California
" of Administration

STATE OF CALIFORNIA
DEPARTMENT OF MOTOR VEHICLES
Division of Administration
John L. McLaughlin, Chief

FHWA 97-2625-9

THE 1964 CALIFORNIA DRIVER RECORD STUDY

PART 9

The Prediction of Accident Involvement From Driver Record and Biographical Data

by

R. S. Coppin, R. S. McBride and R. C. Peck

RESEARCH AND STATISTICS SECTION

in cooperation with
DIVISION OF DRIVERS LICENSES

APRIL 1967

SUPPLEMENTAL INFORMATION
FHWA DOCKET 71-1

47 pp.

C2A-17

PREVIOUS DRIVER RECORD STUDY RELEASES

Part 1

An Introduction and Methodological Description

Part 2

Accidents, Traffic Citations and Negligent Operator Count by Sex

Part 3

Drivers by Age, Sex and Area of Residence

Part 4

The Relationship Between Concurrent Accidents and Citations

Part 5

Driver Record By Age, Sex and Marital Status

Part 6

The Stability of Reported Accidents and Citations

Part 7

The Relationship Between Types of Convictions and Accidents

Part 8

The Prediction of Accident Involvement Using Concurrent Driver Record Data

TABLE OF CONTENTS

1/P - 2000

LIST OF TABLES

CHART

EXHIBITS

ERRATUM TO PREVIOUS PARTS OF THE DRIVER RECORD STUDY

Part 5

In Chart III, the plotting of the mean accident frequency for single males, aged 36-40, is incorrect. It should be lowered to correspond to an accident value of .102 on the vertical axis.

Part 8

The values for the speed and major conviction categories in the list at the top of page 10 are incorrect and should read as follows:

Speed convictions	1	0
Major convictions	0	2

FOREWORD

"The Prediction of Accident Involvement from Driver Record and Biographical Data" represents the ninth and final part of the 1964 California Driver Record Study. This final part, together with its immediate predecessor, Part 8, are concerned with the prediction of accident involvement and the isolation of important driver record relationships. As such, they are the logical culmination of the study, since they involve a consideration of every aspect of the data and their interrelationship.

Part 9, as all previous parts, is the product of the Department's Research and Statistics Section and was accomplished under the general direction of Ronald S. Coppin, Statistical Research Officer. The principal investigators on Part 9 were Robin S. McBride and Raymond C. Peck, Research Analysts.

As might be expected in a study of this magnitude, the number of persons who made significant contributions are too numerous to be acknowledged individually. Special mention, however, is due Gareth Ferdun, formerly a research analyst with this department, for coordinating much of the computer processing of the regression analysis and descriptive tabulations, as well as directing the questionnaire phase of the study. In addition, we must express appreciation to Ronald V. Thunen, Administrator, Division of Drivers Licenses, for his many worthwhile suggestions and constructive criticism of all phases of the study. The Department is also indebted to Dr. Harry M. Hughes, Chief of Data Processing, USAF School of Aerospace Medicine, for contributing the Regression Analysis program used in Parts 8 and 9 of the study.

The Department is confident that the 1964 California Driver Record Study represents a significant contribution to the science of driving behavior and will prove useful to all persons, professional or lay, having an interest in traffic safety and drivers license administration.

John L. McLaughlin

JOHN L. McLAUGHLIN, Chief
Division of Administration

SUMMARY

The more important and revealing findings discussed in this report are summarized as follows:

1. The best overall accident predictor for both concurrent and non-concurrent events is the total number of one count convictions on file. Beyond a knowledge of the number of one count convictions, the contribution of the specific types of violations is very small. The data strongly suggests that the total number of convictions (regardless of type or count) accrued by a driver would be as efficient in predicting accidents as the more complex method of assigning different weights to the various types of violations.

2. Biographical information about a driver (age, marital status, area of residence, physical stature, etc.) slightly increases the accuracy of accident prediction beyond that achieved through knowledge of driver record variables alone. Further additional data obtained from a questionnaire (mileage, occupation, etc.) resulted in a two-fold increase in predictive accuracy.

3. The relationships proved higher for events occurring in the same time period (concurrent) than for events occurring in different time periods (non-concurrent).

4. The relative association between the predictor variables and accident involvement differs for males and females. Also, in both concurrent and non-concurrent prediction, the overall magnitude of the relationships is higher for males than females.

5. A theoretical analysis of the accident distributions indicated that the maximum in predictive efficiency was not achieved with any of the equations; therefore, the inclusion of additional data about the driver would probably increase accident predictability.

INTRODUCTION

This report is the ninth part of a series relating to the California driving population; its topic is the combined relationship between various driver characteristics and accidents. In contrast to Part 8, the characteristics under consideration are not limited to violation patterns but include such variables as age, marital status, physical stature and exposure to accidents. The report also analyzes a variety of non-concurrent relationships -- that is, an attempt is made to actually predict accident involvement from prior events. In all previous parts of the Driver Record Study, the analysis has been confined largely to concurrent relationships.

In all cases the driver record events under consideration occurred during 1961 through 1963 and are based on a random sample of 148,000 drivers with complete 36 month driver records. Additional data was derived by questionnaire on a small sub-sample of this total.

The identification of accident-related traits is, of course, a subject of much concern to everyone, particularly to persons and agencies involved in driver-related activities. Drivers license administrators, traffic judges, insurance underwriters, safety educators and driver improvement personnel all ask: "If I know certain information about a driver (violation record, age, etc.), how accurately can I predict his accident involvement?" This type of question will be explored herein, with the hope of isolating relationships that may prove to have both theoretical and practical significance. In so doing, three specific questions will be posed for consideration:

1. Does knowledge of a driver's age, marital status, area of residence, physical stature, etc., increase our ability to predict his accident involvement?

2. How do predictions made from non-concurrent events differ from those made from events occurring during the same time interval (concurrent predictions)?

3. To what extent does data not ordinarily available from the official driver record file (e.g. annual mileage, occupation, type of driving, number of dependents, etc.), increase one's ability to predict accident occurrence?

METHODOLOGY

Before presenting the results, a few pages will be devoted to a definition of concepts, terms and methods relating to the data analysis. This methodological description will be rather brief, since both the computer processing and statistical analysis (multiple regression) are almost identical to those of Part 8, where the matter was discussed in detail. The sampling design and data collection, having been discussed in Parts 1 and 2, will not be repeated.

Stated briefly, multiple regression analysis is a technique for predicting the frequency of an event (e.g. accident involvement) from a pool of data describing a group of individuals. As applied to accidents, the outcome of the procedure enables one to estimate the accident-likelihood of an individual or unified group by plugging the data describing the individual or group into a mathematical equation. The appropriate equation is initially obtained by a mathematical analysis of identical data on a similar group of individuals.

For this study, data comprising the predictor variables (age, marital status, violation record, etc.) and accident involvement record (criterion variable) were transcribed from each drivers' official file record. With the aid of a generalized regression analysis program and a Philco 2000 series computer, the unique association between each predictor variable and the criterion variable (accident involvement) was determined.

The unique association between the predictor variable and accident involvement is generated in the form of weights (regression coefficients) which, when applied as multipliers to the data obtained from each driver's record, produce an estimate of each driver's accident involvement likelihood. The coefficients indicate how many units the predicted accidents increase for every increase in the units of a predictor variable. For a more complete discussion of the regression analysis technique, the reader is referred to pages 7-10 of Part 8 of this study.

The final regression equations are presented separately for males and females. The rationale for separating the equations is evident when recognizing that the underlying relationships between the predictor variables and accident involvement differ by sex (2); therefore, combining males and females would obscure meaningful relationships.

Another important aspect of regression analysis is that of cross-validation. Cross-validation involves applying the equation to an independent group of drivers and comparing the predictions thus derived for each driver with his actual recorded accident frequency. The purpose of this procedure is to derive an unbiased estimate of an equation's true predictive power or validity, which tends to be overestimated by the original multiple correlation coefficient (R).

In order to achieve cross-validation, it was necessary to split the 148,000 subject sample into two groups -- an equation construction group and a validation group. In a fashion similar to that of Part 8, the equations were generated from the construction sample and applied to the validation sample by each sex.

For the reader who is already familiar with Part 8 of the Driver Record Study, it should be noted that Part 9 differs in three important respects, and will be structured around these extensions:

1. Additional variables from the Division of Drivers License files have been included on all drivers. (Definitions in Exhibit A.)

2. Data obtained from a questionnaire was collected (on a small sub-sample) and included as a subsidary analysis. (Exhibit C, Appendix A.)

3. Non-concurrent relationships will be considered in addition to concurrent relationships.

RESULTS AND FINDINGS

The reader is referred to Chart I on page 6 where the regression equations, predictor variables and multiple correlations are depicted for both concurrent and non-concurrent relationships. The equations are also presented separately by sex. It should be noted that the bases of distinction among the equations are (1) concurrent vs. non-concurrent relationships, (2) males vs. females, (3) number of variables and differences in types of variables, and (4) three year record vs. one year record. All of the equations represent variables which were found to be significantly related to accidents at the .05 level of significance. Those variables found not to be significant were dropped as predictor variables on successive computer runs until each regression equation included only significant variables. Section A will be devoted to comparisons of the multiple correlations among the categories previously mentioned. A discussion of the unique contribution of specific variables is presented in Section B.

Section A. Multiple Correlation

Concurrent Prediction

The multiple correlation coefficients (R), located near the upper margin of equations IA and IB, are .226 and .186 for males and females, respectively. By squaring these coefficients, a measure of the percentage of variability in accident involvement explained by the total contribution of the predictor variables is obtained. The percentage of accident variability explained by the variables in equations IA and IB is 5.1 percent and 3.4 percent for males and females, respectively. One can see that there is a higher relationship for males than females which, incidentally, has been indicated in several previous studies (2, 3, 11).

