US could hardly have a substantive impact on driver behavior, as it does not affect the expected cost of an auto accident to the at-fault driver by any significant amount. Moreover, he provides empirical tests that suggest that no-fault had no statistically significant effect on the number of fatalities, the overall accident rate as measured by property damage claims and the rate of driver negligence in fatal accidents.

Other changes in liability rules
Apart from the switch to no-fault in a selection of states, there have been two other, more widespread changes in liability rules that received some attention in the empirical literature.

The first of these is the shift from negligence with the defense of contributory negligence to some form of comparative negligence. Following White’s (1989) calculations such a shift may lower incentives to avoid accidents. Sloan, Reilly and Schenzler (1995) report from survey data that the switch may have increased binge drinking among drivers, but Sloan, Reilly and Schenzler (1994) do not find a significant impact when it comes to traffic fatality rates.

On the other hand, dram shop liability rules appear to have a rather robust, significant, positive effect on traffic safety (Chaloupka, Saffer and Grossman, 1993; Sloan, Reilly and Schenzler, 1994; Ruhm, 1996; Mast, Benson and Rasmussen, 1999; Young and Likens, 2000; Sloan et al., 2000; Whetten-Goldstein et al., 2000; Liang, Sloan
Other safety measures

Ever since motor vehicle accidents became a major cause of death, governments have been engaged in various kinds of safety-enhancing measures. Starting in the mid-1960s US federal regulation has made seat belts, energy-absorbing steering columns, penetration-resistant windshields and so on, standard elements of vehicle design. The construction and maintenance of roads and intersections have been dealt with. And public policies have been directed toward changing drivers’ behavior, by regulating and enforcing speed limits and seat belt usage and by discouraging drunk driving.

There is a vast literature on these safety measures, which cannot be surveyed within the scope of this chapter. However, it may be useful to give a short introduction to the relevant empirical law and economics studies.

Firstly, the no-fault studies discussed above generally control for other factors that might affect traffic safety. Among the control variables one comes across are seat belt laws, speed limits and so on. Alas, the set of variables varies from study to study and does not yield very robust results. Seat belt laws, for example, are sometimes reported to have a significant negative effect on traffic fatalities (cf. Swan, 1984), but this result is not confirmed in other places (see McEwin, 1989 or Cohen and Dehejia, 2004).

A second strand of the literature is centered on alcohol control policies. Cook and Tauchen (1984) found that reductions in the minimum legal drinking age (MLDA) during the early 1970s from 21 to 18 caused a substantial increase in auto fatality rates among youngsters. Males (1986), on the other hand, concluded that raising the MLDA does not produce any net savings of lives, as the lives that are saved among drivers below the MLDA is counterbalanced by the additional lives that are lost among drivers in the hazardous first year of legal drinking after the MLDA has been reached. Saffer and Grossman (1987a, 1987b) reported that a uniform drinking age of 21 does help to reduce youth motor vehicle accident deaths, but an increase in beer taxes may be even more effective. That beer taxes and MLDA policies, along with other measures, may deter careless driving has subsequently been confirmed in a number of studies (Wilkinson, 1987; Chaloupka, Saffer and Grossman, 1993; Sloan, Reilly and Schenzler, 1994 and 1995). The literature since then is much less concerted. For one thing, the field has changed. By the late 1980s, the MLDA in every state had been set at 21. So, one started to look for other similarly straightforward and relatively costless policies. Furthermore, the 1980s have witnessed the rise of substantial grass-roots activity (such as Mothers Against Drunk Driving), which may affect both potential drinkers and policy making. But this activity and its effects are not easily accounted for in empirical analyses. In a similar vein, researchers have become aware of methodological shortcomings in earlier work. However, introducing additional controls for all kinds of relevant factors such as interstate disparities, fluctuating economic conditions, a full(er) set of alcohol policies and policy endogeneity as yet has produced no robust results, beyond the effectiveness of MLDA and dram shop liability rules. For instance, the result by Ruhm (1996) that higher

---

16 See, for example, Crandall and Graham (1984) who consider the Peltzman (1975) hypothesis that the installation of safety devices in cars will be offset by a behavioral response in driving intensity. They find that the net effect is clearly positive. Feber, Feldmeier and Crocker (2003) demonstrate the feasibility of exploiting insurance claims data to estimate the benefits of highway infrastructure improvements.
beer taxes are associated with reductions in crash deaths is not confirmed by Dee (1999), Mast, Benson and Rasmussen (1999), Young and Likens (2000) or Whetten-Goldstein et al. (2000). Similarly, the suggestion by Benson, Rasmussen and Mast (1999) to intensify criminal law enforcement efforts for driving under the influence does not have much support from other studies.

