U.S. Department
of Transportation

**National Highway
Traffic Safety
Administration**

People Saving People

http://www.nhtsa.dot.gov

DOT HS 809 199 **December 2000**
NHTSA Technical Report

Fatality Reduction by Safety Belts for Front-Seat Occupants of Cars and Light Trucks

Updated and Expanded Estimates Based on 1986-99 FARS Data

Technical Report Documentation Page

1. Report No. DOT HS 809 199	2. Government Accession No.	3. Recipient's Catalog No.

4. Title and Subtitle FATALITY REDUCTION BY SAFETY BELTS FOR FRONT-SEAT OCCUPANTS OF CARS AND LIGHT TRUCKS Updated and Expanded Estimates Based on 1986-99 FARS Data	5. Report Date December 2000 6. Performing Organization Code
7. Author(s) Charles J. Kahane, Ph.D.	8. Performing Organization Report No.
9. Performing Organization Name and Address Evaluation Division, Plans and Policy National Highway Traffic Safety Administration Washington, DC 20590	10. Work Unit No. (TRAIS) 11. Contract or Grant No.
12. Sponsoring Agency Name and Address Department of Transportation National Highway Traffic Safety Administration Washington, DC 20590	13. Type of Report and Period Covered NHTSA Technical Report 14. Sponsoring Agency Code

15. Supplementary Notes

16. Abstract

The National Highway Traffic Safety Administration estimated in 1984 that manual 3-point safety belts reduce the fatality risk of front-seat occupants of passenger cars by 45 percent relative to the unrestrained occupant. The agency still relies on that estimate. Shortly after 1985, the prime analysis technique for Fatality Analysis Reporting System (FARS) data, double-pair comparison, began producing inflated, unreliable results. This report develops an empirical tool to adjust double-pair comparison analyses of 1986-99 FARS data. It validates the adjustments by comparing the belt use of fatally injured people in certain types of crashes to belt use observed on the road in State and national surveys. These methods reconfirm the agency's earlier estimates of fatality reduction by manual 3-point belts: 45 percent in passenger cars and 60 percent in light trucks. Furthermore, they open the abundant 1986-99 FARS data to additional analyses, permitting point-estimation of belt effectiveness by crash type, occupant age and gender, belt type, vehicle type, etc.

17. Key Words safety belt; occupant protection; fatal crash; crashworthiness; FARS; fatality reduction; statistical analysis; evaluation; frontal impact; side impact; rollover; ejection; belt use; automatic belt	18. Distribution Statement Document is available to the public through the National Technical Information Service, Springfield, Virginia 22161

19. Security Classif. (Of this report) Unclassified	20. Security Classif. (Of this page) Unclassified	21. No. of Pages 77	22. Price

Form DOT F 1700.7 (8-72) **Reproduction of completed page authorized**

TABLE OF CONTENTS

1. HISTORY OF THE EFFECTIVENESS ESTIMATES

In 1984, the National Highway Traffic Safety Administration (NHTSA) issued its automatic protection requirement for passenger cars. The agency's regulatory impact analysis estimated that manual 3-point safety belts, when used by drivers or right-front passengers of cars, reduce fatality risk by approximately 45 percent relative to the unrestrained occupant[1]. The effectiveness was also stated as an interval estimate: 40 to 50 percent. These numbers became, and still remain[2] the agency's "official" estimates of belt effectiveness in cars. They were a retrenchment from the agency's 1976 estimate of 60 percent[3] and even higher numbers elsewhere in the literature that had been based on relatively simple comparisons of fatality rates per 100 belted and unrestrained occupants.

The 45 percent estimate (or 40-50 percent range) was the agency's consensus and best judgement based on two types of analyses:

- Recognition that people who buckled up were involved in less severe crashes than people who did not use belts (at least in those days). Conscientious efforts to "adjust" or "control" fatality rates per 100 occupants for differences in crash severity produced point estimates of overall effectiveness in the 39-49 percent range[4] (with a substantially wider range if sampling error is included). In other words, a belted occupant was 39-49 percent less likely to die than an unrestrained person in a crash of the same severity.

- A reality check based on 11 countries and Canadian provinces that had enacted belt use laws. In each case, the observed increase in belt use and the actual reduction in occupant fatalities after the law were employed to estimate the implicit belt effectiveness. These estimates varied considerably (because belt laws often coincided with economic up- or down-swings in those volatile times) but averaged to 47 percent[5].

---

[1] *Final Regulatory Impact Analysis, Amendment to Federal Motor Vehicle Safety Standard 208, Passenger Car Front Seat Occupant Protection,* NHTSA Publication No. DOT HS 806 572, Washington, 1984, pp. IV-1 - IV-16.

[2] *Fourth Report to Congress, Effectiveness of Occupant Protection Systems and their Use,* NHTSA Publication No. DOT HS 808 919, Washington, 1999, pp. 11-12.

[3] *Safety Belt Usage, A Review of Effectiveness Studies, Suggestions for State Programs,* NHTSA Publication No. DOT HS 801 988, Washington, 1976, p. 1.

[4] Partyka, Susan C., "Seat Belt Effectiveness Estimates Using Data Adjusted for Damage Type (January 1984)," *Papers on Adult Seat Belts - Effectiveness and Use,* NHTSA Publication No. DOT HS 807 285, Washington, 1988, pp. 1-12. Kahane, Charles J., *Addendum to "Seat Belt Effectiveness Estimates Using Data Adjusted for Damage Type"*, NHTSA Docket No. 74-14-N35-229-05, 1984.

[5] *Final Regulatory Impact Analysis* (1984), pp. IV-14 - IV-15.

These analyses made the previous 60+ percent estimates unrealistic and supported the 40-50 percent range.

Within two years, Leonard Evans published his influential **double-pair comparison** analyses of 1975-83 FARS data, showing a 41 percent fatality reduction by 3-point belts in passenger cars, with 2-sigma confidence bounds ± 8 percent[6]. Double-pair comparison (which will be defined with examples in Section 3) is valuable because it allows the direct use of FARS data that have a much higher N of fatalities than NASS or state files. A second major advantage is that double-pair comparison implicitly "adjusts" or "controls" for the differences in the severity of crashes involving belted and unrestrained occupants. The 41 percent effectiveness estimate was within the agency's 40-50 percent range and, together with the two preceding analyses provided a strong foundation for the agency's position.

Analysts at NHTSA and elsewhere quickly adopted double-pair comparison for analyzing belts and other safety devices. However, Susan Partyka and others soon noted that belt effectiveness estimates rose substantially as more recent FARS data were fed into the analyses. For example, analyses of 1982-87 FARS data produced a belt effectiveness estimate of 55 percent for passenger cars[7]. After perhaps a little wishful thinking that earlier estimates might have been low by chance alone, or even that belts might have become more effective, NHTSA staff soon concluded that something had gone wrong with belt use reporting on FARS (and other files) and had biased effectiveness estimates upwards[8].

Specifically, New York was the first state to enact a belt use law, effective December 1, 1984. After a brief "wait and see," 21 states, including 9 of the 10 most populous states had belt laws effective by August 1986 for front-seat occupants of passenger cars[9]. For the first time, unbelted people had a tangible incentive - avoidance of a fine - to report that they were belted. NHTSA hypothesized that:

- Uninjured or slightly injured occupants are often up and about before police arrive at the crash scene. Since the investigating officer is not an eye-witness to their belt use, they have an opportunity, and now also a motive, to say they wore belts, even if they hadn't.

---

[6] Evans, Leonard, "Double Pair Comparison - A New Method to Determine How Occupant Characteristics Affect Fatality Risk in Traffic Crashes," *Accident Analysis and Prevention,* Vol. 18, June 1986, pp. 217-227. Evans, Leonard, "The Effectiveness of Safety Belts in Preventing Fatalities," *Accident Analysis and Prevention*, Vol. 18, June 1986, pp. 229-241.

[7] Partyka, Susan C., "Belt Effectiveness in Pickup Trucks and Passenger Cars by Crash Direction and Accident Year (May 1988)," *Papers on Adult Seat Belts - Effectiveness and Use,* NHTSA Publication No. DOT HS 807 285, Washington, 1988, pp. 99-102.

[8] *Ibid*., p. 99 and p. 102.

[9] *Traffic Safety Facts 1998,* NHTSA Publication No. DOT HS 808 983, Washington, 1999, p. 186.

- Mortally injured occupants may be in their original post-crash location when police arrive, often allowing direct observation of belt use.

Thus, NHTSA believes belt use of fatalities is reported without net biases on FARS before and after belt laws[10]. However, after the laws, belt use of survivors is overreported. A bias has apparently been introduced in the reporting of this one data element, for survivors, as a consequence of belt use laws. It has occurred despite the long-term, ongoing efforts by NHTSA and the states in data quality control and analyst training, which have resulted in more accurate, complete and consistent information on most FARS data elements.

When survivors who were actually unrestrained are reported as belted, it lowers the fatality odds in the "belted" population, raises the odds in the "unrestrained" population, and bloats the effectiveness estimate. The following hypothetical example shows how. For simplicity, it is based on fatality rates per 100 crash-involved occupants, as might be derived from state crash files. However, the same type of bias would occur in a double-pair comparison analysis.

First, if belt use had been accurately reported by everyone, there would have been a population of 100 unrestrained and 100 belted occupants, with a fatality rate 45 percent lower for the belted occupants than for the unrestrained:

Based on Actual Belt Use

	Not Belted	Belted	Fatality Reduction
Fatalities	20	11	
Survivors	80	89	
Total	100	100	
Fatality rate	.20	.11	45%

The "fatality reduction," 45 percent = 1 - (.11/.20) = 1 - (belted fatality rate/unbelted fatality rate). If 15 of the unrestrained survivors misreported themselves as "belted" (while all the fatalities continue to be correctly reported), the fatality rate for reportedly "unrestrained" people increases and the "belted" fatality rate decreases:

---

[10] Partyka, Susan C., *Lives Saved by Seat Belts from 1983 through 1987*, NHTSA Publication No. DOT HS 807 324, Washington, 1988.

Based on Reported "Belt Use"

	"Not Belted"	"Belted"	Fatality Reduction
Fatalities	20	11	
Survivors	80 - 15 = 65	89 + 15 = 104	
Total	100 - 15 = 85	100 + 15 = 115	
Fatality rate	.24	.10	59%

The fatality reduction is inflated from a true 45 percent to an observed 59 percent.

The agency reached a decision point in 1989 when it extended automatic protection to light trucks. For the regulatory impact analysis, the agency needed to estimate the effectiveness of manual belts. Partyka's double-pair comparison (based on FARS data through 1987) showed a 69 percent fatality reduction for belts in pickup trucks[11]. The regulatory impact analysis asserted that this result was inflated. Since Partyka's data showed a 55 percent reduction in passenger cars, whereas the agency believed 45 percent was the true reduction, the light truck estimate ought to be scaled back by a similar amount: from 69 to 60 percent[12]. In the process, the agency:

- Reconfirmed the 45 percent estimate for cars and established a 60 percent estimate for light trucks.
- Asserted that FARS analyses producing estimates higher than those were biased and ought not be accepted at face value.

2. GOALS OF THIS REPORT

Eleven years later, as of December 2000, the agency continues to rely on 45 and 60 percent estimates that are essentially based on 1975-85 data and 1975-85 vehicles. Abundant later FARS data, with much higher N's of belted fatality cases, remain untapped. The numbers could have become outdated as belt systems, vehicles and the crash environment changed. The old data do not allow estimates of post-1985 belt configurations, such as automatic belts or belts in vehicles with air bags. The old data are too small a sample for accurate estimation of belt effectiveness in

---

[11] Partyka, "Belt Effectiveness in Pickup Trucks and Passenger Cars by Crash Direction and Accident Year," *op. cit.*

[12] *Preliminary Regulatory Impact Analysis, Proposed Extension of the Automatic Restraint Requirements of FMVSS 208 to Trucks, Buses and Multi-Purpose Passenger Vehicles with a Gross Vehicle Weight Rating of 8,500 Pounds or Less and an Unloaded Vehicle Weight of 5,000 Pounds or Less*, NHTSA Docket No. 74-14-N62-001, 1989, p. 15.

important subgroups of crashes, such as specific crash types, occupant age groups, vehicle types, etc.

The objectives of this paper are:

- To develop an empirical tool to adjust for the biases in double-pair comparison analyses of later FARS data, and open up those FARS files for point estimates of fatality reduction consistent with pre-1986 results.

- To obtain detailed point estimates of belt effectiveness by crash mode, occupant age group, etc., needed for NHTSA regulatory analyses and evaluations, and not really available from the limited pre-1986 data.

- To obtain point estimates of belt effectiveness for configurations that did not exist before 1986, such as automatic belts, or manual belts in vehicles with dual air bags. These estimates, too, are needed for regulatory analyses and evaluations.

- To see if belt effectiveness has changed in the newer vehicles, or has changed over time in response to an evolving crash environment.

- To see if NHTSA's long-standing estimates of 45 percent fatality reduction in cars and 60 percent in light trucks are still appropriate.

However, the point estimates of this report, relying on several critical assumptions, are not like customary statistical estimates derived directly from the data. The uncertainty in our estimates, although it can be discussed to some extent, cannot be fully quantified, based on statistical theory, as "confidence bounds."

3. "CLASSIC" DOUBLE-PAIR COMPARISON: PASSENGER CARS IN CY 1977-85

Evans, Partyka and others have provided detailed examples of double-pair comparison analyses in the literature, but let us run through one case here from start to finish, both as a review and to demonstrate the specific estimation procedure used in this report.

The starting point for this analysis is FARS data for CY 1977-85. Records of passenger cars of model years 1975-86 are extracted (1975 is the first model year with "Type 2" 3-point belt systems, not counting 1974 where cars were also equipped with the ignition interlock). "Passenger cars" in MY 1975-80 are all FARS vehicle records with the variable BODY_TYP = 1-9, and in MY 1981-86 are the cases with decodable VINs that are passenger cars according to the VIN decode program developed for NHTSA evaluations[13]. The analysis is limited to:

---

[13] Kahane, Charles J., *Evaluation of FMVSS 214 - Side Impact Protection: Dynamic Performance*, NHTSA Publication No. DOT HS 809 004, Washington, 1999, pp. 15-16.

- Cars with a driver and a right front (RF) passenger (and perhaps other passengers). When two or more people occupy the same seat, according to FARS, only the occupant with the lowest PER_NO (person number) is included.

- The driver, or the RF passenger, or both were fatally injured.

- The driver and the RF passenger both have known reported belt use: MAN_REST has to be 0 (unrestrained) or 1, 2, 3, 8 or 13 (belted, perhaps incorrectly).

- The driver and the RF passenger are both 14 to 97 years old.

There are 30,665 cars in CY 1977-85 with a driver and a RF passenger, at least one fatal, both with known belt use and 14-97 years old. The vehicle cases tabulate as follows, based on each occupant's belt use and survival:

Vehicles	Driver Died RF Survived	Driver Survived RF Died	Both Died
Both unrestrained	11,186	11,469	5,317
Driver unrestrained, RF belted	300	152	74
Driver belted, RF unrestrained	186	487	102
Both belted	497	653	242

This can be tabulated as fatality rather than vehicle cases, by adding the "both died" column to each of the preceding columns:

Fatalities	Driver Fatalities	RF Fatalities	Driver/RF Risk Ratio
Both unrestrained	16,503	16,786	0.983
Driver unrestrained, RF belted	374	226	1.655
Driver belted, RF unrestrained	288	589	0.489
Both belted	739	895	0.826

In CY 1977-85, it is clear that (1) the overwhelming majority of people killed in crashes were unrestrained; (2) unrestrained drivers and RF passengers are at nearly equal risk in the same crash; and (3) whoever buckled up substantially reduced their risk.

The four rows of data allow a total of four double-pair comparisons, two for computing the effectiveness of belts for drivers, and two for RF passengers. The first comparison for the driver is based on the first and third rows of data:

		Driver Fatalities	RF Fatalities	Driver/RF Risk Ratio
Driver unrestrained	RF unrestrained	16,503	16,786	0.983
Driver belted	RF unrestrained	288	589	0.489

In both pairs, the driver's fatality risk is compared to the same control group: the unrestrained RF passenger. The unrestrained driver has essentially the same fatality risk as the unrestrained RF in the same crash, the belted driver about half. The fatality reduction for belts is

$$1 - (0.489/0.983) = 50.3 \text{ percent.}$$

The other comparison for the driver is based on the second and fourth rows of data:

		Driver Fatalities	RF Fatalities	Driver/RF Risk Ratio
Driver unrestrained	RF belted	374	226	1.655
Driver belted	RF belted	739	895	0.826

Here, the control group is the belted RF passenger. The unrestrained driver has higher fatality risk than the belted RF in the same crash, the belted driver, lower. The fatality reduction is:

$$1 - (0.826/1.655) = 50.1 \text{ percent.}$$

It is important that the effectiveness estimates are nearly identical with the two control groups: it suggests the estimates are robust and not affected by the choice of control group.

The first double-pair comparison for estimating belt effectiveness for the RF passenger is obtained by using the first two rows of data, reversing the order of the columns and computing the RF/Driver rather than the Driver/RF risk ratio:

		RF Fatalities	Driver Fatalities	RF/Driver Risk Ratio
RF unrestrained	Driver unrestrained	16,786	16,503	1.017
RF belted	Driver unrestrained	226	374	0.604

The control group is the unrestrained driver. The fatality reduction for the belted RF passenger is:

1 - (0.604/1.017) = 40.6 percent.

The second estimate uses the last two rows of data:

		RF Fatalities	Driver Fatalities	RF/Driver Risk Ratio
RF unrestrained	Driver belted	589	288	2.045
RF belted	Driver belted	895	739	1.211

The control group is the belted driver. The fatality reduction for the belted RF passenger is:

1 - (1.211/2.045) = 40.8 percent.

Again, the two control groups produce nearly identical estimates. Also, as in earlier studies, belt effectiveness is lower for the RF passenger than for the driver.

The next task is to develop a weighting procedure that combines the two driver estimates into a single number, and likewise for the two RF estimates.

In the 1977-85 FARS data, the actual number of driver fatalities is

Actual driver fatalities = 16,503 + 374 + 288 + 739 = 17,904

The first two numbers in that sum are unrestrained drivers, the last two, belted. However, if every driver had been unrestrained, that sum would have increased to

All-unrestrained driver fatalities = 16,503 + 374 + (0.983 x 589) + (1.655 x 895) = 18,937

(Here, 589 was the number of unrestrained RF fatalities that accompanied the 288 belted drivers and 0.983 is the risk ratio of unrestrained driver to unrestrained RF fatalities; 895 is the number of belted RF fatalities that accompanied the 739 belted drivers and 1.655 is the risk ratio of unrestrained drivers to belted RF fatalities.)

On the other hand, if every driver had buckled up, the sum would have dropped to

All-belted driver fatalities = (0.489 x 16,786) + (0.826 x 226) + 288 + 739 = 9,421

The overall effectiveness of belts for drivers is

(18,937 - 9,421) / 18,937 = 50.25 percent, which is between the results of the two separate double-pair comparisons for drivers (50.1 and 50.3 percent).

Similarly, the actual number of RF passenger fatalities is

Actual RF fatalities = 16,786 + 226 + 589 + 895 = 18,496

If every RF passenger had been unrestrained, that sum would have increased to

All-unrestrained RF fatalities = 16,786 + (1.017 x 374) + 589 + (2.045 x 739) = 19,267

(Here, 374 was the number of unrestrained driver fatalities that accompanied the 226 belted RF passengers and 1.017 is the risk ratio of unrestrained RF to unrestrained driver fatalities; 739 is the number of belted driver fatalities that accompanied the 895 belted RF and 2.045 is the risk ratio of unrestrained RF to belted driver fatalities.)

But if every RF passenger had buckled up, the sum would have dropped to

All-belted RF fatalities = (0.604 x 16,503) + 226 + (1.211 x 288) + 895 = 11,442

The overall effectiveness of belts for RF passengers is

(19,267 - 11,442) / 19,267 = 40.61 percent,

which is between the results of the two separate double-pair comparisons for RF passengers (40.6 and 40.8 percent).

Finally, for an estimate of the overall effectiveness of 3-point belts for front-outboard occupants of passenger cars, we must note that drivers have over the years typically outnumbered RF passengers by very close to 3 to 1 in the general crash-involved population (as opposed to these special cases that were limited to cars with the RF seat occupied). The preceding statistics for drivers need to be weighted by 3 and the statistics for RF passengers by 1. If all drivers and RF passengers were unrestrained, that sum would have increased to

All-unrestrained front-outboard fatalities = (3 x 18,937) + 19,267 = 76,078

If they had all buckled up, the sum would have dropped to

All-belted front-outboard fatalities = (3 x 9,421) + 11,442 = 39,706

The overall effectiveness of 3-point belts for front-outboard occupants is

(76,078 - 39,706) / 76,078 = 47.81 percent, which is between the estimates for drivers and RF passengers, but closer to the driver estimate, as it should be, given the higher weight factor for drivers.

4. INFLATED RESULTS FOR PASSENGER CARS IN CY 1986-99

Let us repeat the double-pair comparison analysis for passenger cars equipped with 3-point belts, but using more recent FARS data, specifically 1986-99.