After a final regression equation has been generated, the validity or degree to which it predicts accident frequency is a most important topic. Since the multiple correlation represents the maximum correlation between the weighted combination of independent variables and a criterion measure (accidents), this estimate capitalizes upon any chance deviations that favor high multiple correlations. Thus, any correlation coefficient computed on the same sample that was used for generating the regression equation may be spuriously high. Highly inflated values are prone to arise in a situation where there are many variables and a comparatively small number of subjects. To check for shrinkage, the regression equations were applied to an independent sample of drivers and the simple correlations between the predicted and actual accident frequencies of that sample were

CHART I

REGRESSION EQUATION BY SETS OF ACCIDENT PREDICTOR VARIABLES

I. Three year concurrent regression equations

A. Males $n = 42{,}228$ $R = .2267$ $R^2 = 5.14\%$

$$\boxed{\begin{aligned}\text{Accidents} = {} & .0818 + .0635(X_1) + .0642(X_2) + .0304(X_3) + .0211(X_4) - .0106(X_5) \\ & - .00768(X_6) + .0613(X_7) - .0000912(X_8) + .00214(X_9) + .0131(X_{10}) \\ & + .0186(X_{11}) + .000367(X_{12})\end{aligned}}$$

Where X_1 = one count convictions, X_2 = two count convictions, X_3 = passing convictions, X_4 = right-of-way convictions, X_5 = equipment convictions, X_6 = miscellaneous technical convictions, X_7 = non-countable convictions, X_8 = age, X_9 = weight, X_{10} = marital status, X_{11}= license restrictions, X_{12} = traffic density

B. Females $n = 30{,}277$ $R = .1860$ $R^2 = 3.46\%$

$$\boxed{\begin{aligned}\text{Accidents} = {} & .173 + .126(X_1) + .0581(X_2) + .194(X_3) + .0371(X_4) - .0258(X_5) \\ & + .0182(X_6) - .0209(X_7) + .0637(X_8) - .0000384(X_9) - .00275(X_{10}) \\ & + .0578(X_{11}) + .000267(X_{12})\end{aligned}}$$

Where X_1 = moving FTA's, X_2 = one count convictions, X_3 = two count convictions, X_4 = right-of-way convictions, X_5 = turning, stopping and signalling convictions, X_6 = speed convictions, X_7 = equipment convictions, X_8 = non-countable convictions, X_9 = age, X_{10} = height, X_{11} = marital status, X_{12} = traffic density

II. Non-concurrent regression equations (Prediction of 1 year from prior 2 year driving record)

A. Males $n = 43{,}509$ $R = .1190$ $R^2 = 1.42\%$

$$\boxed{\begin{aligned}\text{Accidents} = {} & .136 + .0317(X_1) + .0135(X_2) + .0209(X_3) + .0213(X_4) + .0194(X_5) \\ & - .0000398(X_6) - .00105(X_7) + .0000730(X_8)\end{aligned}}$$

Where X_1 = FR accidents, X_2 = CHP accidents, X_3 = FR-CHP accidents, X_4 = one count convictions, X_5 = non-countable convictions, X_6 = age, X_7 = height, X_8 = traffic density

B. Females $n = 30{,}673$ $R = .0961$ $R^2 = .924\%$

$$\boxed{\text{Accidents} = -\ .00279 + .0318(X_1) + .0205(X_2) + .0195(X_3) + .0000834(X_4)}$$

Where X_1 = FR accidents, X_2 = one count convictions, X_3 = marital status, X_4 = traffic density computed. The sample used to generate the regression equation will be referred to as the construct sample. The new sample to which the regression equation is applied for a determination of reliability and validity of the regression equations will hereafter be referred to as the cross-validation sample.

When the regression equations IA and IB were applied to the cross-validation group, no significant differences were found between the multiple correlations derived from the construct and cross-validation samples. These results, as shown in Figure 1, indicate that the regression equations are valid reliable predictors of accident involvement.

Fig. 1 A COMPARISON OF MULTIPLE CORRELATIONS FOR CONSTRUCT SAMPLE AND CROSS VALIDATION SAMPLE BY SEX
(Concurrent prediction)

Sex	Construct sample	Cross validation sample	Probability
Males	.2269	.2343	$P > 0.05$
Females	.1860	.1793	$P > 0.05$

The amount of predictability is rather disappointing and the equations appear to have rather limited application to the practical problem of selecting the accident-liable driver. Before rendering a final evaluation, a description of how the regression equation is translated into a usable screening device will be given.

Since the accident phenomenon is measured as a discrete variable (e.g. 0, 1, 2, etc.), while the values generated by a regression equation are on a continuous scale, it is necessary to find an exact cutoff point on the regression scale which best discriminates between accident and non-accident involved subjects. Individuals with values generated by the equation falling above that cutoff point are predicted to become involved in accidents, whereas those individuals falling below this point are predicted to be accident free. There are two kinds of errors in making such predictions, namely: predicting an individual to be accident involved when he is not (false negative) or predicting an individual to be accident free when he is not (false positive). The percentage of errors in making such predictions is indicative of the practical efficiency of the regression equation. The following fourfold table represents the number of correct and incorrect predictions from regression equations IA and IB when applying an optimally efficient[a] cutoff point of .298 for males

(a)The cutoffs for the regression equation were empirically selected in order to maximize the association between the predicted and actual outcomes (phi coefficient).

and .194 for females. In constructing these tables, any males or females whose predicted accident value exceeded their respective cutoff scores (.298 and .194) were predicted to have one or more accidents.

Fig. 2 PREDICTED ACCIDENT INVOLVEMENT BY ACTUAL ACCIDENT INVOLVEMENT BY SEX

(Concurrent data)

Males

Actual	Predicted 0	Predicted 1
0	28,256	5,364
1	5,290	2,669
Percentage correct	84.2	33.2

Females

Actual	Predicted 0	Predicted 1
0	24,763	2,285
1	2,329	657
Percentage correct	91.4	22.3

Note: Shaded areas are errors of prediction.

For males, the proposed cutoff would result in rejecting or removing from the driving population approximately 8,033 drivers in order to reduce accident involvement by 2,669. For females, 2,942 drivers would be removed to reduce accident involvement by 657.[(a)] In other words, three drivers would be rejected per accident involvement for males and four drivers would be rejected per accident involvement for females. Thus, the equation is more efficient in predicting male accident involvement than it is in predicting female accident involvement. This fact is also borne out by the validity coefficients, on the preceding page.[(b)] This may seem surprising in view of the fact that the female predictions are more accurate on an overall percentage basis. That is, the percentage of misclassifications (false positives and false negatives) indicated by the shaded areas in Figure 2 average about ten percent less for females. This, however, is merely a statistical artifact -- attributable to the fact that females are involved in far less accidents than males and are therefore more predictable on a purely a priori basis. In other words, since about 90 percent of the female driving population is accident free in any three year period, one can be 90 percent correct in predicting their accident involvement merely by predicting all females not to have an accident. The corresponding accident free figure for males is only 81 percent, which renders them less predictable on a purely "blind" basis.

(a) In this particular context, the authors actually mean "drivers involved in accidents," but for simplicity have used the phrase "accident involvement." Actually, had the contingency tables included each accident frequency as a category, the predictive efficiency and accident reduction would have been increased, since it is easier to predict more extreme scores and each such successful prediction would "save" two or more accidents.

(b) For the purpose of verifying the significance of the validity coefficients computed on non-normal data, phi coefficients were computed from the fourfold tables. In all cases, the phi coefficients were statistically significant beyond the .001 level:
Males phi = .175, Females phi = .035
$x^2 = 1275$ $x^2 = 546$

It is quite clear from Figure 2 and the size of the validity coefficients that the predictive efficiency of the equations is low. Consequently, they are of limited utility to the licensing administrator, who must be concerned with both types of prediction errors -- false positives and false negatives. That is, legal, social and public opinion considerations require that the licensing administrator not reject large numbers of potentially "safe" applicants or pass large numbers of "unsafe" or accident prone subjects (16). However, in certain contexts one need only be concerned about false positive errors -- i.e. passing subjects who are potentially unsafe. One such example is the insurance underwriter, who must determine whether a given applicant is a good risk for a certain type of policy. In such a situation, a more stringent cutoff score could be applied in determining whether a subject is a "pass" or "fail." For example, instead of "passing" all males who achieve a predicted accident score of .298 or less (as was done in Figure 2), the underwriter might require that an applicant score less than .100 to qualify for a certain type of policy. This would result in more subjects being failed, but those passed would be a much more select group -- one whose expected accident rate was almost three times lower than that of the group passed on the basis of the .298 accident cutoff.

In actual practice, then, the assignment of a cutoff point is dependent upon the nature of the problem, the type of error one wishes to minimize and the number of potential applicants (selection ratio). In a situation where both types of errors carry equal weight, the most discriminating cutoff point is arrived at empirically, and the selection ratio concept ceases to be a factor.

Non-Concurrent Prediction

The discussion up to this point has been concerned with concurrent prediction or the relationship of events occurring together in the same interval of time -- in the above case, 1961-1963. However, the prediction of accident involvement in one period from that of a prior period is more relevant to driver licensing agencies and insurance companies. It is often the concern of these agencies to estimate the future accident risk of individuals based on their past performance and characteristics. One would expect the relationship in non-concurrent prediction to be lower than that for concurrent prediction since the behavior of the individual and environmental situations are not occurring together over a segment of time. Also, in the non-concurrent equations presented here, the time period of accident involvement to be predicted is only one year, whereas the concurrent equations were based on a three year accident record. The instability of a one year accident record compared to a three year accident record is discussed in detail in Part 6 of the Driver Record Study.

The reader is referred to equations IIA (males) and IIB (females) of Chart I for the non-concurrent regression equations and multiple correlations. As in concurrent prediction, the squared multiple R's, are greater for males (1.42 percent) than for females (0.92 percent). The reader should also note that these multiple R's are in fact much lower than concurrent multiple R's.[(a)] In testing the validity of the non-concurrent R's, the regression equations were applied to the cross-validation sample. No significant difference between the construct and cross-validation samples was found.

Fig. 3 A COMPARISON OF MULTIPLE CORRELATIONS FOR CONSTRUCT SAMPLE AND CROSS VALIDATION SAMPLE BY SEX
(Non-concurrent prediction)

Sex	Construct sample	Cross validation sample	Probability
Males	.1190	.1127	P> 0.05
Females	.0961	.0859	P> 0.05

These results indicate that the non-concurrent regression equations offer statistically significant predictive techniques but are of limited practical utility.[(b)] By utilizing these equations to predict accident involvement in a manner similar to that discussed on page 8, one can partition the drivers into correctly and incorrectly predicted accident categories. The fourfold table at the top of page 11 depicts the number of drivers predicted to become accident-involved by the actual record of accident involvements using the optimal cutoff points of .135 and .083 for males and females respectively.

For males, 2,653 drivers would be rejected to reduce driver accident involvement by 356. For females, 1,021 drivers would be rejected to evidence a reduction in accident involvement of 73. In other words, of those drivers predicted to have one or more accident involvements, 7 males and 13 females would be rejected per accident involvement. As was the case with concurrent prediction, the prediction formulae are more efficient for males even though the overall percentage of misclassifications is larger. Also, as stated earlier, the efficiency of the non-concurrent

(a) This is due in part to the shorter criterion interval used in the non-concurrent case as well as other factors to be discussed later.

(b) For the purpose of verifying the significance of the validity coefficients computed on non-normal data, phi coefficients were computed from the fourfold tables. In all cases, the phi coefficients were statistically significant at the .005 level:
Males phi = .075, Females phi = .038
x^2 = 233 x^2 = 44

Fig. 4

PREDICTED ACCIDENT INVOLVEMENT
BY ACTUAL ACCIDENT INVOLVEMENT BY SEX

(Non-concurrent data)

Males

Actual	Predicted 0	Predicted 1
0	36,024	2,297
1	2,302	356
Percentage correct	94.1	13.4

Females

Actual	Predicted 0	Predicted 1
0	28,053	948
1	960	73
Percentage correct	96.7	7.1

Note: Shaded areas are errors of prediction.

predictions is considerably lower for both sexes than that achieved with the concurrent data. Unfortunately, it is the non-concurrent case which is most applicable from a practical standpoint, since true predictions are inherently non-concurrent in nature.

These results indicate that accidents are only slightly predictable from driver record variables alone. Even though the equations may have some application in situations where the selection ratio can be low, they have only limited utility for the motor vehicle administrator. The implications of these results will be discussed in more detail in a later section of this report.

