© The Journal of Risk and Insurance, 1993 Vol. 60, No. 2, 230-264

The Stochastic Dominance of No-Fault Automobile Insurance

J. David Cummins
Mary A. Weiss

ABSTRACT

This article presents a rigorous analysis of no-fault automobile insurance in terms of stochastic dominance theory. In the baseline case, with identical drivers and actuarially fair insurance, no-fault is stochastically dominant as long as the no-fault insurance premium exceeds the tort premium. In this case, no-fault brings a higher proportion of accident costs under insurance, increasing driver welfare. When expense charges are introduced (actuarially unfair insurance) no-fault may still be stochastically dominant if the expense charge is less under no-fault, even if no-fault weakens incentives for good driving and leads to higher accident rates. Elective no-fault is unlikely to reduce auto insurance costs, because drivers with high propensities toward moral hazard are likely to retain their right to sue by choosing tort.

Introduction

Automobile Insurance: What Went Wrong?

Private passenger automobile insurance traditionally has been a relatively "safe" line of property-liability insurance, insulated from the sharp changes in prices and profits that are prevalent in many commercial coverages. Private passenger auto insurance is the single largest source of revenue for the property-liability industry, with premiums written of $83 billion in 1991, accounting for 37 percent of total industry premium volume. Although private passenger auto insurance long served as a source of stability and strength in the insurance industry, this situation changed during the 1980s. From 1984 to 1989, the auto insurance consumer price index (CPI) grew at an annual rate of 9 percent, while the all items CPI grew at 3.5 percent and the medical care CPI at 6.5 percent (Cummins and Tennyson, 1992a). This high inflation occurred

---

J. David Cummins is Harry J. Loman Professor of Insurance and Risk Management at the Wharton School, University of Pennsylvania. Mary A. Weiss is Associate Professor of Risk Management at Temple University. The authors acknowledge the assistance of Stephen P. D'Arcy.

during a period of declining auto accident rates.[1] The result has been growing dissatisfaction among consumers and growing pressure on insurers from legislators and regulators. More stringent regulation has led to inadequate prices and declining profits for the auto insurance industry (Cummins and Weiss, 1992b), which threatens the stability of the insurance market.[2]

What went wrong? Consumer activists and many regulators blame insurers, saying they are inefficient, that they incur excessive marketing and administrative expenses, engage in lax claims settlement practices, and then pass along the costs to consumers. In fact, there is little or no evidence of inefficiency in the auto insurance market. The market is competitively structured (Cummins and Weiss, 1992b), and claims and expenses for a given policy cohort are paid *after* premiums are collected. Thus, excessive expense or loss payments would directly reduce insurer profits for a given cohort. Insurers attempting to reclaim these lost profits by charging higher premiums in the future would lose market share to more efficient insurers who did not charge such loadings.

The primary cause of the auto insurance problem is inflation in claim costs. The cost increases have been most significant in bodily injury liability (BIL), property damage liability, and collision severity. In addition, BIL frequency has risen in tort states, even though auto claims frequency generally has been declining (Cummins and Tennyson, 1992a). The cost increases have been driven in part by escalating medical care expenses and the increasing costs of repairing automobiles. However, moral hazard and insurance fraud also play an important role.

Insurance Lotteries

Survivors of automobile accidents have the opportunity to participate in an *insurance lottery* (Cummins and Tennyson, 1992b). One does not have to be injured in order to play the lottery. It is necessary only to find a cooperative physician and an attorney. The former conducts unnecessary medical treatments and carefully documents the results, while the latter files a liability claim against an insurer. Those with legitimate injuries can increase the stakes in the lottery by building up an injury to look more serious than it is. Like outright fraud, *build-up* involves conducting and documenting unnecessary medical treatments. The claimant's primary motivation for fraud and build-up is the pain and suffering (general damage) award (lottery prize) that will be received under the tort system if the claim is successful. General damages are frequently determined as a multiple of economic losses.

---

[1] We use the property damage liability claim frequency rate as a proxy for the accident rate (see Cummins and Tennyson, 1992b).

[2] Some states, including Massachusetts and New Jersey, have experienced market failure, with more than half of the drivers insured through the residual market. The market failure in these states actually pre-dated the hyper-inflation of the 1980s. The same economic and socio-legal forces that are now causing problems in other states first surfaced in Massachusetts and New Jersey because of the high accident rates, urbanization, and claims consciousness in those states.

Weisberg and Derrig (1991a, 1991b) provide evidence of fraudulent activity in auto insurance. Their analysis of more than 500 BIL closed claims from Massachusetts insurers by experienced claims personnel revealed that almost one-third involved some degree of fraud or build-up (2.6 percent involved outright fraud, 8.5 percent involved a combination of fraud and build-up, and 20.7 percent involved build-up alone). The percentage of fraudulent or built-up claims was much higher in economically depressed areas such as Lawrence, Massachusetts. Minor injuries such as strains and sprains were much more common in fraudulent and built-up claims than in the sample as a whole.

Because payments for fraudulent and built-up claims serve no legitimate social or economic purpose, reducing such payments seems to be a reasonable way to attack the auto insurance cost problem (Cummins and Tennyson, 1992b). The pain and suffering awards that motivate insurance lotteries are a feature of the tort liability system, which is the traditional mechanism in the United States for compensating automobile accident victims for personal injuries. In addition to motivating insurance lotteries and fraudulent claims, the tort system has significant limitations as a method for compensating accident victims. The tort system has been criticized as being slow, inefficient, and inequitable (e.g., Conard et al., 1964; Keeton and O'Connell, 1965), and support for these allegations has been provided by studies of auto insurance closed claims (All-Industry Research Advisory Council, 1979, 1989).

Tort vs. No-Fault

Since the tort system evolved primarily as a mechanism for deterring negligent behavior, it is not surprising that it is less than optimal as a compensation system. However, the system's effectiveness as a deterrent to negligent driving also has been questioned. Keeton and O'Connell (1965) argued that, because auto accidents are largely random, the expensive process of assessing fault does not provide significant deterrent effects. They argue instead for *no-fault* automobile insurance, which provides compulsory first-party insurance covering economic losses from auto accidents.[3] In contrast to the tort system, no-fault restricts the ability to sue for personal injury losses and general damages. In order to qualify for tort, a claim must exceed a threshold, defined either verbally or as a dollar amount.[4] The savings from the elimination of relatively minor lawsuits by the threshold are reallocated to offset the

[3] With rare exceptions, no-fault laws address only personal injuries—that is, medical bills and lost wages—rather than property damage losses. This is because most of the problems with the tort system are in personal injury compensation. First-party insurance is readily available for collision and other property damage claims, and subrogation through arbitration works well and is relatively inexpensive in the property damage area. Only two states, Michigan and Massachusetts, have no-fault laws affecting property damage losses.

[4] For example, under the Michigan verbal threshold, drivers remain subject to tort liability for general damages only if the victim of an auto accident has suffered death, serious impairment of bodily function, or permanent serious disfigurement (see American Insurance Association, 1987).

costs of additional first-party benefits. No-fault was first enacted in Massachusetts in 1971, and 15 states now have some form of no-fault law.[5]

Although critics claim that no-fault auto insurance weakens incentives for careful driving, evidence of this has been mixed. Landes (1982) found a positive relationship between no-fault insurance and fatal accident rates, but subsequent research contradicts this finding (Kochanowski and Young, 1985; U.S. Department of Transportation, 1985; Zador and Lund, 1986). Cummins and Weiss (1993) point out methodological flaws in these articles and provide evidence consistent with Landes's finding that no-fault is associated with higher fatal accident rates. This adverse effect should be balanced against the beneficial aspects of no-fault when considering compensation system design.[6]

The primary argument for no-fault auto insurance is that it is a more equitable compensation mechanism than tort. Under tort, only victims who can show that someone else negligently caused the accident are able to collect their economic losses. Claimants who are compensable under tort also can sue for general damages (pain and suffering). Thus, under the tort liability system, some victims receive far more than their economic losses and others receive nothing. No-fault auto insurance coverage, on the other hand, compensates all victims for their economic losses, and only those with serious injuries can sue for pain and suffering. Because compensation under no-fault does not require proof of negligence, the system is more efficient than tort in terms of administrative costs and speed of payment (Harrington, 1987; Grabowski, Viscusi, and Evans, 1989). The expensive process of assessing fault is reserved for relatively serious accidents. No-fault auto insurance also helps to control the fraud and build-up problem, because most fraudulent and built-up claims are relatively small and would be eliminated by an effective threshold.

The public policy debate over no-fault insurance has focused on three principal issues: the effects that no-fault insurance might have on incentives to drive safely, insurance costs under tort and no-fault, and the relative equity of the two systems. As mentioned above, no-fault does appear to have a small but significant adverse incentive effect. However, empirical evidence suggests that no-fault effectively reduces auto insurance costs if a stringent verbal threshold is adopted that eliminates a high proportion of the tort claims (Cummins and Weiss, 1991). Inflation is higher in states with low thresholds than it is in tort states and verbal threshold no-fault states (Cummins and Tennyson, 1992a).[7]

---

[5] Twelve additional jurisdictions have enacted so-called add-on laws, which provide first-party medical expense coverage but place no restrictions on lawsuits.

[6] All U.S. no-fault laws permit tort lawsuits under some circumstances, and none of the laws eliminates tort claims for accidents involving serious injuries or death. The effects of no-fault auto insurance on accidents may be more significant in a pure no-fault system, such as the one presently in force in Quebec (Devlin, 1992).

[7] Dollar thresholds act as a magnet for claimants playing the insurance lottery. For example, when the Massachusetts threshold was raised from $500 to $2,000 in 1989, medical costs for small claims rapidly converged to the new threshold (Weisberg and Derrig, 1991b).

Purpose of This Article

This article focuses on the equity issue. In particular, our objective is to determine the conditions under which drivers will prefer no-fault over tort. We rigorously interpret the primary argument for no-fault, that is, that it shifts resources from relatively low value activities (compensating small general damage awards) to relatively more important activities (compensating economic personal injury losses). Conditions are developed under which no-fault stochastically dominates tort, that is, where no-fault would be preferred by any risk-averse expected utility maximizer. This finding has important implications for the public policy debate over no-fault. Proponents of the tort system argue that the adoption of no-fault forces motorists to surrender an important right (the right to sue) without adequate compensation. This argument is difficult to sustain if no-fault is stochastically dominant, and the case becomes even weaker if no-fault is dominant for legitimate claims *and* reduces costs by preventing insurance lotteries.

Our discussion of the stochastic dominance of no-fault interacts with several broader public policy issues regarding automobile insurance that are potentially as significant as deterrence of negligent driving and stochastic dominance. For example, there is the question of whether driving is a privilege or an (almost) inalienable right. Because driving is virtually a necessity in order to commute to work in many parts of the country, there are compelling reasons to view driving as a right. On the other hand, making driving inexpensive as a means of transportation reduces political pressures for improvements in public transportation systems. A second important question is whether the ability to sue is a right or a privilege. Jurisdictions adopting no-fault have opted to place limitations on the ability to sue, implicitly stating that lawsuits are not an inalienable right. Finally, there is the question of the social equity of compulsory or quasi-compulsory auto liability insurance for drivers with very low net worth. Given the high auto insurance premiums in many urban areas, requiring liability insurance imposes a severe regressive tax on economically disadvantaged drivers, forcing many to drive without insurance. A well-designed no-fault law would make especially good sense for these drivers.

The first section of this article focuses on the pure theory of the equity argument by considering the case where insurance is actuarially fair and accident rates are the same for both tort and no-fault. We obtain the seemingly counter-intuitive finding that no-fault is stochastically dominant if the no-fault premium is higher than the tort premium. No-fault cannot be dominant for good risks in this case. Because no-fault may raise accident rates and reduce administrative expenses, the next section assumes that accident rates are higher under no-fault and that insurance is not actuarially fair. Again, no-fault is shown to be stochastically dominant in several relevant cases. No-fault may be stochastically dominant even if it leads to higher accident rates, because administrative expenses are lower. No-fault also may be dominant for good risks in this case, highlighting the importance of expense efficiencies. Bad drivers are more likely than good drivers to select no-fault, suggesting that cost savings under elective no-fault plans may be less than anticipated.[8] The third section presents some empirical evidence on the key parameter values and discusses the impact on the analysis of real-world problems such as duplicate recoveries.