Records of passenger cars of model years 1975-99 equipped with 3-point belts are extracted from 1986-99 FARS files. "Passenger cars" in MY 1975-80 and 1999 are all FARS vehicle records with BODY_TYP 1-9, and in MY 1981-98 are the cases with VINs that decode as passenger cars. "Three-point belts" include manual or automatic (door-mounted) 3-point belts, in cars with no air bags or dual air bags. Cars with only a driver air bag are excluded to preserve the symmetry (nearly equal fatality risk) of the driver and the RF positions in the analysis. As in Section 3, the analysis is limited to cars with a driver and a right front (RF) passenger, both with known reported belt use, both age 14-97, at least one and perhaps both fatally injured. There are 70,668 cars in the 1986-99 files meeting those criteria. The basic tabulation of fatalities is:

Fatalities	Driver Fatalities	RF Fatalities	Driver/RF Risk Ratio
Both unrestrained	23,476	23,579	0.996
Driver unrestrained, RF belted	3,934	1,622	2.425
Driver belted, RF unrestrained	1,815	4,820	0.377
Both belted	11,225	12,901	0.870

Relative to the CY 1977-85 data in Section 3, (1) the number of cases in the cells with belted drivers and/or passengers is an order of magnitude larger - there are a lot more data to work with here; (2) the effect of belts appears far more dramatic at first glance - the ratio of unrestrained driver to belted RF fatalities increased from 1.655 to 2.425 while the ratio of belted driver to unrestrained RF decreased from 0.489 to 0.377. Working through the double-pair comparisons and weighted averages as in Section 3 produces fatality reduction estimates of

63.26 percent for drivers
57.71 percent for RF passengers
61.89 percent for all front-outboard occupants

These are substantially higher than the corresponding reductions in 1977-85: 50 percent for drivers, 41 percent for RF passengers and 48 percent combined. They raise three questions:

- When, and how quickly did the observed effectiveness escalate?

- Could a substantial part of the increase be due to real improvements in the life-saving effectiveness of belts in later-model cars?

- Could a substantial part of the increase be due to changes in the crash environment that have increased the types of crashes where belts are most effective?

When a separate double-pair comparison analysis is run on each individual calendar year of FARS data, the observed overall fatality reductions for belts are the following:

1977	49 percent	1986	61	1993	60
1978	28	1987	58	1994	64
1979	44	1988	61	1995	63
1980	38	1989	63	1996	65
1981	52	1990	69	1997	58
1982	53	1991	62	1998	62
1983	38	1992	60	1999	59
1984	46				
1985	55				

The effectiveness results are also graphed in Figure 1.

During 1977-84, observed belt effectiveness varies a fair amount from year to year, due to the small N's of belted cases on FARS, but arguably centers on about 45 percent with little or no time trend. In 1986, the first year with belt use laws covering a large proportion of occupants (including 9 of the 10 most populous states) the fatality reduction has already reached 61 percent, essentially the 1986-99 average, and it stayed close to that year after year, with no evidence of any time trend within 1986-99. The year 1985 is hard to place: the 55 percent is higher than any preceding year, but just barely higher, for example, than the 53 percent in 1982. It is lower than any subsequent year, although not much lower than the 58 percent in 1987 and 1997. Since belt use laws were just getting started in a few states in 1985, but were well established in 1986[14], it seems most appropriate to include the 1985 data with the "pre-law" period[15].

In any case, the escalation in the belt effectiveness estimate obviously coincided with the inception of belt use laws. The escalation came all at once, with little subsequent change. That not only answers the first question (when and how quickly) but essentially the other two. If any substantial

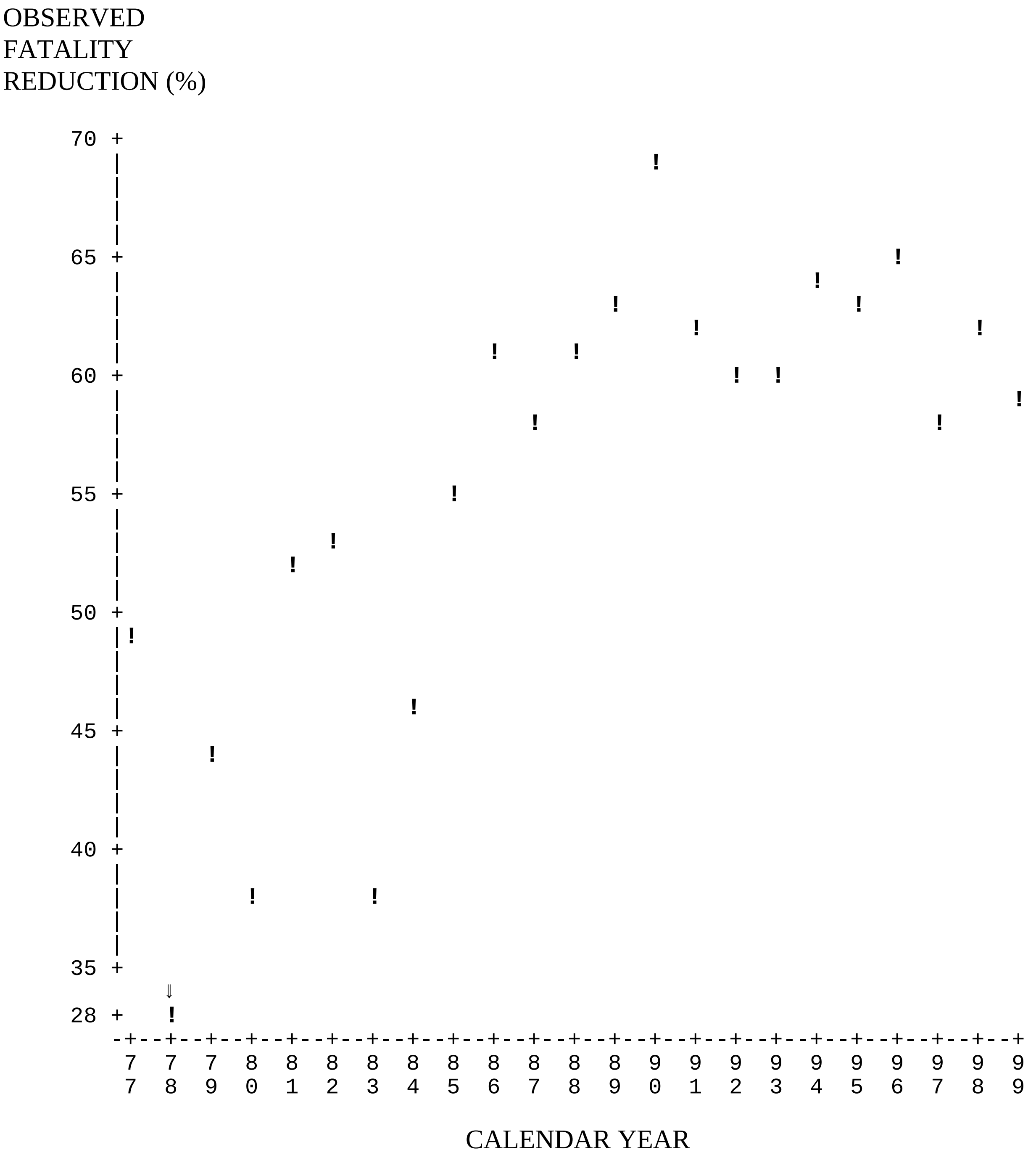
---

[14] *Traffic Safety Facts 1998, op. cit.*, p. 186.

[15] A possible alternative approach would be to limit the "pre-law" period to 1977-84, retain the "post-law" period as 1986-99, and to exclude the 1985 data from the analysis. The observed fatality reduction for 3-point belts in passenger cars is 44.69 percent in 1977-84, as compared to 47.81 percent in 1977-85. This would have raised the UEF, as defined in Section 5, from 1.369 to 1.452.

FIGURE 1

OBSERVED EFFECTIVENESS OF 3-POINT BELTS IN CARS BY CALENDAR YEAR



part of the escalation had been due to genuine improvements in belts, that part would have been gradual, since late-model cars with the improved belts only gradually replace the older cars in the overall vehicle population. The impact of changes in the crash environment would also have been gradual, not abrupt. It is most plausible to conclude that the escalation beginning in 1986 is in fact due to belt use laws resulting in overreporting of belt use by crash survivors, and that ordinary double-pair comparison analysis stopped producing accurate estimates of belt effectiveness in 1986.

A comparison of observed effectiveness by calendar year and model year provides additional evidence that the escalation is due to changes in the data rather than changes in the vehicles:

Observed Fatality Reductions for 3-point Belts in Passenger Cars

	Effect in 1977-85 FARS	Effect in 1986-99 FARS
Manual belts in MY 1975-79 cars	48	63
Manual belts in MY 1980-85 cars	47	63
Manual belts in MY 1986-90 cars		60
Automatic 3-point belts (MY 1987-95)		64
Manual belts in cars with dual air bags (MY 1987-99)		63

In the 1977-85 FARS data, belt effectiveness is nearly the same for MY 1975-79 and MY 1980-85 cars. In the 1986-99 FARS data, belt effectiveness is essentially the same in all cohorts of vehicle model years (and higher than the corresponding numbers for 1977-85 FARS).

5. THE “UNIVERSAL EXAGGERATION FACTOR” (UEF) AND ITS ROBUSTNESS

The two basic results so far are that 3-point belts reduced fatality risk in passenger cars by 47.81 percent in 1977-85 FARS data and were observed to “reduce” fatality risk by 61.89 percent in 1986-99 FARS data. The hypothesis is that the first estimate (give or take some “fine tuning” described in Section 6, and sampling error) is an unbiased estimate of the genuine fatality reduction for belts, whereas the second is biased upwards by inaccurate belt use reporting of survivors in FARS in response to belt use laws.

Let us define the “Universal Exaggeration Factor” (UEF) to be the relative difference of the two estimates:

$$\text{UEF} = (100 - 47.81) / (100 - 61.89) = 1.369$$

It is the adjustment factor that has to be applied to the inappropriately low 1986-99 ratio of "belted" to "unbelted" fatality risk to obtain the accurate 1975-85 ratio of actual belted to unbelted fatality risk:

$$1.369 \times (100 - 61.89) = 100 - 47.81$$

$$47.81 = 100 - [1.369 \times (100 - 61.89)]$$

This UEF derives from the effectiveness estimates for 3-point belts in passenger cars, in all types of crashes, based on direct double-pair comparison analyses[16]. Our hypothesis is that this same UEF = 1.369 is also **empirically valid for other double-pair comparison analyses based on 1986-99 FARS data**, including other types of vehicles or belts, subgroups of crashes or occupants, and more complex weighted averages of double-pair comparisons. In other words, if the analysis of 1986-99 data yields an effectiveness estimate E*, the true effectiveness E is close to

$$E = 100 - [1.369 \times (100 - E^*)]$$

If this working hypothesis is acceptable, it would greatly increase the utility of double-pair comparison analysis. Since the 1986-99 FARS data contain an order of magnitude more belted cases than the 1977-85 data, we will have enough data to obtain effectiveness estimates for specific subgroups of interest (crash modes, occupant age groups, etc.). We will also be able to obtain effectiveness estimates for current vehicle types that did not exist in 1977-85 (e.g., belt effectiveness in vehicles with dual air bags). The UEF should be viewed as an empirical tool for generating needed point estimates and not as a statistical method that will produce confidence bounds for those estimates.

The hypothesis can be tested (not in a formal, statistical sense) by performing double-pair comparison analyses for selected subsets of the 1977-85 data and the corresponding subsets of the 1986-99 data. For each of the subsets, we will compute the exaggeration factor EF of the 1986 - 99 effectiveness over the 1977-85 estimate, and compare it to the UEF = 1.369. We can accept the hypothesis that 1.369 is a "universal" exaggeration factor if the EF's for the various subsets are all "relatively" close to 1.369, but not if they differ "a lot" from subset to subset. One factor that complicates the testing is that the EF's themselves, including the UEF, are statistics and subject to sampling error. Specifically, the 1977-85 FARS data contain relatively few belted cases. When the data are subdivided the 1977-85 effectiveness estimates become quite imprecise, and so will the EF's. After all, if the 1977-85 data were adequate for precise estimation of belt effectiveness in small subsets, it would take away a prime motivation for analyzing the 1986-99 data.

---

[16] A possible alternative approach, since the CY 1977-85 data contain almost exclusively MY 1975-85 cars [with just a few early MY 1986 cars] would be to limit the analysis of 1986-99 to MY 1975-85 cars, too. The observed fatality reduction for 3-point belts in MY 1975-85 passenger cars is 62.90 percent in CY 1986-99, as compared to 61.89 percent for all MY 1975-2000 cars with 3-point belts in CY 1986-99. This would have raised the UEF from 1.369 to 1.407.

The first step in testing the UEF is to measure its variation in **randomly** selected subsets of the data. For example, FARS data can be split into 10 systematic random subsamples of equal size based on the last digit of the case identification number ST_CASE. If separate double-pair comparison analyses are performed on the 10 subsamples of 1977-85 and 1986-99 data, the effectiveness estimates and UEF are:

ST_CASE Ending in	Effect in 1977-85 FARS	Effect in 1986-99 FARS	UEF
0	44.95	64.00	1.5290
1	52.10	62.83	1.2888
2	66.22	60.93	0.8646
3	47.05	60.59	1.3436
4	48.25	59.35	1.2730
5	44.15	64.66	1.5804
6	51.15	62.12	1.2898
7	24.40	58.97	1.8424
8	41.44	62.01	1.5414
9	52.81	63.54	1.2944
1 Std. Dev.	10.55	1.93	0.2591

Belt effectiveness in the 1977-85 data varies considerably across these 10 subsamples, ranging from 24.4 to 66.2 percent, with a standard deviation of 10.55. But in the 1986-99 data, with much larger N's of belted occupants, the inflated effectiveness estimate only ranges from 59.0 to 64.7 percent, with a standard deviation of 1.93. The sampling error in the 1977-85 effect drives the error in the UEF, which ranges from 0.86 to 1.84, with a standard deviation of 0.26.

These are the variations in the UEF from subset to subset, due to sampling error alone, even when the data are randomly split into subsets, and it is the minimum variation that can be expected. If the data were instead split into 10 subsets of approximately equal size based on specific criteria (e.g., occupant age, crash mode), we can accept the "universality" of the UEF if its variation were just moderately larger than the random case, and we would reject it if it were much larger.

The variations of the 1977-85 effect, the 1986-99 effect and the UEF can also be computed, as above, when the data are split into more than 10 systematic random subsets, based on ST_CASE:

n of Subsets	ó for 1977-85 Effect	ó for 1986-99 Effect	ó for UEF	ó / √n
10	10.55	1.93	0.2591	0.0819
15	9.89	3.58	0.3701	0.0955
20	11.40	2.67	0.2702	0.0604
25	11.81	2.79	0.3401	0.0680
30	15.75	4.10	0.4961	0.0906
35	17.30	4.85	0.4966	0.0840

As the number of subsets increases from 10 to 35 and the sizes of the subsets decrease, the random variation of UEF also generally increases. (With 40 subsets based on ST_CASE, one of them has no belted cases in CY 1977-85, making it impossible to calculate the UEF.)

The "standard error" of the subsample variation, ó / √n, remains fairly constant. The arithmetic average of the above six readings of ó / √n, 0.080 is a good estimate of the standard deviation of the UEF for **the entire data set**. Its coefficient of variation (CV) is 0.080 / 1.369 = 5.8 percent.

Now, let us look at the variation of the exaggeration factors across various key subsets of the crash population.

We might expect the 1986-99 effectiveness estimates and the exaggeration factor to vary a lot from **state to state**. Starting in 1986, the public were perhaps more likely to overreport their belt use in states where penalties were higher and more frequently enforced. Police might be especially skeptical in some states of belt use initially reported by the driver and instructed to follow up with additional questions and investigations. Thus, we might expect one group of states with 1986-99 effectiveness close to the true 45 percent, another group with exaggerated fatality reductions near the nationwide 62 percent average, and yet another group with still more exaggerated estimates ranging above 70, 80 or even 90 percent. Here are the 1977-85 and 1986-99 effectiveness estimates and exaggeration factors for the 20 states with the most fatalities of passenger cars occupants in 1977-99 (in order of decreasing N of fatalities):

State	Percent of U.S. Fatals	Effect in 1977-85 FARS	Effect in 1986-99 FARS	UEF
California	8.13	67.21	68.17	1.03
Texas	7.34	42.05	60.43	1.46
Florida	6.00	45.15	58.44	1.32
New York	4.51	52.01	66.10	1.42
Pennsylvania	4.33	26.34	62.85	1.98
Illinois	4.19	49.35	66.71	1.52
Michigan	4.16	43.99	60.27	1.41
Ohio	4.13	54.38	58.70	1.10
North Carolina	3.63	36.66	62.95	1.71
Georgia	3.59	36.22	62.06	1.68
Tennessee	2.93	30.39	65.29	2.01
Alabama	2.75	31.10	55.40	1.54
Missouri	2.68	63.15	67.21	1.12
Indiana	2.49	- 59.74	55.39	3.58
Virginia	2.28	65.38	61.17	0.89
New Jersey	2.25	46.12	63.07	1.46
South Carolina	2.19	- 2.09	44.79	1.85
Louisiana	2.10	30.84	60.48	1.75
Kentucky	2.10	34.32	66.09	1.94
Wisconsin	2.05	35.52	65.22	1.85
1 Std. Dev.		27.55	5.41	0.56

The 1986-99 effectiveness estimates are amazingly consistent from state to state, contrary to what we might have expected. Only South Carolina's is in the 45 percent range - and that could be due to chance alone, since it is based on a relatively small sample. Not a single estimate exceeds 70 percent. Nineteen states range from 55 to 69 percent, suggesting that the tendency of 1986-99 FARS data to produce exaggerated estimates is rather universal and rather equal across the United States. The standard deviation of the 1986-99 estimate for the 20 states is 5.41. The median state in this list has about 3 percent, i.e., 1/33 of the nation's passenger car occupant fatalities. In the preceding table, the standard deviation for 30 **random** subsets was 4.10, and for 35 random subsets, 4.85. Thus, the state-to-state variation (5.41) is only slightly larger than for random subsets.

The 1977-85 effectiveness estimate varies greatly from state to state (ó = 27.55), because some states had only a handful of belted fatality cases in the early years, resulting in large sampling errors. Nevertheless, the standard deviation of the UEF is only 0.56, and this is just slightly larger than the standard deviations of the UEF for 30 or 35 random subsamples (both 0.50 in the preceding table). The UEF does not vary substantially more from state to state than in random subsamples of comparable sizes. (One or two extreme outliers, such as the Indiana data in the preceding table, could also be expected even in random subsamples of comparable sizes.)

We might expect the UEF to vary considerably depending on the driver's behavior. Antisocial behavior such as driving under the influence of alcohol or drugs, driving without a valid license, a history of violations or crashes, reckless driving, attempting to escape police, hit-and-run, or racing (for exact definitions, see Section 9) might be associated with misreported belt use, raising the UEF. Or, conversely, it might spur the police to be extra skeptical about reported belt use, reducing the UEF. Neither of these possibilities can be seen in the actual UEF's:

	Effect in 1977-85 FARS	Effect in 1986-99 FARS	UEF
Drinking or other antisocial behavior	53.97	66.08	1.36
No antisocial behavior	42.98	58.80	1.38

Belt effectiveness is higher for drivers with antisocial behavior than for law-abiding drivers, because many antisocial drivers are involved in rollover crashes and frontal impacts with fixed objects, where belts are especially effective. However, this is true in 1986-99 just as in 1977-85. The UEF's for antisocial and law-abiding drivers are virtually identical.

The UEF might be higher in single-vehicle crashes, which often have no witnesses, than in multivehicle crashes, witnessed by occupants of the other vehicle(s). In fact, the observed UEF is just slightly lower:

	Effect in 1977-85 FARS	Effect in 1986-99 FARS	UEF
Single-vehicle crashes	63.77	71.32	1.26
Multivehicle crashes	35.29	51.90	1.35

The UEF could vary by crash mode, for two reasons. One is that the crash modes themselves indicate different types of driver behavior (frontal or rollover = aggressive driver, side impact = nonaggressive driver). The second is that effectiveness is much higher in some crash modes (rollovers) than others (side impacts), and the UEF might be confounded with the magnitude of the effectiveness, as a mathematical artifact. Again, the UEF only varies to a modest extent:

	Effect in 1977-85 FARS	Effect in 1986-99 FARS	UEF
All frontal impacts	43.31	63.52	1.55
Frontals with another car	47.54	61.92	1.38
All side impacts	34.46	47.96	1.26
Side impacts by another car	26.49	47.63	1.40
All rollovers	75.31	82.07	1.38

Finally, the UEF could differ for younger and older occupants, or for male and female occupants, if age or gender had any association with the accuracy of belt use reporting. In fact, the UEF's show some variation, but no obvious trend in one direction or the other:

		Effect in 1977-85 FARS	Effect in 1986-99 FARS	UEF
Driver ≤ 30	RF ≤ 30	55.37	63.85	1.23
Driver ≤ 30	RF ≥ 31	38.65	65.15	1.76
Driver ≥ 31	RF ≤ 30	35.10	58.98	1.58
Driver ≥ 31	RF ≥ 31	40.24	56.00	1.36
Male driver	Male RF	56.28	64.09	1.22
Male driver	Female RF	36.18	56.43	1.46
Female driver	Male RF	56.23	65.81	1.28
Female driver	Female RF	50.51	63.65	1.36

These analyses demonstrate that the UEF = 1.369 is quite robust across crash modes, driver demographics and driver behaviors. They encourage extensive double-pair comparisons for 1986-99 FARS data, each time correcting the observed result by the UEF. This procedure will be applied first to situations where effectiveness could also have been estimated directly from 1977-85 data alone: large subgroups of crashes of passenger cars with manual 3-point belts. It will then be applied where 1977-85 data are available, but in samples too small for statistically meaningful results: smaller subgroups of crashes of cars with 3-point belts, and all subgroups of light truck crashes. Finally, it will be applied even where no 1977-85 data exist: automatic belts, vehicles equipped with air bags.