Section B. Components of Regression Equations

The reader is referred to Tables I and II for an inspection of the initial and final regression equations to be discussed in the following section. In general terms, the final equations represent the weights of each variable which contributed significantly to a prediction of accident frequency. To use them as predictive tools in determining an individual's likelihood of accident involvement, one would merely insert values from his driver record, multiply each value by the regression weight and add to the constant term the product of all variables represented in the equation. The constant for a given equation always remains the same, irrespective of an individual's driving record. The driver record data to be entered into the equations should be within the range of values on which the equation was constructed. In other words, the equation might not be applicable to fourteen year old drivers since individuals this young were not represented in the sample.

The reader may note that some of the regression coefficients contain negative signs. A seemingly paradoxical situation may arise when the simple correlation (r)[a] between a variable and accident involvement is positive, but the variable is negative in the final regression equation, and thus subtracts from the predicted criterion score (accidents). When a variable is combined with others, complex relationships occur and a variable may actually "take away" or reduce the effect of another, and in some cases may even change the direction of the relationship. It is often difficult to explain such phenomenon and speculate as to what events are occurring in the "real" world to produce the various statistical relationships. There are so many possibilities that further research and knowledge in the area are often required for a definitive assessment.

The following discussion will be devoted to an analysis of some of the more important predictive variables and their relationship to accident frequency, as indicated by the appropriate beta and regression coefficients. The regression coefficient indicates the steepness (slope) and direction of the predictor-criterion relationship, but does not reflect the relative contribution of each variable in predicting accidents. Rather, the contribution of each variable is represented by the beta coefficient and F ratio.[b] The beta coefficient is a measure of the unique contribution of each variable and, when squared, indicates the <u>percentage</u> of accident variability that can be accounted for by a given independent variable. The F ratio, on the other hand, indicates the degree of statistical significance one can attach to the relationships.

<u>Concurrent Regression Equations</u>

An inspection of Tables I and II reveals that a one count conviction is by far the most statistically significant variable in predicting accidents for both males and females, as indicated by the F ratios. One count convictions, incidentally, are based on total incidents rather than the sum of individual violations. That is, each citation incident (court abstract) was counted once, regardless of how many violations were cited on it. Even though the predictive power of this variable is small in absolute magnitude (see beta coefficients in Figure 5 at the top of page 13), the one count conviction variable is highly significant in a statistical sense, and its contribution far exceeds that of all other variables in the equation.

The accident prediction profiles are quite similar between males and females in that one count and non-countable convictions are highly associated with accidents compared to specific types of convictions.

(a)The simple correlation matrices are depicted in Tables 6 and 7 for concurrent prediction and in Tables 8 and 9 for non-concurrent predictions.

(b)An F ratio is a statistical measure of the degree to which a variable is significant. For the sample size in this study, an F equal to or greater than 3.8 is significant at the .05 level of confidence.

Fig. 5 UNIQUE CONTRIBUTION OF SIGNIFICANT VARIABLES TO THE PREDICTION OF ACCIDENTS - CONCURRENT PREDICTION BY SEX

(Percentage = beta squared X 100)

Variables	Males	Females
	Equation IA	Equation IB
Total explained variability	5.143	3.462
One count convictions	2.693	1.388
Two count convictions	0.042	0.116
Passing convictions	0.004	--
Right-of-way convictions	0.012	0.029
Turning, stopping and signalling convictions	--	0.026
Speed convictions	--	0.046
Miscellaneous technical convictions	0.018	--
Equipment convictions	0.022	0.016
Non-countable convictions	0.607	0.112
Age	0.100	0.034
Height	--	0.039
Weight	0.010	--
Marital status	0.011	0.426
License restrictions	0.024	--
Traffic density	0.425	0.370
Moving FTA's	--	0.024
Joint contribution factor	1.175	0.836

There are sex differences, however, in the types of convictions which are related to accident involvement. For females Turning, stopping and signalling and Speed are statistically significant, although they contribute little beyond knowledge of one count convictions. Their respective percentage contributions are only .026 and .046. On the other hand, Right-of-way and Passing are statistically significant predictors for males, but with negligible contributions of .012 percent and .004 percent. Equipment convictions are significant for both males and females. The existence of these sex interactions supports the use of separate equations. Otherwise, relationships such as these would have been obscured.

The reader may be surprised to find that speed convictions did not make a statistically significant contribution to accident prediction for males, since this violation is the subject of much attention among safety authorities. This finding may seem particularly paradoxical in view of the fact that speed convictions are the predominant component of one count convictions, which in turn are the best single predictor of accidents. How can such a situation arise? Since speed accounts for a large proportion of all convictions, it is highly correlated with one count convictions ($r = .731$). Because of this one fact alone, the unique contribution of speed convictions would be reduced by the inclusion of one count convictions more than would be the other conviction categories. The fact still remains, however, that knowledge of a driver's "speed record" does not increase the ability to estimate his accident propensity, whereas knowledge of other aspects of his violation profile does increase one's ability to predict accident involvement. It should not be inferred from this that speeding is a completely innocuous violation. It does appear, however, that certain other recorded violation tendencies are more highly associated with accidents.

The significance of the non-countable conviction variable[a] further supports the hypothesis that violation of traffic laws in general is a key factor in accident involvement, irrespective of the nature of the violations. The two count conviction variable, generally considered to be the most dangerous, is also a highly significant predictor of accident involvement for each sex, but again the contribution is small compared to one count and non-countable convictions. The results of Part 8, incidentally, indicated that the assignment of two points to these types of violations adds little to accident prediction, and this finding seems to be substantiated by the present study. The reader interested in a detailed discussion of the limitation in assigning points to various conviction categories is referred to Part 8 of the Driver Record Study.

Another highly significant variable for both males and females is traffic density. This measure was based on the ratio of the number of vehicles registered to the total number of linear miles of roads and highways in the county in which the driver resided. Although density is certainly not a substitute for total miles driven, it is an important aspect of exposure. The direction of the relationship is what one might expect on analytical grounds, in that increased traffic density is associated with increased accident involvement.

Biographical characteristics such as age and marital status were significantly related to accident involvement for both males and females, although the percentage contribution indicated by the beta coefficients was very small. The young, single drivers tended to become more accident involved than the older, married drivers. The significance of age and marital status is, of course, a well established fact (1, 2, 5, 9). Surprisingly, the driver's weight turned out to be positively related to accident involvement for males, whereas height was inversely related to accident involvement for females. Why this should be true is difficult to arrive at even speculatively. The relationships are very slight, however, and should be interpreted with caution. Another example of sex interaction concerns the license restriction variable, which is a significant predictor for males but not for females. The fact that males with restricted licenses have more accidents than unrestricted males is at least partly due to the fact that persons with poor driving records are more likely to be restricted by the Department of Motor Vehicles in various manners. In other words, a poor accident record can precipitate or "cause" the restriction.

(a) The non-countable conviction variable is defined as the number of times a person was convicted of violations not countable in accordance with the California Vehicle Code. These are primarily equipment and miscellaneous technical violations.

The effect of each variable has been discussed in terms of its unique contribution to the prediction of accidents in a three year concurrent interval of time. To determine how these relationships hold up under conditions of "true" prediction, non-concurrent relationships will be discussed in the next section.

Non-Concurrent Regression Equations

In equations IIA and IIB, it is noted that various types of reported accidents are used as predictor variables. In non-concurrent relationships, it is possible to use the same measure (accidents) both as a predictor and criterion measure since they occur in two different periods of time. In this case, three independent accident categories were isolated. The classification of an accident into one of these categories in some cases may be artificial, in that rather secondary factors may determine its placement. The reader is referred to Exhibit A for a detailed definition of these types of accidents (FR and CHP and FR-CHP).

A non-concurrent analysis also provides an opportunity to assess the contribution of traffic convictions resulting from accidents. Such incidents have been termed "spurious" elements in previous parts of the Driver Record Study because their inclusion results in an inflated correlation between accidents and citations which occur in the same time interval.[a] In the non-concurrent case, however, the paired accident-conviction can be utilized as a predictor variable since it did not occur in the same time interval as the accidents which the equation is attempting to predict.

Another difference between the concurrent and non-concurrent analysis concerns the treatment of conviction data. Since the conviction data was not coded by type for each individual year, it was impossible to use types of convictions as predictor variables in the non-concurrent case. This limited the non-concurrent analysis to four conviction variables: one count convictions, two count convictions, non-countable convictions and convictions directly associated with an accident. The inability to categorize the convictions into specific violation types for the non-concurrent analyses was not considered to be an important limitation since type of violation data contributed very little to the efficiency of the concurrent prediction equations (see figure 5).

As regards biographical data (age, marital status, etc.), the concurrent and non-concurrent analyses were identical in the variables

(a) The term "spurious" as used here refers to citations which, while being fully authentic, were the result of an accident investigation, instead of normal traffic observations. In other words, they are "accident caused."

employed as predictors.

Fig. 6 UNIQUE CONTRIBUTION OF SIGNIFICANT VARIABLES TO THE PREDICTION OF ACCIDENTS - NON CONCURRENT PREDICTION BY SEX

(Percentage = beta squared X 100)

Variables	Males	Females
	Equation IIA	Equation IIB
Total explained variability	1.411	0.923
FR accidents	0.122	0.130
CHP accidents	0.014	--
FR-CHP accidents	0.026	--
One count convictions	0.710	0.379
Non-countable convictions	0.139	--
Age	0.070	--
Height	0.012	--
Marital status	--	0.162
Traffic density	0.057	0.126
Joint contribution factor	0.261	0.126

The results of the non-concurrent regression analysis are depicted in Chart I and Tables 3 and 4, while the percentage contributions of the significant predictors are shown in Figure 6. Consistent with the concurrent regression results, total one count convictions proved to be by far the most significant and powerful accident predictor for males and females. As can be seen from Figure 6, the percentage contribution of all other variables is small relative to total one count convictions. It is interesting to note that the two count conviction variable does not appear as a significant predictor for either sex, which is in direct contrast to the concurrent analyses. It may also seem surprising to many that the number of spurious elements per se (paired accident conviction) failed to make a significant contribution to accident prediction for either sex. Thus, the number of times a person was convicted of a traffic violation in connection with an accident is not uniquely associated with his future accident frequency, and does not increase one's ability to predict future accident involvement. The non-countable conviction variable, while statistically significant in the concurrent case for both males and females, is only significant for males in the non-concurrent case.

As was the case with the concurrent analyses, these findings indicate that a tendency to violate traffic laws in general is the best prognosticator of future accident involvement. The apparent superiority of the one-count conviction variable is perhaps due primarily to its higher frequency of occurrence which would tend to make it a more reliable and stable index of a driver's violation propensity. Again, one is tempted to speculate that a summation of all convictions into a total conviction variable might be as efficient as the more molecular breakdown used in this study. Unfortunately, the form of the data did not allow for a test of this hypothesis, as it did in Part 8.

Turning now to the independent accident variables, it can be seen from Figure 6 that their contribution to future accident prediction, though statistically significant, is extremely small. Of the three accident types, FR accidents are the most efficient predictors of future accidents for both sexes. The most likely explanation for this is that property damage accidents (FR) occur much more often than the other types and are therefore a more stable measure of a driver's overall accident propensity. The superiority of conviction data over accidents as predictors of future accidents is consistent with previous studies by these authors and others (3, 4, 11).