5.2 Industrial accidents
Society has several institutions available for the control of industrial injuries (Chelius, 1974). If safety control is left only to private markets, with no remedy for employer negligence, competition among informed workers would yield higher wages for more hazardous jobs. The compensating wage differentials would give the employer an incentive to deal with the risks faced by his employees, as expending resources on accident prevention might lower his wage bill. The government can also centralize control, by installing a regulatory body (such as OSHA in the US) that would promulgate and enforce safety standards for the organization of the workplace. Under a strict liability regime, it is the obligation to pay for all the accidents costs suffered by his employees which would give the employer an incentive to engage in accident prevention. And under negligence, it is the standard of care that would govern the employer, as taking due care would shield him from liability for the accident costs of his employees. Of course, if safety control is organized in such a manner (regulation, tort) that employees cannot expect their full accident costs to be covered by the employer, they will still require a corresponding wage premium to equalize the net remuneration with other, less hazardous jobs.

For a proper understanding of the field, at least two additional elements have to be taken into account. Firstly, the frequency and severity of industrial accidents in general not only depend on the preventive measures undertaken by the employer, but also on the level of care exercised by employees. If information and bargaining costs are sufficiently small, employer and employees may come to an understanding, trading obligations that follow from the existing safety control institution, such that in the end the least cost preventer will take the efficient amount of action. In full accordance with the Coase theorem, the choice of safety control institution would be irrelevant to the allocation of resources. However, wage bargaining is not without transaction costs, nor is monitoring the daily behavior of employees at the workplace. But then, workplace safety will after all depend on the control mechanism.

Secondly, insurance may play a role. An employer who carries insurance for his (share in) employees’ accident costs has an incentive to be less careful, to the extent that his level of care is not reflected in the premium. Similarly, insured employees may be subject to moral hazard. Either they might take less care, ex ante, which would result in more accidents and injuries. Or they might be tempted, ex post, to report more injuries and/or increase the size and duration of their claims.

---

17 The Occupational Safety and Health Administration, established in 1971. See for more details [www.osha.gov](http://www.osha.gov). Empirical analysis by Viscusi (1979) failed to indicate any significant OSHA impact on industry health and safety investments and on injury rates. He points out that OSHA enforcement efforts are too weak to create an effective financial incentive for firms. They are, moreover, dwarfed by - the safety effect from - workers’ compensation premiums (Moore and Viscusi, 1992).
From negligence to workers’ compensation
Between 1911 and 1949, the US states successively switched from negligence to a system of shared strict liability known as workers’ compensation (WC). Under WC, an employer must compensate his employees for work-related accidents, regardless of the cause. The compensation, however, is substantially less than the employee’s full accident costs. At the same time, the employee who is covered by WC can no longer sue his employer for negligence.

The shift from negligence to WC led to a substantial increase in the likelihood of compensation, but also (as a result of basic rates set by legislatures) in the median amount. If transaction costs prevented wages from adjusting fully, employers’ incentives to prevent accidents can be expected to have increased, while employees’ incentives to exercise care decreased. Two empirical studies have tried to find out which, if either, of these effects dominated. Chelius (1976) analyzed non-motor vehicle machinery fatalities between 1900 and 1940, and found that the switch to WC was associated with a relative decline in industrial accidents. Fishback (1987) looked at coal-mining fatal accident rates and reached the opposite conclusion. The difference between the results may, of course, be due to measurement errors in data. But it might also reflect a real difference, as the costs of accident prevention and monitoring workers may well have been relatively high in mining.

Anyway, transaction costs apparently matter.

Trends in workers’ compensation
Numerous studies examine the impact of workers’ compensation benefits in recent times. The setting for the analysis is provided by a substantial variation in workers’ compensation laws across states, but also over time as states started in the 1970s to improve benefit levels.

