In Australia, is injury less in recent cars than in earlier cars? Evidence from comparing the injury severities of two drivers in the same collision

Anderson, R. W. G., and Hutchinson, T. P.
Centre for Automotive Safety Research, University of Adelaide, South Australia 5005, Australia
email: robert@casr.adelaide.edu.au

Abstract

Comparison of the severities of injury to the two drivers in the one collision is useful because speed of the impact is the same for the two drivers. Using the dataset of routinely-reported crashes in South Australia, 1991-2008, a multiple logistic regression was carried out, the dependent variable being the ratio of the probabilities of the drivers of car 2 and car 1 being killed, conditional on exactly one of them being killed. The independent variables were the difference between the two cars in their build years, the difference between the drivers' ages, and an allowance for whether the vehicles fell within a narrow definition of car. Statistically significant effects were found for all of these. In a similar regression with the probabilities referring to the drivers being seriously injured, an effect of car year was again found.

Keywords

Crashworthiness, Car year, Fatalities, Injury severity, Matched pairs, Secondary safety

Introduction

The "crashworthiness" of vehicles, or perhaps it should be the uncrashworthiness, will in this paper refer to the probability of an occupant's injury reaching at least some criterion level, given that a crash occurs. This has been improving over the last 20 or 30 years (Newstead et al. [1, 2]). The approach of Newstead et al. has been to consider serious injuries as a proportion of crashes, and to regress this on the year of the vehicle and several other variables (including year of crash, age of driver, speed limit, size of vehicle, State). The regression coefficient of the year of the vehicle is found to be negative: serious injuries are less likely for more recent vehicles.

The completeness and reliability of crash data tend to be lower for less serious crashes than for the more serious. It is possible that the reporting of damage-only crashes is more likely for more recent, more valuable, vehicles than for older vehicles. If so, underreporting of damage-only crashes for older vehicles, rather than a genuine improvement in crashworthiness, could be responsible for the negative slope found by Newstead et al. This seems unlikely, as Newstead et al. also show that there is a decline in serious injuries per injury crash. But as it potentially important, it is nevertheless worth investigating using a different method.

The analysis below considers collisions between two cars, and compares the severities of injury to the two drivers in the one collision. We relate this to the difference in the years of manufacture of the two vehicles: thus we ask, given that the severities are different, is the difference associated with the difference between the years of manufacture? Damage-only crashes are excluded because the severities of injury are not different. The strategy of matching two casualties and comparing their injury severities has been used before; the earliest example we know of was a comparison of front seat occupants wearing a lap belt with those not wearing a belt by Campbell and Kihlberg [3]. We are using routine crash data, and the fact that the two drivers are in the one crash matches for the speed of the crash, which is unavailable in routine data.

Method of data analysis

Simply inspecting the frequencies in a cross-tabulation of the two injury severities gives the answer straightforwardly. But it is possible that an effect of car year could be obscured or distorted by associations, chance or otherwise, of car year with variables such as driver age that are also associated with injury severity. To allow for this, multiple logistic regression is carried out. Examples quite similar to the present study include those of Crandall et al. [4] and Martin et al. [5].

This Section covers the following: the database; selection of crashes; summarising a cross-tabulation; interpretation; multiple regression.

The database

In South Australia, the road crash information database originating with the police is known as TARS (Traffic Accident Reporting System). The severity of injury of each casualty is coded as one of the following (from most to least severe): fatal, admitted to hospital, treated at hospital, treated by doctor, none. The vehicle year recorded in TARS will be referred to as the vehicle's "build" year, though this may not be strictly correct. Vehicles will be restricted to 1990 and later, in order for the results to reflect what has been happening in recent years and exclude any differences between very old and present-day vehicles.

The TARS dataset does not include the make and model of the vehicle, and consequently the mass of the vehicle is not known. In a collision between two vehicles, the lighter vehicle will undergo a greater velocity change, and its occupants will tend to be more seriously injured. Being unable to allow for this is a limitation of the analysis. Further, if there has been a trend to heavier or lighter vehicles, a trend to respectively less or more severe injury would be expected.