Dominance with Actuarially Fair Insurance

A Model of Auto Accident Compensation

The model considered in this article is similar to the bilateral case discussed by Shavell (1987), among others, where each accident is assumed to involve two drivers, both of whom suffer personal injury losses. The analysis is conducted from the perspective of driver 1. The other driver involved in the accident is called driver 2.

We assume that drivers have equal initial wealth $W_1 = W_2 = W$, which is nonstochastic, and that they select a level of expenditures on automobile safety, or choose a level of care. The level of care expenditures (x) selected by the driver affects the accident probability λ, where $\lambda'(x) < 0$, and $\lambda''(x) > 0$, and primes indicate derivatives.[9]

In this section, we assume equal care levels and accident rates under tort and no-fault. This is, in effect, a rigorous analysis of the case for no-fault presented by Keeton and O'Connell (1965) and other early proponents. It is a reasonable case to consider both because the evidence on the relationship between no-fault and accident rates is ambiguous and because it provides insights that are helpful in understanding the more complicated cases to follow. This case provides the rationale for the counterintuitive finding that no-fault is dominant if the premium is *higher* than under tort.

When an accident occurs, drivers sustain two types of losses: economic losses and pain and suffering (general damages). Economic losses (denoted ℓ) include medical bills and lost wages and have an unambiguous monetary value. Pain and suffering, on the other hand, can be modeled either as a monetary loss or as an "irreplaceable commodity" (Cook and Graham, 1977). Most of this analysis assumes that pain and suffering can be treated as a replaceable commodity (denoted g), and it is shown that treating general damages as irreplaceable does not materially affect the results. Both ℓ and g are assumed

[8] Elective no-fault (discussed in more detail below) is a relatively new concept whereby drivers are given the option to choose between tort and no-fault. Thus, both systems coexist in the same jurisdiction.

[9] Implicitly, we also assume that accident probability λ asymptotically approaches some non-zero constant as the level of care expenditures (x) approaches infinity; that is, accident rates can never be reduced to zero. The accident rate λ is also a function of the level of care expenditures selected by the state and municipalities where driving takes place (e.g., expenditures on highway design, law enforcement, etc.). Thus, one would expect the function λ, and hence the derivatives of λ with respect to individual choices of the level of care x to differ by jurisdiction. Globally, drivers may choose their accident rates in part through their decisions on where to live, etc. Our model is thus *conditional* on a driver's being located in a particular jurisdiction; that is, our drivers can choose their own care expenditures but not the function λ.

to be nonstochastic. For the present, we assume that all claims are legitimate, involving no-fraud or build-up.

The assignment of negligence is the probabilistic result of a two-stage process. The first stage is the occurrence of an accident, while the second involves the assignment of liability. Liability assignment is summarized in Table 1. It is assumed that p_1, $p_2 > 0$, and $p_1 + p_2 \leq 1$.[10]

Table 1
Liability Assignment

Event	*Probability*
Driver 1 Negligent	p_1
Driver 2 Negligent	p_2
Neither or Both Negligent	$1-p_1-p_2$

The negligence probabilities, p_1 and p_2, are functionally related to the care levels of drivers 1 and 2; that is, $p_i = p_i(x_1,x_2)$, where $\partial p_i(x_1,x_2)/\partial x_i < 0$, $\partial p_i(x_1,x_2)/\partial x_j > 0$.[11] Thus, by taking more care, drivers can reduce their probabilities of being held negligent (p_i) and increase the probability that the other driver is found to be negligent.

We do not need to assume that the assignment of negligence is accurate; there is some probability of being held negligent even if one's care expenditures are relatively high. We consider this a more realistic description of real-world liability systems than the threshold negligence model sometimes used in the literature. According to that model, individuals whose care levels exceed some threshold have a zero probability of being held negligent (see Shavell, 1987). In Cummins and Weiss (1992a), we show that care levels will be higher in systems where the probabilities of being held negligent are relatively more sensitive to changes in care expenditures. Incentives to take care are lowest where negligence assignment is arbitrary ($\partial p_i/\partial x_i = 0$).

The model focuses on the principal differences between tort and no-fault compensation systems. In particular, the drivers' wealth distributions under the two systems are assumed to differ as follows: (1) Under tort, the driver is allowed to sue for general damages and thus recovers these losses in some states of the world in which an accident has occurred. When the driver recovers general damages g, he or she also is assumed to recover ℓ, where ℓ is interpreted as economic losses that are not covered by any other type of

[10] P_1 represents the probability that driver 1 is negligent and driver 2 is not negligent, p_2 is the probability that driver 2 is negligent and driver 1 is not negligent, and $1-p_1-p_2$ is the probability that neither or both are negligent. There is no loss of generality from using the compound probabilities p_1 and p_2.

[11] It is also assumed that $\partial^2 p_i(x_1,x_2)/\partial x_i^2 > 0$, $\partial^2 p_i(x_1,x_2)/\partial x_j^2 < 0$; that is, that there are decreasing returns to care expenditures, and also that p_1 and p_2 approach nonzero positive constants as functions of the x_i. This model is explained in more detail in Cummins and Weiss (1992a).

insurance.[12] (2) Under no-fault, the driver is not allowed to sue for general damages and thus loses g in all states of the world where accidents occur (accident states). The tradeoff is that ℓ is covered by first-party automobile insurance in all accident states.

The objective here is to create an abstraction of the real-world automobile insurance public policy-making environment. Obviously, in a purely theoretical world, it is unlikely that significant economic losses would remain uninsured, thus removing an important part of the rationale for no-fault insurance. In the real world, however, uninsured economic losses from auto accidents clearly exist (evidence on this point is discussed below), and their existence is one of the cornerstones of the case for no-fault automobile insurance.

The no-fault system considered here does not allow lawsuits for auto accidents under any circumstances and thus might be termed "pure" no-fault. The Quebec no-fault system is close to pure no-fault (Devlin, 1992), but the no-fault laws enacted in the United States permit lawsuits under some circumstances, as explained above. Pure no-fault is used here to keep the discussion as straightforward and intuitive as possible. There is no significant loss of generality in considering pure rather than modified no-fault in this model. A modified no-fault model presented in the appendix leads to essentially the same conclusions.

The legal responsibility of drivers under the two systems is defined more precisely as follows: Under the tort system, driver 1 is responsible for paying driver 2's economic losses (ℓ) and general damages (g) if driver 1 is negligent; driver 1 collects from driver 2 if driver 2 is negligent. If neither or both are negligent, each driver bears his or her own economic losses and general damages ($g+\ell$). A negligent driver also bears his or her own losses.[13] As indicated, in the no-fault model, each driver collects economic losses from his or her own insurer and no lawsuits are permitted.

---

[12] The model allows for the possibility that some economic losses are covered by other types of first-party insurance such as health insurance. This coverage is assumed to be the same under tort and no-fault; that is, we are modeling the difference between the two systems at the margin. There are possible secondary effects of the adoption of no-fault that are not modeled here. For example, because no-fault is usually primary with respect to employer-provided health insurance, covering economic losses under no-fault may reduce employer health insurance costs and have some compensating impact on wage rates or on the benefits package firms use to attract some types of workers such as married males. In theory at least, workers whose employer-provided benefits are affected the most by no-fault should receive the highest wage/benefits adjustment. On the other hand, to the extent that losses are moved out of the employer-provided health care system and into the auto insurance system, the driver's overall costs may increase because auto premiums are not tax deductible and transactions costs are higher for individual than for group insurance.

[13] The tort system analyzed here is a pure contributory negligence system where any degree of negligence on the part of driver i prevents recovery. The analysis could easily be extended to consider comparative negligence. It is also possible to model a tort system where first-party economic loss coverage is available, but at this stage in the discussion this would add more complexity than insight.

The assumptions regarding the availability of insurance are the following: Liability (third-party) insurance is available under tort, and first-party economic loss (personal injury protection) insurance is available under no-fault. As in the real world, first-party insurance against general damages is not available under either pure tort or pure no-fault.[14] Drivers are assumed to fully insure.[15] Insurers are assumed to know accident and negligence probabilities and to use this information in pricing.

Drivers are assumed to be expected utility maximizers with identical utility functions, U(W). The usual assumptions are made with regard to the derivatives of the utility function, that is, that $U' > 0$, and $U'' < 0$. The driver's decision is whether to choose tort or no-fault. The driver is assumed to know the relevant accident and negligence probabilities as well as the loss amounts.[16]

Stochastic Dominance

The proponents of no-fault argue that resources devoted to general damages and system expenses could better be used to pay claimants for their economic losses on a first-party basis. The rationale is that receiving reimbursement for one's own economic losses with certainty is better than recovering both economic losses and general damages when the other driver is negligent and (possibly) nothing if that driver is not negligent (Keeton and O'Connell, 1965). Under the assumption that significant economic losses are covered under no-fault but not under tort,[17] the argument for no-fault can be formulated in terms of second-degree stochastic dominance.

Consider the possible economic outcomes for driver 1 under tort and no-fault when insurance is actuarially fair and accident rates are the same under the two compensation regimes, shown in Table 2. Under pure no-fault, wealth is reduced by the dollar value of general damages when an accident occurs, because general damages are not recoverable. Economic losses are fully compensated in return for a premium payment of π_N. Under tort, there is no loss if no accident occurs or if an accident is negligently caused by driver 2.. If driver 1, neither, or both drivers are negligent, driver 1 loses his or her own economic losses and general damages (ℓ+g). If driver 1 is negligent, driver 1's

[14] The most likely reason for the unavailability of this type of coverage in real-world systems is the high cost of controlling moral hazard.

[15] When insurance is not actuarially fair, the full-insurance assumption is equivalent to assuming that insurance is compulsory because full coverage is not optimal in the presence of a loading. This is a reasonable abstraction of reality because no-fault insurance is nearly always compulsory, while auto liability insurance is compulsory or quasi-compulsory through financial responsibility laws.

[16] A useful extension of this article would be to consider how various types of information asymmetries would affect the desirability of no-fault. For example, insurers may not be able to distinguish good risks from bad, or the payoff matrix may not be known to motorists, particularly under tort.

[17] This assumption is likely to be met in practice for a significant proportion of drivers. Empirical evidence on this point is discussed below.

obligation to driver 2 is discharged by driver 1's insurer under his or her liability insurance policy, which is purchased for a premium of π_T.

Table 2
Driver 1's Economic Position Under No-Fault and Tort

Accident State	*Probability*	*Driver 1's Wealth Under No-Fault*	*Driver 1's Wealth Under Tort*
No Accident	$1-\lambda$	$W-x-\pi_N$	$W-x-\pi_T$
Driver 2 Negligent	λp_2	$W-x-\pi_N-g$	$W-x-\pi_T$
Driver 1, Neither, or Both Negligent	$\lambda(1-p_2)$	$W-x-\pi_N-g$	$W-x-\pi_T-\ell-g$
Premium		$\pi_N = \lambda\ell$	$\pi_T = p_1\lambda(g+\ell)$
Expected Wealth		$W-x-\pi_N-\lambda g$	$W-x-\pi_T-\lambda(1-p_2)(\ell+g)$

Note: W = wealth, x = care expenditures, λ = accident probability, p_i = probability that driver i is at fault, ℓ = economic losses, g = general damages, π_j = the insurance premium, j = N (no-fault), T (tort).

The three conditions for second-degree stochastic dominance are (as in Levy and Sarnat, 1984):

Risk Aversion—Decision-makers are risk-averse expected utility maximizers.
Wealth Condition—The expected wealth under the dominant distribution must be at least as large as under the alternative.
Distribution Function Condition—$A_N(W^*) \leq A_T(W^*)$ for all W^*, with strict inequality for at least one W^*, where W^* = wealth, and $A_N(W^*)$ and $A_T(W^*)$ are, respectively, the areas under the dominant and dominated distribution functions up to W^*. In this application, $A_N(W^*)$ is the distribution function of wealth under no-fault, and $A_T(W^*)$ is the distribution function under tort.

