FATAL ERRORS WITH FATALITIES DATA: A COMMENT ON ROBERTSON'S "AN INSTANCE OF EFFECTIVE LEGAL REGULATION"

ROBERTA S. COHEN
HARVEY S. COHEN

We noted with considerable interest Leon Robertson's article (1976), in a recent issue of this journal, dealing with statistical measures of the effectiveness of helmet and daytime headlight laws for motorcyclists. For some time, proponents and opponents of such legislation have been trading enthusiastic arguments and impassioned denials, often including misleading statistical presentations. The result has been that most followers of this debate are left with their personal views and little in the way of substantive support for either side. Robertson uses what appear to be promising data bearing on the safety controversy but, in our view, he does not perform the analyses required to clear the muddled waters and he seriously misinterprets the statistics he employs.

Robertson presents data bearing on three related issues in the motorcycle safety controversy: the levels of compliance with existing safety regulations, the impact of helmet use in reducing traffic fatalities, and the impact of daytime headlight use in reducing fatalities. In a limited observational study in four localities, Robertson finds that helmet and daytime headlight use are considerably greater in those states that require such use than in those that do not. Robertson's discussion of these data goes no further, yet the data can serve to illuminate the impact of legal change. First, the data indicate that the maximum possible impact of introducing helmet or headlight legislation may vary widely from state to state. States without such laws have widely divergent rates of helmet and headlight use. If Robertson's data are reflective of safety measure use throughout the state, and if the high rates of compliance with legally required safety measures that he observes are the norm for all states, the percentage of additional motorcyclists who would use the safety measure in response to the enactment of a law could range from 75 percent to only 40 percent. If such safety measures do reduce fatalities the reduction should vary, at least in part, as a function of the number of additional riders using the measure.

Second, Robertson neglects to note that there may be an interaction in compliance when the two laws—helmet and daytime headlight—are examined simultaneously. This is suggested by his data showing that states that require either helmets or daytime headlights obtain greater compliance with the single measure than the state that requires both measures (Georgia). Moreover, the use of both helmets and headlights is greater in the state that requires neither (California) than in those states that require one but not the other. Apparently, use of a safety measure is a function not merely of whether it is required, but also of whether other measures are required.

The data brought to bear on the other two questions raised in Robertson's paper are not as easily understood. In order to measure fatality reduction due to helmet use, Robertson employs a three-year comparison of eight states adopting helmet laws, beginning with the year before adoption, and eight "matched comparison" states not adopting such laws. Geographical contiguity appears to be the basis for matching, although this is not specified. Usually, an investigator matches sampled units to control for one or more variables, constant within matched units, which are not included in the formal analysis. Since Robertson's analysis included a test for differences in average temperature, it would seem more sensible for matching to have been done on some criterion not related to temperature but also thought to affect riding and/or traffic accident rates, such as the ratio of urban to rural driving, highway congestion, precipitation, or number of motorcycles. The matching procedure, moreover, suggests that the states are to be compared pairwise; in fact, they are not. Robertson performs an analysis of variance (ANOVA) using a 2 x 3 x 2 design of two law conditions by three years by two climate conditions (degree days), a procedure that pools the fatality rates of states that fall into the same cell of the design.

This design may not be the correct one for these data because the same states are examined in each of the three years. In Robertson's ANOVA, two sources of variation are examined—the variation between cells (accounted for by the variables used in the analysis) and the variation within cells (seen as the residual or error term). The test for a significant effect of any variable is performed by taking a ratio of the variance associated with that variable to the variance found within the cells. There are other ANOVA designs available which handle the variance differently. These designs, known as repeated measures or split plot designs, divide the between-cell and within-cell variation into two parts—variation that occurs between subjects (states in this analysis) and variation that occurs within subjects. This is most readily under-

stood with regard to the residual, the within-cell variation. Part of the within-cell variation measures the extent to which the mean rate over years for a state differs from the mean rate for the group of states sharing the same cells. This is the variation within cells not accounted for by between-state factors. An ANOVA using the within- and between-subject notion of variation is carried out in much the same way as the standard ANOVA, except that the variable effects are assigned to that part of the variation that can be attributed to them—the within-subject or the between-subject—and are tested against the residual for only that part of the variation.