Selection of crashes

The starting point is all crashes (whether head-on, rear end, right angle, or other type) involving exactly two cars and no other units. This excludes the active involvement of other vehicles, pedestrians, roadside objects, etc.

- The cars to be considered are those of build years 1990-2008. If the year of either car is unknown, the crash is excluded.
- The time period for the crashes is 1991-2008.
- The term "car" will include, to begin with, vehicles coded as station wagons, panel vans, and utilities and thus include sports utility vehicles and other four-wheel drive vehicles.
- Drivers of all ages are included.

Some variations on these four decisions will also be considered. Vehicles that are used similarly to cars or that are usually variant body styles of them --- station wagons, SUV's, utilities, panel vans, and taxis --- will be described as falling within a broad definition of "car", but outside a narrow definition that includes sedans only. No consideration is given to crash type (frontal, side impact, rear impact etc.), because of the difficulty of identifying in TARS records the positions of the vehicles in asymmetrical crashes.

Summarising a cross-tabulation

A possible way of summarising the difference between the two drivers is to use the numbers in which the driver of the newer car was more severely injured or was less severely injured --- calculate the ratio of these, for example.

However, the following may be more easily interpretable. A threshold of injury severity is specified. For example, the threshold might contrast hospital-admitted or fatal cases with lower severities. Then:

n_1 = the number of crashes in which the driver of the more recent vehicle was seriously injured (i.e., above the threshold) and the driver of the older vehicle was not (i.e., below the threshold), and

n_2 = the number of crashes in which the driver of the more recent vehicle was not seriously injured (i.e., below the threshold) and the driver of the older vehicle was seriously injured (i.e., above the threshold).

Thus n_1 and n_2 are counts of what might be called discordant crashes. If the secondary safety of cars has been improving over time, n_2 will be expected to be greater than n_1.

Table 1 shows the joint distribution of the injury severities of driver 1 and driver 2 in the selection of crashes described above. Making a within-crash comparison of drivers' severities of injury has the advantage that the speed of the crash is the same for both drivers. The process of assigning injury severities to the two drivers is likely to be more consistent than for two randomly-chosen casualties, too. Of course, presenting the data aggregated like this is only appropriate if change has been in the same direction (e.g., improvement) throughout the time period considered.

Further refinement, such as making allowance for, or matching on, other variables may be desirable. For example, the impact directions may be different for the two vehicles, one vehicle may be bigger than the other, and the drivers may differ in their susceptibilities to injury. A positive correlation between the two severities of injury is evident in Table 1. Our inability to allow for the mass ratio of the two cars tends to produce a negative correlation between the injury severities. Variability between crashes in the relative velocities of the vehicles is sufficiently great to overwhelm this.

The numbers n_1 and n_2 in Table 1 are as follows.

- If a contrast is made between fatally and non-fatally injured drivers, n_1 and n_2 are respectively 10 (= 4 + 6) and 37 (= 13 + 24).
- If a contrast is made between drivers killed or admitted to hospital and those with a less severe injury, n_1 and n_2 are respectively 225 (= 221 + 4) and 304 (= 291 + 13).

Obtaining these numbers and noting that the ratio n_2 / n_1 exceeds 1 is the most important aspect of the present paper. If desired, a statistical test of the null hypothesis of equality of these numbers could be conducted. A refinement is to use a regression equation to relate differences in outcomes to differences between the cars and between their drivers, as will be described below.

Table 1. Cross tabulation of driver injury severity of older and new cars (broad definition) for all two-car injury crashes (1991-2008), where both cars were manufactured in 1990 or later. Coding of injury severity is 1 = no injury, 2 = treated by doctor, 3 = treated at hospital, 4 = admitted to hospital, 5 = fatal.

		Severity of driver injury in newer car		
		1-3	4	5
Severity of driver injury in older car	1-3	10170	221	4
	4	291	77	6
	5	13	24	8

Interpretation

The numbers n_1 and n_2 are interpretable using the following model. Injury severity has, in principle, a continuous scale, not just five categories. Suppose that in any one crash the two distributions of injury severity to the two drivers are respectively logistic with mean = 0 and standard deviation = 1, and logistic with mean = m and s.d. = 1. (As the s.d.'s have been specified as 1, the difference of means is in units of the common s.d.) Then, for logistic distributions, the ratio of the discordant counts is the same whatever boundary is chosen: it is related to the difference of means, m, not to the choice of boundary (McCullagh [6]; Hutchinson [7, 8]).