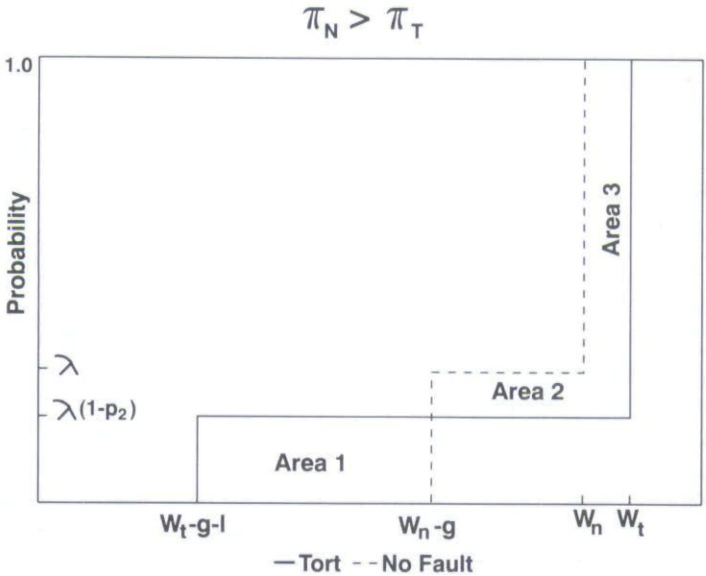
If the three conditions hold, then any risk-averse expected utility maximizer will prefer the dominant probability distribution.

The existence of dominance depends upon assumptions made about the two drivers. We consider two cases: First, that drivers 1 and 2 are identical, and, second, that one of the drivers is a "good" driver and the other is a "bad" driver.

Identical Drivers. First, assume that drivers are identical (i.e., $p_1 = p_2$). Because drivers are assumed to be risk-averse expected utility maximizers, dominance is established by considering the wealth condition and the distribution function condition for dominance. The assumptions of identical drivers, actuarially fair insurance, and equal accident probabilities (λ) under tort and no-fault are sufficient to satisfy the wealth condition, because expected wealth under both systems is equal to $W-x-\lambda(g+\ell)$. The effect of the compensation systems is to allocate a fixed amount of expected wealth across the possible states of the world.

Consideration of the distribution function condition requires the analysis of two cases: *Case 1:* $\pi_N \geq \pi_T$, and *Case 2:* $\pi_N < \pi_T$. We obtain the seemingly counter-intuitive result that no-fault is dominant in Case 1 (where the no-fault premium is at least as large as the tort premium) but not in Case 2 where the no-fault premium is lower. To see this for Case 1, consider the distribution function graphs under the two systems, shown in Figure 1. The dashed line is the distribution function for no-fault, and the solid line is the distribution function for tort. In Figure 1, it is assumed that $\pi_N > \pi_T$.

Figure 1
Dominance With Fair Insurance: Case 1



The no-fault distribution function is below the tort distribution function in the interval W_T-g-ℓ to W_N-g and above the tort distribution function from W_N-g to W_T, where $W_T = W\text{-}x\text{-}\pi_T$, and $W_N = W\text{-}x\text{-}\pi_N$. Thus, $A_N(W^*)$ is strictly less than $A_T(W^*)$ from W_T-g-ℓ to W_N-g; and no-fault is dominant if the area under the no-fault distribution function never exceeds the area under the tort distribution function. The latter will be true if Area 1 $\geq$ Area 2 + Area 3 in Figure 1. Since expected wealth is the same under the two systems, this condition must hold as an equality so that no-fault is dominant.

To check that Area 1 = Area 2 + Area 3, consider the following:

Area 1 = $\lambda(1\text{-}p_2)(\ell + \pi_T - \pi_N)$.

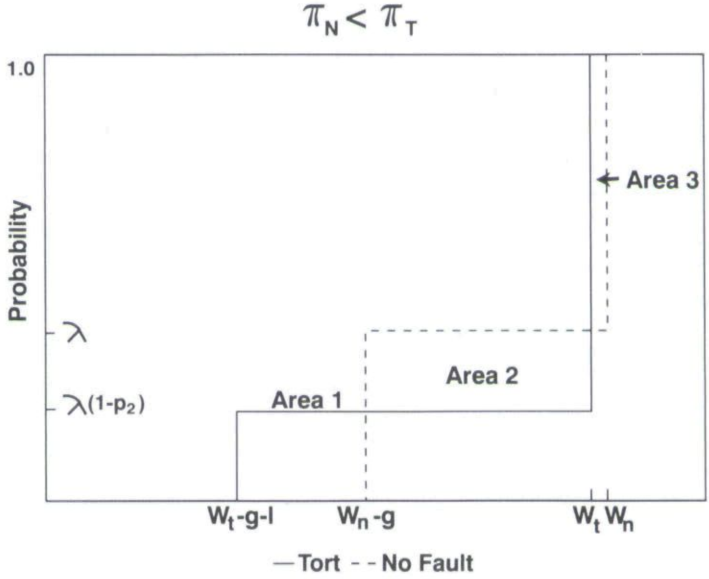
$$
\begin{aligned}
\text{Area 2 + Area 3} &= g\lambda p_2 + [1-\lambda(1-p_2)](\pi_N - \pi_T) \\
&= g\lambda p_2 + \pi_N - \pi_T + \lambda(1-p_2)(\pi_T - \pi_N) \\
&= \lambda\ell(1-p_1) + \lambda(1-p_2)(\pi_T - \pi_N).
\end{aligned}
$$

The result is established by the equality of the driver negligence probabilities p_1 and p_2, which is true if the drivers are identical.

Dominance occurs because no-fault transfers wealth from Areas 2 and 3 to Area 1. That is, in comparison with tort, no-fault transfers wealth from relatively high-wealth states of the world (Areas 2 and 3) to the state of the world with the lowest wealth (Area 1). Because of diminishing marginal utility, wealth in Area 1 has a higher utility value than wealth in Areas 2 and 3. Thus, no-fault is dominant. This is precisely the no-fault proponents' argument that it is advantageous to trade the possibility of receiving general damages for the certainty of recovering economic losses.

No-fault is not dominant if $\pi_N < \pi_T$ (Case 2). To see this, consider Figure 2. In this case, the no-fault distribution function is below the tort distribution function from W_T-ℓ-g to W_N-g and from W_T to W_N but above the tort distribution function from W_N-g to W_T. Equality of wealth under the two systems implies that Area 1 + Area 3 = Area 2. Thus, $A_N(W^*) < A_T(W^*)$ up to W_N-g but $> A_T(W^*)$ at W_T, violating the distribution function condition for dominance.

Figure 2
Dominance With Fair Insurance: Case 2



The dominance of no-fault auto insurance depends upon the relative magnitude of the insurance premiums under the two systems, π_N and π_T. If total expected losses are the same under the two systems and insurance is actuarially fair, drivers prefer the system that covers the largest amount of losses by insurance. To explore this result in more detail, consider the condition for dominance, $\pi_N \geq \pi_T$:

$$\pi_N \geq \pi_T \Rightarrow \lambda \ell \geq \lambda p_1(g+\ell) \Rightarrow \ell(1-p_1) \geq g p_1. \tag{1}$$

If an accident occurs under tort, the driver is insured only against lawsuits, not for his or her own losses. Thus, only $\lambda p_1(\ell+g)$ is covered by insurance. With no-fault auto insurance coverage, expected economic losses $\lambda\ell$ are covered. Thus, by choosing no-fault insurance, the driver gains the expected amount $\ell(1-p_1)$ in accident states but gives up the expected amount $gp_1 = gp_2$ in these states because of the exclusion of general damage suits under no-fault. If the expected gain in insured losses by choosing no-fault exceeds the expected loss in general damage recoveries, then no-fault is dominant. The driver gives up a small amount (the premium) in low marginal utility states of the world (represented by Area 3) in return for a larger amount in the highest marginal utility state (Area 1).

Further examination of the condition $\ell(1-p_1) \geq gp_1$ reveals that no-fault is more likely to be dominant if ℓ (economic losses) is relatively large, p_1 (driver 1's negligence probability) is relatively small, and/or g (general damages) is relatively small. Since p_1 is likely to be less than 0.5, g would have to be greater than ℓ in order to break the dominance condition (g frequently exceeds ℓ in practice.

Good and Bad Drivers. Next consider the case where drivers are not identical; that is, where one driver is safer than the other. Bad drivers are assumed to have higher accident probabilities (λ_B) than good drivers (λ_G); that is, $\lambda_B > \lambda_G$. Good drivers are assumed to be more likely to collect under tort, so that $p_{G1} < p_{G2}$ for good drivers and $p_{B1} > p_{B2}$ for bad drivers, where subscripts G and B stand for good and bad drivers, respectively.[18]

The wealth condition for dominance requires that expected wealth under the dominant distribution be at least as large as expected wealth under the dominated distribution. If drivers are not identical, expected wealth differs under the two compensation systems. For no-fault,[19]

$$E_N(W) = W - x - \pi_{iN} - \lambda_i g = W - x - \lambda_i(\ell + g)\,, \qquad i = G, B, \tag{2}$$

where $E_N(W)$ = expected wealth under no-fault and the additional subscript on π reflects the difference in premiums paid by driver type, with $\pi_{GN} < \pi_{BN}$ and $\pi_{GT} < \pi_{BT}$. Under tort, expected wealth is

[18] Other permutations of λ_i and p_{i1} and p_{i2}, i = G, B, could be used to characterize good and bad drivers. For example, if the negligence system is unable to discriminate between the two groups or if drivers differ only in terms of miles driven, drivers would differ only in their accident probabilities. In the latter case, driver B would not actually be a "bad" driver, just a driver with higher automobile usage. We think the case discussed here is most representative of the real-world meaning of "good" and "bad" driving.

[19] It is assumed here that the level of care (x) is the same under the two systems. Although this will probably not be the case, x is unlikely to differ by a sufficiently large amount to affect the wealth condition for dominance.

$$E_T(W) = W - x - \pi_{iT} - \lambda_i(1 - p_{i2})(\ell + g)$$
$$= W - x - \lambda_i(\ell + g) - \lambda_i(\ell + g)(p_{i1} - p_{i2}), \qquad i = G, B. \tag{3}$$

Thus, no-fault cannot be dominant for good drivers, because $p_{G2} > p_{G1}$ for this group.[20] Dominance can occur only if driver 1 is a bad driver.

Assume that driver 1 is a bad driver. To test for dominance, two cases must be considered: $\pi_{BN} \geq \pi_{BT}$ and $\pi_{BN} < \pi_{BT}$. If $\pi_{BN} \geq \pi_{BT}$, Figure 1 applies, and no-fault is dominant if Area 1 ≥ Area 2 + Area 3. This condition can be written as follows:

$$(\ell + \pi_{BT} - \pi_{BN})\lambda(1 - p_{B2}) \geq g\lambda p_{B2} + (\pi_{BN} - \pi_{BT})[1 - \lambda(1 - p_{B2})]. \tag{4}$$

This expression simplifies to $p_{B1} \geq p_{B2}$, which is satisfied for bad drivers. Thus, no-fault is dominant for this group if $\pi_{BN} \geq \pi_{BT}$. The intuition is that bad drivers have a relatively low probability of compensation under tort and thus gain more than good drivers by switching to no-fault.