As a general result (Chelius, 1982; Butler and Worrall, 1983; Ruser, 1985; Chelius and Kavanaugh, 1988; Krueger, 1990; Butler, 1994), injury and claims frequency are found to increase as WC benefits increase, but severity rates may be lower. Claims frequency also increases as the length of time an injured worker has to wait before receiving benefits is shorter. Butler and Worrall (1983) observe that wages do adjust, but only to some extent. Ruser (1985) notes that the effect of higher benefits on injury rates is smaller, the greater the degree of experience rating in employers’ insurance. Relatively small firms are not perfectly experience rated, while large firms are sometimes allowed to self-insure.18

The findings nicely fit the theoretical argument above, given the presence of transaction costs. Moral hazard on the employees’ side as a result of more generous benefits, ex ante plus ex post, apparently outstrips employers’ reaction to intensify accident prevention. However, the findings do not allow us to disentangle these effects. How much did employers actually invest in safety measures, improving on workplace risks? Which part of employees’ moral hazard is attributable to ex ante willingness to take more risk? And which part is the result of ex post filing of unwarranted claims or dragging claim duration?

That challenge is taken up in the literature. Moore and Viscusi (1989, 1992) study

---

18 Danzon and Harrington (2001) discuss WC insurance premiums and the role of regulatory price controls.
death rate data, where ex post reporting effects are not very plausible. Ex ante moral hazard effects cannot be ruled out a priori, but it is rather unlikely that workers will endanger their lives if only for a more generous compensation to their surviving heirs. The authors conclude that, in the absence of WC, fatality rates would be substantially higher. Thus, WC provides powerful incentives for safety to firms that outweigh moral hazard effects, at least for the more serious accidents.\textsuperscript{19,20}

To place moral hazard in its proper perspective, it is useful to end this section with a reference to Biddle and Roberts (2003). They analyzed a sample of workers who were reported by physicians as having repeated trauma injuries known or suspected to be work-related. A significant proportion of these workers did not make use of the WC system. Even among those who missed more than seven consecutive days of work, almost 40 percent made no attempt to receive wage-loss benefits, and 27 percent did not file for any sort of WC benefits. Many of the non-filers reported they had access to alternative programs offered by their employers, such as sick leave. Apparently in this field too, we should be aware of many ‘false negatives’ alongside ‘false positives’ (cf. section 3.2).

5.3 Product liability
The control of defective products has much in common with the control of industrial accidents. Again, several social institutions are available (Spence, 1977; Higgins, 1978). Safety control might be left to the market, with no remedy for producer negligence. If consumers are well informed about the safety characteristics of each product, prices would adjust to reflect the relative hazardousness. Products that are designed and/or manufactured with inefficiently low (or high) care would be driven out of the market. Under tort, the producer has a stimulus to follow the negligence standard of due care, as it would shield him from liability costs. Under strict liability, the obligation to pay for all his customers’ accident costs would give the producer an incentive for adequate investments in product safety. But the government can also centralize control, by directly imposing product standards (think of FDA and CPSC in the US).\textsuperscript{21}

Once again, under perfect information with no bargaining costs producers and consumers will contract their way to an efficient solution, regardless of which social institution has been put in place. If safety control is organized in such a manner that consumers can claim their accident costs from the producer, this will be reflected in the price. Anyhow, consumers will be well aware of the ‘full’ price of the product. However, when consumers are ignorant of varying levels of product safety or underestimate the

\textsuperscript{19} Moore and Viscusi (1989, 1992) also analyze the process of wage adjustment in some detail.

\textsuperscript{20} In this context one may also refer to Kötz and Schäfer (1993), who study the introduction of a system of rebates and surcharges in accident insurance in the German sugar industry. Their analysis provides another example of how economic incentives to take preventive measures may result in a reduction in the number of accidents.

\textsuperscript{21} The Food and Drug Administration, dating back to 1906, and the Consumer Product Safety Commission, established in 1972. See for more details www.fda.gov and www.cpsc.gov. Empirical studies of the impact of product safety regulation yield mixed results. Viscusi (1985) found no beneficial effect from CPSC actions. Magat and Moore (1996), however, report that safety standards significantly reduced bicycle accident rates. With respect to FDA activities in the period before WW II, Law (2006) observes that the threat of (ex post) enforcement was an ineffective deterrent. The FDA, nevertheless, made several important contributions to product safety where it had the capacity to offer (ex ante) benefits to compliant firms by way of quality certification or direct technical research assistance.
risks, they will not want to pay a higher price for a safer product version. And they will buy too much of a seemingly cheap, but unsafe product. In that case, strict liability may be preferable to negligence, as it informs consumers of the actual risks by adding an (implicit) insurance premium to the product price.

