

**Methods for Estimating Driver Death
Rates by Vehicle Make and Series**

Charles M. Farmer

June 2011

**INSURANCE INSTITUTE
FOR HIGHWAY SAFETY**

1005 NORTH GLEBE ROAD ARLINGTON, VA 22201

PHONE 703/247-1500 FAX 703/247-1678

www.iihs.org

ABSTRACT

Driver death rates per million registered vehicles per year were calculated for 168 passenger vehicle models for sale in 2008. Rates ranged from 0 to 179 deaths per million registrations per year. Minivans and SUVs had driver death rates significantly lower than those for other vehicle types. Within vehicle type, larger vehicles generally had lower driver death rates than smaller vehicles. To account for potential differences by calendar year, driver age, gender, and driving environment, all of which can affect motor vehicle crash and injury experience, rates were standardized to a common distribution of exposure. These standardized rates ranged from 0 to 143 deaths per million registrations per year, and the relative differences among vehicle types were lessened considerably.

INTRODUCTION

Periodically since 1989 the Insurance Institute for Highway Safety (IIHS) has calculated and published driver death rates per registered vehicle. These rates provide an indication of the overall risk to drivers, including both the likelihood of being in a crash and the likelihood of being fatally injured in that crash. Assuming similar amounts and types of driving per year, comparisons can be made among the various vehicle models. Such exposure, however, may vary widely across vehicle models for a number of reasons, including vehicle cost and marketing, geography, and economic conditions. Driver characteristics, especially age and gender, affect both the likelihood of crash involvement and the likelihood of injury in a crash. For example, young drivers are overinvolved in serious crashes, and elderly drivers are more likely to be killed if they are in crashes.

To compensate for the differences in exposure, real-world crash injury rates often are standardized to a common age and gender distribution of involved drivers. For example, standardized relative injury rates per reported crash are published in Australia (Newstead et al., 2010), Sweden (Folksam, 2009), and the United States (Highway Loss Data Institute (HLDI), 2009).

Similar standardization of fatal crash rates for age and gender has not been possible. There is no national database of all reported motor vehicle crashes in the United States, so it is not possible to calculate driver death rates per crash. And, although there is a database of all fatal crashes and a corresponding database of all registered vehicles — so that it is possible to calculate driver death rates per registered vehicle for each vehicle model (IIHS, 2007) — the vehicle registration database does not include information on driver age and gender. Absent information on the exposure of each vehicle model by driver age and gender, driver death rates cannot be directly standardized to a common driver age and gender distribution.

One indirect method of lessening the effect of driver differences is to group vehicle models by size and body style. The driver population within each resulting vehicle group is likely to be much more

homogeneous than the overall driver population. Therefore, comparisons among cars within a market group should be less affected by driver differences. Nevertheless, comparisons between vehicles in different market groups still are affected by driver differences.

A mathematical adjustment of death rates based on vehicle wheelbase and the proportion of occupant deaths in cars with male drivers or drivers younger than 30 was included as part of the first few IIHS (1989) evaluations. In addition to occupant death rates per 10,000 vehicle registrations, the differences between actual death rates and those predicted based on wheelbase, driver age, and gender were presented. About one-sixth of vehicles evaluated had actual death rates more than 40 percent lower or higher than those predicted by the mathematical model. These vehicles then were rated as performing much better or much worse than expected.

In 2001 a new method was introduced for adjusting the rates to differences in driver age and gender (Farmer, 2001). This adjustment made comparisons of vehicles more meaningful by standardizing each vehicle's driver death rate to a population with a common proportion of 25-64-year-old female drivers — that is, the group of drivers with the lowest fatality rate. Because the true proportion of 25-64-year-old female drivers for a given vehicle was unknown, the standardization procedure was indirect, based on an assumed mathematical relationship between overall and group-specific driver death rates.

A sharp decline in fatal crash rates beginning in 2008 necessitated further adjustments to the procedure for estimating driver death rates. Vehicle models newly designed in 2008 tended to have lower death rates than earlier models partly because their exposure period was less risky. To adjust for these differences in exposure, a statistical model was devised for estimating deaths that would have occurred if the new models had been around in earlier years. In addition, the statistical model formulated a relationship of death rates to driver age, gender, and driving environment. Rates then could be standardized to a common distribution of calendar years, age, gender, and driving environment. This paper presents the methodology for computing standardized driver death rates per million registrations per year by vehicle model for model years 2005-08 during calendar years 2006-09.