From Chart 1 and Figure 6, it can be seen that several biographical variables are statistically significant predictors but that their contribution is relatively minor. Perhaps most notable is the influence of sex upon the relationships. For females, only marital status is significant, with single females tending to have more accidents. For males, age and height are significant factors, in addition to marital status. The direction of these relationships indicates that older, taller and married males have fewer accidents than their respective counterparts. The findings with respect to physical stature, incidentally, are in contrast to the concurrent analyses, where only weight was significant for males and height significant for females. An interpretation of these reversals is difficult and would require considerable speculation to produce even tentative hypotheses.

The last variable, traffic density, is a significant predictor for both sexes, with increased traffic density associated with increased accident involvement. The true contribution of traffic density is, of course, underestimated by the data since traffic density is merely based on the average density within each county. Thus, drivers residing in the same county were assigned the same value, regardless of where and in what kinds of traffic they actually drove.

In concluding this section on non-concurrent accident prediction, it must be frankly admitted that the results are disappointing, in that only a small proportion of accident involvement could be accounted for by the predictor variables represented in the equations. For reasons which will be developed in the next section, the prospect of predicting accident involvement is not quite as hopeless as it may appear at this point in the analysis.

DISCUSSION

It may be recalled that three questions were posed in the Introduction. These questions will now form the basis for further discussion of the study findings.

1. Does knowledge of a driver's age, marital status, physical stature, etc., increase our ability to predict his accident involvement?

   As was the case with Part 8, one count convictions proved to be by far the most powerful predictor of accident involvement for both sexes. The validity coefficients associated with the concurrent regression equations (all variables) were .226 and .186 for males and females respectively. Compared with the Part 8 validity coefficients of .219 and .160, those of Part 9 are significantly higher in a statistical sense, though, in absolute terms, the increase is slight. Percentage-wise, the gain in efficiency is less than one percent for both males and females. The major source of the gain resides in the age and marital status and traffic density variables, with other biographical factors playing a lesser role. Thus, there is some basis for weighting age and marital status when assigning negligent operator points or, say, establishing auto insurance premiums. Whether the small gain is of sufficient magnitude to warrant this added complexity, is, of course, a decision for the program administrator.

2. How do predictions made from non-concurrent events differ from those made from events occurring during the same time interval (concurrent predictions)?

   It was readily apparent from the results that predictions made on a concurrent basis were more accurate than those made on a non-concurrent basis. A large part of the superiority of the concurrent predictions, however, was due to the longer criterion interval (three years vs. one year) employed on the concurrent data. A longer criterion interval results in a more reliable measure of accident involvement, which in turn renders them more predictable. Even had the criterion intervals been of the same length, the concurrent equation would have been more efficient, since it is able to capitalize on any circumstances and idiosyncracies which may be common to events occurring in the same time interval. With additional variables and a longer criterion interval, the efficiency of the non-concurrent predictions would undoubtedly improve over the disappointing results achieved in the present study. Because of this, the reader should refrain from concluding that accident prediction is as hopeless as the low correlations may have indicated.

3. To what extent does data not ordinarily available from the official driver record file (e.g. annual mileage, occupation, etc.) increase one's ability to predict accident occurrence?

Since the number of variables available from driver record files alone is rather limited, the question arises as to what influence additional variables might have on the predictability of accidents. In order to reflect on this question, a small pilot survey was undertaken in which a random sub-sample of 2,000 drivers was extracted from the total sample and contacted for additional information (see Exhibit C). Of those contacted, useable responses were received for 536 males and 363 females. Because the respondent female sample was involved in only 12 accidents, it was considered too unreliable for further analysis. Even the male data is based on only 45 accidents and is therefore subject to rather large sampling variation. In addition, this questionnaire sample cannot be generalized to all male drivers, since those responding to a questionnaire may differ in certain respects from those who chose not to respond. In the present study, no significant difference was found between the groups on accidents or convictions, although the conviction difference approached the .10 level of confidence.[a] Any bias would generally suppress relationships, which would tend to make the multiple R obtained from such a sample an underestimate of the value which could have been obtained had all subjects responded to the questionnaire.

The questionnaire responses for males were subjected to a regression analysis similar to that performed on the total sample. The initial and final equations are presented in Table 5. These results clearly indicate that variables other than driver record contribute to accident prediction. Presented in Figure 7 are the contributions of the significant predictors classified by source -- driver record vs. questionnaire.

(a) Convictions: $t = 1.52$, $P > .10 < .15$.
Accidents: $x^2 = .06$, 1 d.f. $P > .80$.

Fig. 7 THE RELATIVE CONTRIBUTION OF DRIVER RECORD AND QUESTIONNAIRE DATA IN PREDICTING ACCIDENTS

(1 year data)

Males (N = 536)	
Predictor variable	Percentage contribution (multiple R^2 X 100)
Total - Driver Record and Questionnaire Contribution combined	12.91
<u>DRIVER RECORD DATA</u>	
1 count convictions, 2 count convictions, non-countable convictions, warning letters, weight, marital status, age	6.41
<u>QUESTIONNAIRE DATA</u>	
Highest grade completed, years driven, make of car, year of car, miles per week, miles 1963, social reputation	6.50

It is readily apparent from Figure 7 that the questionnaire variables made a strong contribution to predicting accidents. In fact, they accounted for over 50 percent of the explained accident variability. These findings, of course, cannot be directly compared with those of the overall analysis, since they are based on a one year concurrent accident frequency instead of a three-year frequency. Despite this and the previously mentioned limitations, it is clear that a number of <u>non-driver</u> record variables <u>are</u> uniquely associated with accident frequency. Furthermore, these additional variables can be measured through conventional questionnaire techniques with a degree of precision sufficient for a substantial increase in accident predictability. Although the 50 percent increase in predictive efficiency is at best only suggestive of the magnitude one might expect in a larger study, it provides a much more optimistic outlook for predicting accidents than was indicated in the Results section.

Now that the study findings have been discussed in relation to the three introductory questions, a more general discussion of the results is in order. This remaining discussion will center largely around the themes of accident predictability and accident proneness. Why are accidents so unpredictable and what implications do the study findings have for a contemporary concept of accident proneness? In reflecting on these two related questions, the reader should consider the following:

1. An important consideration in any prediction study is to determine the specific nature of the phenomenon represented by the equation and the measurements associated with it. One can immediately see the marked limitations in obtaining a measure of total accident and violation behavior. A large percentage of accidents -- particularly the minor ones -- are probably never reported and even a smaller percentage of traffic violations

are detected or cited. Consequently, the predictions in this study are restricted to reported accidents and detected violations in which the driver was cited and convicted. Though reported accidents and traffic convictions are far from a perfect reflection of underlying driver behavior, they are the measure which is most meaningful to the licensing administrator, since his decisions must be based on officially recorded incidents. The administrator is faced with making decisions and developing programs based on information that is available from driver records. However, because the recorded data is far removed from the population of events that are occurring in the actual driving environment, the regression weights are rather limited indicators of the relative importance of the various factors to driving performance. The forecasting efficiency and generality of the equations would increase with a knowledge of the "near accident," accident severity, all violations, accident culpability, and exposure, to name a few. Though collection of all the preceding data is impossible from a practical standpoint, the purpose of this discussion is to clarify the reason for the low forecasting efficiency based on the driver record data alone.

2. Empirical evidence and logic both suggest that accidents are greatly influenced by factors outside the control of the subjects comprising a given driving population. In fact, the very term "accidents" connotes randomness or chance. By "chance," the authors do not necessarily mean that an accident can be a capricious, unlawful event, for one can take almost any accident and conceive of a way in which a driver could have prevented it. Even those which somehow seem completely independent of individual control can often be attributed to some deficiency in the driving system -- or at least a failure of that system to correspond to the "ideal" system. Obviously, such a concept of chance is too broad to be meaningful, and could ultimately lead one to the sterile conclusion that traffic accidents are caused by people who drive cars. The only meaningful conception of chance is achieved by delimiting the term to that which is measureable and exhibits variability. With this restriction, any accident variation which is not associated with persistent, measureable differences in individuals and environments is defined as chance. Among possible accident-related, person-centered traits are variations in basic personality structure, values, attitudes, psychophysical skills, reflexes and perceptual-motor integration. Such variation would include both habitual ways of responding as

well as reactions to atypical situations and stress. Accident-related variations in exposure would consist of such variables as annual mileage, traffic density, weather conditions and types of roads, to name only a few. Variations in vehicles could also be conceived as a type of exposure variable, since they affect one's accident probability and are an aspect of environment.

The concept of chance is an important one, since it places limitations on the predictability and control of phenomena. To the extent that accident involvement is influenced by random environmental and random person-centered factors (chance), it is not predictable even from a perfectly reliable set of predictor variables. Furthermore, since all measurement contains error, the amount of predictability is further reduced by the degree of unreliability inherent in measuring the predictors. When one considers the error involved in measuring phenomena through questionnaires and driving records, combined w th the large amount of random variation inherent in accident involvement it should not be at all surprising that a high degree of accident prec ction is unattainable. Because of this, it is perhaps more meaningfu o evaluate the merits of the accident prediction equation by comparing s validity (correlation) coefficient with the estimated correlational ling imposed upon prediction by the nature of the accident phenome . Toward this end, Newbold and Cobb have constructed an accident ability model which generates an estimated correlational ceiling betwe a given accident distribution and a perfectly reliable prediction eq ion. Shown below in Figure 8 are the amounts of accident predictic associated with the various predictor equations expressed in relation the theoretical Newbold-Cobb ceilings.

Fig. 8 THE RELATIONSHIP OF ACHIEVED ACCIDENT ICTION TO A POSTULATED ACCIDENT PREDICTION LING

Prediction equations	A Actual predicted (percentage)	B Ne d-Cobb th etical c ng of p ction (p ntage)	A/B (percentage)
Concurrent three year record			
Males	5.1	7	37.2
Females	3.5	2	42.7
Non-concurrent one year record			
Males	1.4	8	15.9
Females	0.9	9	23.1
Questionnaire sub-sample (one year concurrent record)			
Males	12.9	8	146.6

These prediction ratios indicate that a c siderable amount of potentially predictable accident variation remains unpredicted by the various equations, particularly in the non-concurrent case. The last equation appears to be a rather anomalous exception, in that the theoretical ceiling has been exceeded. This, however, is due to the inflated nature of the obtained multiple R rather than to any strange and unique talents residing in the present authors' approach.[a] If this questionnaire-sample equation could have been cross-validated against an independent sample of drivers, the obtained correlation coefficient would probably have shrunk below its theoretical ceiling.

These results are encouraging and suggest that prediction ratios approaching one are attainable, particularly for concurrent prediction. They also provide evidence for at least a statistical concept of accident proneness, in that some people are more likely to be involved in accidents than others, even with exposure (as measured by annual mileage and traffic density) controlled. If this were not the case, all non-exposure variables should have dropped from the equation as non-significant contributors. In all cases, however, a number of non-exposure variables were significantly associated with accident involvement, particularly those most indicative of how a driver drives -- the violation/conviction variables. Furthermore, the significance of several non-exposure variables extended to the non-concurrent case.