In Robertson's ANOVA, three main effects are tested: a law effect, which measures the differences in rates that occur between the law states and the nonlaw states summed over all three years; a year effect, which measures the differences in rates that occur among the first, second, and third years of measurement, summed over all states (law and nonlaw); and a climate effect, measuring the differences in rates between states above the heating-degree day median and those below the median summed over years, regardless of law "status". Applying the within-state/between-state distinction to these main effects, the law effect is clearly a between-state effect—states enacting the law must differ significantly from states not enacting the law for this effect to be significant, because the difference is measured across all three years. The effectiveness of the law in reducing fatal involvement rates is an effect that we will find *within* states and one that is dependent upon year. The year effect is a within-state effect, but it only tells us whether rates summed across law conditions differ from year to year. If the law is effective in reducing fatal involvement rates, we would expect rates to decline by year among law states, not to decline by year among nonlaw states, and we would expect law states to differ from nonlaw states only in the years of and after enactment of the law. This is measured by the *law by year* interaction effect, a within-state effect, and an effect that is *not* significant in Robertson's analysis. A significant *law by year* effect would be unlikely even if we were to divide the residual and the effects according to type of variation and test the effects against only the error associated with that type of variation, since less than 5 percent of the total residual would have to be attributed to the within-state variation.

The analysis performed by Robertson indicates that states that passed helmet laws differed from states that did not pass such laws over all three years (including the period before the laws went into effect), but his ANOVA fails to show that this difference changed significantly over time. Thus, while Robertson claims that

his analysis shows the efficacy of helmet laws, it does not. It merely shows that states with helmet laws are and were significantly different from states without such laws.

Most investigators, on obtaining this result, would perform a year by year analysis to discern whether the *law by year* effect is significant over any two years, such as the year before and the year of enactment, or the year of and the year after enactment. The former is strongly suggested by Robertson's Figure 1 plotting over years the average fatal involvement rates per 10,000 registered motorcycles for the eight states that adopted helmet laws and the eight comparison states that did not. It would appear from this figure that the major impact of the law takes place immediately. The average fatality rate for law states drops from almost 11 the year before enactment to just over 7 the year of enactment, while the average rate for nonlaw states increases slightly from just below 10 to just above 10 in those same two years. The figure may be misleading, however, for several reasons. First, Robertson reports that the average rates of law and nonlaw states for the year before the laws went into effect are not significantly different from one another ($t = 0.45$). Using rough estimates of these average rates derived from the figure, we computed the range of standard deviations that could be associated with those rates to yield a t of 0.45. At least one of those standard deviations is very high since the sum of the standard deviations squared is over 45. If the standard deviations are approximately the same for law and nonlaw states, they are on the order of 4.7 On the other hand, one could be equal to 9 and the other less than one. The larger the standard deviation of the rate within a year, the greater must be the change in rate before it is considered statistically significant. Thus, while the figure appears to show dramatic differences between law and nonlaw states in fatalities history, the points shown may represent nothing more than expected statistical fluctuations entirely within the limits of confidence.

A second problem with Figure 1 arises from the data themselves. If changes in average fatality rates are significant changes (given the ANOVA results, however, and the probable sizes of the standard deviations we must think that they are not), there exists a basic inconsistency between what appears to be happening and what we would normally expect to happen. The helmet laws went into effect at different times during the year of enactment for the eight states adopting helmet laws—two states required use as of January 1, one state as of May 1, three states as of July 1, one state as of July 31, and one state as of September 1. On the average, the laws were in effect for just over 7 months of the enactment year.

The fatality rate for law states in the year before enactment measures 12 months of fatalities with no helmet law. The fatality rate of the same states for the year after enactment measures 12 months of fatalities with a helmet law. But the rate for the year of enactment measures, on the average, about 7 months of fatalities occurring with a helmet requirement and about 5 months of fatalities without such a requirement. If helmet use significantly reduces fatalities, the year of enactment should certainly show some improvement in fatal involvement rate, but the greater reduction in rate should come the year after enactment, when riders are required to wear helmets all year. Robertson's figure does not bear this out. The average fatal involvement rates for states enacting laws are approximately 10.85 the year before enactment, 7.2 the year of enactment, and 7.1 the year after enactment. The only difference is found between the year of no law and the year of only partial implementation.