METHOD

Counts of driver deaths for each make and series of 2006-08 passenger vehicles were obtained from the U.S. Department of Transportation's Fatality Analysis Reporting System (FARS), an annual census of fatal motor vehicle crashes. A vehicle series is defined as the combination of vehicle model and body style (e.g., the two-door and four-door Honda Civic are different series).

Counts of registered vehicles for each make and series were obtained from the National Vehicle Population Profile (NVPP), a compilation of data from state registration files produced by R.L. Polk & Company. NVPP registration counts are a snapshot of vehicles registered as of July 1 of each year, so

they tend to misrepresent annual registrations of the current model year. For example, registration counts on July 1, 2009 did not include any of the yet unsold 2009 models, nor did they provide information on how many months any new vehicle had been registered. In this analysis, counts of both driver deaths and vehicle registrations for each model year were restricted to calendar years later than the model year. Nearly all such vehicles registered on July 1 would have been on the road for the whole year. Also, because NVPP does not include government-owned vehicles, driver deaths in police vehicles or vehicles with government tags were excluded from the analysis.

Estimates of the proportions of drivers of each vehicle series who were younger than 25, 65 and older, male, or living in an area with at least 500 vehicles per square mile were derived using a database of automobile insurance policy information maintained by HLDI. The HLDI database covers more than 150 million individual passenger vehicles, amounting to about 80 percent of all privately insured vehicles in the United States.

Analyses were restricted to model years not significantly different in design from the 2008 model year. Design changes include changes in engineering design, such as the dimensions or weight of the vehicle, or the addition of electronic stability control (ESC) or head protection side airbags. For example, the Toyota Camry four-door car was redesigned in 2007, so only model years 2007-08 were included in this analysis. There were 25 driver deaths and 561,250 registrations for 2007 model Camrys during 2008, 19 driver deaths and 556,458 registrations for 2007 models during 2009, and 5 driver deaths and 188,347 registrations for 2008 models during 2009. Thus the totals for the 2007-08 model Camry cars were 49 driver deaths and 1,306,055 registration-years.

Data for each make, series, model year, and calendar year with at least 10,000 vehicle registrations were entered into a Poisson regression model. The model predicted the number of driver deaths based on the vehicle make and series, calendar year, vehicle age (i.e., calendar year minus model year), number of registrations, proportion of HLDI exposure for which the rated driver was younger than 25, proportion of exposure for which the rated driver was 65 or older, proportion of exposure for which the rated driver was male, and proportion of exposure for which the garaging zip code had a vehicle density of at least 500 vehicles per square mile.

By changing the values of all predictor variables other than make, series, calendar year, and vehicle age, the predicted death counts were standardized to a common exposure distribution. For example, the model parameter estimates were used to predict the number of driver deaths that would have been expected if the registrations of each vehicle series were distributed across the 10 model year/ calendar year combinations according to proportions from the overall vehicle population: 6 percent of registrations from model year 2005 in each of calendar years 2006-09, 9 percent of registrations from model year 2006 in each of calendar years 2007-09, 15 percent of registrations from model year 2007 in

each of calendar years 2008-09, and 20 percent of registrations from model year 2008 in calendar year 2009. It also was supposed that each series had about 7 percent of exposure from drivers younger than 25, 13 percent from drivers 65 and older, 49 percent from male drivers, and 44 percent from drivers in areas with high vehicle density.

Standardized driver death rates were computed by dividing predicted deaths by registered vehicle-years (in millions). Approximate 95 percent confidence limits for the standardized rates were based on a Taylor series estimate of the variance of a logarithm (Snedecor and Cochran, 1980). If X represents the predicted deaths under standardized exposure, then $\text{Var}\{\log X\} \approx \text{Var}\{X\} / [E\{X\}]^2$. That is, $\text{Var}\{X\} \approx [E\{X\}]^2 \text{Var}\{\log X\}$. For each of the 10 model year/calendar year combinations, the regression model produced estimates of $E\{X\}$ and $\text{Var}\{\log X\}$, which in turn yielded estimates of $\text{Var}\{X\}$. Summing the estimates of $\text{Var}\{X\}$ gave an approximation to the variance of the sum of predicted deaths.