There is one important ‘but’ in the matter (Calfee and Rubin, 1992). Consumers might not generally be willing to insure for non-economic losses. If non-economic losses are actually covered under strict liability, prices would be unwarrantedly high, distorting consumer choices. This distortion may be particularly disturbing in the case of a product that, overall, serves to reduce risk (for example, a vaccine). The price premium for undesired insurance for non-economic damages might lead well-informed consumers to avoid the product, yielding more risk instead of less.

**Empirical findings**

Let us now turn to empirics, where both the shift from negligence to strict liability in the course of the 1960s and the tort reform since the 1980s constitute interesting research topics. Apart from anecdotal evidence, however, firmly grounded results are scarce, mostly because of a paucity of relevant data. Data on accidental injuries in and around the home, for instance, have shortcomings because these accidents occur for all kinds of reasons, many of which are unrelated to defective or ill-designed products. In general, the specific product at stake and the intensity of its use are also unknown.

Manning (1994, 1997) analyzes the interplay between liability rules and product prices, especially for childhood vaccines. The shift from negligence to strict liability appears to have resulted in a dramatic price increase in the DPT vaccine, which has a small risk of very serious side effects, of well over 2,000 percent. A tentative calculation points out that $5-7 are expended in the process of transferring $1 of compensation to an injured consumer.

The (threatening) liability burden has also been the cause of products being withdrawn from the market. Garber (1993) presents a list that includes the Dalkon shield, Bendectin, the Bjork-Shiley heart valve, and the Markham silicone-gel breast implant. He also points to a series of already developed products whose market introduction has been delayed or forestalled by product liability concerns. Although hard evidence is lacking, it would seem that not all these decisions were equally appropriate. Bendectin is reported to have had much support in the medical community.

Viscusi and Moore (1993) analyze the empirical relationship between product liability costs and R&D expenditures. Product liability costs increase product R&D intensity initially, but there is a point beyond which the effect becomes negative. The authors infer from their findings that the development of new, safer products is the primary outcome of product liability pressure.

That leaves the question of whether product liability affects accident frequency. Higgins (1978) examined the non-transport accidental death rate at home in 1960 and 1970. The shift from negligence to strict liability reduced the accident frequency in states with low levels of educational attainment and raised it in states with high levels. Rubin

---

22 Takaoka (2006) examines how the shift from negligence to strict liability in Japan affected stock market prices. His results imply that firms can not fully recover the additional costs by increasing product prices. The shift to strict liability imposes liability and litigation costs on producers that exceed the value that consumers attach to the additional protection.
and Shepherd (2007) study the impact of tort reform since 1980. Their evidence suggests that reform of the collateral source rule was associated with an increase in non-motor vehicle accidental death rates, but several other several tort reform measures, such as caps on non-economic damages, had the opposite effect.

In conclusion, strict liability would seem to stimulate R&D toward safer products, to drive defective products out of the market and to raise prices. The safety effect of the latter can be positive, when it curtails the purchase of qualitatively inferior products by imperfectly informed consumers. But the findings by Rubin and Shepherd suggest that liability can go too far in its cover of non-economic damages. Lowering the liability burden may then result in lower prices, which enables consumers to buy more risk-reducing products, such as safety equipment or medicines.

Asbestos
This section cannot do without at least a few words about asbestos. Legal claims for injuries from asbestos involved more plaintiffs, more defendants and higher costs than any other type of product liability case in US litigation (White, 2004). Although in itself a highly useful product for its excellent fire-retardant capabilities, asbestos may cause a variety of diseases among production and insulation workers who breath in its fibers, but only after a long latency period of 20 to 40 years. Physicians recognized as early as the 1920s that exposure to asbestos caused diseases. But is was not until the early 1970s that an insulation worker won the first trial from a large asbestos producer, and that OSHA regulation started to impose limits on workplace asbestos exposure.

A central question for the current survey is how liability has fared in this instance. Not so well, initially. In the early 1930s, asbestos disease was not yet recognized as an industrial accident to be covered by workers’ compensation. Under tort, some workers brought negligence claims against asbestos firms, but these were settled quietly. Asbestos producers reacted by accepting workers’ compensation. As the workers’ compensation system is explicitly designed to cover less than the employee’s full losses, it provided inefficiently low incentives for the control of asbestos hazards (Dewees, 1986; Boden and Jones, 1987). Estimates, moreover, suggest that wage differentials were too small to be in any way compensating.