Admittedly, the absolute magnitude of the prediction achieved herein is small and indicates that the accident contribution of the accident-liable driver is a small one. However, the existence of even small predictable relationships justifies continued research and the development of optimally effective programs geared toward the identification, rehabilitation and control of all drivers who represent statistically greater safety risks to themselves and the public. The California Department of Motor Vehicles will therefore continue its research effort in the areas of basic driver record relationships, and negligent driver control. Hopefully, this effort will result in important advancements to the science of driving behavior and its application to the licensing administrators' decision-making functions.

(a) Unless huge sample sizes are used, the sample correlation coefficient is highly unstable when computed on extremely skewed data. This could give rise, by chance, to a coefficient which greatly exceeded the theoretical ceiling of the parameter.

BIBLIOGRAPHY

1. Campbell, B. J. "Drivers Age and Sex Related to Accident Time and Type," Traffic Safety Research Review, Vol. 10, 2, 1966.

2. Department of Motor Vehicles, State of California, The 1964 California Driver Record Study, Part 5, Sacramento, California: The Department, 1965.

3. Department of Motor Vehicles, State of California (Coppin, McBride, and Peck), The 1964 California Driver Record Study, Part 6, Sacramento, California: The Department, 1965.

4. Department of Motor Vehicles, State of California (Coppin and Peck). The 1964 California Driver Record Study, Part 8, Sacramento, California: The Department, 1966.

5. Department of Motor Vehicles, State of New York, "Fatal, Non-fatal, and Property Damage Accidents by Age and Sex, Hour of Day and Day of Week," Statistical Bulletin #6, New York: 1964.

6. Ezekiel, M. and Fox, K. Methods of Correlation and Regression Analysis, New York: Jogh Wiley and Sons Inc., 1959.

7. Forbes, T. W. "The Normal Automobile Driver as a Traffic Problem," Journal General Psychology, 1939, 20, 471-474.

8. Haddon, William, Jr. et al Accident Research: Methods and Approaches, New York: Harper and Row Publishers, 1964.

9. Lauer, A. R. "Age and Sex in Relation to Accidents," Highway Research Bulletin 60, Road-User Characteristics, NAS-NRC, Publication 244, Washington, D. C., 1952.

10. Levonian, Edward Prediction of Negligent Operators, Institute of Transportation and Traffic Engineering, UCLA: The Institute, 1963.

11. Little, Arthur D., Inc. The State of The Art of Traffic Safety, Cambridge: Arthur D. Little, Inc., 1966.

12. McNemar, Q. Psychological Statistics, New York, New York: John Wiley and Sons, Inc., 1962.

13. Schuster, D. H. Prediction of Follow-up Driving Accidents and Violations. University of Iowa: mimeo. unpublished manuscript; August 1965.

14. Schuster, D. H. and Guilford, J. P. "The Psychometric Prediction of Problem Drivers," Human Factors, 6(4). Pp 393-421, August 1964.

15. Thorndike, R. S. The Human Factor in Accidents with Special Reference to Aircraft Accidents, Project No. 21-30-001, Report #1, USAF, School of Aviation Medicine, Randolph Field, Texas, 1951.

16. Uhlaner, J. E. and Drucker, A. J. "Selection Tests - Dubious Aid in Driver Licensing," Highway Research Record, 1965, No. 84, Pp 611-53.

Exhibit A.-DRIVER RECORD DATA DEFINED

1. ACCIDENTS - Number of accidents reported to the department during the year in question.

2. CALIFORNIA HIGHWAY PATROL ACCIDENTS (CHP) - Number of above total which were reported by/or to the California Highway Patrol. This includes all fatal and injury accidents and CHP investigated property damage accidents and property damage accidents reported to the CHP by local authorities.

3. FINANCIAL RESPONSIBILITY ACCIDENTS (FR) - Number of total accidents which were reported in accordance with California Financial Responsibility law but which were not reported by or through the CHP.

4. FR-CHP ACCIDENTS - Number of accidents which were reported both by CHP and by the driver reporting in accordance with California's Financial Responsibility Law.

5. ONE COUNT CONVICTIONS - Number of convictions which contribute one point toward an individual's negligent operator point total. This includes all violations involving the safe operation of a motor vehicle as defined in Section 12810 of the California Vehicle Code, with the exception of certain designated "two-point" violations.

6. TWO COUNT CONVICTIONS - Number of convictions which count double in the negligent operator point system: drunk driving, hit and run, reckless driving, and driving with a suspended or revoked license.

7. MOVING FAIL TO APPEAR (FTA) - Refers to a citation for a potentially countable traffic violation in which the driver failed to keep his signed promise to appear in court.

8. NON-MOVING FTA'S - Number of uncleared "Failure to Appear" stops which are for non-moving violations (i.e. for non-countable violations).

9. NON-COUNTABLE CONVICTIONS - Number of convictions which do not contribute toward a driver's negligent operator count as defined in Section 12810 of the California Vehicle Code.

10. ACCIDENT CONVICTION (SPURIOUS CONVICTION) - The number of countable convictions which were issued upon completion of an accident investigation.

11. WARNINGS - Number of warning letters issued.

12. LICENSE RESTRICTIONS - Number of restrictions noted on license, e.g. must wear glasses, daytime driving only, etc.

13. TRAFFIC DENSITY - Total number of vehicles registered ÷ total number of linear miles of roads and highways for each county.

14. AGE - Months (nearest birthday).

15. HEIGHT - Inches

16. WEIGHT - Pounds

17. RATIO - Height ÷ weight

18. MARITAL STATUS - Coded at time of latest renewal (if married, marital status was coded 1, if single coded 2).

19. TYPES OF CONVICTIONS - See following pages.

Exhibit A (cont.)

Sections Falling Under Various Violation Categories
California Motor Vehicle Code (1963 Edition)

Speeding

22349	Maximum speed limit, 65 miles per hour.
22350	Unsafe speed for prevailing conditions.
22356	Maximum speed 70 mph, when sign posted.
22400a	Minimum speed, impeding normal flow of traffic.
22400b	Minimum speed, below signposted limit.
22405a	Unsafe speed (signposted for condition of bridge, structure, tube or tunnel).
22406	Truck or tractor exceeding 50 mph.
22407	Truck speed on downgrade, exceeding posted limit.
22408	Passenger car or bus towing any vehicle, exceeding 50 mph.
22409	Solid tire vehicle, speed restricted by weight.
22410	Metal tire, vehicle exceeding 6 mph.
22412	School bus, exceeding 45 mph with passengers.
22414	Labor bus or truck, exceeding 45 mph with passengers.
23109a	Speed contest, engage in, aid, or abet.
23109b	Speed contest, blocking or obstructing highway.

Traffic Signs, Signals, Markings

21451a	Green or Go, shall proceed but shall yield to vehicles lawfully within intersection. No U-turn unless permitted by sign.
21453a	Red or Stop, vehicles stop at limit line or X-walk.
21453b	After stopping, may turn right (unless sign posted) but shall yield to vehicles lawfully within intersection.
21453c	After stopping, may turn left (unless sign posted) from one-way to one-way street, but shall yield to vehicles on cross street.
21454	Green Arrow, make only restricted movement indicated, yield to vehicles lawfully within intersection. No U-turn unless permitted by sign.
21457a	Flashing Red, failing to stop for.
21457b	Flashing Yellow, proceed only with caution.
21460a	Double solid lines, driving to left of, except driveway, intersection, or U-turn.
21460b	Solid-broken lines, driving to left when solid line placed on right.
21461	Traffic control sign, failure to obey regulatory provisions.
21462	Traffic control signals, all traffic to obey.
22450a	Stop sign, failure to stop at limit line or crosswalk.
22450b	Stop sign, failure to stop where indicated within intersection.
22450c	Stop sign, failure to stop at posted RR crossing.
22451a	Railroad crossing, failure to stop for signal device.
22451b	Railroad crossing, failure to stop for human flagman.
22452b	Railroad crossing, certain vehicles must stop.
22454	Passing school bus, stop when red lights flashing.

Turning, Stopping and Signalling[(a)]

22100a	Right turn at intersection, improper position.
22100b	Left turn at intersection, improper position.
22101b	Required turn, failure to obey official sign.
22101c	Prohibited turn, failure to obey official sign.
22102	U-turn in business district, other than at intersection, or opening in divided highway.
22103	U-turn in residence district, vehicle approaching within 200 feet.
22104	U-turn at fire station, in front of or using entrance.
22105	U-turn at curve or grade, vision obstructed within 200 feet.

(a) The law does not require that Sections 22500, 22502, 22503, 22508 and 22514 be reported to the Department of Motor Vehicles. If reported, Section 22500h was the only parking violation which was counted in determining an individual's negligent operator count at the time of this study. Generally speaking, ordinary parking violations are not reported to the Department of Motor Vehicles.

21711	Towed vehicle, whipping, swerving, or failing to track properly.
21712	Unlawful riding on portion not intended for passengers or load.
21715	Passenger vehicle, towing more than one other vehicle.
21750	Overtaking vehicle, failure to pass safely to left.
21751	Overtaking vehicle, passing without sufficient clearance.
21752a	Driving left of center, when view limited by curve or hill crest.
21752b	Driving left of center, when view limited by approaching bridge, viaduct or tunnel.
21752c	Driving left of center, traversing intersection or RR crossing.
21753	Overtaken vehicle, not moving to right on audible signal, or increasing speed.
21754	Passing on right when unlawful.
21755	Passing on right, when unsafe, or on shoulder.
21756a	Passing streetcar when receiving or discharging passengers.
21756b	Passing streetcar at unsafe speed.
21756c	Passing trolley coach at unsafe speed.
21757	Passing streetcar on left.
21758	Passing too slowly on grade (10 mph faster, complete pass 1/4 mile).
21759	Passing animals, stop or reduce speed as necessary.

Right-of-Way

21800a	Uncontrolled intersection, yield to first vehicle within.
21800b	Uncontrolled intersection, yield to vehicle on right.
21801a	Left turns, yield until reasonably safe.
21801b	Failure to yield, turning vehicle having yielded (lane by lane).
21802a	Entering through highway, yield until reasonably safe.
21802b	Failure to yield, by vehicle presenting a hazard.
21802c	Proceeding from stop sign or flashing red (within intersection), yield until reasonably safe.
21802d	Failure to yield, by vehicle not a hazard.
21803a	Yield signs, yield until reasonably safe.
21803b	Failure to yield, by vehicle not a hazard.
21804a	Private property, drive or alley, yield to approaching vehicles.
21804b	Left turn into private property, drive or alley, yield until reasonably safe.
21804c	Failure to yield, turning vehicle having yielded.
21805b	Equestrian crossings, failure to yield by driver.
21806a	Emergency vehicles, other driver failing to yield.
21950	Crosswalks, failure to yield to pedestrians within.
21951	Crosswalk, overtaking and passing vehicle stopped for pedestrian within.
21952	Sidewalk, failure to yield to pedestrian on.

Major

14601a	Driving privilege suspended or revoked, driving when.
20001	Hit-run, injury or death, immediate report of fatal.
20002a	Hit-run property damage.
20002b	Hit-run property damage, by runaway vehicle.
20007a	Hit-run, unattended vehicle damaged.
23101, P.C. 367e	Intoxicated driver, causing injury to other than self.
23102a, P.C. 367d	Under influence of alcohol (or combined with drug), driving on highway.
23103	Reckless driving, no injury.
23104	Reckless driving, causing injury.
23105	Narcotics, driving under influence, or by addict.
23106	Other drugs, driving under influence.
23108	Dangerous drugs, driving under influence causing injury.
P.C. 192.3	Manslaughter.