RESULTS

Table 1 summarizes the results of the regression model (except for most vehicle series parameters). Death rates were lowest in calendar year 2009. Compared with 2009, driver death rates were approximately 11 percent higher in 2006, 8 percent higher in 2007, and 5 percent higher in 2008. Death rates were highest for 4-year-old vehicles, about 7 percent lower for 3-year-old vehicles, 1 percent

Table 1
Poisson Regression of Logarithm of Driver Deaths

Parameter	Estimate	Chi-square	p-value
Intercept	-8.8178		
Make and series			
Chevrolet Cobalt 4d	-0.3393		
...			
Calendar year			
2006	0.1069	0.67	0.4142
2007	0.0732	0.67	0.4114
2008	0.0460	0.67	0.4129
2009	0		
Vehicle age			
1 year	-0.0363	0.13	0.7198
2 years	-0.0118	0.02	0.8964
3 years	-0.0755	0.69	0.4064
4 years	0		
Log(Registrations)	1		
Proportion of drivers younger than 25	0.0261	0.71	0.3997
Proportion of drivers 65 and older	0.0025	0.01	0.9078
Proportion male drivers	0.0087	0.26	0.6091
Proportion of drivers in areas with high vehicle density	-0.0121	0.58	0.4462

lower for 2-year-old vehicles, and 4 percent lower for 1-year-old vehicles. Vehicles with a higher percentage of young, old, or male drivers had higher driver death rates. Vehicles garaged in high-density areas had lower driver death rates than vehicles garaged in lower density areas.

The 2005-08 Chevrolet Cobalt had a very high proportion of young drivers (16-19 percent), an average proportion of older drivers (12-16 percent), a low proportion of male drivers (36-40 percent), and a low proportion of drivers in areas with high vehicle density (36-41 percent). So the Cobalt's 175 driver deaths and 1,128,364 registration-years gave it a relatively high rate of 155 driver deaths per million registration-years. When the registrations were redistributed (more to the later years) and the age, gender, and density distributions were standardized, the count of predicted driver deaths was reduced. For example, registrations of the 2006 Cobalt in 2007 were reduced from 129,098 to 108,235, the proportion of young drivers was reduced from 18.9 to 6.9 percent, the proportion of older drivers was increased from 12.1 to 12.7 percent, the proportion of male drivers was increased from 37.1 to 49.3 percent, and the proportion of drivers in areas with high vehicle density was increased from 37.5 to 44.0 percent. The predicted driver deaths for the 2006 Cobalt in 2007 was therefore reduced from 27 to 13.2 — that is, $\exp\{-8.8178 - 0.3393 + 0.0732 - 0.0363 + \log(108235) + 0.0261(6.914) + 0.0025(12.740) + 0.0087(49.329) - 0.0121(44.042)\}$. Overall, 132.4 driver deaths were predicted under the new exposure distribution. The estimated variance of the driver death count was 20.83.

The standardized rate for the Cobalt was 117 driver deaths per million registration-years (compared with 155 before standardization). The 95 percent confidence limits for the standardized rate were computed as $117 \pm 1.96 \sqrt{20.83} / 1.128364$. That is, the confidence interval includes values from 109 to 125.

Driver death rates for the Cobalt and 167 other vehicle series that had at least 100,000 registration-years of exposure were listed in an earlier publication (IIHS, 2011). Standardized rates for the other vehicle series ranged from 0 to 143 deaths per million registrations per year. Rates ranged from 0 to 179 before standardization. Vehicle series with the lowest rates mostly were minivans, SUVs, and luxury cars. The overall rate was 48 driver deaths per million registrations per year.

Table 2 lists driver death rates for passenger vehicles included in the regression model by vehicle type and size category. Columns 4-5 are based on raw counts of driver deaths, and columns 6-7 are based on standardized estimates. Death rates for small and large cars were lowered, whereas rates for midsize cars were increased, but the overall rate for cars did not change much. Death rates for minivans and SUVs tended to increase when standardized, whereas rates for pickups were decreased by quite a bit.