However, the gradual shift in product liability after World War II from negligence under privity of contract to strict liability opened the door for workers in the insulation industry to sue the asbestos producers. By the late 1970s, the pressure of litigation caused US producers to eliminate asbestos from most products, causing overall US consumption to decline sharply. Eventually, then, tort did its job of safety regulation, and did it better than government regulation, which had been captured by the large asbestos producers. The US system also performed better than Europe, where neither liability nor regulation was effective until the 1990s (White, 2004). But that is not to say that the job could not have been done better. Tort litigation has involved massive costs. By the end of 2002, total litigation costs for plaintiffs and defendants have been estimated to amount to $40 billion, to obtain total net compensation of about $30 billion (Carroll et al., 2005).

5.4 Medical malpractice
Medical injuries are a costly by-product of medical care, an otherwise beneficial activity. They have some specific characteristics, due to asymmetric information between health
care providers and patients. Firstly, medical injuries are generally assumed to be unilateral. Secondly, patients, although in a contractual relationship with providers, cannot really be expected to monitor the level of care and to bargain on the price they would like to pay for safer treatment. In theory, then, tort could create incentives for optimal care per procedure.

Under US common law, health care providers are liable for injuries that are attributable to negligence. Courts generally define due care as the customary practice of practitioners of good standing or a significant minority of such professionals (Danzon, 2000). These custom-based standards may be systematically biased, relative to first-best efficient care. One possibility is that the standard for a specific treatment is set too low, for instance because it lags behind new medico-technical developments. But it is at least as likely that the standard is too high, as a result of the financial incentives created by provider fee-for-service reimbursement on the one hand and patient insurance for medical care, typically with modest co-payment and premiums unrelated to own use, on the other hand. However, these incentives for overuse may be corrected by capitation forms of payment that more recently have been adopted under managed care in the US.

As to incentives, it should be added that physicians are nearly universally insured against medical malpractice claims, with only a minimum of experience rating (Sloan, 1990). This does not necessarily imply that deterrence incentives are non-existent, but they mostly seem to hinge on reputation loss and (uninsured) time spent in litigation. For hospitals, the situation is somewhat different. After the tort crisis of the mid-1970s, many hospitals found it difficult to obtain insurance and turned to self-insurance or to mutual companies with at least some experience rating (Mello and Brennan, 2002).

Finally, liability is not the only mechanism to counter medical malpractice. Professional ethics, referral networks, state licensing, disciplinary boards and hospital credentialing committees may also motivate physicians to act with proper care.

Defensive medicine
The central concept in the empirical literature on the incentive effects of medical liability is ‘defensive medicine’. Positive defensive medicine is defined as the use of extra-medical resources (tests, procedures, office visits) due to rising malpractice liability pressure. Negative defensive medicine, on the other hand, refers to decisions by doctors to cut back on the services they offer, to retire or to move to other states.

As to the latter, some studies examine how the medical malpractice crises of the past decades with rapidly rising liability insurance premiums and the ensuing tort reform measures have affected the supply of health care services. Dranove and Gron (2005) report that neurosurgeons in Florida significantly cut back their volume of operations as medical malpractice premiums rose; but there was no similar effect for obstetricians. Mello et al. (2005) present the results of a survey among medical specialists in Pennsylvania, which suggest that the supply of services in some areas is substantially reduced as a result of the cost of liability insurance. Kessler, Sage and Becker (2005) compare the trends in the supply of physicians in states that adopted and did not adopt tort reform laws. They found greater growth in states that adopted reforms directly limiting liability, such as caps on damage awards. A similar result is obtained by Klick and Stratmann (2004, 2007). But note that, even if all these studies point in the same direction, the net effect on public health is not clear. A large(r) supply of physicians in
itself can be presumed to contribute positively to social welfare, but there may also be
offsetting effects in terms of the level of care provided. Indeed, when Klick and
Stratmann (2004, 2007) examine infant mortality, they can only find mixed results. And
Dubay, Kaestner and Waidmann (2001), who report that malpractice liability pressure
was associated with less prenatal care, cannot find evidence that it affected infant health
adversely.