Finally, it is worth noting that Figure 1 does not represent the data analyzed by Robertson in his ANOVA. The ANOVA used rates corrected for unequal variance by weighting according to the size of the state's cycling population. Such a procedure could considerably alter the raw rates shown in Figure 1 as well as the relative sizes of those raw rates. Moreover, the ANOVA included, and therefore controlled for, a factor, not shown in Figure 1, which had a significant effect on fatal involvement rates—heating-degree days. Were the figure accurately to represent the ANOVA, we should see the weighted rates plotted over time for four sets of states: warmer states that enacted a helmet law, warmer states that did not, colder states that enacted a law, and colder states that did not. Then we could rule out the possibility that the difference between law and nonlaw states is due to a difference between warm and cold states.

The analysis of the headlight data presented by Robertson is also suspect on methodological and interpretive grounds, some of which Robertson himself acknowledges. In this analysis, four "matched comparison" pairs of states are examined over four years. Fatal involvement rates are not used; rather, for each pair, the actual number of motorcycle related fatalities in the state that enacted a daytime headlight requirement is divided by the total number of motorcycle related fatalities in both states for each of four years. The fourth year examined for each pair measures 12 full months of fatalities occurring when the daytime headlight law is in effect. In two pairs, the laws became effective during the third year of the analysis (for 6 months in one case, 3 months in the other).

Geographical contiguity appears to be the matching criterion used, yet in this analysis no attempt is made to control for the effects of heating-degree days. This omission is serious in light of Robertson's helmet results showing degree days to be a significant factor in fatalities. But the basic problem in the daytime headlight analysis is that the impact of the laws is measured against a nonexistent trend. Robertson argues that the "before" years set a trend of increasing fatalities which the "after" year reverses. But it is not clear to us why any sort of trend should appear before the laws went into effect in the *proportion* of fatalities occurring in the four law states relative to the four nonlaw states. Robertson is not dealing with the numbers of fatalities, or even the rates. If a state's fatal involvement rate is increasing relative to that of another state, at a time when neither has a daytime headlight law, other factors would seem to be at fault.

A simpler and more reasonable test of the effectiveness of daytime headlight laws, given Robertson's data, would compare the total "before" proportion, computed on the basis of all three "before" years, to the "after" proportion. We attempted to use the unweighted data in Robertson's Table 3 to make a tentative test of this sort, but typographical errors in the table made that impossible (two of the proportions reported do not correspond to the numbers that purportedly produced them). Using the most likely values for the numbers in error, and computing the proportion of fatalities over all three "before" years for each of the four pairs, we find that the proportion increases in two and decreases in two. Averaging the pairs together, we find a slight decrease in the fatality proportion between the "before" and "after" years, too small to attain statistical significance.

We must conclude that neither Robertson's helmet analysis nor his headlight analysis yields results that demonstrate the effectiveness of such safety measures in reducing fatal accident involvement. We must point out, however, that these data do not negate the possibility that such a result could be forthcoming from a more straightforward and more carefully controlled set of analyses. We strongly disagree with Robertson when he states that "motorcyclist helmet use laws represent social policy that has been effective in achieving the purpose of reducing fatal injuries" (1976:475). His analysis, as reported, simply does not justify this statement. But we endorse his call for better data on types of motorcycle crashes (he refers specifically to headlight use, but we would extend that to include helmet use) and we would endorse efforts to improve research on the effectiveness of highway safety legislation generally. With repeal of that provision of the Federal

Highway Safety Act requiring states to adopt helmet laws if they use federal highway monies, it seems all the more important that sound research inform states about effectiveness of helmet and headlight laws in reducing injury and death on the highway.

REFERENCE

ROBERTSON, Leon S. (1976) "An Instance of Effective Legal Regulation: Motocyclist Helmet and Daytime Headlamp Laws." 10 *Law & Society Review* 467.

https://doi.org/10.1017/S0023921600016157 Published online by Cambridge University Press