Figure 2 applies when $\pi_{BN} < \pi_{BT}$. Because the wealth condition is satisfied for bad drivers, dominance occurs if the distribution function condition, $A_{BN}(W^*) \leq A_{BT}(W^*)$ everywhere, is satisfied. This will be the case only if Area 1 ≥ Area 2, which implies

$$(\ell + \pi_{BT} - \pi_{BN})\lambda(1-p_{B2}) \geq g\lambda p_{B2} - \lambda p_{B2}(\pi_{BT} - \pi_{BN}).$$

This simplifies to

$$\ell(1 - p_{B2}) + (\pi_{BT} - \pi_{BN}) \geq p_{B2}\, g. \tag{5}$$

The left side of this inequality is the expected gain in the accident state realized by switching from tort to no-fault. The gain consists of the economic losses not compensated under tort ($\ell(1-p_{B2})$) plus the savings in insurance premiums. The right side is the expected loss of pain and suffering recoveries in the loss state given up by adopting no-fault. Because there is a premium savings from choosing no-fault ($\pi_{BT}-\pi_{BN} > 0$), a sufficient condition for dominance is $\ell(1-p_{B2}) \geq gp_{B2}$. This is more likely to be satisfied for bad risks since p_{B2} is relatively low for this group. Thus, there seems to be a reasonable range of parameter values for which bad risks would find no-fault dominant, particularly if the amount of otherwise uninsured economic losses is relatively high.[21]

The dominance results are summarized in Table 3. No-fault is dominant when $p_1 \geq p_2$ and $\pi_N \geq \pi_T$ and may be dominant for driver 1 if $p_1 > p_2$ and $\pi_N < \pi_T$. Thus, average and bad drivers strictly prefer no-fault when insured losses under no-fault are at least as large as insured losses under tort; and bad drivers

---

[20] Tort can never be dominant because the no-fault distribution function is always below the tort distribution function in the range W_T-ℓ-g to W_N-g, thus violating the wealth condition for dominance.

[21] Considering the premium condition $\pi_T > \pi_N$, which implies $(g+\ell)p_1 > \ell$, and the sufficient condition $\ell(1-p_2) \geq p_2 g$, no-fault will be dominant for bad risks if the ratio of economic losses to general damages satisfies the following interval condition: $p_2/(1-p_2) \leq \ell/g < p_1/(1-p_1)$. For example, if ℓ and g are equal, then no-fault is dominant if $p_2 \leq 0.5$ and $p_1 > 0.5$.

may strictly prefer no-fault even when the insurance premium is less under no-fault.[22] No-fault is not dominant when $p_1 = p_2$ and $\pi_N < \pi_T$ and is not dominant for good drivers ($p_1 < p_2$) regardless of the premium relationship. Of course, good drivers may still *prefer* no-fault coverage even though no-fault is not dominant for this group.

Table 3
Dominance Results with Actuarially Fair Insurance

Driver Types	*Premium*	*No-Fault Dominant?*
Bad Drivers	$\pi_N \geq \pi_T$	Yes
Bad Drivers	$\pi_N < \pi_T$	Ambiguous
Identical Drivers	$\pi_N \geq \pi_T$	Yes
Identical Drivers	$\pi_N < \pi_T$	No
Good Drivers	π_N <, =, or > π_T	No

Identical (average) and bad drivers are likely to choose no-fault under an elective no-fault law, and good drivers are more likely to choose tort. (Because the dominance results for "identical" drivers are determined primarily by the relationship between negligence probabilities (i.e., $p_1 = p_2$), the identical drivers can also be termed "average" drivers.) Thus, the results imply that elective no-fault may not reduce overall premiums as much as policy-makers might predict by analyzing average drivers because of the adverse selection effect (more bad drivers choosing no-fault).

If General Damages Are Irreplaceable. Pain and suffering losses are not inherently monetary but are given monetary value through the operation of the tort system. Thus, modeling such losses as monetary losses may not seem correct. This section demonstrates that the principal results are unchanged if pain and suffering is treated as irreplaceable, provided that an additional condition is satisfied.[23] The analysis in this section focuses on identical drivers (with equal λ and $p_1 = p_2$).

Cook and Graham's (1977) analysis of irreplaceable losses provides a model of consumer decision making with regard to assets that have unique attributes—that is, assets for which perfect market substitutes are not readily available. An important example of an irreplaceable commodity is one's good health. Pain and suffering awards can be viewed as the economic valuation of the temporary or permanent loss of health due to the actions of a tort feasor. In essence, the tort system steps in to provide an evaluation of a commodity (health or absence of pain) for which a market price does not exist.

Cook and Graham's model of irreplaceable assets assumes that the loss of such an asset shifts the decision-maker from his or her original utility function

---

[22] An average driver is assumed to be one for whom $p_1 = p_2$. The term "strictly prefer" as used here is equivalent to stochastic dominance.

[23] We retain the assumptions that insurance is actuarially fair and that care expenditures and accident rates are the same under tort and no-fault.

$U_b(W)$ to a new utility function $U_a(W)$, where $U_b(W) > U_a(W)$, for all W. Cook and Graham show that the shift from $U_b(W)$ to $U_a(W)$ can be expressed as monetary amounts using either the owner's value (c) or the ransom value (g), defined as

Owner's value = c(W), such that $U_a[W+c(W)] = U_b(W)$.

Ransom value = g(W), such that $U_a(W) = U_b[W-g(W)]$.

Thus, c(W) represents the amount of money the accident victim must receive in order to restore his or her utility to the pre-loss value, and g(W) is the amount the victim would be willing to pay for the certainty of avoiding the loss state. In this article, we assume that both the ransom value and the owner's value are increasing in wealth; that is, $g'(W) > 0$ and $c'(W) > 0$.

The effects of irreplaceability on the dominance analysis depend upon the relationship of general damage awards to the owner's value and ransom value of pain and suffering. As an introduction to this type of analysis, it is interesting to assume that the legal system under tort is accurate and objective in assessing general damages. Accuracy is defined such that the damage award under pure tort moves the victim exactly back to his or her pre-loss utility level. An accident involving personal injury moves the driver from utility function $U_b(W)$ to utility function $U_a(W)$ and reduces wealth from W_T to $W_T-\ell$, where $W_T = W-x_T-\pi_T$; that is, general damages are expressed as a change in utility rather than a monetary amount. If the driver makes a successful negligence claim, he or she receives a monetary award of ℓ so that wealth returns to W_T. Because the driver is still on utility function U_a, the owner's value condition implies that a general damage award of $c(W_T)$ will restore the driver's pre-loss utility. If the driver has an accident under tort and is not successful in making a negligence claim, wealth remains at $W_T-\ell$. With this new value of wealth, the ransom value condition becomes $U_a(W_T-\ell) = U_b[W_T-\ell-\hat{g}(W_T-\ell)]$, where $\hat{g}(W_T-\ell)$ = the ransom value applicable to wealth level $W_T-\ell$.

Under no-fault, the injured driver moves to utility function U_a and wealth $W_N-\ell$ following an accident, where $W_N = W-x_N-\pi_N$. The driver receives ℓ under his or her no-fault insurance policy, restoring wealth to W_N. However, because there is no pain and suffering award, the ransom value condition shows that utility becomes $U_a(W_N) = U_b[W_N-g(W_N)]$, where $g(W_N)$ = the ransom value for wealth level W_N. Table 4 summarizes the driver's utility outcomes under tort and no-fault.

Under the assumption that c(W) and g(W) are increasing in W, $\hat{g}(W-\ell) < g(W) < c(W)$ (Cook and Graham, 1977, p. 148). Since the wealth arguments may differ in our analysis, it is helpful to establish the slightly stronger condition that $\hat{g}(W_T-\ell) < g(W_N) < c(W_T)$. Because c(W) and g(W) are increasing in wealth, the first inequality, $\hat{g}(W_T-\ell) < g(W_N)$, holds if $W_T-\ell < W_N$. It is easy to show that $W_T-\ell < W_N$ with certainty. The second inequality, $g(W_N) < c(W_T)$, holds if $W_N-g < W_T$. While this is not a certainty,[24] it is very likely to hold, and this is assumed in the following discussion.

[24] A sufficient condition for $W_N-g < W_T$ is that $g/c > p_1\lambda$. This is not very stringent under reasonable assumptions about accident rates.

Table 4
Driver 1's Economic Position Under No-Fault and Tort with Pain and Suffering Irreplaceable

State	*Probability*	*Utility: No-Fault*	*Utility: Tort*
No Accident	$1-\lambda$	$U_b(W-x-\pi_N)$	$U_b(W-x-\pi_T)$
Driver 2 Negligent	λp_2	$U_a(W-x-\pi_N) =$ $U_b(W-x-\pi_N-g)$	$U_a(W-x-\pi_T+c)$ $= U_b(W-x-\pi_T)$
Driver 1, Neither, or Both Negligent	$\lambda(1-p_2)$	$U_a(W-x-\pi_N) =$ $U_b(W-x-\pi_N-g)$	$U_a(W-x-\pi_T-\ell) =$ $U_b(W-x-\pi_T-\ell-\hat{g})$
Premium		$\pi_N = \lambda\ell$	$\pi_T = p_1\lambda(c+\ell)$
Expected Wealth Under $U_b(W)$:		$W-x-\pi_N-\lambda g$	$W-x-\pi_T-\lambda(1-p_2)(\ell+\hat{g})$

Note: W = wealth, x = care expenditures, λ = accident probability, p_i = probability that driver i is at fault, ℓ = economic losses, and g = general damages. Also, $g=g(W-x-\pi_N)$, $\hat{g} = \hat{g}(W-x-\pi_T-\ell)$, $c = c(W-x-\pi_T)$.

Since all outcomes can be stated in terms of $U_b(W)$, it is possible to conduct the dominance analysis based on the arguments of $U_b(W)$ shown in Table 4. We consider the case of identical drivers. First, it is necessary to establish that the expected wealth is at least as large under no-fault as under tort. Computing expected wealth under the two systems reveals that the expected outcome is at least as large under no-fault if the following condition holds: $g \leq \hat{g}(1-p_2) + p_1c$; that is, if the ransom value of pain and suffering lost under no-fault is less than the weighted average of the ransom value and the owner's value under tort, with weights of $(1-p_2)$ and p_1, respectively. For convenience, this condition is called the *ransom/owner's value condition*. Intuitively, the condition states that the loss of the ransom value under no-fault (g) should be no greater than the expected ransom value loss under tort, $(1-p_2)\hat{g}$, plus the expected payment to the other driver if driver 1 is held negligent. This condition is more likely to be satisfied for bad drivers than for good drivers.

If the ransom/owner's value condition $g \leq \hat{g}(1-p_2)+cp_1$ is satisfied, then dominance analysis proceeds as above. As before, two cases must be considered: that $\pi_N \geq \pi_T$ and $\pi_N < \pi_T$. If the ransom/owner's value condition holds and $\pi_N \geq \pi_T$, then no-fault is dominant, regardless of whether the driver is a good or bad risk. Thus, unlike the replaceability case, dominance of tort by no-fault is now possible for good drivers. However, the ransom/owner's value condition is more likely to be satisfied for bad drivers since $p_1 > p_2$ for these drivers and $c > \hat{g}$. If $\pi_N < \pi_T$, the results are indeterminate, but no-fault is unlikely to be dominant for reasonable values of the parameters.

Dominance with Transactions Costs and Unequal Accident Rates

The preceding analysis was based on the assumptions that insurance is actuarially fair and that there is no difference in accident rates between the no-fault and tort regimes. Real world insurance systems clearly are not actuarially fair due to administrative expenses, costs of moral hazard (i.e., fraudulent and exaggerated claims), and insurer risk-bearing charges, and no-fault *may* weaken incentives, leading to higher accident rates. Thus, it is of interest to compare tort and no-fault with these assumptions removed.

To recognize the possibility of higher accident rates under no-fault, define the accident rates under no-fault and tort as λ_N and λ_T, respectively, with $\lambda_N > \lambda_T$. Assume that the no-fault accident rate is higher because of lower care expenditures—$x_N < x_T$—resulting from the weakening of the tort deterrent. Administrative costs, insurer risk-bearing charges, and moral hazard costs are introduced through a proportionate loading so that gross premiums are $G_N = \lambda_N \ell(1+e_N)$ for no-fault, and $G_T = \lambda_T p_1(\ell+g)(1+e_T)$ for tort,[25] where e_N and e_T are the loading factors for no-fault and tort, respectively. Thus, legitimate claims are reflected in the pure premium, while the costs of fraudulent and exaggerated claims are incorporated into the loading. It is assumed that $e_N > 0$, $e_T > 0$, and $e_N < e_T$. The latter assumption is justified because the administrative costs under a liability-based system are significantly higher than those under first-party coverages such as no-fault. In addition, the pain and suffering awards available under tort increase the incentive for drivers to file fraudulent or exaggerated claims, further increasing the dead-weight costs of the system for honest drivers.