The literature on positive defensive medicine is somewhat more extensive. Most
attention has gone to obstetrics, the field which has one of the highest levels of premiums,
claim frequency and damage payments. Several studies examined the impact of
malpractice claim risk on cesarean section rates. Cesarean sections are a treatment
thought to be more frequently adopted in equivocal situations under the influence of
malpractice liability fears. But the empirical evidence is mixed. According to the results
from Localio et al. (1993), higher malpractice claim risk increased the probability of
delivery by cesarean section, but only at the hospital level, not for individual doctors. See
also Rock (1988). The opposite effect is reported by Tussing and Wojtowycz (1992),
while Baldwin et al. (1995) and Sloan et al. (1997) did not observe any effect. Dubay,
Kaestner and Waidmann (1999) provide evidence that higher malpractice risk does result
in increased use of cesarean sections, most notably for women with the lowest
socioeconomic status. But when they look at the impact on infant health at birth, there
isn’t any. Thus, it is concluded that the observed defensive response to malpractice claim
risk is socially wasteful. That result also seems to be in line with a separate study by
Sloan et al. (1995), which showed no systematic improvement in various indicators of
infant health at birth when obstetricians faced a higher malpractice litigation threat.

Kessler and McClellan (1996) focus on a rather different field of medicine,
cardiac illness in the elderly. Their findings indicate that tort reform measures that
directly limit liability, such as caps on damage awards, may reduce hospital expenditures
by 5-9 percent. The effects are somewhat smaller for actual heart attacks (AMI) than for a
relatively less severe form of heart disease (IHD), for which more patients may have
marginal indications for treatment. Reforms, on the other hand, did not lead to any
consequential differences in mortality or the occurrence of serious complications. Kessler
and McClellan (2002) update the analysis, by extending it into the era of managed care.
The reduction in defensive practices that can be achieved with direct tort reform is found
to be smaller in areas with high managed care enrollment. Managed care apparently
reduces physicians’ incentives and ability to engage in defensive treatment for IHD. In
that sense, managed care and direct tort reform are substitutes.

The results by Kessler and McClellan, however, have not gone uncontested.
Beider and Hagen (2004) apply their methods to a broader set of ailments, but can find no
evidence that direct tort reforms reduce medical spending. Dhankhar, Khan and Bagga
(2007) study the effect of medical malpractice risk on the application of three different
procedures for treating AMI patients. Their results suggest that higher medical
malpractice risk leads to a choice of procedure that is less invasive for the patient, and
cost saving. Interestingly enough, health outcomes of patients improve too.23

23 In this context one may also refer to results by Fenn, Gray and Rickmann (2007) for the UK. They
examine the impact of risk-sharing arrangements in medical liability insurance on the use of certain
diagnostic procedures. Hospitals facing higher expected costs per claim as a result of higher deductibles use
these tests more frequently.
Taking stock of the available evidence, it can be concluded that medical malpractice risk does affect the behavior of health care providers. Increasing malpractice pressure appears to have a negative impact on the supply of medical services, and it would seem to affect the choice of tests and procedures in the direction of defensive medicine. Tort reform measures like caps on damage awards might restore the balance. It can also be concluded that the incentives from medical liability interact with the financial incentives from the provision of medical services (fee-for-service versus managed care).

However, when it comes to efficiency, the findings should be interpreted with caution. If malpractice risk drives some physicians out of the market and makes other health care providers defensively adapt the treatment of their patients, this might be detrimental to social welfare if initially they all exercised due care. But there are two problems here. Firstly, it is not altogether impossible, as we know from section 3.2, that the doctors concerned performed below standard. Secondly, there is no guarantee that the standard of due care as it is applied in tort litigation equals the efficient level. The results for health outcomes should provide us with the material for a more decisive answer. But these results are still rather scarce and mixed.

6. Cost-benefit analysis

There is a lot of disagreement in society over how well the tort system is performing the functions it is supposed to play. The record clearly has some black spots: the costs of administering the system are substantial; large numbers of negligently injured people don’t claim and don’t receive compensation; safety and welfare-enhancing products and services are sometimes barred from the market; defensive medicine unwarrantedly drives up health care costs. But the system is, on the other hand, an intrinsic part of the whole fabric of social institutions set up to promote safety. The literature surveyed here has made clear that financial incentives from liability rules definitely play their role as part of that fabric. Overall, for the tort system to be worth retaining, one would hope at a minimum that the benefits, in terms of injuries deterred, exceed the costs of litigation and other associated costs.