We should note, however, that in Table 1 and in other similar tables that we have constructed, the ratio of the discordant counts has usually been biggest when contrasting fatal with non-fatal injury, intermediate when contrasting serious or fatal injuries with slight or no injury, and smallest when contrasting injury with non-injury. The same is true in the data of Martin et al. [5].

- It may be that the model is not a good description of what is happening. After all, it would be no surprise if more effort were put into reducing the higher severities of injury than the lower.
- A specific possibility is that random variation in the position of the boundary between one level of injury and the next distorts the results. The model specified in the previous paragraph assumed that it is possible to divide injury severity at some sharp boundary. If instead this boundary is a random variable, perhaps because one driver chooses to seek medical attention and another driver with the same injury does not, it will be equivalent to the s.d.'s of the distributions increasing from 1 to some greater value s. As the difference of means, m, is expressed is units of

the s.d., this is equivalent to the difference of means being reduced to m/s. Thus the ratio of the discordant counts becomes closer to 1.

If indeed there is variability in assessment of injury severity, this suggests that the ratio of the discordant counts for the contrast of fatal vs. non-fatal is more nearly correct than the others, and that this would apply at lower injury severities if it were possible to impose a sharp boundary to distinguish more severe from less severe injuries. (This is all based, of course, on a model and some assumptions, and thus it should apply approximately rather than exactly.) In particular, it seems likely that the ratio for the most severe of the admitted cases will be similar to that of the fatals.

Multiple regression

The form of the multiple logistic regression is $\ln(p_2/p_1) = \beta_1.x_1 + \beta_2.x_2 + \beta_3.x_3$.

- Probabilities p_1 and p_2 refer respectively to the drivers of car 1 and car 2 being killed, conditional on exactly one of them being killed (and thus $p_1 + p_2 = 1$).
- The β's are coefficients to be estimated.
- x_1 is the independent variable of chief interest, the difference between the two cars in the years they were built. (The difference is positive if car 2 is the more recent. If there tends to be less serious injury in more recent cars, β_1 will be negative.)
- x_2 is the difference between the two drivers in their ages. (The difference is positive if driver 2 is older. If the probability of death is higher for older drivers, β_2 will be positive.)
- x_3 is the difference between the cars in whether or not they fell within the narrow definition of "car". (The difference is +1 if car 2 is outside the narrow definition and car 1 is within, and -1 if the reverse is the case. If, as we suspect, vehicles that are outside the narrow definition of car tend to be the heavier, β_3 will be negative.)
- There is no constant to be estimated, because when the differences x_1, x_2, and x_3 are all 0, the two cars are equivalent, and so p_1 and p_2 must both be 0.5.

As will be noted below, other predictor variables were also used. However, in the case of the fatal crashes, the dataset seemed to be too small to support an expanded set of predictors.

Results

Selection of a subset of crashes

In the period 1991-2008, there were 121880 crashes reported in which at least one person was injured. In 21421 crashes, no person was killed but at least one person involved was admitted to hospital, and in 2491 crashes, at least one person was killed (in all, 2819 people were killed). Only a fraction of these were collisions between two passenger vehicles of the type considered for this study: 41436 crashes were reported from 1991-2008 that involved two vehicles that fell into the broad definition of "car" and no other units. In 3033 crashes, the most serious injured driver required admission to hospital or was killed.

When crashes are considered in which the build year of the oldest vehicle was 1990 or later, the crash population can be described as follows:

- 11742 injury crashes occurred that involved precisely two units (vehicles) that fell into the broad definition of "car", and no other units. (In some of these crashes, the cars were of the same year, and thus are excluded from Table 1.)
- 700 resulted in at least one of the drivers being killed or admitted to hospital.
- 57 caused the death of one or both of the drivers.
- 26 of these fatal crashes involved vehicles both of which fell into the narrow definition of "car".
- In 21 of these crashes only one driver was killed.
- Restricting consideration to crashes in which both drivers were under 65 years of age produces a set of 14 fatal crashes in which only one driver was killed, in 12 of which there was a difference in vehicle years.
- There were 49 crashes included in the multiple regression, consisting of those where one driver was killed and the other was not, where drivers were of any age, and where the vehicles fell into

the broad definition of "car". The 47 of these for which the cars differed in year are included in Table 1.