In this section, we consider dominance for both replaceable and irreplaceable general damages under the revised assumptions. The discussion applies to honest drivers, that is, those who do not file false or exaggerated claims.

When General Damages Are Replaceable

The economic position of driver 1 when general damages are a replaceable commodity is summarized in Table 5. In order for no-fault to be dominant, the three conditions for dominance must be satisfied. Again, drivers are assumed to be risk-averse expected utility maximizers so the risk aversion condition is satisfied. The wealth condition requires that expected wealth under the no-fault regime must be at least as large as that under tort. The wealth condition holds if

$$x_N + G_N + \lambda_N g \leq x_T + G_T + \lambda_T (1 - p_2)(\ell + g). \tag{6}$$

Condition (6) requires that the costs of the accident compensation system be no larger under no-fault than under tort. The costs consist of (1) care expenditures (x_N and x_T), (2) premiums (G_N and G_T), and (3) expected uncompensated losses ($\lambda_N g$ for no-fault and $\lambda_T(1-p_2)(\ell+g)$ for tort).

Substituting the premium expressions from Table 5 into condition (6), the wealth condition becomes

$$\begin{aligned} x_N + \lambda_N (g + \ell) + \lambda_N \ell e_N \leq x_T + \lambda_T (g + \ell) + \\ \lambda_T p_1 (g + \ell) e_T + (p_1 - p_2) \lambda_T (\ell + g). \end{aligned} \tag{7}$$

[25] When general damages are irreplaceable, g is replaced with c in the tort premium.

Table 5
Driver 1's Economic Position Under No-Fault and Tort with Expenses and $\lambda_N > \lambda_T$

State	*Probability: No-Fault, Tort*	*No-Fault*	*Tort*
No Accident	$1-\lambda_N$, $1-\lambda_T$	$W-x_N-G_N$	$W-x_T-G_T$
Driver 2 Negligent	$\lambda_N p_2$, $\lambda_T p_2$	$W-x_N-G_N-g$	$W-x_T-G_T$
Driver 1, Neither, or Both Negligent	$\lambda_N(1-p_2)$, $\lambda_T(1-p_2)$	$W-x_N-G_N-g$	$W-x_T-G_T-\ell-g$
Premium		$G_N = \lambda_N\ell(1+e_N)$	$G_T = p_1\lambda_T(g+\ell)(1+e_T)$
Expected Wealth		$W-x_N-G_N-\lambda_N g$	$W-x_T-G_T-\lambda_T(1-p_2)(\ell+g)$

Note: W = wealth; x_N, x_T = care expenditures under no-fault and tort, respectively; λ_N, λ_T = accident probabilities under no-fault and tort, respectively; p_i = probability that driver i is at fault; ℓ = economic losses; g = general damages, and e_N, e_T = proportionate loadings for administrative, risk-bearing, and moral hazard costs under no-fault and tort, respectively.

Condition (7) indicates that the wealth condition is satisfied if care expenditures (x_N) plus expected accident losses ($\lambda_N(\ell+g)$) plus compensation system expenses ($\lambda_N\ell e_N$) under no-fault are less than total compensation system costs under tort. The tort costs include care expenditures, expected accident losses, and system expenses plus a penalty for being a bad driver or a reward for being a good driver, $(p_1-p_2)\lambda_T(\ell+g)$.[26] Even if no-fault increases the accident rate, the wealth condition may still be satisfied if the higher expected accident costs are offset by lower care expenditures ($x_N < x_T$) and system expenses ($\lambda_N \ell e_N < \lambda_T p_1 (g+\ell) e_T$). This is more likely to be the case for bad drivers than for good drivers. We now assume that the wealth condition is satisfied and consider the distribution function condition for identical drivers and for good and bad drivers.

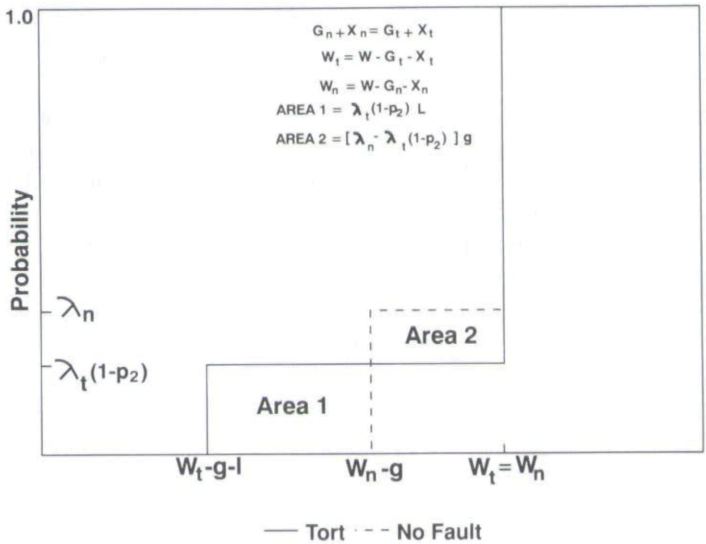
Identical Drivers ($p_1=p_2$). If the wealth condition (expression [6]) is satisfied, no-fault is dominant if the distribution function condition is also satisfied. Three cases are considered: (1) $G_N + x_N = G_T + x_T$, (2) $G_N + x_N < G_T + x_T$, and (3) $G_N + x_N > G_T + x_T$. In the first case, depicted in Figure 3, the premium plus care expenditures are equal under tort and no-fault. The distribution function condition is satisfied if Area 1 ≥ Area 2. The distribution function condition is the same as the wealth condition in this case. With the equality of the sum of premiums and care expenditures under the two systems, the wealth/distribution function condition reduces to the following necessary and sufficient condition for no-fault to be dominant:

$$\lambda_T \ell (1-p_2) \geq \lambda_T g p_2 + g(\lambda_N - \lambda_T). \tag{8}$$

Condition (8) is analogous to condition (1).

[26] Good drivers are assumed to have lower accident rates than bad drivers under both tort and no-fault. In addition, $p_1 < p_2$ for good drivers and $p_1 > p_2$ for bad drivers, as before. To conserve notation, the subscripts for driver class (G and B) are suppressed in this section.

Figure 3
Dominance with Administrative Expense



The left side of condition (8) is the expected value of the uncompensated economic loss under tort. The condition implies that no-fault is dominant if this uncompensated loss is greater than the expected general damages given up by moving from tort to no-fault ($\lambda_T g p_2$) plus the expected increase in pain and suffering costs, $g(\lambda_N - \lambda_T)$, attributable to the higher accident rate under no-fault. As above, this means that the gain in the lowest wealth state from moving to no-fault (the left side of condition [8]) exceeds the loss in the higher wealth state (the right side of condition [8]). Because there is an expected increase in pain and suffering due to the higher accident rate under no-fault, the condition for the dominance of no-fault is more stringent than it would be with equal accident rates.[27]

The second case we consider is when $G_N + x_N < G_T + x_T$, or when the wealth and distribution function conditions are not identical. This case is analogous to the one portrayed in Figure 2. The distribution function condition is

$$\lambda_T \ell (1 - p_2) \geq \lambda_T p_2 g + g(\lambda_N - \lambda_T) + \lambda_N (x_N + G_N - x_T - G_T). \tag{9}$$

Because the last expression on the right side of condition (9) is negative in this case, this expression is more likely to be satisfied than condition (8). In fact, condition (8) is sufficient but not necessary for dominance when $G_N + x_N < G_T$

---

[27] Although condition (8) itself does not depend upon the loading, one should keep in mind that the equality of premiums and care expenditures under the two systems is unlikely to occur when $\lambda_N > \lambda_T$ unless $e_N < e_T$.

$+ x_T$. This case, which provides the strongest possibility for dominance by no-fault, helps to explain why public policy discussions focus so often on the effect of no-fault on premiums and reinforces the argument that fraudulent and exaggerated claims have an adverse welfare effect.

In the third case we consider, when $G_N + x_N > G_T + x_T$, the wealth and distribution function conditions are again identical. However, because the sum of premiums and care expenditures is now larger under no-fault than under tort, condition (8) is no longer sufficient for dominance. Instead, the wealth condition (6) is necessary and sufficient for dominance. No-fault is least likely to be dominant in this case.

The conclusions of this section are as follows: (1) Condition (8) is sufficient for dominance of no-fault as long as the sum of premiums plus care expenditures under no-fault does not exceed that under tort. This provides a strong rationale for policy-makers' preoccupation with the effects of no-fault on premiums. (2) Because the expense ratio is lower under no-fault ($e_N < e_T$), no-fault may be stochastically dominant even if it leads to higher accident rates.

Good and Bad Drivers. The analysis in the good and bad driver case parallels that in the preceding section, with $p_1 > p_2$ for bad drivers and $p_1 < p_2$ for good drivers. The discussion focuses on the good drivers, since bad drivers are more likely to gain by switching from tort to no-fault. The driver-type subscript (G) is suppressed, and p_1, p_2, and λ are assumed to apply to good drivers.

When $G_T + x_T = G_N + x_N$ for the good drivers, expression (8) is still a necessary and sufficient condition for dominance.[28] However, condition (8) is less likely to be satisfied for good drivers, who give up more by switching from tort to no-fault. However, unlike the actuarial neutrality case, it is now possible that no-fault is dominant for the good drivers. Under actuarial neutrality, no-fault cannot be dominant for the good risks because the wealth condition is violated. With the introduction of system expenses, it is possible for the wealth condition to be satisfied for the good risks.[29]

As above, condition (8) is sufficient but not necessary for dominance when $G_N + x_N < G_T + x_T$ for the good drivers. However, G_N+x_N is less likely to be $< G_T+x_T$ for good drivers, and condition (8) is less likely to hold for this group.

It is more likely that $x_N + G_N > x_T + G_T$ for good drivers than for bad drivers. Recall that, in this case, the wealth condition is necessary and sufficient for dominance. The effects on good drivers are most clearly identified when the wealth condition is written as equation (7). Equation (7) shows that the wealth condition is less likely to be satisfied for the good risks because they lose their reward for good driving, $(p_1-p_2)\lambda_T(\ell+g)$, if no-fault is adopted. For bad drivers,

---

[28] In fact, the assumption that $p_1 = p_2$ was not used in deriving condition (8).

[29] More technically, no-fault cannot be dominant for good risks under actuarial neutrality because the condition $\pi_T+x_T = \pi_N+x_N$ and the distribution function condition cannot hold simultaneously for the good drivers. With actuarial neutrality and equal accident rates under tort and no-fault, $\pi_T+x_T=\pi_N+x_N$ implies that $\ell(1-p_1) = gp_1$, whereas the distribution function condition is that $\ell(1-p_2) \geq gp_2$. For good drivers, $p_1 < p_2$ so $\ell(1-p_2) < gp_2$ when $\ell(1-p_1) = gp_1$.

it is more likely that $G_N+x_N \leq G_T+x_T$ and less likely that $G_N+x_N > G_T+x_T$. The chance that condition (8) holds is also higher for bad drivers, so no-fault is more likely to be dominant for this group.

The conclusion is that no-fault is less likely to be dominant for good drivers than for bad drivers. However, unlike the actuarially neutral case, with the introduction of system expenses, it is possible for no-fault to be dominant for good drivers even if it increases the accident rate.

The discussion must be tempered somewhat if experience rating is not perfect so that insurance premiums under tort are not based on each driver's negligence probability p_1 but, say, on a weighted average of the driver's p_1 and the overall average negligence probability. In this case, it is more likely that $G_N+x_N \leq G_T+x_T$ for the good drivers so that condition (8) is sufficient for dominance. Thus, imperfect experience rating would make no-fault more attractive to the good drivers since they do not receive full credit in the premium for their lower negligence rate.