Table 2
Driver Death Rates per Million Registration-Years (standardized
using Poisson regression), 2005-08 Models during 2006-09

Vehicle style	Vehicle size	Registration-years	Raw		Standardized	
			Driver deaths	Death rate	Driver deaths	Death rate
All	All	65,078,867	3,371	52	3,114	48
Car	Mini	1,390,941	117	84	103	74
	Small	12,295,227	911	74	829	67
	Midsize	12,333,881	538	44	596	48
	Large	8,603,132	429	50	424	49
	Very Large	1,686,396	79	47	73	43
Minivan	All	36,309,577	2,074	57	2,024	56
	All	2,835,972	54	19	70	25
SUV	Small	3,432,915	100	29	121	35
	Midsize	5,793,887	124	21	154	27
	Large	2,443,757	40	16	49	20
	Very Large	206,257	6	29	6	31
	All	11,876,816	270	23	331	28
Pickup	Small	4,236,333	346	82	233	55
	Large	6,734,169	483	72	348	52
	Very Large	1,295,307	97	75	60	46
	All	12,265,809	926	75	641	52

DISCUSSION

Standardization for driver age and gender greatly reduced the variability of driver death rates among vehicle types. Vehicle types popular with male drivers and/or young drivers, such as sports cars and pickups, had standardized death rates that were much lower than the raw rates. However, within vehicle type, size, and body style, standardization of driver death rates had less effect. With only a few exceptions, those vehicles with the highest and lowest raw driver death rates in the class also had the highest and lowest standardized rates. Vehicles similar in size and body style seem to appeal to similar types of drivers. Of course, there are driver characteristics other than age and gender that affect crash risk.

Differences in when and where vehicles are driven may lead to differences in driver death rates, even if the driver characteristics are similar. The standardization for vehicle density at the garaging location was meant to account for vehicles driven more often on rural roads, but it was an imperfect surrogate. Also, some vehicles may be driven less often at night or in poor weather conditions. The procedure described here did not address such differences in exposure.

Some vehicles are just driven less than others. Such differences in overall exposure were not accounted for directly, but may have been a factor in the adjustments for vehicle age and calendar year. National estimates of vehicle miles traveled declined in 2008 after increasing consistently for 25 years (Longthorne et al., 2010). Thus vehicles first sold in 2008 might be expected to have been driven fewer

miles per year than vehicles sold earlier. Vehicles with exposure only in the later calendar years tended to have driver death rates that were increased by the standardization procedure.

In conclusion, although the standardization procedure led to a much cleaner comparison of driver death rates by vehicle make and series, other effects of driving behavior and environment still may exist. Note also that even vehicle series with millions of registration-years of exposure had rates with wide confidence intervals. So the standardized rates remain as somewhat imprecise measures of vehicle crashworthiness.

ACKNOWLEDGMENT

This work was supported by the Insurance Institute for Highway Safety.

REFERENCES

- Farmer, C.M. 2001. Driver death rates by vehicle make and series with adjustments for driver age and gender. Arlington, VA: Insurance Institute for Highway Safety.
- Folksam. 2009. How safe is your car? Stockholm, Sweden.
- Highway Loss Data Institute. 2009. Collision losses. Insurance special report A-82. Arlington, VA.
- Insurance Institute for Highway Safety. 1989. Computing fatality rates by make and series of passenger cars. Arlington, VA.
- Insurance Institute for Highway Safety. 2007. Special issue: driver death rates. *Status Report* 42(4). Arlington, VA.
- Insurance Institute for Highway Safety. 2011. Dying in a crash. *Status Report* 46(5). Arlington, VA.
- Longthorne, A.; Subramanian, R.; and Chen, C. 2010. An analysis of the significant decline in motor vehicle traffic crashes in 2008. Report No. DOT HS-811-346. Washington, DC: National Highway Traffic Safety Administration.
- Newstead, S.; Watson, L.; and Cameron, M. 2010. Vehicle safety ratings estimated from police reported crash data: 2010 update. Australian and New Zealand crashes during 1987-2008. Report no. 297. Clayton, Victoria: Monash University Accident Research Centre.
- Snedecor, G.W. and Cochran, W.G. 1980. *Statistical Methods*, 7th Edition, p. 288. Ames, Iowa: Iowa State University Press.