Application of our selection criteria has resulted in a very small set of crashes.

Multiple regression

Multiple logistic regression was carried out as described earlier. There were 49 relevant crashes (the 47 in Table 1 for which one driver died and the other did not, plus 2 crashes for which the cars were of the same year). The effects β_1, β_2, and β_3 were estimated to be -0.33 (per car year), 0.065 (per year of driver age), and -1.52. All are of the expected sign. All are statistically significant ($p < .05$), the respective standard errors being estimated as 0.11, 0.023, and 0.68. However, it might be thought that β_1 = -0.33 (per year) is too large an effect to be credible: if the cars differ by 10 years, for example, the ratio of p_1 to p_2 would be 27.

Further evidence

The following points are offered in support of the contention that serious injury tends to be less in more recent cars.

A. A strong effect can be seen simply by tabulating the data. Consider collisions between exactly two cars (narrow definition), with both drivers being under 65, with the cars being of different years, and in which only one driver was killed. There were 12 such car driver deaths. Of these, 9 were in the earlier car and 3 were in the more recent car, a ratio of 3 to 1. (In this subset of crashes, the earlier car was on average 4 years older than the more recent).

B. If the broad definition of "car" is used, and collisions are not restricted by driver age, the number of crashes is 47, of which 37 driver deaths were in the earlier car and 10 in the more recent car (see Table 1).

C. A similar regression was carried out with the probabilities p_1 and p_2 referring to the drivers of car 1 and car 2 being seriously injured, conditional on exactly one of them being seriously injured. ("Seriously injured" refers to being either admitted to hospital or killed.) The effect of car year was appreciable, though weaker than for fatalities, and statistically significant: β_1 was estimated to be -0.062 (per car year) (the estimated standard error was 0.017). β_2 and β_3 were estimated to be 0.084 (per year) and -0.58 (the estimated standard errors were 0.020 and 0.16). (In this case, x_2 refers to the difference in years of age over 65.) The estimated standard errors of β_2 and β_3 were 0.020 and 0.16. Additionally, the difference between the drivers in whether or not they were male was included in the regression. (The difference is +1 if driver 2 is female and driver 1 is male, -1 if the reverse is the case, and 0 if they are of the same sex.) This was motivated by the possibility that male drivers tend to drive larger cars, in which case the coefficient would be positive. The coefficient had not been statistically significant when analysing fatalities, but in this case it was, and was estimated to be 0.91 (the estimated standard error was 0.15).

The equation fitted uses specific predictor variables with a specific subset of crashes. And even if the predictor variables are appropriate ones, they could be implemented in different ways: difference in the vehicles' years within the 1990's might have a different effect from difference in the vehicles' years within the 2000's if the rate of progress has been uneven, or years of driver age over 65 might have a different effect from years of driver age under 65. Consequently, other regression equations were also fitted, using different predictor variables with different subsets of crashes. This is not altogether good practice, as it gives many opportunities for variables to appear significant by chance. However, in the present instance, focus is on a single independent variable, x_1: when β_1 remains significant despite variations in the exact regression performed, confidence in this result is increased. There still remains the limitation that there are rather few crashes, and regressions on more than about three or four variables sometimes do not give meaningful results.

In the case of the fatal crashes, we attempted including the masses of the cars in the dataset (using the registration number to identify the vehicle). Both cars' masses could be found in 33 of the 49 cases. However, inclusion of the difference between the logarithms of the cars' masses in the regression led to a substantial increase in the standard errors of the estimates of the β's. None of the predictors (not even mass) were now significant, though the increased standard errors lead us to consider this negative result unreliable, the dataset now being a little too small for meaningful results.