Irreplaceable General Damages

This section considers the case where general damages are irreplaceable under the assumptions that $x_N < x_T$, $\lambda_N > \lambda_T$, and $e_N < e_T$; that is, in comparison with tort, the no-fault regime is characterized by lower care expenditures, a higher accident rate, and lower expense loadings. The driver's economic position can be expressed analogously to Table 4, inserting G_N for π_N and G_T for π_T. General damages are compensated under tort at the owner's value, $c = c(W\text{-}x_T\text{-}G_T)$. Under no-fault, the driver loses the ransom value $g = g(W\text{-}x_N\text{-}G_N)$ in the accident states; while under tort, the pain and suffering loss in the uncompensated accident state is $\hat{g} = \hat{g}(W\text{-}x_T\text{-}G_T\text{-}\ell)$. As above, we make the non-restrictive assumption that $\hat{g} < g < c$.

With identical drivers, three cases are considered: $x_N + G_N =, <,$ and $> x_T + G_T$. If $x_N + G_N = x_T + G_T$, the following condition is necessary and sufficient for no-fault to be dominant:

$$\lambda_T (1 - p_2) \ell \geq \hat{g} \lambda_T p_2 + g \lambda_N - \hat{g} \lambda_T . \tag{10}$$

This is directly analogous to relation (8) for the replaceable loss case. Consequently, relation (8) can be viewed as an approximation to the irreplaceable loss condition (10). Condition (10) is necessary and sufficient for no-fault to be dominant if $x_N + G_N < x_T + G_T$. Finally, if $x_N + G_N > x_T + G_T$, the following condition, analogous to condition (6), is necessary and sufficient for dominance:

$$\lambda_T (1 - p_2) \ell \geq \lambda_T p_2 \hat{g} + \lambda_N g - \lambda_T \hat{g} + (x_N + G_N - x_T - G_T). \tag{11}$$

Condition (11) is obviously more difficult to satisfy than condition (10), so no-fault is less likely to be dominant when the total of care expenditures and insurance premiums is higher under no-fault than under tort. The conditions for dominance under irreplaceability are thus analogous to those in the replaceable loss case. The replaceable loss conditions can be viewed as approximations to those derived under irreplaceability.

Summary

No-fault auto insurance may stochastically dominate tort even if no-fault leads to a higher accident rate. This occurs primarily because the system expense ratio is lower under no-fault. Because the empirical evidence suggests that the expense ratio difference may be *substantial* (see below), this analysis should have considerable relevance for public policy regarding auto accident compensation.

Empirical Evidence and System Flaws

Empirical Evidence

The preceding analysis has derived conditions under which no-fault would stochastically dominate tort as a compensation mechanism for automobile accidents. To gauge the appeal of no-fault in a real world setting, it is useful to consider the available empirical evidence about the key parameter values to determine whether the dominance conditions are likely to hold in practice.

The analysis is conducted from the perspective of one who is considering a move from tort to no-fault. A pure tort system is compared to modified no-fault. As in the appendix, modified no-fault is modeled by introducing the parameter F^T = the probability that a claim satisfies the threshold, thereby qualifying for tort. For pure tort, $F^T = 1$; for pure no-fault, $F^T = 0$; and for modified no-fault, $0 < F^T < 1$. The tort premium is G_T from Table 5, while the modified no-fault premium is $G_N = \lambda_N[\ell(1+e_N) + gF^Tp_1(1+e_T)]$. For modified no-fault, the wealth condition and revised condition (8) become

$$\textit{Wealth Condition}: x_N + G_N + \lambda_N(1 - p_2F^T)g \leq x_T + G_T + \lambda_T(1 - p_2)(g + \ell) \tag{12}$$

$$\textit{Revised Condition (8)}: \lambda_T(1 - p_2)(\ell + g) \geq \lambda_N g(1 - p_2F^T). \tag{8*}$$

By analogy with the results of the preceding section, condition (8*) is necessary and sufficient for dominance if $x_T+G_T = x_N+G_N$ and sufficient but not necessary if $x_N+G_N < x_T+G_T$. In the latter case, the necessary condition is the analogue of condition (9). If $x_N+G_N > x_T+G_T$, the wealth condition must be evaluated, because condition (8*) is not sufficient for dominance.

Evidence presented in Cummins and Weiss (1991) indicates that states with verbal threshold no-fault laws had lower total premiums for personal injury coverage (personal injury protection plus bodily injury liability) than tort states, but that states with low dollar thresholds tended to have higher personal injury premiums than tort states. Thus, condition (8*) is more likely to be sufficient for dominance when a verbal threshold rather than a weak monetary threshold is being considered. It is also important when $x_N+G_N > x_T+G_T$, because the difference between the left and right sides of condition (8*) must be > 0 and sufficiently large to offset the higher premiums under no-fault.

Data compiled by the U.S. Federal Highway Administration (1986) reveal that the frequency rate for accidents involving personal injury in tort states is about 0.018, or 1.8 accidents for every 100 driver years of experience.[30] We adopt this rate as an estimate of the accident probability under tort, λ_T. Evidence about the effects of no-fault on accident rates is ambiguous. An important study by Landes (1982) suggests that no-fault may have increased fatal accident rates by as much as 5 percent. We adopt this figure for purposes of evaluating no-fault and assume that it also applies to non-fatal personal injury accidents so that the estimated accident rate under no-fault is $\lambda_N = 0.0189 = 0.018 \times 1.05$.

To estimate the expense ratios under tort and no-fault, we use state insurance data from the A. M. Best Company. The average pure loss ratio in auto liability insurance in tort states for the period 1975 through 1982 was 0.66. This implies that the loading— that is, the inverse loss ratio minus one—was about 0.5 or 50 percent of losses for tort coverages. We adopt 0.5 as an estimate of e_T, the expense loading under tort. Using similar reasoning, the estimate of e_N, the expense loading under no-fault, based on personal injury protection coverage in no-fault states is approximately 0.2 or 20 percent. Thus, it is clear that no-fault is the more efficient auto insurance system in terms of costs required to deliver \$1 of losses.[31]

The All-Industry Research Advisory Council (AIRAC) (1989) provides a source for the economic losses and general damages arising from auto accidents. AIRAC's findings indicate that general damages on the average are approximately equal to economic losses. Thus, we assume that $\ell = g$. This permits the elimination of ℓ and g from condition (8^*). An estimate of the effectiveness of the threshold is also needed. AIRAC's results reveal that only about 10 percent of personal injury claims qualified for tort under Michigan's verbal no-fault threshold, and we use this as the initial estimate of F^T; that is, $F^T = 0.1$. There were four no-fault states with F^T in the neighborhood of 0.3 and two near 0.5. Consequently, we also use these values of F^T in the simulations. The most difficult parameter to estimate is p_2, the probability of being able to collect from the other driver under tort. Because there are no completely satisfactory data on this parameter, we estimate condition (8^*) for values of p_2 ranging from 0.16 to 0.5.

---

[30] This is the sum of the injury and fatality rates in U.S. Department of Transportation (1986) for 1986. The rate is stated as a frequency rather than a probability because auto accidents have been shown to follow either a Poisson or negative binomial distribution so that it is possible for a given driver to have more than one accident.

[31] The expense ratios are obtained as one minus the pure loss ratio, which is the ratio of incurred losses to premiums earned. Thus, the expense ratios include loss adjustment expenses. Consistent with our theoretical model, we express the loading as a proportion of the loss. This approach differs from the usual expense ratio, which is a proportion of premiums. Because premiums are a function of the present value of losses and the loss ratio is the ratio of undiscounted losses to premiums, the expense ratio is understated relative to the present value of losses. In addition, our expense ratio reflects only insurer expenses. We are not taking into account the fact that a significant part—30-45 percent—of the pure loss payment under tort goes to the plaintiff's attorney. Thus, our assumptions are conservative in that they are biased against the dominance of no-fault.

Based on these parameter assumptions, we test for the dominance of no-fault by evaluating condition (8^*) and, where appropriate, condition (12) and the analogue of condition (9) for modified no-fault. The results of the simulations are shown in Table 6. The top panel of the table assumes that accident rates are higher under no-fault than under tort. Various values of F^T, p_1, and p_2 are tested, holding the other parameters constant. For example, the upper left box in Table 6 uses $F^T = 0.1$ and varies p_1 from 0.2 to 0.5 for a range of values of p_2. As expected, when $p_1 = 0.5$ so that the decision-maker is a bad driver ($p_1 < p_2$) for nearly every case tested, no-fault is dominant in all cases. As the decision-maker's probability of negligence (p_1) declines, no-fault is less likely to be dominant. Interestingly, no-fault is rarely dominant for very good drivers (e.g., those with $p_1 = 0.2$) even when the probability of collecting from the other driver (p_2) is very high. This is because increases in p_2 have a larger (inverse) impact on expected uncovered losses under tort than under no-fault.

The middle box in the upper panel runs the same set of experiments, changing F^T to 0.3. The number of cases where no-fault is dominant declines, reflecting the higher premiums under no-fault, which lead to violations of the wealth condition for some sets of parameter values. Changing F^T to 0.5 (see the right-most box in the top panel of Table 6) has little effect on the results; there is only one less dominant case than for $F^T = 0.3$. The reason is that increasing F^T has offsetting effects on wealth: the premium increases but the chance of being uncompensated for general damages is reduced.

We then explore the effects of changes in the ratio of economic losses to general damages and of changes in the relationship of the accident rates under the two systems. The baseline case is the middle box in the lower panel of Table 6 (identical to the middle box in the top panel), where the key assumptions are $\lambda_N > \lambda_T$, $F^T = 0.3$, and the ratio of economic losses (ℓ) to general damages (g) is equal to 1.0. The lower left box in Table 6 shows the effect of increasing the ratio of economic loss to general damages from 1.0 to 1.5.[32] No-fault is dominant in fewer cases when economic losses are higher relative to general damages primarily because the accident rate is higher under no-fault. With higher accident severity, the higher assumed accident rates under no-fault make a more significant difference. This is confirmed by results in the lower right box in Table 6, where $\ell/g = 1.5$ but $\lambda_N = \lambda_T$. No-fault is *more likely* to be dominant with equal accident rates because more of the expected accident losses are brought into the insurance system, reducing uncertainty and improving utility. This occurs even though insurance is actuarially unfair, in part because the tort expense ratio exceeds the no-fault expense ratio.

As a final experiment (not shown in the table) we increase the tort expense ratio to 0.6 and then to 0.7. In both cases, no-fault is dominant in more

[32] The change in the ratio of economic loss (ℓ) to general damages (g) affects the results in two primary ways: accident severity increases from ℓ+g=2g to 1.5ℓ+g=2.5g, and the relative magnitude of economic losses increases relative to general damages. Because no-fault does a better job of compensation for economic losses than for general damages, both the absolute and relative effects play an important role.