Although the UEF opens up the 1986-99 data for many analyses, it has an element of "assuming what we are trying to prove": it assumes the overall 1977-85 result is accurate, and that all 1986-99 results should be adjusted down to 1977-85 levels. For full confidence in the 1986-99 findings, they should be validated by another analysis method not dependent on the UEF. That method is described in Section 9, where belt effectiveness is inferred by comparing the belt use of fatally injured people in FARS to belt use observed on the road in state and national surveys.

6. A REFINED EFFECTIVENESS ESTIMATE - CONTROLLING FOR CRASH MODE6

Double-pair comparison analysis has been criticized because the data are limited to vehicles occupied by a driver and a RF passenger. Unaccompanied drivers might have different types of crashes, and different belt effectiveness than accompanied drivers. The fatality reductions based on a single application of double-pair comparison, as in the preceding sections, might not be accurate for the entire occupant population including unaccompanied drivers.

A procedure to mitigate this possible source of bias involves separate double-pair comparison analyses of 1986-99 data in 8 crash modes, resulting in 16 individual estimates: 8 for the driver and 8 for the RF passenger. A weighted average of these estimates is calculated, weighted by the number of cases in the 16 cells in the entire 1977-99 population of unrestrained front-outboard fatalities including unaccompanied drivers, corrected by the UEF.

This procedure yields a "best" estimate of 45.02 percent overall fatality reduction for 3-point belts in passenger cars, slightly lower than the 47.81 percent from the simple double-pair comparisons of Sections 3 and 4. The procedure works like this:

Five crash modes are defined as follows:

Frontal impacts	IMPACT2 = 1,11,12 and HARM_EV ≠ 1 - 6
Left side impacts	IMPACT2 = 8,9,10 and HARM_EV ≠ 1 - 6
Right side impacts	IMPACT2 = 2,3,4 and HARM_EV ≠ 1 - 6
Primary rollovers	IMPACT2 = 13 or HARM_EV = 1
Rear or other	IMPACT2 = 5,6,7 or HARM_EV = 2 - 6 or (HARM_EV = 7 and not one of the above)

The frontal, left-side and right-side impacts are further subdivided into single- and multivehicle crashes, producing a total of eight crash modes. For passenger cars, the "rear or other" crash mode consists of 64 percent rear impacts by another vehicle, 1 percent rear impacts by trains, 23 percent skidding rear-first into fixed objects, 6 percent immersions, 2 percent falling out of moving vehicles and 4 percent fires and other noncollisions. (For light trucks, they are 55% rear impacts by vehicles, 1% rear impacts by trains, 17% rear-first into fixed objects, 6% immersions, a substantial 15% falling from moving vehicles and 6% fires and other noncollisions.)

Next, double-pair comparisons are used to compute belt effectiveness for drivers and RF passengers in each of the eight crash modes, exactly as in Section 3 - i.e., in 1986-99 FARS, for MY 1975-99 passenger cars equipped with 3-point belts (manual belts and no air bags, or manual belts and dual air bags, or 3-point automatic belts and no air bags), occupied by a driver and a RF passenger both age 14 to 97 (and possibly other occupants). The 16 effectiveness estimates, **not corrected with the UEF**, are:

Crash Mode		Observed, Uncorrected Belt Effectiveness (%) Drivers	RF Passengers
Frontal	Single-vehicle	71.92	66.29
Frontal	Multivehicle	58.60	55.06
Left side	Single-vehicle	40.10	58.88
Left side	Multivehicle	35.90	45.77
Right side	Single-vehicle	60.60	48.14
Right side	Multivehicle	54.06	20.09
Rollover (primary)		81.52	79.78
Rear & other		68.86	64.53

Next, the 214,560 cases of fatally-injured, **unrestrained** front-outboard occupants of MY 1975-99 passenger cars in 1977-99 FARS, age 14-97, are tabulated by crash mode and seat position (all unbelted front-outboard occupants are included, regardless of the type of belt system and/or air bags installed in the car, and regardless of how many occupants were in the vehicle):

Crash Mode		Actual unrestrained fatalities in 1977-99 Drivers	RF Passengers
Frontal	Single-vehicle	42,082	10,581
Frontal	Multivehicle	50,222	14,095
Left side	Single-vehicle	9,126	1,430
Left side	Multivehicle	17,767	2,549
Right side	Single-vehicle	7,342	3,880
Right side	Multivehicle	11,089	8,878
Rollover (primary)		21,970	6,103
Rear & other		5,374	2,072

The unrestrained fatality counts are multiplied by [1 - Effectiveness], the uncorrected belted-to-unrestrained fatality risk ratio, to obtain uncorrected estimates of how many fatalities there would have been in each crash mode if the unrestrained occupants had been belted. The unrestrained and belted fatalities are summed over the crash modes. The belted fatalities are multiplied by the UEF to obtain the overall effectiveness estimate:

			Unrestrained Fatalities	Risk Ratio (1 - Effectiveness)	Belted Fatalities
Frontal	Single-veh	Driver	42,082	0.2808	11,817.5
		RF	10,581	0.3371	3,566.7
	Multiveh	Driver	50,222	0.4140	20,793.4
		RF	14,095	0.4494	6,334.7
Left	Single-veh.	Driver	9,126	0.5990	5,466.8
		RF	1,430	0.4112	588.0
	Multiveh	Driver	17,767	0.6410	11,388.8
		RF	2,549	0.5423	1,382.3
Right	Single-veh.	Driver	7,342	0.3940	2,892.7
		RF	3,880	0.5186	2,012.0
	Multiveh.	Driver	11,089	0.4594	5,094.1
		RF	8,878	0.7991	7,094.7
Rollover (Primary)		Driver	21,970	0.1848	4,059.1
		RF	6,103	0.2022	1,234.2
Rear & Other		Driver	5,374	0.3114	1,673.3
		RF	2,072	0.3547	734.9
UNCORRECTED FATALITIES			214,560		86,133.1
UEF Correction					x 1.369
CORRECTED FATALITIES			214,560		117,970
FATALITY REDUCTION FOR 3-POINT BELTS					45.02 %

Based on the uncorrected effectiveness estimates, the 214,560 unrestrained fatalities would have decreased to 86,133.1 if all occupants had worn belts. Application of the UEF (1.369) increases that estimate to 117,970. The corrected estimate of belt effectiveness is

$$1 - (1.369 \times 86{,}133.1 / 214{,}560) = 1 - (117{,}970 / 214{,}560) = 45.02 \text{ percent}$$

This is our "best" estimate based on 1986-99 FARS data and it is indeed very close to NHTSA's 1984 estimate of 45 percent.

7. SAMPLING ERROR CONSIDERATIONS

The discussion that follows presents formulas for calculating a "textbook" sampling error for effectiveness estimates. Since the formulas do not capture the unknown and basically unquantifiable non-sampling error that could be present in our estimates, they are insufficient for generating "confidence bounds" - i.e., measuring the upper bound of possible uncertainty. But they are quite useful for assessing the minimum or lower bound of possible uncertainty due to small N's. They help distinguish between those estimates that are based on a lot of data, yet may have unknown non-sampling error, and those estimates that are statistically meaningless under any circumstances because they are based on small N's.

The effectiveness estimate has just been defined as

$$E = 1 - (1.369 \times 86{,}133.1 / 214{,}560) = 1 - (\text{UEF} \times \text{Belted} / \text{Unrestrained})$$

In other words, the corrected fatality risk ratio

$$R = 1 - E = \text{UEF} \times \text{Belted} / \text{Unrestrained}$$

is the product and quotient of three statistics that are uncorrelated for all **practical** purposes: the variance of UEF [to the extent that its variance is a meaningful concept] is essentially the variance of the 1977-85 effectiveness estimate for passenger cars in all crashes, the variance of "Belted" is primarily the uncertainty of the 1986-99 effectiveness estimate in the types of crashes we are studying (here it happens to be all crashes), and the negligible variance of "Unrestrained" derives from the variability of the crash-mode distribution of 1977-99 FARS (including cars with unaccompanied drivers). Under these circumstances it is acceptable[17] to approximate

$$\text{Var}(R) / R^2 = [\text{Var}(\text{UEF}) / \text{UEF}^2] + [\text{Var}(\text{Belted}) / \text{Belted}^2] + [\text{Var}(\text{Unres.}) / \text{Unres.}^2]$$

and the standard deviation of effectiveness

$$\text{s.d.}(E) = \text{s.d.}(1 - E) = \text{s.d.}(R) = R \times [\text{CV}(\text{UEF})^2 + \text{CV}(\text{Belted})^2 + \text{CV}(\text{Unrestrained})^2]^{½}$$

where CV is the coefficient of variation (the standard deviation divided by the mean).

We already estimated the CV of the UEF in Section 5, based on subdividing the FARS data into 10-35 systematic random subsamples and noting the variation of the observed UEF among the subsamples. It is 5.8 percent.

It would be futile to compute the CV of "Belted" and "Unrestrained" by breaking the data into 10 or more subsamples and computing these statistics within the subsamples. Since there are relatively few belted fatality cases in some of the smaller cells (e.g., RF passengers in rear & other impacts), some of the subsamples are likely to have zero cases, making it impossible to calculate the statistics.

---

[17] Hansen, Morris H., Hurwitz, William N., and Madow, William G., *Sample Survey Methods and Theory, Volume I*, John Wiley & Sons, New York, 1953, pp. 512-514.

Instead, it is appropriate to use a **jackknife** procedure[18]. Based on the last digit i of ST_CASE, the FARS data are allocated to ten overlapping subsamples, each containing 9/10 of the cases (by removing 1/10 of the cases whose ST_CASE ends in i). “Belted” and “Unrestrained” are estimated for the 9/10 subsample, as in Section 6, without the UEF correction. “Pseudo-estimates” for the remaining 1/10 of the data are obtained by subtracting the 9/10 estimates from the uncorrected full-data-set estimates (86,133.1 for “Belted” and 214,560 for “Unrestrained”). The variation of the pseudo-estimates is used to compute the CV’s for the whole data set, as follows:

	Unres.	Belted	Uncorrect. Fat. Red.	Unres.	Belted	Uncorrect. Fat. Red.
	For the Full Data Set					
	214,560	86,133.1	59.86			
	For 9/10 of the Data Excluding ST_CASE Ending in:			Pseudo-Estimate for Remaining 1/10 of the Data		
0	193,317	77,772.8	59.77	21,243	8,360.2	60.64
1	193,033	78,048.0	59.57	21,527	8,085.1	62.44
2	192,999	77,618.6	59.78	21,561	8,514.4	60.51
3	192,826	77,517.1	59.80	21,734	8,616.0	60.36
4	193,235	77,247.1	60.02	21,325	8,885.9	58.33
5	193,204	77,958.7	59.65	21,356	8,174.3	61.72
6	192,879	77,086.2	60.03	21,681	9,046.9	58.27
7	193,235	76,711.3	60.30	21,325	9,421.7	55.82
8	193,100	77,254.4	59.99	21,460	8,878.7	58.63
9	193,212	77,990.0	59.64	21,348	8,143.1	61.86
s = standard deviations of the pseudo-estimates:				165.1	441.3	
S = $\sqrt{10}$ s = standard deviations for the full data set:				522.1	1,395.6	
X = totals for the full data set				214,560	86,133.1	
CV = S / X = coefficients of variation				0.24%	1.62%	

The “Unrestrained” counts vary by negligible amounts among the pseudo-estimates, implying a CV of only 0.24 percent for the “Unrestrained” total of 214,560 for the entire data set. The “Belted” counts (and the uncorrected effectiveness estimates) vary a bit more, but still not much: even with

---

[18] Mosteller, F. and Tukey, J.W., *Data Analysis and Regression: A Second Course in Statistics*, Addison-Wesley, Reading, MA, 1977.

subsamples 1/10 as large as the full data set, uncorrected effectiveness ranges only between 55.82 and 62.44 percent. The CV for the “Belted” total of 86,133.1 is 1.62 percent. Both are substantially lower than the CV of the UEF, 5.8 percent.

The standard deviation of the corrected effectiveness estimate (45.02 percent) is

$$\text{s.d. (E)} = (1 - \text{E}) \text{ x } [\text{CV(UEF)}^{2} + \text{CV(Belted)}^{2} + \text{CV(Unrestrained)}^{2}]^{½}$$

$$= (1 - 0.4502) \text{ x } [0.058^{2} + 0.0162^{2} + 0.0024^{2}]^{½} = 3.31 \text{ percentage points}$$

and the ± 1.96 ó sampling-error bounds[19] are 45 ± 6.5 percent, or 38.5 to 51.5 percent.

In the preceding calculations, over 90 percent of the measurable sampling error in the effectiveness estimate derives from the UEF term. The CV’s of “Belted” and “Unrestrained” are small by comparison. This situation prevails as long as we use the procedure to estimate effectiveness of 3-point belts, for all front-outboard occupants in all types of passenger car crashes. After all, we defined the UEF to correct the 1986-99 effectiveness to make it the same as the 1977-85 effectiveness. But the 1977-85 effectiveness is itself fairly uncertain, since it is based on relatively few belted FARS cases. Clearly, our corrected 1986-99 estimate, no matter how many FARS cases it is based on, cannot have less uncertainty than the 1977-85 estimate that we demand it must equal. Essentially, for the overall effectiveness estimate in passenger cars, the 1986-99 data do not add any new information, because we basically continue to rely on the 1977-85 estimate.

A different situation will prevail, however, when we use this procedure to estimate effectiveness for subgroups of crashes, vehicles, or occupants. The hypothesis in Section 5 is that the same UEF can be used repeatedly for the various subgroups. It will continue to have CV = 5.8 percent. Only the CV’s for “Belted” and “Unrestrained” will grow as the subgroups shrink. We can now obtain many effectiveness estimates, with relatively small sampling errors based on our formula, that could not have been obtained from the 1977-85 data alone.

But this empirical tool that generates credible point estimates and reduces measurable sampling error - repeated application of the UEF - introduces an unknown non-sampling error. Even though we showed in Section 5 that the UEF is quite robust in a variety of subgroups, we cannot prove it is valid for every subgroup considered in the next section, or quantify its error. Thus, the error bounds generated by the preceding formulas, while useful for indicating the minimum range of error in the point estimates that follow, cannot be considered confidence bounds that encompass all sources of uncertainty in the estimates.

---

[19] In textbook analyses with normally distributed statistics, ±1.96 standard deviations correspond to “two-sided 95 percent confidence bounds.” In this report, the interval is not presented as “confidence bounds” but as an indicator of minimum sampling error.

8. RESULTS: BELT EFFECTIVENESS ESTIMATES

First, let us estimate the overall effectiveness of 2-point automatic belts in passenger cars and 3-point belts in light trucks - i.e., pickup trucks, vans and sport utility vehicles (SUV). These are the counterparts to the 45 percent effectiveness estimate for 3-point belts in passenger cars. (There are hardly any light trucks with 2-point automatic belts, certainly not enough to estimate effectiveness.)

The analysis for cars with 2-point belts is similar to the one in Sections 6 and 7[20]. It comprises motorized as well as non-motorized shoulder belts. "Belted" occupants include anybody with MAN_REST or REST_USE = 1, 2, 3, 8 or 13 on FARS. Although these include separate codes for "lap-only," "shoulder-only" and "lap+shoulder" belt use, FARS is not reliable for such precise distinctions[21]. Thus, the "belted" occupants are a mix of people who used the manual lap belt as well as the automatic shoulder belt, and those who did not. The fatality reduction for 2-point belts is 32 percent (the minimum sampling error range, as defined in Section 7, is ± 13.0 percentage points). This is a lower point estimate than the 45 percent for 3-point belts in passenger cars. Based on the 1986-99 FARS data alone, and without considering the UEF or other possible biases in these data, 3-point belts appear to be significantly more effective than 2-point belts[22]. It is unclear from FARS to what extent, if any, the effectiveness of 2-point belts is lowered because some occupants use only the shoulder belt and not the lap belt[23].

---

[20] Double-pair comparisons are performed, by crash mode, using only the 2-point belt cases on the 1986-99 FARS. However, the uncorrected effectiveness at each crash mode/seat position is weighted by **all** 1977-99 unrestrained passenger-car fatalities, including cars with 3-point belts, exactly as in Section 6. The rationale is to estimate the fatality reduction that would occur if every occupant of every car on the road used 2-point belts, relative to being unrestrained. Thus, "Unrestrained" is 214,560 as in Sections 6 and 7 and its CV is again 0.24 percent. The UEF = 1.369 and its CV = 5.8 percent are also unchanged. However, the CV for "Belted" increases from 1.62 to 7.82 percent since there are far fewer cases with 2-point belts.

[21] For example, in those 1987-89 Volkswagens, Hyundais, Mitsubishis and Yugos that were not even equipped with a lap belt, FARS reports that 65 percent of the fatally injured belt users "wore lap and shoulder belts" and 7 percent wore the "lap belt only." Only 11 percent were correctly reported "shoulder belt only" and the remainder "used belts - unknown type."

[22] For 2-point belts, the uncorrected belted fatalities, as defined in Section 6, are 106,342, with standard deviation 8,316. For 3-point belts, these numbers are 86,133 and 1,396, respectively. The difference is 20,209 and its standard deviation is 8,432. This is statistically significant ($z = 2.40$, $p < .05$).

[23] An analysis of the FARS cases of occupants who were certain or likely not to have used the lap belt - occupants of the 1987-89 cars that were not equipped with lap belts plus occupants coded "shoulder belt only" in FARS - led to inconclusive results. The effectiveness estimate was 28%, not really different from the 32% estimate for all 2-point belts.

The fatality reduction for 3-point belts in light trucks - pickups, vans and SUVs - is estimated to be 60 percent[24] (minimum sampling error range ± 5.1 percentage points). That coincides exactly with the agency's 60 percent estimate of 1989[25]. Based on the 1986-99 FARS data alone, belt effectiveness is significantly higher in light trucks than in cars (45 percent)[26]. The rest of this section will show some of the reasons why.

Table 1 summarizes the overall effectiveness of belts in cars and light trucks[27].

---

[24] Double-pair comparisons are performed, by crash mode, on the MY 1975-2000 light trucks, with no air bags or with dual air bags, in CY 1986-99 FARS. The effectiveness estimates at each crash mode/seat position are weighted by all 1977-99 unrestrained **light-truck** fatalities. "Unrestrained" is 88,119 and its CV is 0.24 percent. "Belted" is 25,638.9 and its CV is 2.97 percent (about double the CV for cars with 3-point belts). The UEF = 1.369 and its CV = 5.8 percent are unchanged from the car analyses. A modest proportion of the MY 1975-80 light trucks were equipped with separate lap and shoulder belts, or with lap belts only, rather than with integral 3-point belts. A possible alternative approach would have been to limit the analysis to MY 1981-2000 light trucks, exclusively equipped with integral 3-point belts. That would have raised the effectiveness estimate to 61% rather than 60%.

[25] *Op. cit.*, NHTSA Docket No. 74-14-N62-001, p. 15.

[26] For cars, the uncorrected belted fatalities, as defined in Section 6, are 86,133, with standard deviation 1,396; unrestrained fatalities are 214,560, with standard deviation 522. The ratio is 0.40144 and its standard deviation is 0.00658. For light trucks, the uncorrected belted fatalities are 25,639, with standard deviation 760; unrestrained fatalities are 88,119, with standard deviation 214; the ratio is 0.29096 and its standard deviation is 0.00862. The difference of the ratios is 0.11048 and its standard deviation is 0.01084. This is statistically significant ($z = 10.19$, $p < .01$).

[27] The three alternative approaches considered in footnotes 15, 16 and 22 would have produced slightly different estimates: (1) Computation of the UEF based on CY 1977-84 rather than CY 1977-85 FARS data would have lowered the estimates to 42 percent for cars [with 3-point belts] and 58 percent for light trucks. (2) Computation of the UEF based on cars of MY 1975-85 rather than on cars of all model years would have lowered the estimates to 44 percent for cars and 59 percent for light trucks. (3) Limiting the light trucks to MY 1981 and later [100% integral 3-point belts] would have raised the light truck estimate to 61 percent and left the car estimate unchanged. All of these alternative procedures generate car estimates in the 40-50 percent range, and light truck estimates close to 60 percent. Basically, the alternative methods do not make an important difference in the results.