Discussion

Interpretation of results

The above results are capable of different interpretations. Our interpretation is that there has been an appreciable improvement in recent decades, but the estimate of β_1 (-0.33, with an estimated standard error of 0.11) seems implausibly large. On the other hand, the estimate based on hospital admission (-0.062) may be smaller than it should be because of inconsistency in the hospital admission decision. In fairness, two alternative interpretations should be noted.

The straightforward, optimistic, interpretation is that recent cars have much improved secondary safety than cars even a few years older. Advocates of this interpretation would point to the consistency of the result for fatalities: a strong effect is seen whether a regression analysis is performed or whether a simple comparison is made as in (A) or (B). The relative weakness of the result for hospital admission would be attributed to inconsistency in the hospital admission decision. The case against this interpretation is that it is not plausible that such a large improvement has taken place in so short a time.

A pessimist might say that there is not much evidence for any effect. The difference between the counts of crashes, 9 and 3 in (A) above, is not statistically significant. Results based on 12 cases from a period of 18 years, selected so as to make the comparison of injury severities as fair as practicable, are in any case such a small proportion of total crashes that results should not be claimed to apply broadly. The broader subset in (B) could be dismissed as not relevant because the crashes are heterogeneous. Similarly, the results from multiple logistic regression could be said to be unconvincing because the estimated β_1 is averaged over a heterogeneous set of crashes (and there are still only 49 cases). Finally, the result for hospital admission could be dismissed with a claim that being admitted to hospital is a poor proxy for injury severity.

As already said, our view is intermediate between the optimist and the pessimist.

Critical consideration of methodology

Four features of this analysis should be acknowledged.

First, the starting point for this study was the concern that damage-only crashes involving old vehicles might be underreported. Analogously, it might be that if it is the older car that has the lower severity of injury, it is more likely for the year to be unknown (and thus the crash excluded from our comparison) than if the older car has the higher severity of injury. However, car year was unknown in only a small proportion of cases: at most 8 per cent for the regression involving fatalities and at most 11 per cent for the regression involving hospital admission cases. (We can only say “at most” because it is possible, perhaps likely, that many of the missing-year vehicles were too early to be included in our analysis.)

Second, it could be that more recent cars are not themselves more crashworthy, but are instead more aggressive (i.e., produce more severe injury in the other car, rather than reduce injury to their own occupants).

- There will indeed be greater aggressivity if recent cars are of higher mass. Analysis of a dataset that includes vehicle mass would establish or refute that there is an effect of vehicle year above and beyond that of vehicle mass.
- As to the hypothesis of other forms of aggressivity such as a high floor pan, it seems impossible to conclusively refute it. However, it seems unwarranted to take the view that aggressivity rather than crashworthiness is the main factor, when design is directed to reducing injury to a vehicle’s occupants rather than increasing it in other vehicles.
- Newstead et al. [9] have reported no long-term trend (within vehicles manufactured 1964-2000) towards either improving or worsening aggressivity.

Third, the difference between the two cars in the years they were first registered is the same as the difference in their ages. It is an assumption that what is important is the year of the car, not the age. An effect of age is not utterly out of the question (e.g., there could be corrosion of the vehicle structure, or increased slack in the seat belts) but it seems unlikely in this set of fairly new cars.

Fourth, the strategy of within-crash comparison compensates for things that are the same for the two drivers (notably the speed of the crash and how the classification of injury severity is carried out), but cannot allow for things that are different. These include the following.

- Frailty of driver. In some of the analyses, we minimised this effect by making allowance for, or restricting the range of, driver's age.
- Mass of vehicle. As noted earlier, whereas the common factor of speed will introduce a positive correlation between the injury severities, the mass ratio will introduce a negative correlation; in practice, there is a positive correlation between injury severities.
- Usage of seat belt. In recent years, the usage of seat belts by drivers has been high in South Australia --- over 90 per cent in surveys since 1998 and about 98 per cent in 2009 (Wundersitz and Anderson [10], especially Table 3). However, the wearing rate may conceivably be a little lower in older cars than in newer, in which case our results would be biased in favour of the new cars.