Table 6
Simulations of Dominance of No-Fault (1 = Dominance)

$\lambda_T = 0.018$ $\lambda_N = 0.0189$ $\ell/g = 1.0$ $e_T = 0.5$ $e_N = 0.2$

	$F^T = 0.1$				$F^T = 0.3$				$F^T = 0.5$			
	p_1				p_1				p_1			
p_2	0.5	0.4	0.3	0.2	0.5	0.4	0.3	0.2	0.5	0.4	0.3	0.2
0.16	1	1	1	1	1	1	1	1	1	1	1	1
0.18	1	1	1	1	1	1	1	0	1	1	1	0
0.20	1	1	1	0	1	1	1	0	1	1	1	0
0.22	1	1	1	0	1	1	1	0	1	1	1	0
0.24	1	1	1	0	1	1	1	0	1	1	1	0
0.26	1	1	1	0	1	1	1	0	1	1	1	0
0.28	1	1	1	0	1	1	1	0	1	1	1	0
0.30	1	1	1	0	1	1	1	0	1	1	1	0
0.32	1	1	1	0	1	1	1	0	1	1	0	0
0.34	1	1	1	0	1	1	0	0	1	1	0	0
0.36	1	1	0	0	1	1	0	0	1	1	0	0
0.38	1	1	0	0	1	1	0	0	1	1	0	0
0.40	1	1	0	0	1	1	0	0	1	1	0	0
0.42	1	1	0	0	1	1	0	0	1	1	0	0
0.44	1	1	0	0	1	1	0	0	1	1	0	0
0.46	1	1	0	0	1	1	0	0	1	1	0	0
0.48	1	1	0	0	1	0	0	0	1	0	0	0
0.50	1	0	0	0	1	0	0	0	1	0	0	0

$\lambda_T = 0.018$ $\lambda_N = 0.0189$ $F^T = 0.3$ $e_T = 0.5$ $e_N = 0.2$

	$\ell/g = 1.5$				$\ell/g = 1$				$\ell/g = 1.5$; $\lambda_N = \lambda_T$			
	p_1				p_1				p_1			
p_2	0.5	0.4	0.3	0.2	0.5	0.4	0.3	0.2	0.5	0.4	0.3	0.2
0.16	1	1	1	0	1	1	1	1	1	1	1	1
0.18	1	1	1	0	1	1	1	0	1	1	1	1
0.20	1	1	1	0	1	1	1	0	1	1	1	1
0.22	1	1	1	0	1	1	1	0	1	1	1	1
0.24	1	1	1	0	1	1	1	0	1	1	1	0
0.26	1	1	1	0	1	1	1	0	1	1	1	0
0.28	1	1	1	0	1	1	1	0	1	1	1	0
0.30	1	1	1	0	1	1	1	0	1	1	1	0
0.32	1	1	0	0	1	1	1	0	1	1	1	0
0.34	1	1	0	0	1	1	0	0	1	1	1	0
0.36	1	1	0	0	1	1	0	0	1	1	1	0
0.38	1	1	0	0	1	1	0	0	1	1	0	0
0.40	1	1	0	0	1	1	0	0	1	1	0	0
0.42	1	1	0	0	1	1	0	0	1	1	0	0
0.44	1	1	0	0	1	1	0	0	1	1	0	0
0.46	1	0	0	0	1	1	0	0	1	1	0	0
0.48	1	0	0	0	1	0	0	0	1	1	0	0
0.50	1	0	0	0	1	0	0	0	1	1	0	0

Note: Dominance is determined by evaluating condition (8^*) and, where appropriate, (12) and the analogue of (12) for modified no-fault. ℓ_T = accident frequency, tort; ℓ_N = accident frequency, no-fault; F^T = probability of exceeding no-fault threshold; ℓ = economic loss; g = general damage loss; e_T = expense ratio for liability coverage; e_N = expense ratio for no-fault economic loss coverage; x_N = care expenditures under no-fault; x_T = care expenditures under tort; p_1 = probability of driver being held negligent; p_2 = probability of collecting from the other driver; x_N was set arbitrarily at 40 percent of expected loss costs under no-fault, while x_T was set at 45 percent of no-fault expected loss costs. Sources of other parameter values are given in the text.

cases than when the tort expense ratio is lower. Such a reduction in expected benefits under the policy would be characteristic of systems with high attorneys' fees. That is, the expense ratios we derived from Best's account only for insurance company expenses, not for the share of the liability award that goes to the plaintiff's attorney. In addition, drivers who do not file fraudulent automobile insurance claims are implicitly paying a higher loading, because part of their premium is used to pay fraudulent or built-up claims from which they receive no benefit (see Weisberg and Derrig, 1991a). Thus, honest drivers in states with relatively serious insurance fraud problems are more likely to choose no-fault auto insurance.

The conclusions of this discussion can be stated as testable hypotheses that could be explored in subsequent research: (1) Bad drivers are more likely than good drivers to choose no-fault. (2) For drivers and/or jurisdictions where economic losses are high relative to general damages, no-fault is more likely to be preferred as long as its impact on accident rates is minimal. (3) No-fault is more likely to be preferred when thresholds are stringent, although this effect is blunted somewhat by the higher chance of collecting for general damages when thresholds are low. (4) No-fault is more likely to be preferred in states with serious fraud problems. These hypotheses could be tested by surveying drivers or by estimating a probit equation with selection of no-fault as an endogenous variable.

Other Sources of Recovery

A real-world complication that must be considered in evaluating tort and no-fault is the presence of other sources of recovery, known as *collateral sources.* If auto accident losses are already effectively covered by other efficient compensation mechanisms (e.g., group health insurance), then the first-party benefits under no-fault are not likely to yield the utility gains necessary for stochastic dominance. Central to the case for no-fault is the contention that significant economic losses are compensated under no-fault that would not be compensated under tort. This section examines some evidence regarding this critical assumption.

In the debate over no-fault, the collateral source issue has been controversial for a number of reasons. The *collateral source rule* under tort prohibits juries from considering the availability of alternative sources of compensation when determining damage awards. Thus, it is possible for some claimants to collect more than once for economic losses under tort; others collect less than their economic loss, because there is no one to sue or because collateral insurance benefits are unavailable. One objective of no-fault was to reallocate money from double or triple recoveries that can occur under tort to compensate those who would otherwise not be reimbursed for their economic losses. As discussed below, no-fault has partially accomplished this objective.

Because of the patchwork nature of the medical care compensation system in the United States, empirical evidence on coverage gaps and collateral sources has been sketchy and difficult to obtain. The Rand Corporation (1985) has conducted the most authoritative study on this issue to determine the adequacy of compensation from all sources. Rand analyzed survey data on 1,500 households where at least one household member had suffered a personal injury loss from an automobile accident. It was found that the probability of receiving compensation for an accident was higher in no-fault than in tort states.[33] Furthermore, the Rand study indicates that "conditional on the accident victim receiving some payment, total payment is on average closer to economic loss in no-fault states than it is in tort states," and, further, that "compared to no-fault states, tort states average higher payment at small losses and lower payment at large losses" (v. 3, pp. 4, 21). Thus, no-fault has reduced the uncertainty associated with personal injury losses from automobile accidents.

Apart from the Rand study, it is well known that a significant proportion of households in the United States (about 13 percent in 1985) are not covered by any type of private or governmental health insurance (U.S. Department of Commerce, 1991) and that coverage is inadequate for many of those with insurance. Even more households have no coverage for wage losses. Furthermore, because of copayment provisions in most health insurance polices, health insurance often covers a smaller proportion of any given loss than auto insurance. Such provisions typically are not present in first-party no-fault auto coverage.[34] Thus, no-fault clearly covers some losses that health insurance does not.[35]

The amount of otherwise uncovered losses that would be insured under no-fault varies by driver depending upon the adequacy of first-party health coverage.[36] As long as some significant amount of economic loss is covered under no-fault and not covered under tort, the analytical techniques presented here can be used to study dominance.[37] The analysis would proceed essentially as above, except that ℓ would be interpreted as the amount of economic loss covered by no-fault and not covered by any other source.[38]

---

[33] Eighty-seven percent of victims received some payment for medical or wage loss under no-fault, compared to 80 percent under tort. For wage loss alone, 77 percent received payment under no-fault, compared to only 59 percent under tort.

[34] Purchasing individual first-party health coverage rather than auto insurance to plug the gaps is usually not a viable option, because the most efficient health coverage is provided by group plans, where individual coverage options are minimal, especially to plug copayment gaps that are intended to control moral hazard. Individual health insurance tends to be very expensive, with a high proportion of the premium accounted for by administrative and sales costs.

[35] In most states, in an attempt to control for duplicate recovery, no-fault is primary over health coverage when the injured victim has both types of insurance.

[36] In tort states, a small amount of first-party personal injury coverage is provided through the medical payments section of the automobile policy. This coverage is very limited compared to the first-party coverage provided by most no-fault laws.

[37] Of course, if the amount brought under insurance by no-fault is very small, the results are technically correct but the practical effects would be limited. The reader should recall, however, that the additional coverage will be quite large for some drivers even though it may be small for others. Some recent no-fault laws have given drivers the option to substitute other health coverage for first-party no-fault benefits.

[38] Recognizing the possibility of double recovery would complicate the analysis but would not significantly change the results, for two reasons: Double recoveries are not free; they lead to

No-Fault and National Health Insurance

This discussion implies that comprehensive national health insurance would weaken the case for no-fault. However, even if health coverage were more widely available, there are still reasons to consider no-fault. For example, no-fault with a strong verbal threshold is an effective mechanism for controlling insurance fraud, because it eliminates from the auto insurance system the relatively small claims that are easier to fake or to build up. Thus, no-fault could still play a role in reducing auto insurance costs. In addition, optimal resource allocation provides an argument for making auto insurance primary with respect to national health insurance in order to accurately reflect the higher expected costs that additional driving places on the insurance system. This would have to be balanced against the potentially higher administrative costs of individual insurance relative to a national health insurance system. In addition, if a national health plan incorporated significant cost control mechanisms—such as peer review—moral hazard would be a more serious problem under auto insurance to the extent that such features are not included in the auto insurance system.

Elective No-Fault

Passing no-fault insurance laws with effective thresholds is difficult politically because of the powerful trial attorneys' lobby. Trial lawyers argue that no-fault forces drivers to give up a valuable right—the right to sue—which most would not surrender voluntarily. To overcome this argument, Pennsylvania and New Jersey have adopted *elective* no-fault laws. In both states, drivers purchase first-party economic loss coverage but have the option of retaining the right to sue or being covered under a verbal threshold no-fault system.[39] Drivers electing no-fault can sue only for accidents whose severity satisfies the verbal threshold, while drivers electing tort retain the right to sue for all accidents.

Drivers who choose no-fault provide a cost savings to the system as a whole equal to the expected value of the liability awards they would have received if they had retained the right to sue. In our notation, the savings is $\lambda p_2 g(1-F^T)(1+e_T)$. In principle, this savings is passed back to the drivers electing no-

---

higher insurance premiums. And double payments take place in relatively low marginal-utility states of the world in comparison with payments for otherwise uncovered economic losses.

[39] Although we consider elective no-fault to be an appropriate name for this system, neither Pennsylvania nor New Jersey uses this term. The New Jersey Insurance Department maintains that New Jersey has a no-fault system with a choice of two thresholds, verbal and zero. No-fault with a zero threshold looks suspiciously like tort. Pennsylvania calls its system "limited tort." The mechanism of choice varies by state. In Pennsylvania, drivers must specifically opt out of the full tort system; in New Jersey, no-fault with a verbal threshold is the default, and drivers must explicitly elect the zero threshold option. Not surprisingly, only 25 percent of drivers in Pennsylvania have elected the limited tort option, whereas 75 percent of New Jersey drivers are insured under the verbal threshold system.

fault through a reduction in their insurance premiums.[40] However, these drivers must still pay for liability insurance because they can be sued by drivers who have elected to retain tort. Thus, the premium for drivers selecting full tort coverage is $G_T = \lambda[\ell(1+e_N) + p_1g(1+e_T)]$, and the premium for drivers selecting no-fault coverage is $G_N = G_T - \lambda p_2g(1-F^T)(1+e_T)$. The system differs from the tort models discussed above in providing first-party economic loss coverage.

Stochastic dominance analysis reveals that wealth is certain to be increased for bad risks who elect no-fault and may be increased for the good risks as well. No-fault will be dominant for the bad risks but is unlikely to be dominant for good risks under reasonable assumptions about parameter values. Thus, bad risks are more likely than good risks to choose the no-fault option, and the amount of savings that will be realized under the system will be less than expected if this adverse selection effect is not taken into account. In addition, elective no-fault is likely to be ineffective in curbing fraud and build-up, because motorists intent on exploiting the insurance system will retain the right to sue. In order to achieve meaningful welfare gains, states planning to adopt elective no-fault should closely examine the incentives created by the system.

Elective no-fault creates an additional source of risk for insurers: the risk that the proportion of their policyholders selecting no-fault will differ from the state average. For example, suppose that proportion α of drivers choose no-fault in the state as a whole, $0 < \alpha < 1$, but proportion $\beta > \alpha$ of a particular insurer's drivers choose no-fault. The system-wide savings will be $\alpha\lambda p_2g(1-F^T)$ per driver, but this insurer's premium revenues will go down by more than this amount because more of its drivers elect no-fault and receive the premium reduction. The insurer's revenues will decrease proportionately to β, but its costs will decline only by proportion α, forcing the insurer into a loss position even if the overall system savings have been estimated correctly. New Jersey's solution to this problem was to set up a pooling arrangement whereby all insurers end up with the average reduction in premiums. This approach is under consideration in Pennsylvania.