TABLE 1 - OVERALL EFFECTIVENESS OF SAFETY BELTS

	Fatality Reduction (%)	Minimum Sampling Error Range[28]
Passenger cars, 3-point belts	45	x
Passenger cars, 2-point automatic belts	32	xx
Light trucks, 3-point belts	60	x

Table 2 addresses a fundamental issue: the effectiveness of safety belts by crash mode[29]:

TABLE 2: FATALITY REDUCTION BY DIRECTION OF IMPACT

	Cars 3-Point Belts		Cars 2-Point Belts		Light Trucks 3-Point Belts	
	Fat. Red.	Mser	Fat. Red.	Mser	Fat. Red.	Mser
Frontal impacts	50	x	30	xx	53	x
Side impacts	21	xx	18	xxx	48	x
Nearside	10	xx	18	xxx	41	xx
Farside	39	x	18	xxx	58	x
Rollovers (primary)	74	x	62	x	80	x
Rear impacts & other crashes	56	x	68	xx	81	x

Three-point belts are quite effective in frontal crashes, and they are more or less equally effective in cars (50%) and light trucks (53%). Two-point automatic belts are effective (30%) but apparently less so than 3-point belts.

---

[28] Mser = minimum sampling error range based on the formula in Section 7: “x” denotes the 1.96ó sampling error is in the ± 4-10 percentage point range - a precise point estimate according to the formula, but with unknown non-sampling error; “xx” = ± 10-20 percentage points; “xxx” = ± 20-50 percentage points; and “xxxx” denotes more than ± 50 percentage points, a statistically meaningless point estimate.

[29] The crash modes are defined at the beginning of Section 6. “Nearside” impacts include left-side impacts for drivers and right-side for RF passengers. “Farside” includes right-side for drivers and left-side for RF. The analysis is the same as in Sections 6 and 7, except that only some of the cells are used in each case - e.g., the frontal effectiveness is the weighted average of single-vehicle frontal and multivehicle frontal.

The difference between cars and light trucks is quite clear in side impacts. In passenger cars, 3-point belt effectiveness in side impacts after the UEF correction is only 21 percent, substantially lower than in the other crash modes. In nearside impacts (left side for the driver, right side for the RF) the point estimate drops to 10 percent, with even the minimum sampling error range larger than that. Only in farside impacts (39 percent) is effectiveness comparable to frontals. In light trucks, on the other hand, the fatality reduction in side impacts is a healthy 48 percent and even in nearside impacts it is still 41 percent. We may surmise that nearside impacts to passenger cars often involve compartment intrusion where belts are unable to prevent fatalities, while the compartments of light trucks, often with higher sills and seating heights, are less vulnerable to intrusion and allow belts to accomplish their benefits of preventing ejection and mitigating impacts with interior components. Tables 3 and 4 will support that idea.

Belts are highly effective in rollovers, where, as we shall see in Table 4, the majority of unbelted fatalities are ejectees. Effectiveness is high in light trucks (80%) and in cars with 3-point belts (74%) and it is slightly lower in cars with 2-point belts (62%).

Belts also have high point estimates of effectiveness in rear impacts and other crashes (immersions, falls from moving vehicles, etc.). That may surprise those who envision rear impacts as simple, moderate-severity crashes where the occupants' primary motion is directly into the seat behind them and belts are unlikely to be important. However, this "typical" rear impact is rarely fatal. We shall see that many of the fatalities involve ejection, and many others undoubtedly have other unusual circumstances, such as multiple or oblique impacts, where belts can be useful.

Table 3 looks at frontal and side impacts in-depth, subdividing them into single- and multivehicle, and the latter by the type of the "other" vehicle (car, light truck, heavy truck)[30].

---

[30] "Fixed object" includes all single-vehicle crashes. "By a car" includes all 2-vehicle crashes where the other vehicle is a passenger car plus all 3-vehicle crashes where both of the other vehicles are cars.

TABLE 3: FATALITY REDUCTION IN FRONTAL AND SIDE IMPACTS BY TYPE OF VEHICLE/OBJECT STRUCK

	Cars 3-Point Belts		Cars 2-Point Belts		Light Trucks 3-Point Belts	
	Fat. Red.	Mser*	Fat. Red.	Mser*	Fat. Red.	Mser*
Frontal impacts						
Fixed object	60	x	45	xx	64	x
Multivehicle	42	x	18	xxx	40	x
With a car	48	x	13	xxx	51	x
With a light truck	39	x	15	xxx	42	xx
With a heavy truck	34	xx	53	xxxx	30	xx
Nearside impacts						
Fixed object	21	xx	23	xxx	47	xx
Multivehicle	5	xx	15	xxx	36	xx
By a car	12	xx	27	xxxx	69	xx
By a light truck	2	xx	23	xxxx	31	xxx
By a heavy truck	2	xxx	- 26	xxxx	- 17	xxx
Farside impacts						
Fixed object	46	xx	52	xx	61	xx
Multivehicle	35	x	- 4	xxx	54	xx
By a car	45	xx	- 25	xxxx	71	xx
By a light truck	36	x	- 4	xxxx	50	xx
By a heavy truck	20	xx	- 28	xxxx	18	xxx

*Minimum sampling error range: x = ± 4-10 percentage points, xx = ± 10-20, xxx = ± 20-50, xxxx = more than ± 50

In all three crash modes, observed belt effectiveness is higher, sometimes substantially higher in impacts with fixed objects than in multivehicle crashes. For example, in cars with 3-point belts, belts reduce fatalities by 60 percent in frontals with fixed objects, versus 42 percent in frontal impacts with other vehicles. In the multivehicle crashes, belt effectiveness is consistently highest when the "other" vehicle is a passenger car and lowest when it is a heavy truck. The findings are not surprising: fixed-object impacts are especially likely to involve ejection (following non-catastrophic violation of passenger compartment integrity) - where belts work best, while impacts by heavy trucks are most likely to involve catastrophic intrusion - where belts help least.

Above all, Table 3 shows the contrast between belt effectiveness in cars and light trucks in **nearside** impacts. In cars with 3-point belts, the fatality reduction is 21 percent in nearside impacts with fixed objects, but a negligible 5 percent in impacts by other vehicles. The point estimate is smaller than even the minimum sampling error regardless whether the striking vehicle is another car, a light truck or a heavy truck. By contrast, belts are extremely effective in light trucks when they are struck in the near side by a car (69%), may still have some benefits when the striking vehicle is another light truck (31%), and become ineffective only when the striking vehicle is a heavy truck. Evidently, the high sills and rigid structures and higher seat heights of light trucks enable them to ward off intrusion when they are struck by passenger cars and allow safety belts to accomplish their life-saving function of mitigating occupant contacts with undamaged interior structures.

Since the preceding discussion mentions repeatedly that many of the lives saved by belts are due to the prevention of ejection, it is appropriate to compare the percentages of unrestrained fatalities who were ejectees, by crash mode and by vehicle type. Table 4 shows how large those percentages are, especially in light trucks. (Table 4 includes all driver and RF fatalities in the 1977-99 FARS. Partial as well as complete ejections are included among the "ejectees.")

TABLE 4: PERCENT OF FATALITIES WHO WERE EJECTEES

	Passenger Cars	Light Trucks
ALL CRASHES	28	48
Frontal impacts	21	33
Fixed object	31	42
Multivehicle	12	24
Nearside impacts	21	39
Fixed object	34	44
Multivehicle	14	36
Farside impacts	26	45
Fixed object	37	48
Multivehicle	20	42
Rollovers (primary)	69	78
Rear impacts & other crashes	37	52

Complete or partial ejection occurred in a substantial 28 percent of unbelted passenger car fatalities and a shocking 48 percent of unbelted light truck fatalities. An estimated 74 percent of ejection fatalities would have survived if they had remained within their vehicle[31]. FARS data

---

[31] Kahane, Charles J., *An Evaluation of Door Locks and Roof Crush Resistance of Passenger Cars*, NHTSA Publication No. DOT HS 807 489, Washington, 1989, pp. 30-32. Evans, Leonard, *Traffic Safety and the Driver,* Van Nostrand Reinhold, New York, 1991, pp. 52-54. Evans, Leonard and Frick, Michael C., "Potential Fatality Reductions through Eliminating Occupant Ejection from Cars," *Accident Analysis and Prevention*, Vol. 21, pp. 169-182. Sikora, James J., *Relative Risk of Death for Ejected Occupants in Fatal Traffic Accidents,* NHTSA Publication No. DOT HS 807 096, Washington, 1986.

suggest that 3-point belts reduce the probability of ejection by at least 91 percent in fatal crashes in cars and also in light trucks[32]. If safety belts had no benefits other than preventing ejection, they would reduce overall fatality risk by

$$0.28 \times 0.74 \times 0.91 = 19 \text{ percent in passenger cars}$$

and

$$0.48 \times 0.74 \times 0.91 = 32 \text{ percent in light trucks}$$

In other words, prevention of ejection accounts for a substantial portion of the benefit of belts, but not nearly all the benefits. The overall effectiveness of belts is 45 percent in cars and 60 percent in light trucks, well beyond the 19 and 32 percent attributable to preventing ejection. Much of their benefit comes from mitigating injuries within the vehicle. These statistics also show one of the main reasons why belts are more effective in light trucks than in cars: a lot more of the fatalities in light trucks are ejectees.

Table 4 also shows substantial differences in the proportions of ejectees by crash mode. In rollovers, of course, the majority of unbelted fatalities are ejectees from passenger cars (69%), but even more so from light trucks (78%). Ejectees account for a large proportion of the fatalities in “rear and other” impacts. However, ejection is also quite common in frontal and side impacts with fixed objects, more so than in multivehicle crashes. For example, in frontal impacts of passenger cars, 31 percent of the fatalities were ejectees in fixed-object collisions but only 12 percent in collisions with other vehicles. That goes a long way to explaining why belts are more effective in fixed-object than in multivehicle frontal collisions (60 vs. 42 percent according to Table 3). Ejectees also account for a larger proportion of the fatalities in farside than in nearside impacts (because a smaller proportion of the farside impacts involve intrusion that endangers the occupant).

The only crash modes in which fewer than 15 percent of the fatalities are ejectees are multivehicle frontal and nearside impacts of passenger cars. In the frontals, belts are still quite effective (42% according to Table 3), even though ejection is not a major factor, because they mitigate occupant injuries within the vehicles. But in the nearside impacts, belts have little effect (5% according to Table 3) because intrusion takes away opportunities for belts to minimize occupant contacts with interior surfaces.

Another factor that makes belts more effective in light trucks than cars is that a large proportion of light-truck fatalities are in rollovers, where belts are most effective, as shown in Table 5.

---

[32] In 1977-99 FARS data, the ratio of ejection to nonejection fatalities is 91 percent lower for belted occupants than for unbelted occupants of cars equipped with 3-point belts, and also in light trucks. Even if belts had no effect on non-ejected fatalities, this would imply a 91 percent reduction of the probability of ejection. To the extent that belts also reduce nonejection fatalities, the reduction of the probability of ejection is greater than 91 percent.

TABLE 5: CRASH MODE DISTRIBUTION OF UNRESTRAINED FATALITIES

	Passenger Cars	Light Trucks
Frontal impacts	55	52
Nearside impacts	19	10
Farside impacts	10	7
Rollovers (primary)	13	27
Rear impacts & other crashes	3	4
	100	100

Rollovers account for 27 percent of the fatalities of unrestrained occupants in light trucks, but only 13 percent in cars. By contrast, cars are especially vulnerable in nearside impacts, accounting for 19 percent of fatalities - where belts are least effective. Only 10 percent of unrestrained light-truck fatalities are nearside.

Tables 2 - 5 demonstrate three reasons why belts are more effective in light trucks than in cars:

- Ejection is substantially more frequent for unbelted occupants of light trucks than cars.
- Belts are much more effective in side impacts of light trucks - especially nearside impacts by light vehicles - because intrusion is much less of a problem in light trucks.
- Light trucks have relatively more rollovers, where belts are most effective, and relatively fewer side impacts, where belts are least effective.

Table 6 aggregates across crash modes to estimate effectiveness for all types of single vehicle crashes and all multivehicle crashes, the latter subdivided by the type of the "other" vehicle.

TABLE 6: FATALITY REDUCTION - SINGLE VS. MULTIVEHICLE CRASHES

	Cars 3-Point Belts		Cars 2-Point Belts		Light Trucks 3-Point Belts	
	Fat. Red.	Mser*	Fat. Red.	Mser*	Fat. Red.	Mser*
Single vehicle	58	x	49	xx	70	x
Multivehicle	32	x	17	xxx	43	x
With a car	41	x	20	xxx	57	x
With a light truck	31	x	15	xxx	45	xx
With a heavy truck	25	xx	19	xxx	28	xx

*Minimum sampling error range: x = ± 4-10 percentage points, xx = ± 10-20, xxx = ± 20-50, xxxx = more than ± 50

Effectiveness is substantially higher in single-vehicle crashes (fixed-object impacts + rollovers + other non-collisions) than in multivehicle crashes: 58 vs. 32 percent in cars with 3-point belts, 70 vs. 43 percent in light trucks. In multivehicle crashes, belt effectiveness is the highest when the "other" vehicle(s) is a passenger car, and lowest when it is a heavy truck. However, 3-point belts have point estimates 25 percent or higher in every configuration in Table 6.

Evans and Partyka both observed higher effectiveness of belts for drivers than RF passengers in double-pair comparisons of pre-1986 FARS data, although Evans' confidence bounds for the two reductions substantially overlap[33]. Table 7 presents estimates of belt effectiveness by seat position based on 1986-99 FARS data, including the effect of the lap belt for the center-front (CF) occupant.

TABLE 7: FATALITY REDUCTION BY SEAT POSITION

	Cars 3-Point Belts		Cars 2-Point Belts		Light Trucks 3-Point Belts	
	Fat. Red.	Mser*	Fat. Red.	Mser*	Fat. Red.	Mser*
In All Crashes						
Driver	48	x	38	xx	61	x
RF passenger	37	x	12	xxx	58	x
CF passenger lap belt[34]	4	xxx	☜		38	xxx
In Frontal Crashes						
Driver	52	x	34	xx	54	x
RF passenger	45	x	14	xxx	48	x

*Minimum sampling error range: x = ± 4-10 percentage points, xx = ± 10-20, xxx = ± 20-50, xxxx = more than ± 50

The data continue to show somewhat higher effectiveness for the driver (48%) than the RF passenger (37%) in cars with 3-point belts. Based on the 1986-99 FARS data alone, this difference is statistically significant. In light trucks there is less difference, 61 vs. 58 percent.

---

[33] Evans, *Traffic Safety and the Driver (op. cit.)*, p. 233. Partyka, *Papers on Adult Seat Belts (op. cit.)*, p. 1-12.

[34] Since cars with 2- and 3-point belts for outboard occupants both have just a lap belt for the CF occupant, they were combined for the CF analysis.

One reason that driver belt effectiveness could be higher is that unaccompanied drivers are slightly overinvolved in rollovers, where belts are especially effective, while a large portion of RF fatalities are in right-side impacts, where belts are least effective. For an analysis unbiased by possible differences in the distribution of crash modes, the lower half of Table 7 estimates effectiveness in **frontal** crashes only. Even in frontals, belt effectiveness is consistently higher for drivers than for RF: 52 vs. 45 percent in cars with 3-point belts and 54 vs. 48 percent in light trucks. The results suggest that the function of the driver's belt (protecting the driver from dangerous impact with the steering assembly) may be even more valuable than the function of the RF belt (protecting the RF from dangerous impacts with the instrument panel).

The 1986-99 FARS data do not contain enough cases of belted CF occupants of passenger cars for a statistically meaningful estimate of the overall effectiveness of lap belts (4%), let alone any analyses of subsets of the data[35]. The effectiveness of the CF lap belt in light trucks has a large point estimate, 38 percent, but it is highly uncertain.

The effectiveness of safety belts by **occupant age**, as shown in Table 8, is of interest. Belts could be less effective for the youngest and oldest occupants, who are least able to tolerate loading from the belt in crashes; young occupants also have difficulty wearing belts correctly.

Table 8 confirms that belts are most effective for occupants 15 to 54 years old, and effectiveness drops off for pre-teen and early-teen passengers, as well as for occupants over 70 and, especially, over 80[36]. For example, in passenger cars, the highest estimate is 50 percent fatality reduction at age 15-29, dropping to 34 percent at age 5-9 and 27 percent at age 80 or older. Light trucks have the same trend.

---

[35] To evaluate belts for the CF passenger, records of MY 1975-99 vehicles with a driver and a CF passenger were extracted from 1986-99 FARS files. Both the driver and the CF had to have known reported belt use, at least one and perhaps both fatally injured. There were only 1,985 cars with a driver and CF, as compared 70,668 cars having a driver and RF. With this small sample, analyses were limited to basic double-pair comparisons and correction by the UEF, as in Sections 4 and 5, with no separate computation and weighting by crash mode. Since cars with 2- and 3-point belts for outboard occupants both have just a lap belt for the CF occupant, they were combined for the CF analysis.

[36] Two sets of double-pair comparisons are performed: (1) drivers subdivided by age group (and RF can be any age 14 and over) to compute driver effectiveness by age group; (2) RF subdivided by age group (drivers any age) to compute RF effectiveness. These estimates were then weighted by unrestrained fatalities by age group, crash mode and seat position. RF age 5-14 were analyzed separately, by double-pair comparison with drivers of any age. When data were insufficient to subdivide and average across 8 crash modes, simplified analyses with just 3 crash modes (frontal, side, rollover + other) or without any subdivision were employed.

TABLE 8: FATALITY REDUCTION BY OCCUPANT AGE

	Cars 3-Point Belts		Cars 2-Point Belts		Light Trucks 3-Point Belts	
	Fat. Red.	Mser*	Fat. Red.	Mser*	Fat. Red.	Mser*
In All Crashes						
5- 9 (passengers)	34	xxx	- 10	xxxx	59	xxx
10-14 (passengers)	35	xx	12	xxxx	63	xx
15-29	50	x	38	xx	63	x
30-54	49	x	44	xx	64	x
55-69	43	x	26	xxx	53	xx
70-79	38	x	17	xxxx	42	xx
80 and older	27	xx	- 11	xxxx	30	xxx
In Frontal Crashes						
5- 9 (passengers)	37	xx	- 46	xxxx	45	xxx
10-14 (passengers)	48	xx	15	xxxx	61	xxx
15-29	52	x	33	xx	54	x
30-54	51	x	38	xxx	58	x
55-69	44	x	33	xxx	41	xxx
70-79	39	xx	2	xxxx	34	xxx
80 and older	36	xx	38	xxxx	28	xxx

*Minimum sampling error range: x = ± 4-10 percentage points, xx = ± 10-20, xxx = ± 20-50, xxxx = more than ± 50

Still, it is noteworthy how **little**, rather than how much effectiveness drops off for the youngest and oldest occupants, considering that a 3-point belt is really not designed to fit the 5-9 year old child, and considering the great vulnerability of people over 80 to impact loads. Although safety belts are not ideal protection for these occupants, they are clearly better than no protection at all.

Moreover, the basic results in the top section of Table 8 overstate the drop-off. Young and old occupants are underrepresented in rollover crashes, where belts are especially effective. All of the children and most of the oldest occupants are RF passengers - whereas belts are more effective for drivers. Thus, when the analysis is limited to frontal crashes, as in the lower half of Table 8, effectiveness in passenger cars varies only from 37 percent at age 5-9, to 52 percent at age 15-29, and back to 36 percent at age 80 and older.

Similarly Tables 8-D (drivers) and 8-P (passengers) confirm that effectiveness remains fairly constant, at each seat position, over a wide range of occupant age. For example, effectiveness for car drivers in frontal crashes drops gradually from 52 percent for age 15-29 to 40 percent for age 80 and above.

TABLE 8-D: DRIVERS - FATALITY REDUCTION BY AGE

	Cars 3-Point Belts		Cars 2-Point Belts		Light Trucks 3-Point Belts	
	Fat. Red.	Mser*	Fat. Red.	Mser*	Fat. Red.	Mser*
In All Crashes						
15-29	51	x	40	xx	63	x
30-54	49	x	46	xx	64	x
55-69	44	x	29	xxx	53	xx
70-79	39	xx	25	xxxx	39	xxx
80 and older	28	xx	- 15	xxxx	28	xxx
In Frontal Crashes						
15-29	52	x	33	xx	54	x
30-54	51	x	38	xxx	58	x
55-69	44	x	33	xxx	41	xxx
70-79	40	xx	10	xxxx	32	xxx
80 and older	40	xx	41	xxxx	28	xxx

*Minimum sampling error range: x = ± 4-10 percentage points, xx = ± 10-20, xxx = ± 20-50, xxxx = more than ± 50

TABLE 8-P: RF PASSENGERS - FATALITY REDUCTION BY AGE

	Cars 3-Point Belts		Cars 2-Point Belts		Light Trucks 3-Point Belts	
	Fat. Red.	Mser*	Fat. Red.	Mser*	Fat. Red.	Mser*
In All Crashes						
5- 9	34	xxx	- 10	xxxx	59	xxx
10-14	35	xx	12	xxxx	63	xx
15-29	49	x	32	xx	63	x
30-54	44	x	33	xxx	63	x
55-69	39	xx	13	xxx	57	x
70-79	35	x	- 5	xxxx	52	xx
80 and older	27	xx	- 2	xxxx	40	xxx
In Frontal Crashes						
5- 9	37	xx	- 46	xxxx	45	xxx
10-14	48	xx	15	xxxx	61	xxx
15-29	55	x	34	xx	57	x
30-54	50	x	33	xxx	54	x
55-69	48	xx	4	xxxx	43	xx
70-79	37	xx	- 21	xxxx	43	xx
80 and older	29	xx	31	xxxx	25	xxx

*Minimum sampling error range: x = ± 4-10 percentage points, xx = ± 10-20, xxx = ± 20-50, xxxx = more than ± 50

Belt effectiveness might also differ for male and female occupants. “Real” anatomical or physiological differences that could influence effectiveness include: greater vulnerability of females to impact loads[37], belts improperly fitting small occupants (primarily female) or exceptionally large occupants (usually male). Furthermore, behavioral differences affect the crash exposure of males and females, with consequences for the overall effectiveness - e.g., young males are overinvolved in rollovers, where belts are especially effective; when a male and female ride together, the male is more often the driver. Table 9 analyzes belt effectiveness by occupant gender[38].