Exhibit A (cont.)

22106	Starting or backing when unsafe.
22107	Unsafe turn, and/or without signalling.
22108	Turning without signalling last 100 feet.
22109	Stopping suddenly without signalling.
22111	Hand signals, improperly given.
22112	School bus signals, misuse by bus driver.
22500a	Parking unlawfully, within intersection.
22500b	Parking unlawfully, on crosswalk.
22500c	Parking unlawfully, adjacent to safety zone.
22500d	Parking unlawfully, within 15 feet of fire station driveway.
22500e	Parking unlawfully, blocking any driveway.
22500f	Parking unlawfully, on a sidewalk.
22500g	Parking unlawfully, blocking excavation.
22500h	Parking unlawfully, double parking.
22500i	Parking unlawfully, in posted bus loading zone.
22500j	Parking unlawfully, in tube or tunnel.
22500k	Parking upon any bridge, unless posted to permit.
22502	Park parallel on right, and/or within 18" if curbed.
22504a	Stopping or parking, on roadway outside city limits.
22505	Parking on state highway where sign posted.
22510	Parking in snow areas, when sign posted.
22514	Fire hydrant, parking unattended vehicle within 15 feet.
22515	Unattended vehicle, motor running and/or brakes not set.
22517	Vehicle doors, opening to traffic when unsafe, leaving open.
22520	Stopping or parking, on freeway having full control of access and no crossings at grade.

Equipment

24002-27907	The vehicle code sections under this category are too numerous to list individually. In general, they relate to inadequacies of lighting and braking equipment, windshields and mirrors, smog devices and exhaust systems. Loading regulations are also subsumed under this category.

Driving, Overtaking and Passing

21650	Right half of roadway, failure to drive on.
21651	Divided highways, driving to left, over, or across dividing section.
21652	Service road, entering or leaving adjacent highway from other than lawful opening.
21653	One-way street, driving against traffic.
21654	Slower vehicle in left lane(s).
21655b	Slow vehicles (22406-22414) using left lane(s), or passing in lane other than adjacent to right lane.
21656	Slow vehicle, failure to use signposted turnout.
21657	Off-center lanes, failure to obey signs designating.
21658a	Laned roadways (2 or more lanes in direction of travel), straddling or changing when unsafe.
21658b	Failure to obey directions of a traffic device on a divided roadway.
21659	3-lane highway, driving in far left lane, or using center lane when unsafe.
21660	Meeting vehicles, failure to pass to right, and/or yield half of roadway.
21661	Descending narrow grade, yield to ascending vehicle.
21662	Mountain driving, keep to right, sound horn when required.
21700	Obstructing driver's view, or control, by passengers or load.
21702a	Driving hours--Persons, not to exceed 10 hours.
21702b	Driving hours--Property, not to exceed 12 hours.
21703	Following too closely, not reasonable and prudent.
21704a	Distance between trucks, 500 feet on 2-lane highway.
21705	Caravan, maintain at least 100 feet distance between vehicles.
21706	Fire department or police vehicles, following within 300 feet.
21707	Fire area, operating vehicle within the block or 300 feet.
21708	Fire hoses, driving over unprotected.
21709	Safety zone, driving through.
21710	Coasting, in neutral or downgrade.

Error

An error occurred while processing this page. See the system log for more details.

Exhibit D

NEGLIGENT OPERATOR REVIEW POINT COUNTS
(Section 12810)

Point	Section	Violation
2	192 pc	manslaughter
	367d pc	drunk driving
	367e pc	drunk driving inj.
1	2800 (no ped. or equip.)	Officer's signal
	2801 (no ped. or equip.)	Fireman's signal
	13360	violation of restr.
	14603	violation of restr.
	16457	violation of restr.
2	14601	driv-susp., rev.
1	14610	misuse of license
2	20001	H&R-death, inj.
	20002	H&R-prop. damage
	20007	H&R-unattend. veh.
1	21451a	yield-green
	21452a	yield-yellow
	21453a	limit lane-stop
	b	rt. turn on red
	c	1-way turn on red
	21454a (no ped.)	yield-green arrow
	21457a	flashing red-stop
	b	flashing yellow
	21459-21460	over double line
	21461-21462 (no ped.)	obey control device
	21650	wrong side of road
	21651	wrong side-div. hwy.
	21652	wrong entrance-hwy.
	21653	wrong side-1 way
	21654	wrong lane(notpass.)
	21655	wrong lane-tr.,tr.
	21656	too slow; pull over
	21657	off-center lane
	21658a	straddling-marked
	b	slow traffic lane
	21659	yield-middle lane
	21660	straddling-unmarked
	21661	yield to ascending
	21662	wrong side-mt. driv.
	21700	obstructed view
	21702	special hours
	21703	following too close
	21704	too close-tr. trail.
	21705	too close-caravan
	21706	too close-fire veh.
	21707	inter.-fire area
	21708	driv. over firehose
	21709	thru safety zone
	21710	coasting on grade
	21711(acc)	towed veh. swerving
	21712	unlawful riding

NOTE: Do not count owner's responsibility: 40000-40003

Point	Section	Violation
1	21750	cutting in
	21751	passing clear-100 ft.
	21752	pass. on grad-rr
	21753	yield if overtaken
	21754	improper pass.
	21755	improper pass. on rt.
	21756	wrong pass. str. car
	21757	wrong pass. str. car
	21758	slow pass. on grade
	21759	wrong pass. animals
	21800a	yield-at intersect.
	21800b	yield-to car at right
	21801	yield-when turning
	21802	yield-thru highway
	21803	yield-yield sign
	21804	yield-alley, driveway
	21805b	yield-horseback rider
	21806a	yield-emergency veh.
	21950	yield-to ped.(crswlk)
	21951	yield-otr. stopped veh
	21952	yield-to ped.(sdwlk)
	21954b	care if ped. yields
	22100a	rt. turn-rt. lane
	b	left turn-left lane
	22101	turns-obey markers
	22102	U turns: business
	22103	U turns: residence
	22104	U turns: fire stat.
	22105	U turns: curve, grade
	22106	starting or backing
	22107	unsafe turn-signal
	22108	give signal, 100ft.
	22109	stop signal
	22110	signal device req.
	22111	hand signals
	22349-22363	speed laws
	22400	too slow
	22405	bridges, etc.
	22406	speed-trucks, trl.
	22407	trucks, trl-desc.grd.
	22408	speed-towing
	22411	speed-lift carriers
	22412	speed-school bus
	22450	stop sign
	22451	stop-train signal
	22452-22453	stop-rr crossing
	22454	stop-pass. school bus
	22455	stop-rr:trucks, bus
	22500(h)	double parking
	22517	opening drs. on traf.
2	23101	fel. dr. driv.-inj.
	23102	drunk driving
	23103	reck. dr.-prop. dam.
	23104	reck. dr.-inj.
	23105	narcotics
	23106	other than narcotics
	23108	dangerous drugs
1	23109	racing
	23253	vehicular crossing
	24008	lowered veh.
	24409	lights
	26300	brakes
	26457	stopping sp.-loads

Appendix A.-DERIVED QUESTIONNAIRE SAMPLE DATA

The following variables are based on the data collected from the questionnaire but the scales of measurement are not dependent upon a direct rating by the respondent.

1. Occupation - Several studies have established that individuals in certain occupational categories have a higher accident and violation rate than individuals in other occupational categories. This could be due to the characteristics of individuals selecting certain occupations and/or the nature of the occupation itself. In an attempt to measure these underlying factors, several different techniques were used to dimenison the characteristics that seemed to discriminate among individuals in different occupations. Each occupation was rated on the following scales (a-d).

    a) Power Scale: The degree to which an individual or group in an occupational class wields economic and political power.

    b) Social Reputation: The likelihood of a person in this occupation being socially accepted and reputable.

    c) Intelligence: The likelihood of a person in this occupation having a high intelligence.

    d) Social Contact: The likelihood of a person in this occupation enjoying a large amount of social contact.

2. Date Received: The number of days which elapsed between the time the first questionnaire was sent and date which the completed questionnaire was received was considered a variable. This data was collected to determine whether individuals who returned their questionnaires more rapidly tended to have better accident records.

Table 1.-RESULTS FOR THREE YEAR CONCURRENT REGRESSION EQUATIONS
(Male subjects; N = 42,228)

Variables	Regression coefficient	Mean square	F(a)
INITIAL EQUATION			
Moving FTA's	.041252	0.81458	3.267
Non-moving FTA's	-.0041427	0.084365	0.338
One count convictions	.072795	29.768	119.402
Two count convictions	.097303	2.2379	8.977
Signal/sign convictions	-.010414	0.48237	1.935
Passing convictions	.023445	1.6172	6.487
Right-of-way convictions	.014078	0.45229	1.814
Turning, stopping and signalling convictions	-.0075759	0.17002	0.682
Speed convictions	-.0092719	0.47310	1.898
Major convictions	-.030097	0.30841	1.237
Equipment convictions	-.012177	1.6781	6.731
Miscellaneous technical convictions	-.0092656	1.3400	5.375
Non-countable convictions	.063422	24.932	100.004
Age	-.000094673	9.3710	37.588
Height	-.0020288	0.36900	1.480
Weight	.0070344	0.38934	1.562
Height/weight ratio	-.0030552	0.15296	0.614
Marital status	.013092	1.1102	4.453
License restrictions	.018688	2.5069	10.056
Traffic density	.00036688	43.934	176.226
Constant term = .21637	-	0.87833	3.523
FINAL EQUATION(b) - 1A			
One count convictions	.063523	197.44	791.969
Two count convictions	.064175	4.5472	18.240
Passing convictions	.030385	4.0069	16.072
Right-of-way convictions	.021079	1.2618	5.061
Equipment convictions	-.010586	1.3096	5.253
Miscellaneous technical convictions	-.0076815	1.0023	4.020
Non-countable convictions	.061309	24.210	97.112
Age	-.000091244	9.2085	36.937
Weight	.0021353	1.0523	4.221
Marital status	.013078	1.1082	4.445
License restrictions	.018647	2.4967	10.015
Traffic density	.00036730	44.357	177.923
Constant term = .081813	-	3.6753	14.742

(a)F ratio of 3.84 required for significance at .05 level.

(b)All variables significant beyond .05 level.

