

TITLE: The California Driver Record Study: A Multiple Regression Analysis of Driver Record Histories from 1969 through 1982

DATE: August 1992

AUTHOR(S): Raymond C. Peck & Michael A. Gebers

REPORT NUMBER: Unnumbered (NRN057)

NTIS NUMBER: None

FUNDING SOURCE: Departmental Budget

PROJECT OBJECTIVE:

To provide a resource for researchers and statisticians having an interest in driver accident correlates and accident prediction modeling.

SUMMARY:

The paper summarizes the results of a large number of regression analyses of driving record variables measured over several time periods. It was planned to use the results from the analyses to guide the development of models using similar data extracted in 1992 as part of the California Driver Record Study database. Because this latter effort would both include and extend the data set used for these analyses, the paper was limited to a presentation of the various multiple regression equations and a very brief summary of the major highlights.

Data for the analyses were obtained from a 1% random sample of approximately 180,000 licensed California drivers, whose records were extracted during the 1983 update of the California Driver Record Study database. The variables comprising the predictor set represented the majority of potentially relevant driving population parameters contained in California driver record files, and they were chosen to be consistent with variables used in previous California driver record studies. Stepwise multiple regression analysis was used for identifying the combination of variables that provided the most accurate prediction of the criterion measure. For most of these analyses, the criterion measure was an accident frequency variable, but in some of them a traffic conviction variable served as the criterion measure.

The results indicated that total accidents were the most predictable of the accident criterion measures, but the highest multiple R (.216) for the nonconcurrent data (independent driver record time periods) was still disappointing. The structure of the equations was very similar to that found in prior California studies. It was recommended that planned future studies explore a number of techniques for increasing predictive accuracy, including lengthening the criterion interval beyond 5 years.

IMPLEMENTATION STATUS OF FINDINGS AND RECOMMENDATIONS:

Not applicable. The results are being used in directing subsequent model development efforts.

SUPPLEMENTARY INFORMATION:

See Gebers and Peck, Report #144.