---

[37] Evans, *Traffic Safety and the Driver (op. cit.)*, pp. 22-25.

[38] The procedures are similar to the analyses by occupant age.

TABLE 9: FATALITY REDUCTION BY OCCUPANT GENDER

	Cars 3-Point Belts		Cars 2-Point Belts		Light Trucks 3-Point Belts	
	Fat. Red.	Mser*	Fat. Red.	Mser*	Fat. Red.	Mser*
In All Crashes						
Male	**45**	**x**	**34**	**xx**	**60**	**x**
Driver	46	x	37	xx	60	x
RF passenger	44	x	19	xxx	61	x
Female	**45**	**x**	**33**	**xx**	**62**	**x**
Driver	51	x	41	xx	66	x
RF passenger	33	x	17	xxx	57	x
In Frontal Crashes						
Male	**49**	**x**	**33**	**xx**	**53**	**x**
Driver	49	x	36	xx	53	x
RF passenger	51	x	18	xxx	54	x
Female	**52**	**x**	**29**	**xxx**	**53**	**x**
Driver	56	x	33	xxx	59	x
RF passenger	43	x	21	xxx	44	x

*Minimum sampling error range: x = ± 4-10 percentage points, xx = ± 10-20, xxx = ± 20-50, xxxx = more than ± 50

Overall belt effectiveness is nearly the same for male and female occupants. For example, in all crashes of passenger cars with 3-point belts, the estimates are identically 45 percent. Estimates are within a few percent, or identical, in frontal crashes, and in light trucks. However, Table 9 shows one pattern that is too strong and consistent to dismiss as coincidence: belt effectiveness is consistently higher for female drivers than RF, whereas among males, belt effectiveness is nearly equal for drivers and RF. The pattern appears even when the analysis is limited to frontal crashes (e.g., 56 vs. 43 percent for female drivers and RF in passenger cars). In fact, belts appear to be more effective for the female than the male driver, but less effective for the female than the male RF. Perhaps, belts are especially effective in protecting the small female, but less so the large male, from harmful contact with the steering assembly - compensating for other disadvantages of the female occupant (greater vulnerability to impact, poor belt fit). On the other hand, the observed pattern could be due to an artifact of the data - e.g., female passengers might be the group least likely to misreport their belt use, resulting in lower observed effectiveness estimates.

One of the most important reasons for analyzing 1986-99 FARS data is to find out if the effectiveness of 3-point belts has changed since the early 1980's - by computing separate estimates for vehicles of older and more recent model-year groups. Modifications that made vehicles safer in general, such as air bags or side impact protection, would not necessarily influence belt effectiveness if they have more or less the same relative benefits for the unrestrained as well as the belted occupant. On the other hand, specific improvements to belt systems, such as pretensioners or load limiters, could augment belt effectiveness, since they have benefits for the belt user and no benefits for the unrestrained occupant. Table 10 presents effectiveness estimates for several generations of vehicles, and for specific types of automatic belts.

TABLE 10: FATALITY REDUCTION BY BELT TYPE AND MODEL YEAR GROUP

	Passenger Cars		Light Trucks	
	Fat. Red.	Mser*	Fat. Red.	Mser*
Manual belts in non-air-bag vehicles	45	x	60	x
MY 1975-79[39]	47	x	49	xx
MY 1980-85	46	x	63	x
MY 1986 +	42	x	60	x
Manual belts in dual-air-bag vehicles[40]	48	x	63	xx
3-point automatic belts	48	xx	n/a	
2-point automatic belts (with or w/o lap belts)	32	xx	n/a	

*Minimum sampling error range: x = ± 4-10 percentage points, xx = ± 10-20, xxx = ± 20-50, xxxx = more than ± 50

The effectiveness of all types of 3-point belts is close to 45 percent in passenger cars. There is little difference in the manual belts of MY 1975-79 (47%), MY 1980-85 (46%), MY 1986-89 (42%) or MY 1990-99 with dual air bags (48%) - these variations are easily within the “noise” range of the sampling error. In other words, when cars were made safer during 1975-99, they became safer by about the same relative amount for the belted and the unrestrained occupant.

For light trucks, the MY 1975-79 estimate (49%) is lower than the others (all close to 60%). This quite possibly is due to a modest proportion of the MY 1975-79 trucks being equipped with separate lap and shoulder belts, or even with lap belts only, rather than with integral 3-point belts.

---

[39] I.e., MY 1975-79 vehicles involved in fatal crashes during 1986-99.

[40] This effectiveness estimate is the fatality reduction for safety belt plus air bag relative to air bag alone.

However, it is hard to draw conclusions because the sampling error of the MY 1975-79 estimate is relatively large (since the N of belted occupants is small in those older trucks).

Observed belt effectiveness in cars of the 1990's with dual air bags is a few percentage points higher than in earlier cars. This is not a significant increase, given the sampling errors. However, the numbers are at least consistent with the possibility that recent improvements to belt systems, such as pretensioners and load limiters, could be a factor.

Three-point automatic belts, when used, are about equally effective (48%) as manual belts in cars, when used (45%). The 1994 National Occupant Protection Use Survey shows equal use rates for automatic and manual 3-point belts[41]. In other words, the automatic and manual 3-point belt systems are basically equivalent in effectiveness and benefits (i.e., effectiveness x use).

Earlier studies did not show large differences in belt effectiveness in light and heavy cars[42]. Table 11 estimates belt effectiveness for three curb-weight groups of passenger cars[43].

TABLE 11: FATALITY REDUCTION BY PASSENGER CAR WEIGHT

	In All Crashes		In Frontal Crashes	
Curb Weight	Fat. Red.	Mser*	Fat. Red.	Mser*
2499 pounds or lighter	48	x	52	x
2500-3149 pounds	44	x	52	x
3150 pounds or heavier	41	x	45	x

*Minimum sampling error range: x = ± 4-10 percentage points, xx = ± 10-20, xxx = ± 20-50, xxxx = more than ± 50

---

[41] The 1994 NOPUS controlled-intersection survey is the latest file of belt use observations plus VINs to allow decoding of the type of belts. The use rates in cars with automatic 3-point belts, and in cars [of about the same age] with manual belts and air bags were both 69 percent.

[42] Evans, *Traffic Safety and the Driver (op. cit.)*, pp. 235-236.

[43] For MY 1981-98, car weights in this analysis are based on decoding the VINs of the FARS cases to determine make-model/body type and linking to the Polk National Vehicle Population Profile, which provides accurate curb weights. Prior to MY 1981, we used FARS VIN_WGT plus 100 pounds, since that variable in the past averaged 100 pounds below curb weight. For MY 1999, VIN_WGT has become a good estimate of curb weight.

The fatality reductions for all types of crashes show slight decreases as the cars get larger, but the differences are in the "noise" range. Belts are perhaps more effective in small cars because they have a higher proportion of rollovers. However, Table 11 shows a similar slight trend even in frontal crashes: 52 percent reductions for small and middle-weight cars, and 45 percent in heavy cars. In other words, there is no clear evidence of differences by car size. Intuitively, both small and large cars have characteristics that favor belts. Small cars tend to have younger occupants and a higher proportion of single-vehicle crashes, but large cars offer better protection against intrusion.

Table 12 estimates belt effectiveness for different types of light trucks[44].

TABLE 12: FATALITY REDUCTION BY LIGHT TRUCK TYPE

	In All Crashes		In Frontal Crashes	
	Fat. Red.	Mser*	Fat. Red.	Mser*
Pickup truck	**58**	**x**	**52**	**x**
Compact pickup	55	x	43	xx
Full-sized pickup	61	x	59	x
Van or SUV	**63**	**x**	**55**	**x**
Minivan	60	x	53	xx
Full-sized van	62	xx	45	xxx
Compact SUV	67	x	61	xx
Full-sized SUV	74	xx	51	xxx

*Minimum sampling error range: x = ± 4-10 percentage points, xx = ± 10-20, xxx = ± 20-50, xxxx = more than ± 50

In general, the various types of light trucks have overall effectiveness close to 60 percent and fatality reductions in frontal crashes close to 53 percent (the averages for all light trucks in Tables 1 and 2).

There are two hints of divergence from the average. Compact pickup trucks have point estimates of belt effectiveness lower than average, especially in frontals. Compact SUVs have point estimates slightly higher than the average. At first glance, it is peculiar that frontal effectiveness is 61 percent in compact SUVs but only 43 percent in compact pickups, given that manufacturers often build similar chassis, frame and body parts into their compact SUVs and pickups. However, the two vehicle types have dissimilar drivers and exposure patterns, resulting in different crash distributions: SUVs have more rollovers, more off-road excursions, and more ejections. In

---

[44] Table 12 is limited to MY 1985-98 light trucks with decodable VINs.

compact SUVs, 65 percent of the unbelted fatalities are complete or partial ejections, the highest for any truck type, but only 49 percent in compact pickups. In frontal crashes, 47 percent of the SUV fatalities are ejectees, but only 34 percent in the pickups. Even in multivehicle frontal crashes, 37 percent of the SUV deaths are ejectees, but just 24 percent in the pickups. It is unknown whether the high ejection rate in SUVs represents vehicle factors (e.g., large side windows) in addition to exposure factors, but in any case this explains the high effectiveness of belts in SUVs.

9. VALIDATION: BELT USE OF FATALITIES VS. OBSERVED ON THE ROAD

History Since 1988, NHTSA has relied on the following assumptions and procedures to estimate the number of lives saved by safety belts at front-outboard seats each year:

- The belt use of **fatally injured** occupants on FARS is accurately reported.
- Belt effectiveness is 45 percent in passenger cars and 60 percent in light trucks.
- Thus, for every 55 belted, fatally injured passenger car occupants on FARS, there must exist 45 other people who were saved by the belt (and these people cannot necessarily be seen on FARS - e.g., if everybody in the crash survived), and for every 40 belted light truck fatalities on FARS there must exist 60 other people who were saved by the belt[45].

While developing the procedure, NHTSA staff identified four **different** use rates for belts:

U_1 = the **actual** use rate of fatally injured occupants as reported (accurately, we believe) on FARS. $U_1 = F_1 / (F_0 + F_1)$ where F_0 is the number of unrestrained fatalities and F_1 is the number of belted fatalities.

U_2 = the **hypothetical** belt use in potentially fatal crashes (UPFC), the use rate for the combined population of fatally injured people plus those who must have been saved by the belt. If effectiveness is E,

$$U_2 = \{F_1 + [E \times F_1 / (1-E)]\} / \{F_0 + F_1 + [E \times F_1 / (1-E)]\}$$

In other words, we add $[E \times F_1 / (1-E)]$, the hypothetical number of people who were saved by the belt, to both the numerator and the denominator of U_1.

---

[45] The basic assumptions and procedure were documented by Partyka, Susan C., *Lives Saved by Seat Belts from 1983 through 1987,* NHTSA Publication No. DOT HS 807 324, Washington, 1988. They were reconfirmed by Blincoe, Lawrence J., *Estimating the Benefits from Increased Safety Belt Use,* NHTSA Publication No. DOT HS 808 133, Washington, 1994, Appendix A; Klein, Terry M. and Walz, Marie C., *Estimating Lives Saved by Restraint Use in Potentially Fatal Crashes,* NHTSA Research Note, Washington, 1995.

U_3 = the **actual** belt use observed on the road for the general traffic population in state surveys, the National Occupant Protection Use Survey (NOPUS), or the 19-City Study.

U_4 = the reported belt use of crash survivors on FARS. This number is not considered meaningful, at least in absolute terms, because crash survivors have been overreporting their belt use since 1986, when belt use laws took effect in most of the large states (see Sections 1 and 4). It is further bloated by cases of occupants saved by the belt who got into the FARS system because some other person was killed in the crash.

U_1, U_2, and U_3 are meaningful numbers. The relationship between them was studied extensively at NHTSA. Specifically, U_1 is less than U_2, by definition, whereas U_2 for the overall potentially-fatal-crash population is empirically demonstrated to be less than U_3 for the general motoring public.

U_1 is less than U_2 because belt use of fatally injured occupants excludes the belt users who were saved by the belt, while UPFC includes them. The relationship between the actual rate U_1 and the hypothetical rate U_2 is deterministic, not statistical, and it can be specified exactly when effectiveness E is known. Indeed, U_2 is a hypothetical rate that cannot be observed, but can only be imputed from U_1 and E by the following formula:

$$U_2 = U_1 / \{1 - [E \text{ x } (1 - U_1)]\}$$

Conversely, $$U_1 = [(1 - E) \text{ x } U_2] / [1 - (E \text{ x } U_2)]$$

For example, if U_1 is 50 percent and E is 45 percent then U_2 is necessarily 64.52 percent.

U_2 is less than U_3 because the type of people who get involved in actually or potentially fatal crashes are less likely to use belts than the average person observed in seat belt surveys. It has long been known that belt users are, on the average, more cautious drivers than non-users, and underinvolved in severe crashes[46]. Alcohol-impaired drivers, risk-takers and people oblivious to safety are less likely to use belts and they are overinvolved in severe crashes. Another factor is that observational surveys are usually during the daytime, when belt use is higher, while many fatal crashes are at night.

However, the relationship of U_3 to U_2 is not deterministic but can vary over time and place. The only certainties are when U_3 is 0 percent, so is U_2 and when U_3 is 100 percent, so is U_2. At all

---

[46] Reinfurt, Donald W., Silva, Claudio Z. and Seila, Andrew F., *A Statistical Analysis of Seat Belt Effectiveness in 1973-75 Model Cars Involved in Towaway Crashes, Volume 1*, NHTSA Publication No. DOT HS 802 035, Washington, 1976, pp. v-x and 13. *Final Regulatory Impact Analysis, Amendment to Federal Motor Vehicle Safety Standard 208, Passenger Car Front Seat Occupant Protection*, NHTSA Publication No. DOT HS 806 572, Washington, 1984, p. IV-6.

intermediate points, U_2 lags behind U_3. NHTSA uses empirical statistical models[47] such as quadratic regressions over time and across states to model the "typical" lag of U_2 behind U_3. These models are needed to achieve two program objectives: (1) Predicting how many additional lives would be saved if on-the-road belt use increased by a specified number of percentage points; (2) Conversely, estimating U_3 for states that did not perform observational surveys that year, based on U_1 observed and U_2 calculated from FARS data for that year.

There are always two known, observable quantities U_1 and U_3 and two hypothetical constructs that cannot be directly observed, E and U_2. All of the above procedures simply **assumed** E was also known (45 percent in cars and 60 percent in light trucks) and used the three "knowns" U_1, U_3 and E to calculate the other unobservable quantity U_2 as well as the important statistic $U_3 - U_2$.

Analysis overview In the remainder of this report, the process will be reversed: we will assume we know U_2, the UPFC, and use it with U_1, the belt use of fatally injured occupants to estimate the effectiveness E. Just as in Sections 3-8, though, these point estimates, relying on several critical assumptions, are not like customary statistical estimates derived directly from the data. Their uncertainty can be discussed to some extent, but not be fully quantified, based on statistical theory, with "confidence bounds." The analysis is another empirical tool to support a qualitative judgement if NHTSA's long-standing estimates of 45 percent fatality reduction in cars and 60 percent in light trucks are still appropriate.

The above formulas for U_1 or U_2 can be solved for E:

$$E = [U_2 - U_1] / [U_2 - (U_2 \times U_1)]$$

Specifically, we will identify a **subset of occupants in potentially fatal crashes whose belt use is arguably the same as the average road-users included in observational surveys** - i.e., $U_2 = U_3$ for this subset. Then we can substitute the known, observed U_3 for the unobservable U_2 in the preceding formula, and together with the belt use U_1 of the fatalities in that subset, estimate the effectiveness:

$$E = [U_3 - U_1] / [U_3 - (U_3 \times U_1)]$$

For example, if 50 percent of the fatalities were belted in this special subset, and belt use on the road were 65 percent, belt effectiveness would be 46.15 percent.

In 1989, Susan Partyka took the first step in the direction of this analysis when she computed U_1 and U_2 for various subgroups in 1987-88 FARS - still assuming E = 45 percent in deriving U_2 from U_1 - and then compared them to U_3 = 44 percent, the observed belt use in passenger cars during

---

[47] Partyka, Susan C. and Womble, Kathleen B., *Projected Lives Savings from Greater Belt Use*, NHTSA Research Note, Washington, 1989. Blincoe (*op. cit.*), Appendix B.

1987-88 in 19 U.S. cities[48]. Although U_2 for the entire FARS was well below 44 percent, U_2 for drivers without police-reported alcohol use was 46 percent and U_2 for drivers in daytime, weekday crashes, without police-reported alcohol use or previous crash involvements was 51.5 percent, actually exceeding belt use observed in the 19 cities. This study confirmed that the agency's 45 percent effectiveness estimate was more plausible than higher numbers proposed by some researchers. It also demonstrated some groups had lower-than-average belt use: alcohol-impaired drivers, people with bad driving records, and nighttime traffic.

The analysis in this report will be based on **1991-99 FARS**, a large data base. Belt use of fatally injured occupants will be compared to belt use on the road, derived from national averages of **1991-99 state observational surveys**. We will define a FARS subgroup that arguably ought to have UPFC equal to observed belt use on the road: **non-drinking, "non-antisocial," daytime involvements in multivehicle crashes** [or, alternatively, non-drinking, "non-antisocial," daytime involvements in all crashes]. This group can be defined on a fairly consistent basis across the states and avoids some biases inherent in FARS data, as will be explained later. U_1 will be obtained from FARS for passenger cars with 3-point belts and for light trucks. The heart of the analysis is an adjustment of the 1991-99 state belt use rates, with the help of NOPUS and other data, to obtain separate U_3 rates for cars with 3-point belts and for light trucks that are comparable to U_1. Finally, U_1 and U_3 will be used to compute the effectiveness E - but with a recognition that effectiveness in this special subset of crashes (e.g., limited to multivehicle) should not necessarily be the same as for all crashes.

Issues that must be considered in making state survey results comparable to FARS are:

- State surveys generate a single use rate each year, aggregating cars with 3-point belts, cars with 2-point automatic belts, and light trucks. We need to tease out separate rates for cars with 3-point belts and for light trucks, based on differentials seen in other data.

- Before 1998, a fair number of state surveys omitted some or all light trucks if the law in that state did not cover light trucks and/or for other reasons. Many surveys omitted local roads, and a few were limited to drivers. Results need to be adjusted to reflect all front-outboard occupants of all types of light vehicles on all roads.

- Observational surveys have a rural-urban mix proportional to VMT, but FARS cases have proportionately more rural crashes.

---

[48] Partyka, Susan C., *Belt Use in Serious Impacts Estimated from Fatality Data*, NHTSA Publication No. DOT HS 807 519, Washington, 1989. Bowman, Brian L. and Rounds, Donald A., *Restraint System Usage in the Traffic Population - 1988 Annual Report*, NHTSA Publication No. DOT HS 807 447, Washington, 1989.

FARS data analysis A basic working assumption here, as elsewhere in this report and other NHTSA studies, is that the belt use of fatally injured occupants is reported accurately in FARS, or at least without a net bias (net over- or underreporting).

Another given is that we cannot prove with existing data that any specific subgroup of people involved in fatal or potentially fatal impacts actually has the same belt use as the average person on the road. The people "saved by the belt," who are a part of the subgroup, simply aren't in the data. We can only make a logical argument, supported by the literature, that this subgroup ought to have the same belt use as the general population.

This is fundamentally a process of excluding groups with behaviors that are typically associated with lower-than-average belt use, leaving a kernel of people "just like you and me" who had the misfortune of becoming involved in a fatal crash. The process needs to be reasonably consistent across states and from year to year (1991-99).

As noted above, an excellent starting point is to exclude alcohol-impiared drivers and their accompanying RF passengers. However, the completeness of FARS/police reporting of alcohol involvement varies a lot from state to state[49]. Nineteen states, including California, Texas, Florida and Illinois have "not reported," "unknown" or missing data on over 40 percent of driver fatalities in cars or light trucks during 1991-99. Moreover, in some states (such as Texas and Missouri) these codes are often a surrogate for "no alcohol" while in others they really mean "unknown."

That raises a question of whether to include only those specifically reported as "no alcohol" (and exclude the unknowns) or merely to exclude those positively reported as "alcohol-involved" (and include the unknowns). Limiting to the "no alcohol" cases is unsatisfactory because it would delete almost all the Texas and Missouri data and over 70 percent of California, Illinois, Georgia, Wisconsin, etc. cases: what remains would not be nationally representative. On the other hand, excluding only the "alcohol-involved" cases and retaining the unknowns would leave a fair number of drinking drivers in the file: 17 states, including New York, Florida, North Carolina and Georgia, list under 20 percent of their cases as positively alcohol-involved (hard to believe - except Utah).

Thus, it is appropriate to exclude all driver records, and their accompanying RF passengers, that were positively coded as "alcohol-involved," but that is only a starting point. We must identify other behaviors that should be excluded, and that will make the exclusion rate more uniform across states.