Table 2.-RESULTS FOR THREE YEAR CONCURRENT REGRESSION EQUATIONS
(Female subjects; N = 30,277)

Variables	Regression coefficient	Mean square	F(a)
INITIAL EQUATION			
Moving FTA's	.13642	0.94540	7.824
Non-moving FTA's	-.0022844	0.012119	0.100
One count convictions	.050309	2.6185	21.671
Two count convictions	.22153	3.0296	25.073
Signal/sign convictions	.0091348	0.083770	0.693
Passing convictions	.012745	0.10771	0.891
Right-of-way convictions	.045750	0.92780	7.678
Turning, stopping and signalling convictions	-.018066	0.21522	1.781
Speed convictions	.025948	0.68608	5.678
Major convictions	-.021263	0.085467	0.707
Equipment convictions	-.022941	0.44255	3.663
Miscellaneous technical convictions	-.015638	0.34081	2.821
Non-countable convictions	.077078	2.7211	22.520
Age	-.000045144	1.3636	11.285
Height	-.0033250	1.1877	9.829
Weight	.0022631	0.048576	0.402
Height/weight ratio	-.00055606	0.0068807	0.057
Marital status	.057942	15.392	127.386
License restrictions	.0011303	0.0065193	0.054
Traffic density	.00026823	14.561	120.505
Constant term = .19344	-	0.98179	8.125
FINAL EQUATION - 1B			
Moving FTA's	.12640	0.88251	7.304
One count convictions	.058062	19.647	162.599
Two count convictions	.19416	4.4239	36.613
Right-of-way convictions	.037085	0.91533	7.575
Turning, stopping and signalling convictions	-.025765	0.74240	6.144
Speed convictions	.018246	0.83541	6.914
Equipment convictions	-.020851	0.39094	3.236(b)
Non-countable convictions	.063668	2.6738	22.129
Age	-.000038368	1.1399	9.434
Height	-.0027467	1.4486	11.989
Marital status	.057824	15.467	128.008
Traffic density	.00026660	14.438	119.493
Constant term = .17275	-	1.3532	11.199

(a) F ratio of 3.84 required for significance at .05 level.

(b) F = about .10 (left in equation because of a significant suppressor effect on the equipment conviction variable.)

Table 3.-RESULTS FOR ONE YEAR NON-CONCURRENT REGRESSION EQUATIONS
(Male subjects; N = 43,309)

Variables	Regression coefficient	Mean square	F[a]
INITIAL EQUATION			
Warning letters....................	.0033303	0.0043895	0.052
Moving FTA's.......................	.020881	0.11385	1.347
Non-moving FTA's...................	-.0036430	0.0020129	0.024
FR accidents.......................	.032658	4.0167	47.530
CHP accidents......................	.016277	0.61939	7.329
FR-CHP accidents...................	.024158	0.99412	11.764
Two count convictions..............	-.0081364	0.067877	0.803
One count convictions..............	.021427	18.308	216.645
Paired accident convictions........	-.016549	0.21149	2.503
Non-countable convictions..........	.019209	4.0477	47.898
Age................................	-.000044307	2.1675	25.649
Height.............................	-.0016969	0.27513	3.256
Weight.............................	.00094382	0.0078748	0.093
Height/weight ratio................	.0010598	0.020917	0.248
Marital status.....................	.00075407	0.0038260	0.045
License restrictions...............	.0024524	0.044470	0.526
Traffic density....................	.000071175	1.7097	20.232
Constant term = .14043	-	0.39687	4.696
FINAL EQUATION[b]-2A			
FR accidents.......................	.031670	3.8297	45.308
CHP accidents......................	.013495	0.45495	5.382
FR-CHP accidents...................	.020881	0.80625	9.539
One count convictions..............	.021328	19.805	234.313
Non-countable convictions..........	.019360	4.1550	49.157
Age................................	-.000039814	2.0539	24.300
Height.............................	-.0010500	0.38349	4.537
Traffic density....................	.000073014	1.8121	21.438
Constant term = .13635	-	1.2744	15.077

(a) F ratio of 3.84 required for significance at .05 level.

(b) All variables significant beyond .05 level.

Table 4.-RESULTS FOR ONE YEAR NON-CONCURRENT REGRESSION EQUATIONS
(Female subjects; N = 30,673)

Variables	Regression coefficient	Mean square	F(a)
INITIAL EQUATION			
Warning letters................	.026062	0.079682	1.980
Moving FTA's....................	.027044	0.016678	0.414
Non-moving FTA's...............	-.064131	0.044634	1.109
FR accidents....................	.031796	1.4539	36.121
CHP accidents..................	.011601	0.10086	2.506
FR-CHP accidents...............	.014675	0.11674	2.900
One count convictions...........	.019871	3.7437	93.010
Two count convictions...........	.019040	0.042601	1.058
Paired accident convictions.....	-.011848	0.032684	0.812
Non-countable convictions.......	-.0022126	0.0035749	0.089
Age............................	-.0000091527	0.057809	1.436
Height.........................	-.00084396	0.077161	1.917
Weight.........................	.00022941	0.00050152	0.012
Height/weight ratio............	.00020460	0.00093677	0.023
Marital Status.................	.019718	1.8162	45.121
License restrictions...........	.00064429	0.0021521	0.053
Traffic density................	.000082848	1.4088	34.999
Constant term = .046629	-	0.57524	1.429
FINAL EQUATION(b)-2B			
FR accidents....................	.031761	1.4812	36.798
One count convictions...........	.020531	4.2489	105.557
Marital status.................	.019511	1.8233	45.297
Traffic density................	.000083366	1.4355	35.661
Constant term = -.0027900	-	0.019500	0.484

(a) F ratio of 3.84 required for significance at .05 level; F = 6.63 required for significance at .01 level.
(b) All variables significant beyond .05 level.

Table 5.-RESULTS FOR ONE YEAR CONCURRENT REGRESSION EQUATIONS -- QUESTIONNAIRE SUBJECTS
(Male subjects: N = 536)

Variables	Regression coefficient	Mean square	F[a]
INITIAL EQUATION			
One count convictions	.14421	3.1535	37.946
Two count convictions	-.45847	0.18896	2.274
Non-countable convictions	-.099688	0.23596	2.839
Warning letters	-.43968	0.35634	4.288
Height	-.0022832	0.010820	0.130
Weight	.0012771	0.28106	3.382
Marital status	.054061	0.14416	1.735
Age	.00027688	0.20582	2.477
Dependents	-.00076644	0.00047374	0.006
Social contact	-.0088067	0.029063	0.350
Intelligence	.013943	0.038792	0.467
Different jobs	.00093395	0.00031554	0.004
Highest grade completed	-.0092204	0.28406	3.418
Economic status	-.016865	0.11196	1.347
Years driven	-.0057851	0.53203	6.402
Make of car	-.022987	0.24816	2.986
Year of car	-.0068230	0.37093	4.464
Miles per week	.00023540	0.53388	6.424
Miles night	.000015474	0.00013550	0.002
Miles freeway	.00046189	0.23014	2.769
Miles 1963	-.00032557	0.071639	0.862
Miles 1964	-.00023139	0.047987	0.577
Condition of car	-.015379	0.14326	1.724
Date received	.00028330	0.0050319	0.061
Traffic density	.0000093302	0.029073	0.350
Occupational power A	-.012351	0.014191	0.171
Occupational social reputation A	.019965	0.057833	0.696
Occupational power B	.028289	0.082866	0.997
Occupational social reputation B	-.053890	0.45420	5.466
Constant term = .65969	-	0.16152	1.944
FINAL EQUATION			
One count convictions	.14679	3.3871	41.168
Two count convictions	-.59338	0.33233	4.039
Non-countable convictions	-.097398	0.23393	2.843
Warning letters	-.45920	0.39413	4.790
Weight	.0012644	0.41344	5.025
Marital status	.057724	0.24102	2.929
Age	.00031105	0.29062	3.532
Highest grade completed	-.0079045	0.25605	3.112
Years driven	-.0055204	0.50355	6.120
Make of car	-.022699	0.25581	3.109
Year of car	-.0052017	0.25484	3.097
Miles per week	.00021759	0.50754	6.169
Miles 1963	-.00023410	0.20109	2.444[b]
Occupational social reputation B	-.021039	0.44316	5.386
Constant term = .33616	-	0.18612	2.262

(a) F ratio of 3.84 required for significance at .05 level; F = 2.71 required for significance at .10 level.
(b) P = about .15 (left in equation because of a significant suppressor effect on the miles 1963 conviction variable.)

Table 6.-MATRIX OF SIMPLE CORRELATIONS FOR VARIABLES USED IN THE REGRESSION ANALYSIS OF 42,228 MALE SUBJECTS

(Three year concurrent driver record)

Predictor variables	Criterion variable (accident involvement)	Predictor variables[a]																			
		One count convictions	Two count convictions	Passing convictions	Right-of-way convictions	Equipment convictions	Miscellaneous technical convictions	Non-countable convictions	Age	Weight	Marital status	License restrictions	Traffic density	Moving FTA's	Non-moving FTA's	Signal/sign convictions	Turning, stopping, and signalling convictions	Speed convictions	Major convictions	Height	Height/weight ratio
One count convictions.....	.204	-																			
Two count convictions.....	.039	.073	-																		
Passing convictions.....	.098	.402	.053	-																	
Right-of-way convictions.....	.060	.251	.023	.046	-																
Equipment convictions.....	.088	.293	.077	.111	.054	-															
Miscellaneous technical convictions.....	.097	.328	.123	.143	.075	.422	-														
Non-countable convictions.....	.111	.273	.079	.118	.049	.668	.680	-													
Age.............	-.081	-.237	-.047	-.071	-.018	-.168	-.140	-.157	-												
Weight..........	.007	.002	-.020	.001	-.008	-.023	.001	-.001	.018	-											
Marital status..	.051	.147	.052	.038	.030	.089	.067	.066	-.290	-.094	-										
License restrictions....	-.005	-.066	-.014	-.023	.007	-.054	-.044	-.052	.262	-.041	.002	-									
Traffic density.........	.080	.099	-.001	.032	.055	-.022	-.009	-.031	-.007	-.017	.048	.031	-								
Moving FTA's....	.019	.050	.026	.056	.019	.103	.245	.070	-.043	-.003	.022	-.017	.004	-							
Non-moving FTA's...........	-.001	.006	.006	.005	.001	.035	.059	.017	-.005	.000	.009	.000	.007	.042	-						
Signal/sign convictions.....	.122	.622	.049	.115	.087	.144	.204	.145	-.106	.005	.082	-.034	.116	.072	.006	-					
Turning, stopping, signalling convictions.....	.077	.394	.036	.105	.061	.102	.125	.088	-.034	.000	.040	-.010	.059	.039	.001	.154	-				
Speed convictions.....	.145	.731	.062	.157	.051	.220	.269	.236	-.255	.004	.133	-.072	.018	.099	.009	.198	.098	-			
Major convictions.....	.033	.063	.887	.102	.020	.069	.107	.070	-.042	-.020	.044	-.012	-.001	.025	.004	.040	.030	.220	-		
Height..........	.011	.033	-.011	.004	-.010	.009	.022	.017	-.192	.489	.023	-.061	-.018	.009	.004	.005	-.010	.057	-.011	-	
Height/weight ratio...........	.003	-.008	-.019	-.001	-.007	-.028	-.006	-.007	.080	.937	-.112	-.028	-.013	-.005	-.001	.004	.004	-.014	-.019	.212	-

(a)An r of .010 is required for significance at .05 level of confidence.