Conclusions

No-fault has been proposed as a compensation system that is superior to tort in handling the costs of automobile accidents. No-fault limits the ability to sue for relatively minor injuries and utilizes the savings in general damages to pay the economic losses of accident victims. Previous research has shown that drivers are paid more promptly under no-fault and that it costs less to administer than tort. The chances of being uncompensated or under-compensated for economic losses are lower under no-fault than under tort.

The principal argument in favor of no-fault can be expressed in terms of stochastic dominance, assuming that no-fault brings under first-party insurance

---

[40] In practice, states have given discounts to motorists based on the estimated overall savings arising from the elective no-fault system. In both Pennsylvania and New Jersey, the elective no-fault laws were accompanied by mandatory premium reductions.

coverage some significant losses that are not covered under tort. Under this assumption, drivers with no-fault coverage give up general damages in some subset of accidents in return for economic loss reimbursement in all accidents. Thus, the probability of losing both general damages and economic losses due to one's own negligence is reduced. No-fault eliminates the possibility of low-wealth states of the world by transferring funds from higher-wealth states. Thus, there is reason to hypothesize that no-fault stochastically dominates tort.

With actuarially fair insurance and equal accident rates under tort and no-fault, no-fault is dominant as long as its premium exceeds the tort premium; that is, as long as a higher amount of the expected accident losses are covered by insurance under no-fault. This can be viewed as the "pure theory" of the equity argument for no-fault. However, in the pure theory case, no-fault cannot be stochastically dominant for good drivers because they sacrifice wealth in moving from tort to no-fault. This obviously reduces the attractiveness of no-fault in this case.

When expense loadings are introduced, however, a stronger case can be made for no-fault as long as it delivers benefits more efficiently than tort. Under these circumstances, no-fault may be dominant, even for good drivers and even if no-fault leads to higher accident rates by weakening the tort deterrent. However, in a practical sense, no-fault is unlikely to be dominant unless a high proportion of tort claims are eliminated by the threshold. No-fault is also unlikely to be dominant for drivers who have adequate first-party health insurance. On the other hand, no-fault also provides an effective mechanism for controlling insurance fraud and build-up because it eliminates from the tort system the relatively minor claims that are easiest to exploit. Thus, even if no-fault is not dominant when one considers only legitimate claims, it may well be dominant for honest drivers in states with serious moral hazard problems.

Although there are persuasive arguments in favor of no-fault, enacting no-fault laws with effective thresholds has proven to be politically infeasible in most states. This has led to the introduction of a relatively new concept, *elective no-fault,* where drivers are given the option of choosing tort or no-fault. Under elective no-fault, drivers for whom no-fault is dominant could choose no-fault, and others could remain within the tort system. However, because bad drivers are more likely to choose no-fault, the cost savings from this system are likely to be less than anticipated. Furthermore, elective no-fault will be ineffective in curbing fraud and build-up because dishonest drivers will simply retain the right to sue. Thus, elective no-fault plans may suffer severe adverse selection and moral hazard problems. This places a more significant burden on anti-fraud enforcement to control auto insurance costs, and enforcement is typically less effective than market mechanisms. Thus, for states interested in maximizing consumer welfare and controlling fraud and build-up, mandatory no-fault with a strict verbal threshold still may be the best option.

Appendix

Stochastic Dominance Under Modified No-Fault

All no-fault laws presently in existence in the United States permit lawsuits under some circumstances. Such systems can be called *modified no-fault*. Under modified no-fault, an injured victim can sue for pain and suffering provided that the loss satisfies a threshold, defined either in terms of a dollar amount of economic loss or in verbal terms. Define the probability of exceeding the threshold as F^T. Then driver 1's economic position can be defined as in Table A1.

Table A1

Driver 1's Economic Position Under Modified No-Fault and Tort

State	*Probability*	*Modified No-Fault*	*Tort*
No Accident	$1-\lambda$	$W-x-\pi_N$	$W-x-\pi_T$
Driver 2 Negligent:			$W-x-\pi_T$
Threshold Satisfied	$\lambda p_2\ F^T$	$W-x-\pi_N$	
Threshold Not Satisfied	$\lambda p_2\ (1-F^T)$	$W-x-\pi_N-g$	
Driver 1, Neither, or Both Negligent	$\lambda(1-p_2)$	$W-x-\pi_N-g$	$W-x-\pi_T-\ell-g$
Premium		$\pi_N = \lambda\ell + p_1\lambda F^T g$	$\pi_T = p_1\lambda(g+\ell)$

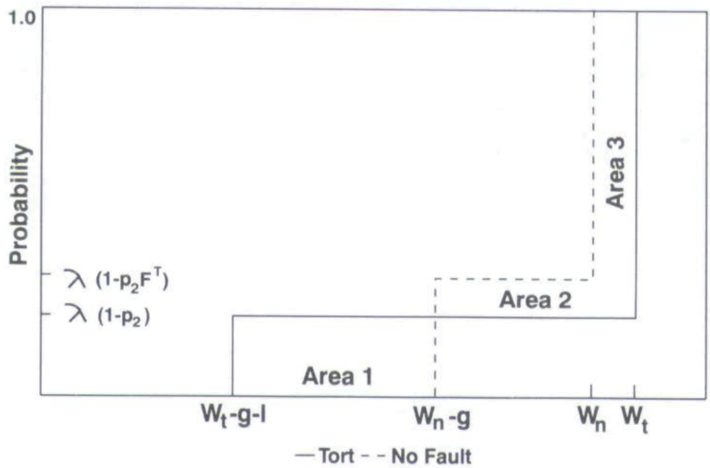
Note: W = wealth, x = care expenditures, λ = accident probability, p_i = probability that driver i is at fault, ℓ = economic losses, and g = general damages.

Retaining the assumption that the accident rate λ is the same under the two systems, the dominance analysis proceeds as in the pure no-fault case presented in the text. In the appendix, we consider only the case of identical drivers. With these assumptions, it is easily demonstrated that expected wealth is the same under tort and modified no-fault. Two cases are considered, (1) $\pi_N \geq \pi_T$, and (2) $\pi_N < \pi_T$. As in the text, no-fault is dominant in the first case but not in the second.

To analyze case 1, consider Figure 1A. The figure is very similar to Figure 1, except that the probability of being in the loss state under no-fault is now smaller: $\lambda(1-p_2F^T)$ rather than λ. Area 1 is the same size as in the pure no-fault case, and Area 1 = Area 2 + Area 3, as before. The difference is that a portion of the expected value under no-fault is shifted from Area 2 to Area 3. Area 2 $= g\lambda p_2(1-F^T)$, and Area 3 $= [1-\lambda(1-p_2)](\pi_N-\pi_T)$. Because of the ability to sue under modified no-fault, the expected value of the general damages lost (Area 2) is smaller than under pure no-fault, and the expected value of the difference between the tort and no-fault premiums is larger. Thus, modified no-fault is utility increasing when compared to tort, and it is easy to show that dominance exists for case 1. The lack of dominance in case 2 also is demonstrated similarly to the pure no-fault model.

Figure 1a
Modified No-Fault: Case 1

$\pi_N \geq \pi_T$



References

All-Industry Research Advisory Council, 1979, *Automobile Injuries and Their Compensation in the United States,* 2 Volumes (Oak Brook, Ill.: AIRAC).

All-Industry Research Advisory Council, 1989, *Compensation for Automobile Injuries in the United States* (Oak Brook, Ill.: AIRAC).

A. M. Best Company, annual, *Best's Executive Data Service* (Oldwick, N.J.: A. M. Best).

American Insurance Association, 1987, *Summary of Selected State Laws and Regulations Relating to Automobile Insurance* (New York: AIA).

Conard, Alfred F., et al., 1964, *Automobile Accident Costs and Payments* (Ann Arbor: University of Michigan Press).

Cook, Philip and Donald Graham, 1977, The Demand for Insurance and Protection: The Case of Irreplaceable Commodities, *Quarterly Journal of Economics,* 91: 143-156.

Cummins, J. David and Sharon Tennyson, 1992a, Controlling Automobile Insurance Costs, *Journal of Economic Perspectives,* 6: 95-115.

Cummins, J. David and Sharon Tennyson, 1992b, The Tort System and Insurance Fraud: Theory and Evidence from Automobile Insurance, Working Paper, Department of Insurance and Risk Management, Wharton School, University of Pennsylvania, Philadelphia.

Cummins, J. David and Mary A. Weiss, 1991, The Effects of No Fault on Auto Insurance Loss Costs, *Geneva Papers on Risk and Insurance,* 16: 20-38.

Cummins, J. David and Mary A. Weiss, 1992a, Incentive Effects of No Fault: Evidence from Insurance Claim Data, in: Georges Dionne, ed., *Contributions to Insurance Economics* (Norwell, Mass.: Kluwer Academic Publishers).

Cummins, J. David and Mary A. Weiss, 1992b, The Structure, Conduct, and Regulation of the Property-Liability Insurance Industry, in: Richard W. Kopcke and Richard E. Randall, eds., *The Financial Condition and Regulation of Insurance Companies* (Boston: Federal Reserve Bank of Boston).

Cummins J. David and Mary A. Weiss, 1993, The Incentive Effects of No-Fault Automobile Insurance, Working Paper, Department of Insurance and Risk Management, Wharton School, University of Pennsylvania, Philadelphia.

Devlin, Rose Anne, 1992, Liability Versus No Fault Automobile Insurance Regimes: An Analysis of the Experience in Quebec, in: Georges Dionne, ed., *Contributions to Insurance Economics* (Norwell, Mass.: Kluwer Academic Publishers).

Grabowski, Henry, W. Kip Viscusi, and William Evans, 1989, Price and Availability Tradeoffs of Automobile Insurance Regulation, *Journal of Risk and Insurance,* 56: 275-299.

Harrington, Scott E., 1987, A Random Coefficient Model of Interstate Differences In the Impact of Rate Regulation on Auto Insurance Prices, *Review of Economics and Statistics,* 69 (February): 167-170.

Keeton, Robert E. and Jeffrey O'Connell, 1965, *Basic Protection for the Traffic Victim* (Boston: Little, Brown).

Kochanowski, Paul S. and Madelyn V. Young, 1985, Deterrent Aspects of No-Fault Automobile Insurance: Some Empirical Findings, *Journal of Risk and Insurance,* 52: 269-288.

Landes, Elisabeth M., 1982, Insurance, Liability, and Accidents: A Theoretical and Empirical Investigation of the Effect of No-Fault Accidents, *Journal of Law and Economics,* 25 (April): 49-65.

Levy, Haim and Marshall Sarnat, 1984, *Portfolio and Investment Selection: Theory and Practice* (Englewood Cliffs, N.J.: Prentice-Hall).

Rand Corporation, 1985, *Automobile Accident Compensation,* 4 Volumes (Santa Monica, Cal.: Rand).

Shavell, Steven, 1987, *Economic Analysis of Accident Law* (Cambridge, Mass.: Harvard University Press).

U.S. Department of Commerce, 1991, *Statistical Abstract of the United States: 1991* (Washington, D.C.: U.S. Government Printing Office).

U.S. Department of Transportation, 1986, *Compensating Auto Accident Victims: A Follow-up Report on No-Fault Auto Insurance Experiences* (Washington, D.C.: U.S. Government Printing Office).

U.S. Federal Highway Administration, 1986, *Highway Statistics* (Washington, D.C.: U.S. Government Printing Office).

Weisberg, Herbert and Richard A. Derrig, 1991a, Fraud and Automobile Insurance: A Report on Bodily Injury Liability Claims in Massachusetts, *Journal of Insurance Regulation,* 9: 497-541.

Weisberg, Herbert and Richard A. Derrig, 1991b, Preliminary Evaluation of Massachusetts Automobile Bodily Injury Tort Reform, Working Paper, Automobile Insurers Bureau of Massachusetts, Boston.

Zador, Paul and Adrian Lund, 1986, Re-Analysis of the Effects of No-Fault Auto Insurance on Fatal Crashes, *Journal of Risk and Insurance,* 53: 226-241.