A more sweeping approach that turns out to be unsuccessful is to rely on the widely-accepted theory of "induced exposure": the non-culpable party(s) in a multivehicle crash is for all practical

---

[49] Klein, Terry M., *A Method for Estimating Posterior BAC Distributions for Persons Involved in Fatal Traffic Accidents,* NHTSA Publication No. DOT HS 807 094, Washington, 1986.

purposes a typical road user who just happened to be in the wrong place at the wrong time[50]. That seems exactly what we need, even if it limits us to certain multivehicle involvements, because we can be confident this group will have UPFC equal to the belt use of typical road users.

The FARS variables, "Related Factors - Driver Level," DR_CF1 - DR_CF4 identify culpable drivers to be excluded: anybody with one or more of these variables equal to 6, 26, 27, 28, 30, 31, 35, 36, 38, 39, 44, 48, 50, 51, 57 or 58. Most of these codes are specific, momentary driving errors that lead directly to collisions: lane departure, failure to obey traffic signal, failure to yield right of way, driving too fast for conditions, following too closely, etc. They are generally not underlying behavioral problems or long-term conditions (although the culpable drivers may have some of those, too). The non-culpable drivers to be included in the analysis are a large group with no DR_CF codes at all, and a smaller group with other codes than the above, generally not pertaining to specific driving errors (e.g., glare, crosswind).

Unfortunately this approach does not work with FARS data, even though it has been quite successful with primarily nonfatal state data. The DR_CF codes in FARS are severely confounded with whether or not the driver survived. A disproportionate share of the driving errors are attributed to the fatally injured driver in a multivehicle crash. This becomes clear, for example, in head-on, front-to-front collisions of two vehicles. Intuitively, which driver dies and which survives ought to have a lot to do with the relative weights of the vehicles and ages of the drivers, and little to do with who "started" the collision by crossing the center-line. Head-on is pretty much head-on, regardless of how or where it happened. But in FARS, based on the DR_CF codes, 65 percent of the fatally injured drivers in these collisions were culpable, but only 32 percent of the surviving drivers. There is a strong tendency to blame the crash on the dead driver, and it persists even when the cases are subdivided by driver age, vehicle type, etc. We certainly would not want to analyze belt effectiveness in a group whose selection factor (culpable or non-culpable) is highly confounded with the principal outcome variable of our analysis (survival or non-survival).

The remedy is to find an intermediate approach that works with FARS data. We need to exclude a wider range of antisocial behaviors than just the drinking driver, but not exclude all culpable drivers. The exclusions should be reasonably consistent across states (unlike the FARS alcohol-involvement variable), and not confounded with crash survival (unlike DR_CF as treated above).

The following drivers are excluded from the analysis because they display forms of antisocial behavior that could indicate a lower belt use rate than the average driver on the road. RF passengers are excluded if their accompanying driver exhibits any of these behaviors:

---

[50] Haight, Frank A., *Indirect Methods for Measruing Exposure Factors as Related to the Incidence of Motor Vehicle Traffic Accidents*, NHTSA Publication No. DOT HS 800 601, Washington, 1971. Cerrelli, Ezio, *Driver Exposure: Indirect Approach for Obtaining Relative Measures*, NHTSA Publication No. DOT HS 820 179, Washington, 1972.

- FARS positively reports alcohol involvement (DRINKING = 1)
- FARS positively reports drug use (DRUGS = 1)
- Driving without a license, or with a suspended, revoked, expired, canceled or denied license (L_STATUS = 0, 1, 2, 3, 4 in 1991-99)
- A sub-par driving record as indicated by any one (or more) of the following during the 3 years preceding the fatal crash involvement:
  - **S** at least 1 DWI conviction, or
  - **S** at least 2 speeding convictions, or
  - **S** at least 2 other harmful-moving-violation convictions, or
  - **S** at least 2 license suspensions or revocations, or
  - **S** at least 2 police-reported crash involvements
- Possibly antisocial behavior(s) at the time of the crash, as evidenced by any of DR_CF1 - DR_CF4 having the following values:

19	driving on suspended or revoked license
36	erratic, reckless, careless or negligent driving
37	attempting to escape police pursuit
46	racing
90	hit and run
91	vehicular homicide

Since Florida, Kansas, North Carolina, Ohio and Utah encode “erratic, reckless, careless or negligent driving” far more often than other states, cases will be excluded from these states if DR_CF1 = 36, but DR_CF2, DR_CF3 or DR_CF4 = 36 will not be grounds for exclusion.

All other drivers, and their accompanying RF passengers will be considered non-antisocial and become candidates for inclusion in the study. Inattentive driving due to a temporary distraction such as talking or eating (DR_CF = 6) is not considered “antisocial behavior” in this analysis.

The proportion excluded is fairly consistent from state to state, ranging from 15 percent of Rhode Island daytime driver fatalities in 1994-99 to 45 percent of Kansas fatalities; however 41 of the 51 states plus D.C. are in a relatively narrow range of 21 to 34 percent excluded. What remains after the exclusions is nationally representative. This is a better filter than the DRINKING variable alone, which resulted in near-zero exclusions from some important states. It is even more consistent year-to-year, showing a gradual downward trend from 51 percent excluded in 1977, to 45 percent in 1991, to 41 percent in 1999, consistent with the real reductions in alcohol-impaired driving and some other antisocial driving behaviors in the United States during the past 25 years.

Unlike “culpability,” FARS just minimally confounds this variable with the probability of surviving a crash. In head-on, front-to-front collisions of two vehicles of any type, 33 percent of the fatally injured drivers were excluded, but only 28 percent of the surviving drivers. In passenger cars, these percentages are an even closer 33 and 31, respectively.

Most important, it is plausible that the remaining “non-antisocial” drivers and their accompanying RF passengers buckle up as much as the average road user. A wide range of habitual risk-taking, aggressive or safety-indifferent behaviors - likely to be associated with a low rate of belt use - has been excluded. Although the "non-antisocial" group still includes a fair number of "culpable" drivers, these are mostly one-time, specific driving errors, without a history of previous errors. Although they were "culpable" on that occasion, they are still "average" road users - not perfect drivers, but susceptible to human error from time to time. Mere "culpability" need not indicate low belt use.

Since belt use surveys are conducted in the daytime, the analysis is further limited to daytime crashes, specifically from 7:00 A.M. to 6:59 P.M.

A question arises whether to include single-vehicle crashes. If running off the road *per se* demonstrates aggressive, risk-taking behavior, it would be appropriate to limit the analysis to multivehicle daytime non-antisocial crash involvements. But if running off the road, in the absence of specific antisocial behaviors, is merely one of the errors occasionally committed by the average driver, it would not be necessary to exclude single vehicle crashes. Both approaches will be used here.

The 1991-99 FARS data included the following N's of front-outboard fatalities in daytime, non-antisocial crash involvements, with belt use reported as either yes (REST_USE = 1, 2, 3, 8, 13) or no (REST_USE = 0), in model year 1975-99 passenger cars and light trucks equipped with 3-point belts:

- 37,713 multivehicle fatalities in passenger cars (57.5 percent belted)
- 11,845 multivehicle fatalities in light trucks (42.3 percent belted)
- 49,612 single + multivehicle fatalities in passenger cars (53.1 percent belted)
- 19,997 single + multivehicle fatalities in light trucks (36.6 percent belted)

These are large N's and they enable precise calculation of U_1, the belt use rate of fatally injured occupants (sampling error based on $1.96 \sqrt{pq/n}$ is less than 1 percentage point in each case). The next tasks are to obtain U_3, belt use observed on the road, that we assert equals U_2, belt use in potentially fatal crashes, and compute E, the fatality reducing effectiveness of belts. But first let us take a step back and use double-pair comparison to estimate what kind of effectiveness we might expect to see in daytime, non-antisocial crash involvements.

Anticipated effectiveness, based on double-pair comparison Many rollovers occur at night and/or involve drivers classified as "antisocial." Table 2 showed belts are especially effective in rollovers. Thus, we cannot expect safety belts to be as effective in daytime, non-antisocial crashes as they are overall (45 percent in passenger cars, 60 percent in light trucks). Table 6 showed even lower effectiveness in multivehicle crashes, where almost all rollovers are excluded. Table 13 estimates belt effectiveness in daytime, non-antisocial crash involvements, based on the methods of Sections 3-8: double-pair comparison of 1986-99 FARS data, corrected by the UEF.

TABLE 13: FATALITY REDUCTION IN DAYTIME, NON-ANTISOCIAL CRASH INVOLVEMENTS - BASED ON DOUBLE-PAIR COMPARISON OF 1986-99 FARS

	Passenger Cars		Light Trucks	
	Fat. Red.	Mser[51]	Fat. Red.	Mser
In MULTIVEHICLE daytime non-antisocial crash involvements	29	± 10	42	± 12
In ALL daytime non-antisocial crash involvements	38	± 8	52	± 8

The point estimate of effectiveness for passenger cars in daytime, non-antisocial multivehicle crashes is just 29 percent, with minimum sampling error range ± 10 percentage points. In other words, if the analysis method of this section produced an E anywhere from 19 to 39 percent it would still be statistically consistent with the double-pair comparison result, and a validation of the method of Sections 3-8, but the closer it is to 29 percent, the better. Similarly, we would anticipate a still-low 38 percent fatality reduction when single-vehicle crashes are included for passenger cars. For light trucks, the anticipated effectiveness levels are 42 and 52 percent, also lower than the overall 60 percent fatality reduction in light trucks.

Based on the formula:

$$U_3 = U_2 = U_1 / \{1 - [E \text{ x } (1 - U_1)]\}$$

the belt-use rate of 57.5 percent for fatally injured occupants of passenger cars in the multivehicle crashes, and the 29 percent effectiveness suggest an anticipated 65.6 percent belt use on the road in cars during 1991-99. The belt-use rate of 53.1 percent in cars for single + multivehicle fatalities, combined with 38 percent effectiveness anticipate a quite similar 64.6 percent belt use on the road. The anticipated belt-use rates on the road for light trucks, calculated the same way, are 55.8 percent and 54.6 percent, respectively. As a first cut, we can see these numbers are reasonably close to belt use actually observed on the road during the 1990's, and suggest the remaining analyses are likely to validate the double-pair comparison estimates. Now, let us proceed with adjusting and interpreting the state observational survey data, so we may calculate a real E from a real U_3 rather than an anticipated U_3 from an anticipated E.

---

[51] Mser = minimum sampling error range (± 1.96 standard deviations) based on the formula in Section 7; does not include possible non-sampling errors from assuming the UEF is applicable for all collision types, etc.

Analysis of state survey data NHTSA has an estimate of statewide belt use observed on the road, for every state plus D.C. in every year[52], 1991-99, as shown in the Appendix. A basic working assumption here is that each of these numbers is an unbiased and reasonably precise estimate of whatever that state purported to survey that year. For example, if a state claimed they surveyed passenger cars on major roads in 1992, the number is indeed a good estimate of belt use in passenger cars on major roads. The assumption is not as simple as it sounds. To be sure, NHTSA is confident that no state bloated its numbers by systematically inaccurate reporting or by deliberately selecting observation sites known or expected to have high belt use. However, before 1998 some of the survey designs departed to varying degrees from probability sampling. The earlier sample designs could have inadvertently created biases in some of the results. Thus, the working assumption is that there is no systematic upward bias, as long as we understand that many states did not pretend to sample all types of vehicles and all types of roadways. We will need to adjust for the non-sampled vehicle types, roadway types, and seat positions.

The starting points for the analysis are the annual U.S. estimates of seat belt use, 1991-99, when the 51 state estimates plus D.C. are **weighted by vehicle miles of travel (VMT)** as shown in Table 14.

TABLE 14: VMT-WEIGHTED NATIONAL AVERAGE
OF BELT USE OBSERVED AND REPORTED BY THE STATES

1991	58.3 percent
1992	60.9
1993	65.0
1994	66.3
1995	67.1
1996	66.5
1997	67.5
1998	68.6
1999	70.1
1991-99 average	65.6 percent

These numbers exaggerate actual belt use by all front-outboard occupants in all passenger vehicles on all roads in the United States, especially in the earlier years, because some states by design excluded RF occupants, light trucks, and/or local roads from their surveys. One goal is to adjust the numbers until they reflect all passenger vehicles on all roads, as in the NOPUS surveys. The above rates combine cars with 3-point belts, light trucks, and cars with automatic 2-point belts. Another goal is to obtain separate rates for each vehicle type, and to adjust these rates for the roadway mix that prevails in FARS cases (more rural and less urban than overall VMT).

---

[52] On occasions when [primarily small] states did not survey in a given year, NHTSA used the previous year's number. For one small state, NHTSA estimated on-the-road belt use from FARS data using Partyka's model before the state commenced observational surveys.

Belt use parameters The strategy is to disaggregate national belt use and model it as a primarily linear function of five discrete variables: vehicle type, belt law, road type, seat position and calendar year:

4 vehicle types
    cars with 3-point belts
    cars with 2-point belts
    pickup trucks
    vans/SUVs

2 belt law conditions (depending on vehicle type)
    covered by law
    not covered by law

4 road types
    major rural (interstate + arterial)
    local rural (local + collector)
    major urban (interstate +freeway + arterial)
    local urban (local + collector)

2 seat positions
    driver
    right front (RF)

9 calendar years: 1991 through 1999

Thus, BELTUSE(vehicle, law, road, seat, CY) will be estimated for the United States for a total of 4x2x4x2x9 = 576 conditions. We will try to make this a linear model without interaction terms. In other words, if the driver has belt use 4 percentage points higher than the RF passenger in pickup trucks with a belt law on major rural roads in 1991 that will also be true in vans/SUVs without a belt law on local urban roads in 1992, etc. (The only exception to this independent linear approach, as we shall see, will be in cars with automatic 2-point belts.) It is necessary to separate pickup trucks and vans/SUVs in the model, because they have different belt use rates and laws, but later they will be combined into a single “light truck” estimate.

The parameters for this model will now be derived from various sources, including NOPUS, the state surveys themselves, and sometimes, survivor belt use on FARS (percentage point differences between rates, not absolute rates).

The effect of seat position is the easiest to derive, since all three full NOPUSes (1994, 1996 and 1998) observed and reported belt use separately for drivers and RF. The average belt use in those three surveys was 4.0 percentage points higher for the driver than the RF.

Belt use in any type of vehicle, except cars with automatic 2-point belts, is 24.3 percentage points lower if there is no belt law than if there is a belt law. This estimate is based on state surveys for CY 1991-94, comparing average belt use in states with laws to use in states without laws, in each calendar year.

Now, let us tackle the issue of belt use in cars with automatic 2-point belts, since this information will be needed for some of the other parameter estimates. Unlike all other vehicle types, belt use in cars with automatic 2-point belts did not change throughout 1991-99, as evidenced by use rates of FARS survivors. (Although these use rates are not credible as absolute numbers, they can be trusted as evidence that belt use did not change from year to year.) It is also unaffected by belt laws. The absolute use rate for drivers was a more-or-less constant 84 percent during 1991-99 (higher than the national-average use of 3-point belts in any vehicle type, road type or year). This is the average of 78.2 percent observed in the 1994 NOPUS (controlled intersections, cars with decodable VINs) and the 89.5 percent observed in 19 cities during 1990-91[53]. Given 84 percent use by drivers, that implies 80 percent for RF passengers (and 83 percent for a 75:25 driver:RF mix) in each CY, on all roads, with or without a belt law

All else being equal, belt use in pickup trucks is 10.8 percentage points lower than in cars with 3-point belts. In vans/SUVs it is 0.6 percentage points higher than in cars with 3-point belts. The differentials are estimated from the 1998 NOPUS, as follows. Belt use in all passenger cars is 71.3 percent in 1998 NOPUS. Since 13.2 percent of passenger cars in 1998 had 2-point belts (see below), used by 83 percent of front-seat occupants, the use rate in cars with 3-point belts would have to be 69.5 percent to get a 71.3 percent average for all cars. That compares to 58.7 in pickups (10.8 percentage points lower than 69.5) and 70.1 in vans/SUVs (0.6 percentage points higher)[54].

The highest belt use rate is on major urban roads. The percentage point decrements in belt use on other roadway types are: 3.6 on major rural, 10.9 on local rural, 2.7 on local urban, as seen in the reported belt use of survivors in daytime non-antisocial multivehicle involvements. These decrements are for passenger cars with 3-point belts, but are also approximately right for pickup trucks, vans and SUVs[55]. Cars with automatic 2-point belts had similar belt use on all roadways.

---

[53] *Evaluation of the Effectiveness of Occupant Protection - Federal Motor Vehicle Safety Standard 208 - Interim Report, June 1992,* NHTSA Publication No. DOT HS 807 843, Washington, 1992, p. 13, weighting the three types of 2-point belts in Table 2-2 at 45:40:15.

[54] Belt use rates among FARS survivors of daytime non-antisocial multivehicle involvements show similar differentials - viz., a 9 percentage point decrement for pickups and a 1 percentage point increment for vans/SUVs over passenger cars with 3-point belts in 1998.

[55] The 1996 NOPUS also estimates belt use separately for “major” and “local” roads in “rural” and “non-rural” areas. It, too, shows highest belt use on “major non-rural” roads (67.1%) and lowest on “local rural” roads (57.4%). However, it is preferable to rely on the road-type definitions and belt-use differentials of FARS survivors, since they obviously correspond exactly to categories that are defined for the FARS fatalities, and also to FHWA *Highway Statistics*.

Finally, belt use could change (generally increase) from year to year during 1991-99 (except in cars with automatic 2-point belts). However, the increments cannot be estimated at this point. That will be the final step in developing the model, and it will be derived from the state survey data, after they have been adjusted to create NOPUS-like estimates.

Weight factors Now that we have deconstructed belt use into 576 cells, with a use rate in each cell, we may reconstruct national overall use rates by assigning a weight factor to each cell and taking the weighted average. Three sets of weight factors will be defined, allowing three types of national estimates:

- A national factor, NATLWGT(vehicle, law, road, seat, CY) that simply weights each cell in a given calendar year according to its proportion of the nation's VMT in that year (or, more correctly, its proportion of front-outboard occupant miles of travel). When all the cells are weighted by NATLWGT, the result should be a "true" estimate of national belt use in each year, comparable (but not necessarily equal) to what NOPUS generated in 1994, 1996 and 1998.

- A state factor, STATEWGT that equals NATLWGT on those cells where all 51 state + D.C. surveys observe belt use - e.g., passenger car drivers on major roads - but is lower than NATLWGT on those cells where some of the states did not observe belt use - e.g., pickup truck passengers, on local roads, in 1991. Each state is weighted by its share of VMT; STATEWGT/NATLWGT equals the share of VMT for the states that did survey belt use in that cell divided by the VMT for all states plus D.C. We will then set the CY parameters for BELTUSE so that when all the cells are weighted by STATEWGT, the results will **exactly** equal the annual VMT-weighted national averages of belt use observed and reported by the states, as shown in Table 14. In other words, the basic state estimates in Table 14 drive the rest of the model.

- A FARS factor, FARSWGT that weights each cell according to its proportion of the nation's fatalities rather than VMT. We will generate one set of FARSWGT for multivehicle daytime non-antisocial involvements, and another for single + multivehicle daytime non-antisocial involvements. When all the cells are weighted by one of these FARSWGT, the result will be an estimate of belt use in potentially fatal crashes (UPFC) we need to complete our analysis.

NATLWGT The procedure for obtaining NATLWGT has two guidelines: (1) Whenever possible, it should use a simple multiplicative model, p(A and B) = p(A) x p(B), with interaction terms (conditional probabilities) only when they are really needed. We will obtain weight factors for each of the parameters (vehicle type, law, road type, seat position) separately, or two-at-a-time, and multiply them to obtain the cell weight. (2) The cell weights in each CY add up to 1.

The weight factors for seat position are 0.75 for the driver and 0.25 for the RF passenger. The 3 to 1 ratio of drivers to passengers has been consistently observed in numerous data bases over many years[56].

Belt laws in many states always applied to all passenger vehicles, but before 1998 there were a fair number of states that exempted all light trucks, or just pickup trucks, from the laws. The percentages of vehicles in the United States covered by state belt laws (i.e., weighting the states by VMT), by calendar year, are shown in Table 15[57]. In other words, the "belt law" weight factor for passenger cars in 1991 is 0.91 and the "no belt law" weight factor is 0.09.

TABLE 15: PERCENT OF VEHICLES COVERED BY BELT LAWS

	Passenger Cars	Pickup Trucks	Vans & SUVs
1991	91.0	64.6	78.1
1992	93.1	65.2	78.6
1993	93.7	69.1	79.3
1994	97.7	73.4	83.6
1995	99.0	74.7	84.9
1996	99.6	78.9	85.4
1997	99.6	78.9	85.4
1998 & 99	99.6	99.6	99.6

The ratio of passenger cars to light trucks in NOPUS (weighted, moving-traffic data) should be the same as their ratio of VMT. In 1994, the NOPUS data were 62 percent cars and 38 percent light trucks[58]. Comparable numbers for the 1996 and 1998 NOPUS are 58 percent and 57 percent cars, respectively. Smoothing out this trend gives approximately 64 percent cars in 1991, and then declining by 1 percentage point each year, down to 56 percent cars in 1999. Light trucks' share of VMT increased 1 percentage point a year from 36 percent in 1991 to 44 percent in 1999.