These issues should not be overlooked, but for the reasons given we feel they are unlikely to affect interpretation of the results.

Commentary

As already stated, our interpretation of these results is that in South Australia, recent cars have appreciably better secondary safety than older ones. If this is correct, there are at least four possible types of reason.

- Increasing inclusion of specific safety features such as superior restraint systems.
- Improved resistance to intrusion into the occupant compartment.
- Sophisticated management of the crushing process during impact.
- More recent cars may tend to have a greater mass.

It would be desirable to establish the total effect of the first three of these, distinct from that of car mass, but we have been unable to disentangle these different reasons with the data available. The most likely improvements feasible with another dataset (e.g., from another jurisdiction) are the availability of the ratio of masses of the colliding vehicles as a regressor, along with a greater number of crashes.

The evidence here is compatible with other reports. In the U.S.A., Ryb et al. [11] have recently found a very substantial improvement in the probability of occupant death from model years 1994-1997 to model years 2005-2007. Earlier, in a paired-driver study by Crandall et al. [4], there was a small but significant effect of the difference in vehicle ages, with vehicle weight, air bag deployment, and restraint use among the variables also present in the equation. In France, a paired-driver study by Martin et al. [5] found a strong effect of vehicle age, whether or not vehicle weight was in the equation.

We consider that the claim (Newstead et al. [1, 2]) that the crashworthiness of cars in Australia has been improving is justified. Our method of analysis, that involved comparing the severities of injury to the two drivers in the one collision, proved practicable though not free of difficulties.

Acknowledgements

We are grateful for financial support of this project from Austroads. The Centre for Automotive Safety Research receives core funding from both the South Australian Department for Transport, Energy and Infrastructure and the South Australian Motor Accident Commission. The views expressed are those of the authors and do not necessarily represent those of the University of Adelaide or the funding organizations.

References

1. Newstead S, Watson L, Delaney A, Cameron M. (2004). Crashworthiness and aggressivity of the Australian light vehicle fleet by major crash type. Report 227, Monash University Accident Research Centre.
2. Newstead S, Watson L, Cameron M. (2008). Vehicle safety ratings estimated from police reported crash data: 2008 Update. Australian and New Zealand crashes during 1987-2006. Report 280, Monash University Accident Research Centre.
3. Campbell BJ, Kihlberg JK (1965). Seat belt effectiveness in the non-ejection situation. In: D M Severy (Editor), Proceedings of the 7th Stapp Car Crash Conference, pp. 177-188. Charles C Thomas, Springfield, Illinois.
4. Crandall CS, Olson LM, Sklar DP (2001). Mortality reduction with air bag and seat belt use in head-on passenger car collisions. American Journal of Epidemiology, 153, 219-224.
5. Martin J-L, Derrien Y, Laumon B (2003). Estimating relative driver fatality and injury risk according to some characteristics of cars using matched-pair multivariate analysis. Paper 364, Proceedings of the18th International Technical Conference on the Enhanced Safety of Vehicles. NHTSA, Washington, D.C.
6. McCullagh P (1977). A logistic model for paired comparisons with ordered categorical data. Biometrika, 64, 449-453.
7. Hutchinson TP (1980). The definition of restraint effectiveness. Accident Analysis and Prevention, 12, 81-93.
8. Hutchinson TP (1982). Statistical aspects of injury severity. Part IV: Matched data. Transportation Science, 16, 83-105.
9. Newstead S, Watson L, Cameron M (2004). Trends in aggressivity of the Australian light vehicle fleet by year of manufacture and market group: 1964 to 2000. Report 214, Monash University Accident Research Centre.
10. Wundersitz LN, Anderson RWG (2009). Results from an observational survey of restraint and child restraint use, 2009. Proceedings of the 2009 Australasian Road Safety Research, Policing and Education Conference and the 2009 Intelligent Speed Adaptation Conference, pp. 295-302. Roads and Traffic Authority of New South Wales, Sydney. (ISBN of USB: 978-1-921692-26-0.)
11. Ryb GE, Dischinger PC, Ho S (2009). Vehicle model year and crash outcomes: A CIREN Study. Traffic Injury Prevention, 10, 560-566.