Some authors try and have a shot. Donohue (1989) started from the $16-19 billion estimate by Kakalik and Pace (1986) of the tort system’s administrative costs. He then wondered whether the tort system yields at least this much in deterrence benefits. For that purpose, he combined a $91 billion NSC estimate of all accident costs in the US during the year 1987 (exclusive of work accidents governed primarily by workers’ compensation) with Landes’ (1982) finding that the switch from tort to no-fault increased road fatalities by 15 percent, at most. Supposing that the tort system produces as much as a 15 percent reduction in all accident costs, the savings from deterrence would equal $16 billion. If so, the tort system would not be cost-effective.

Schwartz (1994) counters this conclusion, by arguing that Donohue may have structured an appropriate inquiry, but used unsatisfactory data. The ‘true’ costs of accidents are definitely much higher than the NSC estimate on which Donohue relied, and should be adjusted upward so as to reach a total of over $690 billion. The overhead costs of the tort system also need revision, as the Kakalik-Pace study did not consider the administrative costs of resolving claims that did not reach the point of an actual lawsuit.
On balance, the annual costs of resolving all claims might well be $25 billion. The minimum reduction in accident costs that the tort system would need to achieve in order to justify its overhead would then be less than 4 percent, which seems rather modest.

The minimum reduction would become even more modest, if one were to realize that without the tort system society would have to organize another way of victim compensation (for example, no-fault), with overhead costs of its own. That line is followed by Danzon (1985) and Devlin (1990), when they narrow the topic and move from the universe of all accidents to a particular field of tort. Devlin points out that the switch from tort to no-fault for automobile accidents in Quebec was an inefficient move. The estimated reduction of administrative costs by $94 million per year is far behind the additional social cost of $247 million as a result of the increase in traffic accidents.

Danzon makes a back-of-the-envelope calculation for medical malpractice, comparing victim compensation through tort and first-party insurance. The tort system should deter at least one injury of comparable severity for every injury currently compensated, in order that the benefits from deterrence outweigh the additional litigation costs. Referring to results from the 1974 California study (CMA and CHA, 1977), which showed that only 1 in 25 incidents of negligent injury receives compensation, she concludes that only a 4 percent reduction in the rate of negligent injury is required to justify the costs of the tort system. Which should be no problem.

Yet, instructive as these calculations may be, they mainly have a heuristic value. Firstly, a full cost-benefit evaluation is impossible in the current state of affairs. Notably lacking are data on ‘indirect’ costs: the costs of precautions by potential injurers; the opportunity costs of goods and services that are withdrawn from the market or whose introduction is forestalled; the opportunity costs of goods and services that are not bought because of liability-induced price increases; and the disruption costs of layoffs and bankruptcies caused by liability problems (CBO, 2003). Secondly, even if the benefits of the current system do outweigh its costs, the search for marginal improvements or more cost-effective alternatives remains an open question.
Bibliography


Beider, Perry and Stuart Hagen (2004), Limiting Tort Liability for Medical Malpractice, Washington DC, Congressional Budget Office.


CBO (Congressional Budget Office) (2003), The Economics of U.S. Tort Liability: A Primer, Washington DC.
CBO (Congressional Budget Office) (2004), The Effects of Tort Reform: Evidence from the States, Washington DC.
CMA (California Medical Association) and CHA (California Hospital Association) (1977), Report on the Medical Insurance Feasibility Study, San Francisco CA, Sutter Publications.


Garber, Steven (1993), Product Liability and the Economics of Pharmaceuticals and Medical Devices, Santa Monica CA, RAND Institute for Civil Justice.


Kakalik, James S. and Nicholas M. Pace (1986), Costs and Compensation Paid in Tort Litigation, Santa Monica CA, RAND Institute for Civil Justice.


Saks, Michael J. (1992), ‘Do We Really Know Anything about the Behavior of the Tort Litigation System - And Why Not?’, *140 University of Pennsylvania Law Review*, 1147-1292.
Shanley, Michael G. and Mark A. Peterson (1987), Posttrial Adjustments to Jury Awards, Santa Monica CA, RAND Institute for Civil Justice.
Studdert, David M., Michelle M. Mello, Atul A. Gawande, Tejal K. Gandhi, Allen Kachalia, Catherine Yoon, Ann Louise Puopolo and Troyen A. Brennan (2006),