The passenger car VMT may be subdivided into cars equipped with 3-point belts or with automatic 2-point belts, based on the distribution of FARS survivors in daytime non-antisocial multivehicle

---

[56] In the 1994, 1996 and 1998 NOPUS, the average ratio was 73 to 27.

[57] Based on the "Key Provisions of Safety Belt Use Laws" tables in *Fatal Accident Reporting System 1991* and *Traffic Safety Facts 1992* through *Traffic Safety Facts 1998,* NHTSA publications. The information is summarized in the Appendix of this report.

[58] The 1994 NOPUS belt use results, 62.8 percent in cars and 50.2 percent in light trucks, for an overall average of 58.0 percent, imply that 62 percent of the vehicles in the sample were cars and 38 percent were light trucks.

crash involvements, by calendar year. The light truck VMT is split into pickup trucks or vans/SUVs, based on registration data from the R.L. Polk National Vehicle Population Profile. The VMT shares for the four vehicle types are shown in Table 16. Above all, vans and SUVs increased at the expense of passenger cars. Pickup trucks remained constant. Two-point automatic belts increased at first but began phasing out after 1994 because no new ones were entering the fleet.

TABLE 16: SHARES OF THE NATION'S VMT

	Cars, 3-Pt. Belts	Cars, 2-Pt. Belts	Pickup Trucks	Vans & SUVs
1991	58.4	5.6	22.4	13.6
1992	56.5	6.5	22.3	14.7
1993	54.3	7.7	22.3	15.7
1994	52.6	8.4	22.8	16.2
1995	51.8	8.2	22.8	17.2
1996	50.9	8.1	22.7	18.3
1997	50.1	7.9	22.7	19.3
1998	49.5	7.5	22.6	20.4
1999	49.2	6.8	22.4	21.6

The Federal Highway Administration (FHWA) collects annual statistics on VMT by roadway system. In 1993 (and these distributions change little from year to year), the VMT shares were[59]:

24.3% major rural (interstate + arterial)
14.4% local rural (local + collector)
47.4% major urban (interstate +freeway + arterial)
13.9% local urban (local + collector)

The FHWA does not count VMT separately by vehicle type. We will use the above distribution for all vehicle types. Intuitively, light trucks should have proportionately somewhat more rural VMT and cars more urban. However, FARS data on crash survivors suggests the differences are relatively small.

NATLWGT is the product of the parameters defined above. For example, the proportion of occupant miles traveled in 1991 that were drivers of passenger cars equipped with 3-point belts, in states with belt laws for cars, on major rural roads is:

$$0.75 \times 0.91 \times 0.584 \times 0.243 = 0.0968549$$

---

[59] Teets, Mary K., *Highway Statistics 1993*, Publication No. FHWA-PL-94-023, Federal Highway Administration, Washington, 1994, p. V-116. By FHWA's definition, a road is either "rural" or "urban"; most roads in suburbs are classified as "urban."

For RF passengers of pickup trucks in states without belt laws on local rural roads in 1996, it is:

$$0.25 \times 0.211 \times 0.227 \times 0.144 = 0.0017242$$

STATEWGT is computed for any cell by multiplying NATLWGT by the VMT-weighted proportion of state surveys that included the cell. We will obtain survey-inclusion factors for each of the parameters (vehicle type, law, road type, seat position) separately, or two-at-a-time, and multiply them to obtain the cell inclusion factor.

D.C., Kentucky, Missouri and West Virginia, comprising 5 percent of the nation's VMT, surveyed only drivers in 1996-97. All other states in 1996-97, and all states in 1998-99 surveyed drivers and RF passengers. In 1994, it is known that six states (or 5 states plus D.C.) surveyed only drivers, but it is unknown which ones. Let us assume those six jurisdictions included D.C., Kentucky, Missouri and West Virginia, and comprised 7.5 percent (5 x 6 / 4) of the nation's VMT. Let us assume that the coverage in 1991-93 was the same as in 1994, but in 1995 it was midway between 1994 and 1996 - i.e., 6.3 percent of the surveys were limited to drivers. In other words, the survey inclusion parameter for drivers is always 1.0. For RF passengers, it is 0.925 in 1991-94, 0.937 in 1995, 0.95 in 1996-97 and 1.0 in 1998-99.

All states surveyed passenger cars, regardless of whether or not the state had a belt law. In 1998-99, all states surveyed all passenger vehicles - cars, pickup trucks, vans and SUVs. During 1996-97, however, the nine states listed in the Appendix surveyed only passenger cars, but no light trucks. Interestingly, some states had belt laws for pickup trucks and/or vans and SUVs but did not survey them, while others did not have belt laws but did survey them. Based on the information in the Appendix, in 1996-97, a VMT-weighted 84.4 percent of the states with belt laws for pickup trucks also surveyed them, and 48.2 percent of the states without belt laws surveyed them. The corresponding inclusion rates for vans and SUVs are 81.2 percent and 50.9 percent.

In 1994, it is known that 24 state surveys were limited to passenger cars only. This is apparently the sum of all 19 states that did not have belt laws for pickup trucks in 1994, plus the 5 states that had laws but were still not surveying light trucks in 1996-97. Therefore, we may assume that also in 1991-93, any state without a belt law for pickup trucks or vans/SUVs also did not survey them, and neither did the 5 states that had belt laws but still did not survey light trucks in 1996-97. We may also assume that the inclusion rates for 1995 were midway between 1994 and 1996. These assumptions result in the survey inclusion parameters shown in Table 17.

TABLE 17: STATE SURVEY INCLUSION PARAMETERS - VEHICLE TYPE AND BELT LAW

Vehicle Type ☞	Passenger Cars		Pickup Trucks		Vans and SUVs	
In States With ☞	Belt Law	No Law	Belt Law	No Law	Belt Law	No Law
1991	1.0	1.0	0.807	0.0	0.792	0.0
1992	1.0	1.0	0.811	0.0	0.795	0.0
1993	1.0	1.0	0.821	0.0	0.797	0.0
1994	1.0	1.0	0.829	0.0	0.805	0.0
1995	1.0	1.0	0.835	0.241	0.810	0.254
1996	1.0	1.0	0.844	0.482	0.812	0.509
1997	1.0	1.0	0.844	0.482	0.812	0.509
1998-99	1.0	1.0	1.0	1.0	1.0	1.0

All states surveyed belt use on “major” roads in one form or another, but quite a few states limited or entirely omitted observation on “local” roads (where belt use is lower). NHTSA believes at least ten states adapted Westat’s design for the NOPUS, which includes and gives appropriate weights to all types of roads[60]. Six states designed other types of surveys that clearly included local as well as major roads. These 16 states comprise just over a of the Nation’s VMT. At least ten states (26% of U.S. VMT), such as Michigan, sampled and observed only at controlled or signalized intersections[61]. NHTSA believes controlled-intersection surveys “miss” about half the VMT on local roads[62]. At least seven states (14% of VMT) limited their surveys to major roads, such as those inventoried in their Highway Performance Management Systems. Survey designs for other states are not currently on file at NHTSA, but many of them are known not to be the Westat type.

On the whole, it appears that approximately a of the state surveys (VMT-weighted) were Westat designs or other designs that included all types of local roads, a were based on controlled intersections and a were limited to major roads. Controlled-intersection surveys miss about half

---

[60] Rood, Debra H., Kraichy, Patricia P. and Carubia, Jean, *Evaluation of New York State’s Mandatory Occupant Restraint Law*, NHTSA Publication No. DOT HS 806 950, Washington, 1985. *National Occupant Protection Use Survey*, Washington Consulting Group, Washington, 1995.

[61] Wagnenaar, A.C., Wiviott, M.B.T. and Compton, C.P., *Direct Observation of Seat Belt Use in Michigan: April 1985*, Publication No. UMTRI 85-26, University of Michigan Transportation Research Institute, Ann Arbor, 1985.

[62] The 1996 NOPUS results for controlled intersections were halfway between the moving-traffic results for “all roads” and “major roads.”

the VMT on local roads. Thus, the three types of state surveys, overall, are equivalent to missing ½ of the local-road VMT but capturing all major-road VMT.

STATEWGT is the product of NATLWGT and the state-survey inclusion parameters defined above. For example, in 1991, drivers of passenger cars equipped with 3-point belts, in states with belt laws for cars, on major rural roads, were fully included in state surveys. Here, all inclusion parameters are 1.0.

STATEWGT = (1 x 0.75) x (1 x 0.91 x 0.584) x (1 x 0.243) = 0.0968549 = NATLWGT

But RF passengers of pickup trucks in states without belt laws on local rural roads were extensively omitted from state surveys in 1996.

STATEWGT = (0.95 x 0.25) x (0.482 x 0.211 x 0.227) x (0.5 x 0.144) = 0.0003947
STATEWGT = 0.229 x NATLWGT

FARSWGT The important difference between VMT and fatal crashes is that a disproportionate number of fatalities are on rural roads. Belt use is somewhat lower on rural than on urban roads. We must assign greater weight to the rural roads in FARSWGT than in NATLWGT.

The percentage distributions of fatalities in **multivehicle**, daytime, non-antisocial crash involvements during 1991-99 by roadway type, by vehicle type, were:

	Cars, 3-Pt. Belts	Cars, 2-Pt. Belts	Pickup Trucks	Vans & SUVs
Major rural	38.0	35.8	46.1	46.6
Local rural	23.2	21.9	31.0	22.1
Major urban	31.9	34.7	19.3	26.4
Local urban	6.9	7.6	3.6	4.9
	100	100	100	100

The corresponding percentages for **all** daytime, non-antisocial crash involvements were:

Major rural	36.4	35.1	41.7	47.2
Local rural	26.4	25.0	37.6	25.4
Major urban	29.4	31.7	16.3	22.3
Local urban	7.8	8.2	4.4	5.1
	100	100	100	100

For the other parameters - seat position, belt law and vehicle type - we may use the same factors as NATLWGT[63].

For example, in the multivehicle crashes, the FARSWGT for drivers of passenger cars equipped with 3-point belts, in states with belt laws for cars, on major **rural** roads in 1991 is:

$$0.75 \times 0.91 \times 0.584 \times 0.380 = 0.15146$$

and it is higher than NATLWGT (0.0968549), because fatality rates per VMT are higher on rural roads.

Calibration of the belt-use-by-CY parameter We may now proceed to fill in the remaining gap in the model. Based on the STATEWGT factors for the 576 cells, and the linear effects of vehicle type, belt law, road type, and seat position on belt use that we have already calibrated, the left column of Table 18 shows exactly what percent driver belt use, on major urban roads, in belt-law states, in cars with 3-point belts, is needed in each calendar year to obtain exactly the overall average belt use reported in the state surveys. Those averages are shown in the right column of Table 18, and they are copied from Table 14.

TABLE 18: CALIBRATION OF THE BELT-USE-BY-CY PARAMETER

	Must have this belt use by drivers, cars, 3-pt. belts major urban roads, belt law	To obtain this actual National average from state surveys
1991	62.8	58.3
1992	64.9	60.9
1993	69.1	65.0
1994	69.8	66.3
1995	71.4	67.1
1996	71.3	66.5
1997	72.5	67.5
1998	73.2	68.6
1999	74.9	70.1

The nine numbers in the left column drive the model, since the belt use rate in every cell can be computed from one of them (or is already known, for cars with 2-point belts), and then averaged

---

[63] For example, drivers constitute 75 percent of fatalities as well as occupants. While it is true that passenger cars are overrepresented and light trucks are underrepresented among occupant fatalities, since we intend to compute UPFC separately for cars and light trucks, and never to average the two, it is not necessary to develop new weights by vehicle type.

using STATEWGT for a National average of the state surveys. For example, BELTUSE in 1991 for drivers of passenger cars equipped with 3-point belts, in states with belt laws for cars, on major rural roads is

$$62.8 - 3.6 = 59.2 \text{ percent}$$

(where 62.8 percent is the 1991 belt use of drivers of cars with 3-point belts on major urban roads in states with belt laws, as shown in Table 18, and 3.6 is the percentage point decrement for major rural roads relative to major urban roads). The STATEWGT for this cell was derived earlier and found to be 0.0968549.

For RF passengers of pickup trucks in states without belt laws on local rural roads in 1996, BELTUSE is

$$71.3 - 4.0 - 24.3 - 10.8 - 10.9 = 21.3 \text{ percent}$$

(where 71.3 percent is the 1996 belt use of drivers of cars with 3-point belts on major urban roads in states with belt laws, as shown in Table 18, and 4.0 is the percentage point decrement for the RF passenger relative to the driver, 24.3 is the decrement for “no law,” 10.8 is the decrement for pickup trucks, and 10.9 is the decrement for local rural roads). The STATEWGT for this cell is 0.0003947.

These examples illustrate the procedure for estimating BELTUSE in cars, pickup trucks and vans/SUVs with 3-point belts. The simpler procedure for cars with automatic 2-point belts was described earlier. When these BELTUSE numbers are averaged, weighted by STATEWGT, they will yield exactly the VMT-weighted national averages of the state surveys, in each year, as shown in the right column of Table 18, or in Table 14.

Estimation of National on-the-road belt use The national average of the state surveys is an overestimate of actual belt use on the road, especially before 1998, because a number of state surveys did not include light trucks, RF passengers and/or local roads - where belt use is lower than for car drivers on major roads. But if the belt use rates in the 576 cells are weighted by NATLWGT rather than STATEWGT, they will yield national estimates of belt use on the road, for each year from 1991 through 1999. Those estimates, at least in theory, ought to be directly comparable to NOPUS, and subject to far less sampling error than NOPUS. Table 19 shows these national estimates for all vehicles, for passenger cars, and for light trucks. It also displays, for comparison purposes, the unadjusted average of the state surveys (copied from Table 14), and the NOPUS results for 1994, 1996 and 1998.

TABLE 19: ESTIMATED NATIONAL ON-THE-ROAD BELT USE

	Estimated Belt Use (Adjusted State Survey Data)			Unadjusted State Survey Data
	Overall	**Cars (2 & 3 Pt.)**	**Light Trucks**	Overall
1991	**54.1**	**59.1**	**45.1**	58.3
1992	**56.6**	**61.8**	**47.6**	60.9
1993	**61.0**	**66.1**	**52.7**	65.0
1994	**62.6**	**67.8**	**54.5**	66.3
1995	**64.3**	**69.5**	**56.6**	67.1
1996	**64.5**	**69.5**	**57.3**	66.5
1997	**65.6**	**70.5**	**58.7**	67.5
1998	**68.0**	**71.1**	**63.9**	68.6
1999	**69.5**	**72.4**	**65.8**	70.1
		NOPUS Data		
1994	58.0	62.8	50.2	
1996	61.3	64.8	56.4	
1998	68.9	71.3	65.7	

Based on this model, belt use increased steadily in the United States, from 54.1 percent in 1991 to 69.5 percent in 1999. Belt use in cars increased from 59.1 percent to 72.4 percent; in light trucks, from 45.1 percent to 65.8 percent. The gains for light trucks, especially from 1997 (58.7%) to 1998 (63.9%), were spurred by states that extended their belt laws to include light trucks, and by a market shift from pickup trucks (with low belt use) to SUVs and vans (with high belt use).

The unadjusted state data overestimated belt use by 4.2 percentage points in 1991, primarily because nearly half the states omitted pickup trucks from their surveys, especially the states whose laws did not apply to the trucks. By 1999, the overestimate was only 0.6 percentage points; many state surveys still omitted or undersampled local roads, but no state omitted light trucks or RF passengers. The unadjusted state data showed little progress from 1994 (66.3%) to 1997 (67.5%), because some states were just beginning to survey light trucks in those years, driving down their reported numbers. The adjusted data correct for that and show gains in the overall number every year.

The relationship between our adjusted national estimates and NOPUS is not so clear. The discrepancies are not huge; however, the1994 NOPUS estimate is 4.6 percentage points lower than the adjusted state data, the 1996 NOPUS is 3.2 percentage points lower, and the 1998 NOPUS is 0.9 percentage points **higher** than the adjusted state data. Perhaps no explanation is required: since the 2-sigma sampling error of NOPUS is ± 4 percentage points[64], the discrepancies can be considered within the "noise" range of NOPUS. They are substantially less than the 8.3 percentage point discrepancy between 1994 NOPUS and the unadjusted state estimate, clearly beyond sampling error bounds, and clearly due to the state surveys that omitted light trucks.

Less welcome explanations of why the adjusted state data exceed the early NOPUS could be that our model does not fully correct for the state survey omissions in the earlier years, or even that the non-probability sample designs of some early state surveys created upward biases that our model does not address. However, if either of these are true - if the early adjusted state numbers overestimate actual belt use - we can expect the effectiveness analysis later in this report to overestimate fatality reduction in the earlier 1990's. As we shall see, it does not.

Belt effectiveness in daytime, non-antisocial crash involvements The final step of the validation analysis is to use the FARSWGT to estimate an on-the-road belt use we believe equal to use in potentially fatal crashes ($U_2 = U_3$), and in combination with the belt use of fatally injured occupants (U_1), compute the fatality reduction by belts:

$$E = [U_3 - U_1] / [U_3 - (U_3 \text{ x } U_1)]$$

and compare it to the reductions obtained by double-pair comparison, adjusted by the UEF (Table 13).

Table 20 computes belt use for our two populations of interest - front-outboard occupants of passenger cars equipped with 3-point belts, and of light trucks - on-the-road (using NATLWGT), and in potentially fatal crashes (using FARSWGT): multivehicle, and single + multivehicle daytime, non-antisocial crash involvements. Unlike Table 19, which combined 3-point and automatic 2-point belts into a single "passenger car" estimate for comparison with NOPUS, Table 20 is limited to cars with 3-point belts.

Belt use on the road increased from 56.8 to 71.0 percent in cars with 3-point belts. These numbers are lower than in Table 19, because cars with automatic 2-point belts are excluded - but by 1999, it is only 1.4 percentage points lower (71.0 vs. 72.4), as the 2-point belts gradually phase out. In light trucks, it increased from 45.1 to 65.8 percent, as in Table 19. However, in potentially fatal multivehicle crashes, belt use is 1.3 percentage points lower than on-the-road in cars, and 1.9 percentage points lower than on-the-road in light trucks, because many of these crashes occur on rural roads, where belt use is lower. In single plus multivehicle crashes, belt use is yet another 0.3-0.5 percentage points lower, since an even greater proportion of the single-vehicle crashes are rural.

---

[64] *Observed Safety Belt Use in 1996*, NHTSA Research Note, Washington, 1997.

TABLE 20: BELT USE ON-THE-ROAD AND IN POTENTIALLY FATAL CRASHES

	On the Road		In Potentially Fatal, Daytime Non-Antisocial Crash Involvements			
			Multivehicle		Single + Multi	
	Cars (3 Pt.)	Light Trucks	Cars (3 Pt.)	Light Trucks	Cars (3 Pt.)	Light Trucks
1991	56.8	45.1	55.5	43.2	55.2	42.7
1992	59.4	47.6	58.1	45.7	57.8	45.2
1993	63.7	52.7	62.5	50.8	62.2	50.3
1994	65.4	54.5	64.2	52.6	63.8	52.1
1995	67.3	56.6	66.1	54.7	65.8	54.2
1996	67.4	57.3	66.1	55.4	65.8	54.9
1997	68.6	58.7	67.3	56.8	67.0	56.3
1998	69.3	63.9	68.0	62.0	67.7	61.5
1999	71.0	65.8	69.7	63.9	69.4	63.4

TABLE 21: FATALITY REDUCTION BASED ON UPFC, CARS WITH 3-POINT BELTS MULTIVEHICLE DAYTIME, NON-ANTISOCIAL CRASH INVOLVEMENTS

	Belt Use (%)		
	Fatalities (U_1)	UPFCs ($U_2 = U_3$)	Fatality Reduction (%) (E)
1991	49.0	55.5	23.1
1992	50.2	58.1	27.4
1993	54.6	62.5	27.8
1994	57.2	64.2	25.3
1995	57.2	66.1	31.4
1996	59.6	66.1	24.4
1997	61.3	67.3	23.1
1998	63.6	68.0	17.9
1999	63.7	69.7	23.8
Average			24.9
Standard Deviation / $\sqrt{9}$			1.3

Table 21 performs the effectiveness analysis for passenger cars with 3-point belts in multivehicle, daytime non-antisocial crash involvements. It is remarkable how the belt use in potentially fatal crashes, based on adjusted state survey data, rises in tandem with the belt use rate of fatally injured occupants, based on FARS. Belt effectiveness, $E = [U_3 - U_1] / [U_3 - (U_3 \text{ x } U_1)]$, is close to 25 percent in this special group of crashes, year after year[65]. The average of the nine annual estimates is 24.9 percent. The standard deviation of the nine individual estimates, divided by $\sqrt{9}$, is the standard error of the nine-year average. The 1.96ó sampling-error range for effectiveness is 24.9 ± 2.5 percent. (This interval estimate is not presented as "confidence bounds" but as an indicator of minimum sampling error that could be augmented by non-sampling errors introduced by the various assumptions in the model.)

In multivehicle, daytime, non-antisocial crash involvements, the effectiveness estimate based on double-pair comparison analysis of 1986-99 FARS data, corrected by the UEF, is 29 ± 10 percent, as shown in Table 13. Thus, our interval estimate based on belt use among fatalities (1991-99 FARS) and UPFC (1991-99 state surveys), 24.9 ± 2.5 percent, is well within the minimum sampling error range of the double-pair comparison analysis, 29 ± 10 percent. It validates the use of double-pair comparison, when corrected by the UEF, on post-1986 FARS data.