Table 7.-MATRIX OF SIMPLE CORRELATIONS FOR VARIABLES USED IN THE REGRESSION ANALYSIS OF 30,277 FEMALE SUBJECTS

(Three year concurrent driver record)

Predictor variables	Criterion variable (accident involvement)	Moving FTA's	One count convictions	Two count convictions	Right-of-way convictions	Turning, stopping, and signalling convictions	Speed convictions	Equipment convictions	Non-countable convictions	Age	Height	Marital status	Traffic density	Non-moving FTA's	Signal/sign convictions	Passing convictions	Major convictions	Weight	Height/weight ratio	License restrictions	Miscellaneous technical convictions
		Predictor variables(a)																		
Moving FTA's	.025	-																		
One count convictions	.150	.045	-																	
Two count convictions	.040	-.002	.023	-																
Right-of-way convictions	.061	.024	.320	-.006	-															
Turning, stopping, signalling convictions	.040	.015	.391	.011	.064	-														
Speed convictions	.108	.091	.642	.018	.047	.048	-													
Equipment convictions	.030	.045	.143	.019	.028	.049	.075	-												
Non-countable convictions	.046	.053	.122	.028	.030	.064	.085	.606	-											
Age	-.021	-.021	-.061	.002	-.002	.043	-.128	-.058	-.068	-										
Height	-.016	-.011	.005	-.015	.001	-.009	.016	-.004	-.007	-.082	-									
Marital status	.087	.008	.137	.027	.043	.064	.072	.028	.026	.103	.036	-								
Traffic density	.079	-.003	.086	.003	.043	.041	.014	-.002	-.002	.016	-.009	.093	-							
Non-moving FTA's	-.002	.000	-.001	.001	-.001	-.001	-.001	.016	.008	.006	-.002	-.002	.004	-						
Signal/sign convictions	.098	.044	.663	.021	.078	.106	.175	.061	.075	.000	-.006	.101	.087	.001	-					
Passing convictions	.050	.042	.330	.008	.034	.048	.069	.044	.048	-.004	-.003	.039	.031	.001	.080	-				
Major convictions	.024	.006	.012	.688	-.004	.005	.011	.011	.020	-.007	-.009	.012	.007	.001	.011	.003	-			
Weight	-.007	-.002	.006	-.008	.010	.014	-.018	.011	-.008	.238	.341	.015	-.050	.014	.016	-.001	-.009	-		
Height/ weight ratio	-.004	.000	.004	-.006	.009	.017	-.023	.013	-.006	.275	.121	.007	-.049	.016	.018	.000	-.008	.937	-	
License restrictions	.006	-.008	-.026	-.009	-.001	.000	-.046	-.023	-.026	.252	-.017	.110	.018	-.003	.004	-.004	-.009	.039	.047	-
Miscellaneous technical convictions	.054	.210	.275	.044	.086	.102	.189	.243	.556	-.075	.001	.042	.013	.025	.184	.084	.033	-.011	-.011	-.020	-

(a)An r of .010 is required for significance at .05 level of confidence.

Table 8.-MATRIX OF SIMPLE CORRELATIONS FOR VARIABLES USED IN REGRESSION ANALYSIS OF 43,509 MALE SUBJECTS

(One year non-concurrent driver record)

Predictor variable	Criterion variable (accident involvement)	Predictor variable[(a)]																
		Warning letters	Moving FTA's	Non-moving FTA's	FR accidents	CHP accidents	FR-CHP accidents	Two count convictions	One count convictions	Paired accident convictions	Non-countable convictions	Age	Height	Weight	Height/weight ratio	Marital status	License restrictions	Traffic density
Warning letters........	.033	-																
Moving FTA's...........	.010	.013	-															
Non-moving FTA's.......	.005	.004	.145	-														
FR accidents...........	.047	.038	.003	.004	-													
CHP accidents..........	.026	.047	.009	.007	.031	-												
FR-CHP accidents.......	.030	.052	.004	.022	.038	.029	-											
Two count convictions..	.006	.028	.013	.022	.027	.080	.080	-										
One count convictions............	.101	.300	.027	.026	.131	.126	.129	.067	-									
Paired accident convictions...........	.018	.045	-.005	.015	.132	.260	.287	.192	.167	-								
Non-countable convictions...........	.060	.115	.051	.074	.045	.080	.066	.070	.247	.043	-							
Age....................	-.047	-.101	-.030	-.023	-.039	-.045	-.052	-.043	-.218	-.032	-.140	-						
Height.................	-.002	.018	.006	.007	.007	.005	.012	-.009	.032	.002	.021	-.191	-					
Weight.................	.009	-.012	-.002	-.005	-.004	-.004	.009	-.017	.001	-.002	.001	.021	.487	-				
Height/weight ratio....	.011	-.019	-.002	-.007	-.008	-.006	.005	-.017	-.008	-.003	-.006	.081	.201	.933	-			
Marital status.........	.022	.078	.012	.012	.038	.034	.038	.053	.137	.040	.055	-.291	.023	-.096	-.114	-		
License restrictions...	-.009	-.021	-.015	-.015	.004	-.006	.001	-.010	-.060	-.001	-.046	.261	-.060	-.040	-.027	.003	-	
Traffic density........	.030	.004	.002	.001	.062	.007	.018	-.008	.082	-.039	-.023	-.005	-.018	-.017	-.012	.044	.030	-

(a)An r of .010 is required for significance at .05 level of confidence.

Table 9.-MATRIX OF SIMPLE CORRELATIONS FOR VARIABLES USED IN REGRESSION ANALYSIS OF 30,673 FEMALE SUBJECTS

(One year non-concurrent driver record)

Predictor variables	Criterion variable (accident involvement)	Predictor variables[(a)]																
		Warning letters	Moving FTA's	Non-moving FTA's	FR accidents	CHP accidents	FR-CHP accidents	Two count convictions	One count convictions	Paired accident convictions	Non-countable convictions	Age	Height	Weight	Height/weight ratio	Marital status	License restrictions	Traffic density
Warning letters........	.016	-																
Moving FTA's..........	.007	.019	-															
Non-moving FTA's.......	-.004	-.000	-.000	-														
FR accidents...........	.045	.021	.016	-.004	-													
CHP accidents..........	.016	.006	.004	-.003	.012	-												
FR-CHP accidents.......	.018	.005	.014	.010	.020	.010	-											
Two count convictions..	.010	.015	-.001	.026	.052	.040	.046	-										
One count convictions............	.070	.110	.031	.028	.109	.091	.095	.013	-									
Paired accident convictions............	.016	.003	-.002	-.002	.147	.241	.209	.148	.181	-								
Non-countable convictions............	.007	.027	.080	.107	.033	.013	.026	.019	.103	.009	-							
Age....................	-.004	-.017	-.019	-.017	-.004	-.008	-.019	.008	-.051	.015	-.061	-						
Height.................	-.008	.008	-.001	-.003	-.013	-.010	-.002	-.013	.002	-.003	-.006	-.082	-					
Weight.................	-.001	-.001	-.001	.000	-.005	.000	-.000	-.008	.005	.012	-.009	.239	.341	-				
Height/weight ratio....	.001	-.006	-.001	.000	-.003	.001	-.002	-.014	.004	.013	-.007	.275	.121	.938	-			
Marital status.........	.052	.016	.014	.012	.065	.034	.039	.031	.125	.032	.026	.103	.035	.016	.008	-		
License restrictions...	.004	.001	-.009	.003	.001	.003	-.003	-.005	-.020	.009	-.023	.255	-.017	.041	.049	.109	-	
Traffic density........	.044	.008	.002	.006	.047	.019	.024	.005	.072	-.020	.001	.015	-.009	-.052	-.051	.093	.017	-

(a)An r of .010 is required for significance at .05 level of confidence.

139

Table 10.-MATRIX OF SIMPLE CORRELATIONS FOR VARIABLES USED IN REGRESSION ANALYSIS OF 536 MALE QUESTIONNAIRE SUBJECTS

(One year concurrent driver record)

Prediction variables	Criterion variable (accident involvement)	Prediction variables[a]																								
		One count convictions	Two count convictions	Non-countable convictions	Warning letters	Height	Weight	Marital status	Age	Social contact	Intelligence	Highest grade completed	Economic status	Years driven	Make of car	Year of car	Miles per week	Miles freeway	Miles 1963	Miles 1964	Condition of car	Traffic density	Occupational power A	Occupational social reputation A	Occupational power B	Occupational social reputation B
One count convictions......	.265	-																								
Two count convictions......	-.012	.124	-																							
Non-countable convictions......	-.005	.187	-.009	-																						
Warning letters..	-.017	.227	-.003	-.013	-																					
Height...........	.014	-.012	.056	-.052	.007	-																				
Weight...........	.042	-.049	.058	-.047	-.016	.553	-																			
Marital status...	.089	.156	.081	.065	.041	.039	-.199	-																		
Age..............	-.053	-.206	-.032	-.124	-.054	-.177	.109	-.387	-																	
Social contact...	-.028	-.066	.053	.052	.012	-.072	-.018	.088	-.036	-																
Intelligence.....	.002	.014	.062	.104	-.006	-.055	.066	-.013	.031	.631	-															
Highest grade completed........	-.020	.036	-.073	-.058	-.011	.094	-.013	.045	-.218	-.471	-.534	-														
Economic status..	-.043	.025	.144	.076	.048	-.047	-.021	.119	-.026	.340	.337	-.316	-													
Years driven.....	-.083	-.195	-.076	-.120	-.072	-.067	.272	-.407	.892	-.064	.014	-.167	-.059	-												
Make of car......	-.078	-.011	-.027	-.025	.024	.035	-.032	.127	-.189	.095	.084	-.009	.106	-.170	-											
Year of car......	-.034	.073	-.048	-.095	.008	.106	.061	-.041	-.061	-.258	-.188	.192	-.180	-.013	.041	-										
Miles per week...	.096	.182	.047	.038	.008	.090	.150	-.129	-.057	-.134	.089	.095	-.058	.024	-.071	.190	-									
Miles freeway....	.006	.020	-.002	-.045	.004	.041	.073	-.088	-.040	-.153	-.006	.114	-.068	.013	.034	.117	.444	-								
Miles 1963.......	.022	.128	.002	.007	-.021	.036	.031	-.102	-.108	-.153	.011	.122	-.102	-.029	-.048	.152	.716	.762	-							
Miles 1964.......	.008	.131	.020	.024	-.006	.054	.089	-.099	-.137	-.176	-.006	.126	-.106	-.070	.008	.139	.634	.733	.916	-						
Condition of car.	-.063	-.062	.166	-.045	.063	-.049	-.127	.094	-.077	.152	.118	-.106	.279	-.128	.056	-.365	-.172	-.067	-.170	-.162	-					
Traffic density..	.041	.024	-.059	-.074	.019	-.075	-.059	.058	.063	-.101	-.162	.065	-.070	.013	.015	.012	-.043	-.065	-.039	.003	-.026	-				
Occupational power A..........	.008	.052	.053	.132	.030	-.016	.075	.083	-.040	.606	.767	-.425	.331	-.051	.071	-.240	.038	.011	.007	-.015	.121	-.066	-			
Occupational social reputation A.....	-.006	.045	.065	.124	.018	-.003	.057	.089	-.023	.552	.789	-.501	.324	-.051	.041	-.207	.032	-.011	-.015	-.029	.104	-.085	.836	-		
Occupational power B..........	-.012	.032	.050	.146	.025	-.003	.069	.084	-.024	.533	.720	-.367	.289	-.023	.090	-.205	.028	.029	.002	-.011	.114	-.058	.911	.775	-	
Occupational social reputation B.....	-.051	.024	.066	.133	.017	.002	.053	.063	-.021	.493	.752	-.469	.287	-.040	.064	-.184	.040	.035	.011	.007	.105	-.090	.781	.909	.818	-

(a)An r of .084 is required for significance at .05 level of confidence.