The remaining three analyses agree even more closely with the double-pair comparison results of Table 13. Table 22 analyzes light trucks in multivehicle, daytime non-antisocial crash involvements.

Again, UPFC keeps a steady lead over U_1, but by a greater margin than in Table 21, resulting in a higher effectiveness. The sampling error range for this estimate, 39.3 ± 3.7 percent is entirely within the interval estimate based on double-pair comparison, 42 ± 12 percent, and the point estimates are quite close[66].

---

[65] If the adjusted state data had overestimated actual belt use in the earlier years, we would have expected higher effectiveness in those years, and low effectiveness later on.

[66] Table 22 shows higher fatality reduction in 1998-99 than in preceding years, due to a 1998 jump in our estimated U_2 that exceeds the steady year-to-year gains in U_1. Two factors may be at work. (1) The accelerating market shift from pickup trucks to SUVs could increase effectiveness over time, since belts appear to be more effective in SUVs than in pickup trucks (see Table 12 and its discussion). (2) *Traffic Safety Facts (op. cit.)* suggests all state belt laws encompassed light trucks in 1998, but quite a few did not in 1997. Our method may have understated belt use in light trucks in the last years before 1998 in the "no law" states, or there may have been additional states that extended their belt laws to light trucks before 1998. Notwithstanding the jump in 1998, we believe the 1991-99 average effectiveness is accurate.

TABLE 22: FATALITY REDUCTION BASED ON UPFC, LIGHT TRUCKS MULTIVEHICLE DAYTIME, NON-ANTISOCIAL CRASH INVOLVEMENTS

	Belt Use (%)		
	Fatalities (U_1)	UPFCs ($U_2 = U_3$)	Fatality Reduction (%) (E)
1991	31.3	43.2	40.0
1992	36.3	45.7	32.3
1993	39.3	50.8	37.2
1994	41.7	52.6	35.4
1995	44.1	54.7	34.6
1996	43.4	55.4	38.4
1997	44.4	56.8	39.3
1998	46.2	62.0	47.4
1999	47.2	63.9	49.5
Average			39.3
Standard Deviation / √9			1.9

TABLE 23: FATALITY REDUCTION BASED ON UPFC, CARS WITH 3-POINT BELTS SINGLE + MULTIVEHICLE DAYTIME, NON-ANTISOCIAL CRASH INVOLVEMENTS

	Belt Use (%)		
	Fatalities (U_1)	UPFCs ($U_2 = U_3$)	Fatality Reduction (%) (E)
1991	44.9	55.2	33.9
1992	46.1	57.8	37.6
1993	50.8	62.2	37.2
1994	52.5	63.8	37.4
1995	53.1	65.8	41.0
1996	55.4	65.8	35.4
1997	56.5	67.0	36.0
1998	59.3	67.7	30.5
1999	58.7	69.4	37.3
Average			36.3
Standard Deviation / √9			0.97

Table 23 returns to passenger cars with 3-point belts, but expands the analysis to all daytime, non-antisocial crash involvements - single plus multivehicle. This analysis is based on the largest N of FARS cases, and shows the highest year-to-year consistency for fatality reduction. The effectiveness estimate, 36.3 ± 1.9 percent is well within the interval estimate based on double-pair comparison, 38 ± 8 percent, and the two point estimates are quite close.

Table 24 analyzes light trucks in single plus multivehicle crashes.

TABLE 24: FATALITY REDUCTION BASED ON UPFC, LIGHT TRUCKS
SINGLE + MULTIVEHICLE DAYTIME, NON-ANTISOCIAL CRASH INVOLVEMENTS

	Belt Use (%)		
	Fatalities (U_1)	UPFCs ($U_2 = U_3$)	Fatality Reduction (%) (E)
1991	27.4	42.7	49.3
1992	32.0	45.2	42.9
1993	33.4	50.3	50.4
1994	35.6	52.1	49.1
1995	37.8	54.2	48.6
1996	37.8	54.9	50.2
1997	39.4	56.3	49.6
1998	40.7	61.5	57.0
1999	39.4	63.4	62.5
Average			51.1
Standard Deviation / $\sqrt{9}$			1.9

The interval estimate of fatality reduction is 51.1 ± 3.7 percent and it is very compatible with the 52 ± 8 percent generated by double-pair comparison. Indeed, all four analyses here validate their counterpart estimates based on double-pair comparison[67].

This method based on UPFC generates effectiveness estimates with less sampling error than the corresponding double-pair comparisons. The reader will be tempted to ask, “Why did we limit the analyses to just these special groups of fatal crashes? Why couldn’t we use this method directly to estimate the overall fatality reduction by safety belts?” The answer, of course, is that this method only works for subgroups of fatalities where it is plausible that the UPFC is equal to belt use observed on the road. For the full set of fatalities, including alcohol-impaired drivers, reckless

---

[67] Table 24, like Table 22, shows higher effectiveness in 1998-99. See the discussion in Footnote 61.

drivers and nighttime crashes, UPFC is much lower than belt use observed in surveys, and this method would greatly overestimate effectiveness[68].

However, we can compare the UPFC-based and double-pair comparison estimates for these subgroups of fatalities and, based on ratios, assess the implications for the full set of fatalities. For cars with 3-point belts in multivehicle, daytime, non-antisocial crash involvements, the fatality reduction is 24.9 percent based on the UPFC analysis (Table 21) and 29 percent based on double-pair comparison (Table 13). Since the overall effectiveness of belts is 45 percent, based on double-pair comparison (Table 1), the implicit effectiveness for all crashes, by the UPFC method would be

$$1 - \{[(1 - 0.249) / (1 - 0.29)] \text{ x } [1 - 0.45]\} = 41.8 \text{ percent}$$

For cars with 3-point belts in single + multivehicle, daytime, non-antisocial crash involvements, the fatality reduction is 36.3 percent with UPFC (Table 23) and 38 percent with double-pair comparison (Table 13). The implicit effectiveness for all crashes, by the UPFC method would be

$$1 - \{[(1 - 0.363) / (1 - 0.38)] \text{ x } [1 - 0.45]\} = 43.5 \text{ percent}$$

Both of these numbers are well within the 40 to 50 percent range that NHTSA estimated for safety belt effectiveness in 1984 and still uses as its “official” estimate[69]. They are also well within the “noise” range of the UEF-corrected double-pair comparison estimate of 45 percent fatality reduction by 3-point belts in passenger cars.

For light trucks in multivehicle, daytime, non-antisocial crash involvements, the fatality reduction is 39.3 percent with UPFC (Table 22) and 42 percent with double-pair comparison (Table 13). Since the overall effectiveness of belts is 60 percent, based on double-pair comparison (Table 1), the implicit effectiveness for all crashes, by the UPFC method would be

$$1 - \{[(1 - 0.393) / (1 - 0.42)] \text{ x } [1 - 0.60]\} = 58.1 \text{ percent}$$

In the single + multivehicle crashes, the fatality reduction is 51.1 percent with UPFC (Table 24) and 52 percent with double-pair comparison (Table 13). The implicit effectiveness for all crashes, with UPFC would be

$$1 - \{[(1 - 0.511) / (1 - 0.52)] \text{ x } [1 - 0.60]\} = 59.3 \text{ percent}$$

---

[68] Specifically, the belt use rate of all passenger-vehicle front-outboard fatalities in 1991-99 was 36.6 percent and the average 1991-99 belt use rate on the road, based on the adjusted state data (Table 19, left column) was 62.9 percent. That works out to 66 percent effectiveness for belts.

[69] *Final Regulatory Impact Analysis* (1984), p. IV-2.

Both of these numbers are close to NHTSA's "official" 60 percent estimate[70] and well within the "noise" range of the UEF-corrected double-pair comparison estimate of 60 percent fatality reduction by belts in light trucks.

Conclusion The relatively high rate of belt use among fatally injured occupants in daytime, non-drinking, non-antisocial crash involvements shows it is very unlikely that belt effectiveness could be 60+ percent in cars, or 70+ percent in light trucks, as suggested by uncorrected double-pair comparison. These analyses of belt use on the road versus belt use of fatally injured occupants support NHTSA's long-standing estimates that safety belts reduce fatalities by close to 45 percent in cars and 60 percent in light trucks. They support the use of the UEF, as defined in Section 5, to adjust downwards the estimates based on double-pair comparison of FARS data collected after 1985.

---

[70] *Preliminary Regulatory Impact Analysis* (1989), p. 15.

REFERENCES

Blincoe, Lawrence J. *Estimating the Benefits from Increased Safety Belt Use*. NHTSA Publication No. DOT HS 808 133, Washington, 1994.

Bowman, Brian L. and Rounds, Donald A. *Restraint System Usage in the Traffic Population - 1988 Annual Report*. NHTSA Publication No. DOT HS 807 447, Washington, 1989.

Cerrelli, Ezio. *Driver Exposure: Indirect Approach for Obtaining Relative Measures*. NHTSA Publication No. DOT HS 820 179, Washington, 1972.

Evaluation of the Effectiveness of Occupant Protection - Federal Motor Vehicle Safety Standard 208 - Interim Report, June 1992. NHTSA Publication No. DOT HS 807 843, Washington, 1992.

Evans, Leonard. “Double Pair Comparison - A New Method to Determine How Occupant Characteristics Affect Fatality Risk in Traffic Crashes.” *Accident Analysis and Prevention,* Vol. 18, June 1986, pp. 217-227.

_________. “The Effectiveness of Safety Belts in Preventing Fatalities.” *Accident Analysis and Prevention*, Vol. 18, June 1986, pp. 229-241.

_________. *Traffic Safety and the Driver*. Van Nostrand Reinhold, New York, 1991.

Evans, Leonard and Frick, Michael C. “Potential Fatality Reductions through Eliminating Occupant Ejection from Cars.” *Accident Analysis and Prevention*, Vol. 21, pp. 169-182.

Final Regulatory Impact Analysis, Amendment to Federal Motor Vehicle Safety Standard 208, Passenger Car Front Seat Occupant Protection. NHTSA Publication No. DOT HS 806 572, Washington, 1984.

Fourth Report to Congress, Effectiveness of Occupant Protection Systems and their Use. NHTSA Publication No. DOT HS 808 919, Washington, 1999.

Haight, Frank A. *Indirect Methods for Measruing Exposure Factors as Related to the Incidence of Motor Vehicle Traffic Accidents*. NHTSA Publication No. DOT HS 800 601, Washington, 1971.

Hansen, Morris H., Hurwitz, William N., and Madow, William G. *Sample Survey Methods and Theory, Volume I*. John Wiley & Sons, New York, 1953.

Kahane, Charles J. *Addendum to "Seat Belt Effectiveness Estimates Using Data Adjusted for Damage Type”*. NHTSA Docket No. 74-14-N35-229-05, 1984.

_________. *An Evaluation of Door Locks and Roof Crush Resistance of Passenger Cars*. NHTSA Publication No. DOT HS 807 489, Washington, 1989.

_________. *Evaluation of FMVSS 214 - Side Impact Protection: Dynamic Performance*. NHTSA Publication No. DOT HS 809 004, Washington, 1999.

Klein, Terry M. *A Method for Estimating Posterior BAC Distributions for Persons Involved in Fatal Traffic Accidents*. NHTSA Publication No. DOT HS 807 094, Washington, 1986.

Klein, Terry M. and Walz, Marie C. *Estimating Lives Saved by Restraint Use in Potentially Fatal Crashes*. NHTSA Research Note, Washington, 1995.

Mosteller, F. and Tukey, J.W. *Data Analysis and Regression: A Second Course in Statistics*. Addison-Wesley, Reading, MA, 1977.

National Occupant Protection Use Survey. Washington Consulting Group, Washington, 1995.

Observed Safety Belt Use in 1996. NHTSA Research Note, Washington, 1997.

Partyka, Susan C. “Belt Effectiveness in Pickup Trucks and Passenger Cars by Crash Direction and Accident Year (May 1988).” *Papers on Adult Seat Belts - Effectiveness and Use*. NHTSA Publication No. DOT HS 807 285, Washington, 1988.

_________. *Belt Use in Serious Impacts Estimated from Fatality Data*. NHTSA Publication No. DOT HS 807 519, Washington, 1989.

_________. *Lives Saved by Seat Belts from 1983 through 1987*. NHTSA Publication No. DOT HS 807 324, Washington, 1988.

_________. “Seat Belt Effectiveness Estimates Using Data Adjusted for Damage Type (January 1984).” *Papers on Adult Seat Belts - Effectiveness and Use*. NHTSA Publication No. DOT HS 807 285, Washington, 1988.

Partyka, Susan C. and Womble, Kathleen B. *Projected Lives Savings from Greater Belt Use*. NHTSA Research Note, Washington, 1989.

Preliminary Regulatory Impact Analysis, Proposed Extension of the Automatic Restraint Requirements of FMVSS 208 to Trucks, Buses and Multi-Purpose Passenger Vehicles with a Gross Vehicle Weight Rating of 8,500 Pounds or Less and an Unloaded Vehicle Weight of 5,000 Pounds or Less. NHTSA Docket No. 74-14-N62-001, 1989.

Reinfurt, Donald W., Silva, Claudio Z. and Seila, Andrew F. *A Statistical Analysis of Seat Belt Effectiveness in 1973-75 Model Cars Involved in Towaway Crashes, Volume 1*. NHTSA Publication No. DOT HS 802 035, Washington, 1976.

Rood, Debra H., Kraichy, Patricia P. and Carubia, Jean. *Evaluation of New York State's Mandatory Occupant Restraint Law*. NHTSA Publication No. DOT HS 806 950, Washington, 1985.

Safety Belt Usage, A Review of Effectiveness Studies, Suggestions for State Programs. NHTSA Publication No. DOT HS 801 988, Washington, 1976.

Sikora, James J. *Relative Risk of Death for Ejected Occupants in Fatal Traffic Accidents*. NHTSA Publication No. DOT HS 807 096, Washington, 1986.

Teets, Mary K. *Highway Statistics 1993*. Publication No. FHWA-PL-94-023, Federal Highway Administration, Washington, 1994.

Traffic Safety Facts 1998. NHTSA Publication No. DOT HS 808 983, Washington, 1999.

Wagnenaar, A.C., Wiviott, M.B.T. and Compton, C.P. *Direct Observation of Seat Belt Use in Michigan: April 1985*. Publication No. UMTRI 85-26, University of Michigan Transportation Research Institute, Ann Arbor, 1985.

APPENDIX: STATE BELT USE SURVEYS AND LAWS, 1991-99

		OBSERVED BELT USE (%)									Belt Law Effective				
	VMT 1998 (Million)	91	92	93	94	95	96	97	98	99	Car	P/U	Van/ SUV	91-97 Survey Omits	Survey Type
Alabama	54,870	53	58	55	55	52	54.0	52.0	52.0	58.1	91½	98	98		ADT Roads
Alaska	4,489	66	66	69	69	69	53.0	56.0	57.0	64.0	*	*	*		Unknown
Arizona	45,209	65	73	73	60	60	55.9	63.2	61.5	71.9	91	98	91		Signalized X-section
Arkansas	28,170	52	55	55	51	51	48.0	50.5	52.6	57.2	91½	91½	91½		ADT Roads
California	284,642	71	70	83	83	85	86.6	86.4	88.6	89.3	*	*	*		Other Local-road
Colorado	39,039	51	50	53	54	56	55.6	59.6	66.0	65.2	*	*	*		Non-WESTAT Unk Type
Connecticut	29,142	61	71	71	72	72	59.4	59.8	70.1	72.5	*	*	*		WESTAT Survey
Delaware	8,165	42	70	68	63	60	62.0	59.0	62.3	64.4	92	98	98		Non-WESTAT Unk Type
D.C.	3,282	49	59	62	62	63	57.0	66.0	79.6	77.9	*	*	*	LT+RF	WESTAT Survey
Florida	136,624	60	57	62	61	59	62.7	60.0	57.2	59.0	*	*	*		Unknown
Georgia	96,430	54	51	57	57	53	58.0	65.0	73.6	74.2	*	96	*		ADT Roads
Hawaii	7,929	85	83	84	84	85	78.3	80.0	80.5	80.3	*	*	*		WESTAT Survey
Idaho	13,337	45	53	59	61	59	50.0	49.0	57.3	57.3	*	*	*		Non-WESTAT Unk Type
Illinois	100,630	51	65	67	68	69	64.0	64.2	64.5	65.9	*	*	*	LT	ADT Roads
Indiana	68,443	52	56	56	56	64	52.7	53.2	61.8	57.3	*	98	98		Non-WESTAT Unk Type
Iowa	28,722	68	71	73	73	76	74.8	74.9	76.9	78.2	*	*	*		Other Local-road
Kansas	26,936	64	70	70	70	54	54.0	56.0	58.7	62.6	*	98	*		Other Local-road
Kentucky	46,285	48	41	40	58	52	55.0	54.0	54.3	58.6	94½	94½	94½	RF	Unknown
Louisiana	40,063	37	47	48	50	59	59.0	64.0	65.6	67.0	*	*	*		WESTAT Survey
Maine	13,468	35	36	36	36	50	50.0	61.0	61.3	59.0	96	96	96		Signalized X-section
Maryland	48,044	72	75	72	69	70	70.0	71.0	82.6	82.7	*	*	*		ADT Roads
Massachusetts	51,510	35	31	34	47	53	54.0	53.0	51.0	52.0	94	94	94		Signalized X-section
Michigan	93,332	64	53	64	66	67	66.1	66.9	69.9	70.1	*	*	*		Signalized X-section
Minnesota	49,331	52	53	55	57	65	64.0	64.8	64.2	71.5	*	*	*		WESTAT Survey
Mississippi	33,999	32	24	25	43	46	46.0	48.2	58.0	54.5	*	98	*	LT	WESTAT Survey
Missouri	64,138	54	70	70	68	71	62.0	66.6	60.4	60.8	*	98	*	LT+RF	Non-WESTAT Unk Type
Montana	9,530	67	71	71	69	70	70.8	72.6	73.1	74.0	*	*	*		Other Local-road

*In effect before 1991

		OBSERVED BELT USE (%)									Belt Law Effective				
	VMT 1998 (Million)	91	92	93	94	95	96	97	98	99	Car	P/U	Van/ SUV	91-97 Survey Omits	Survey Type
Nebraska	17,459	33	33	54	63	64	64.6	62.9	65.1	67.9	93	93	93		Other Local-road
Nevada	17,196	68	63	70	71	71	70.1	69.4	76.2	79.8	*	98	98		ADT Roads
New Hampshire	11,499	49	50	51	54	57	56.0	57.7	55.6	56.0	NEVER				Unknown
New Jersey	64,111	58	68	71	64	61	60.3	62.0	63.0	63.3	*	98	98	LT	WESTAT Survey
New Mexico	22,053	67	66	75	79	86	85.0	88.0	82.6	88.4	*	*	*		Non-WESTAT Unk Type
New York	122,605	68	69	72	72	72	73.6	75.2	75.3	76.1	*	98	98	LT	WESTAT Survey
North Carolina	84,747	60	70	80	81	81	80.0	82.0	76.7	78.1	*	93	*		Signalized X-section
North Dakota	7,299	30	30	30	32	42	41.8	49.4	40.0	46.7	94½	94½	94½		Non-WESTAT Unk Type
Ohio	104,280	50	58	62	62	62	62.4	65.2	60.6	64.8	*	*	*	LT	Non-WESTAT Unk Type
Oklahoma	41,770	37	44	47	45	46	47.5	60.0	56.0	60.7	*	*	*		Unknown
Oregon	33,158	70	72	73	77	80	81.5	82.1	82.6	83.4	*	94	94		Signalized X-section
Pennsylvania	99,291	60	63	68	72	71	65.0	65.0	67.8	69.7	*	*	*		Other Local-road
Rhode Island	7,929	28	32	32	58	58	59.0	59.0	58.6	67.3	91½	98	98		ADT Roads
South Carolina	42,557	60	53	59	64	64	61.1	60.8	64.8	65.0	*	*	*		Unknown
South Dakota	8,034	33	42	26	40	40	47.0	68.0	45.7	45.7	95	95	95		Non-WESTAT Unk Type
Tennessee	62,169	51	58	58	60	64	63.0	61.0	56.7	61.3	*	*	*	LT	WESTAT Survey
Texas	204,726	68	69	69	71	72	74.0	74.6	74.4	74.0	*	*	*		Signalized X-section
Utah	21,134	45	50	50	53	56	60.1	62.9	66.7	67.4	*	*	*		Non-WESTAT Unk Type
Vermont	6,563	40	47	54	68	67	68.5	70.9	62.7	69.8	94	98	98		Non-WESTAT Unk Type
Virginia	70,255	58	72	73	72	70	69.6	67.1	73.6	69.9	*	*	*		Signalized X-section
Washington	51,615	69	73	78	81	83	83.2	81.4	79.1	81.1	*	*	*	LT	WESTAT Survey
West Virginia	18,561	43	34	52	58	58	64.0	67.0	56.5	51.9	94	98	98	RF	Signalized X-section
Wisconsin	55,448	58	59	64	64	64	58.1	51.6	61.9	65.1	*	*	*		Signalized X-section
Wyoming	7,981	66	66	67	70	71	58.5	59.5	50.1	50.1	*	*	*		Unknown

*In effect before 1991