Accidents, Convictions, and Demerit Points: An Ontario Driver Records Study

A. SMILEY, B. PERSAUD, E. HAUER, AND D. DUNCAN

A sample of 827,955 records of Ontario drivers containing information about age, gender, convictions, accidents, demerit points, and suspensions for 1981–1984 has been examined. On this basis 16 alternate models to estimate a driver's accident potential have been formulated. It appears that the currently used demerit point system, wherein the number of points associated with an offense reflects the perceived seriousness of the offense, is not a good predictor of accident potential. One can predict better by relying on the driver's record of accidents and convictions and still better by making use of a model for which the "regression weights" have been rigorously estimated. The performance of alternative models for the estimation of drivers' accident potential is described in terms of "hits" and "false alarms." It is shown, for example, that if the top 10,000 drivers are selected by the best model, 3,757 of these are expected to have an accident potential in excess of four times the population average; these are the "hits." Of the same 10,000, one should expect 792 to have an accident potential that is below the population average. These are the "false alarms." The best model uses age, gender, total accidents, and 14 conviction categories. This model identifies approximately twice as many high accident potential drivers as the current demerit point system. Even the simplest model, which uses total convictions as the only variable, predicts 50 percent more high accident potential drivers than the current system.

The current demerit point system in use in Ontario allocates points to offenses on the basis of the perceived seriousness of the offense. An offense is considered serious if it is thought to be associated with a relatively large chance of precipitating an accident. This is why a nonmoving violation, such as not having a trailer permit, receives no points but running a red light receives several points. The goal of the work described in this paper was to allocate points to offenses with a different purpose in mind. The purpose here was to use a driver's record of convictions and accidents to predict, as well as possible, which drivers, based on their past record, are most likely to have an accident in the near future. This is determined by estimating "accident potential" for each driver, namely, how many accidents per year a driver is likely to have, on the average.

Because a person's "accident potential" can be only indirectly estimated (not directly measured) and because, mercifully, it is rare for any one driver to be involved in an accident, the accuracy with which a driver's accident potential can be estimated is bound to be severely circumscribed. Thus, the aim of this work is not only to produce for each driver an estimate of his or her accident potential but also to say how accurate that estimate is.

Such estimates are the kind of knowledge that might then be used in the determination of post-licensing-control action. Thus although a nonmoving violation may not be a threat to traffic safety, such a conviction on a driver's record may be an important clue about that person's likelihood of future accidents.

DRIVER RECORD SAMPLE

The analysis examined driver records over a recent 4-yr period. Of the 5.5 million Ontario drivers, 827,995 qualified for inclusion in the sample. Driver record data included the following information: age and gender; for each conviction: type, date, and demerit points assigned; for each accident: degree of severity and date; and for each suspension: type and time period.

PREPARATION FOR ANALYSIS

Making sense of large data sets requires careful preparation. First, it was established that the sample statistics correspond to what is known about the population of Ontario drivers. Next, several consistency checks were performed on a sample of the data. Inconsistencies could not be removed in all cases. For example, of 45 drivers convicted for "failing to remain" at the scene of an accident, only 28 show an associated accident. Also, the count of a certain conviction changed from 0 in 1981 to 68 in 1984. This must reflect a change in law or enforcement practices. Following these preliminaries the main preparatory task—grouping the multitude of offenses into a smaller number of categories—was begun.

Selection of Conviction Categories

A preliminary analysis of a sample of about 8,000 drivers indicated that, during the 4-yr period 1981–1984, these drivers were convicted of approximately 200 different traffic offenses. Speeding accounted for some 60 percent of non-accident-related

A. Smiley, Human Factors North Inc., 118 Baldwin Street, Toronto, Ontario M5T 1L6, Canada. B. Persaud, Department of Civil Engineering, Ryerson Polytechnical Institute, 350 Victoria Street, Toronto, Ontario M5B 2K3, Canada. E. Hauer, Department of Civil Engineering, University of Toronto, Sir Sanford Fleming Building, St. George Street, Toronto, Ontario M5S 1A4, Canada. D. Duncan, Ontario Ministry of Transportation, 1201 Wilson Avenue, West Building, Downsview, Ontario M3M 1J8, Canada.

convictions; seat belt offenses accounted for about 10 percent and failure to stop at an intersection, for about 5 percent. Most of the other offense types had very few convictions. It was obviously impractical to assign different weights for each of 200 conviction types. Nor was it feasible to obtain statistically reliable estimates of how much these "leaner" convictions add to a driver's expected number of accidents. It was therefore necessary to place offenses with few convictions into larger groups. It should be noted that the same approach is already present in the current Ontario demerit point system. There are essentially 7 categories of offenses, that is, those assigned 0, 2, 3, 4, 5, 6, and 7 demerit points. For example, all the nonmoving violations as well as some of the moving ones fall into the 0-point category.

The clustering of all possible offense types into a manageable number of categories was accomplished in several steps. The first step was to combine the offense types that are quite similar in nature. In this way the various conviction types were combined into 45 categories. This first step was based on judging which offenses were of a similar nature. In the second step, groups of offenses were identified that were similar in their contribution to a driver's average number of accidents. To obtain results that are unambiguous and free of confounding, the only records used were those for drivers who in any one of the 4 yr had a single conviction. The results are given in Table 1. For example, there were 12,337 drivers who in 1 yr were convicted for not wearing a seat belt (offense *m*1) and had no other conviction in that year. During the

TABLE 1 ACCIDENTS FOR DRIVERS WITH ONE CONVICTION IN 1 YR

Category	Brief Description	No. of Drivers	3 yr. Accs.	Wghtd Mean	95% Upper	Limits Lower
n1	Minor neglect, licenses, permit	6495	2918	0.434	0.445	0.422
n2	Neglect, insurance, permits, etc.	1589	719	0.414	0.438	0.392
n3	License suspended, HTA	874	454	0.424	0.456	0.394
n4	Learners	34	18	0.343	0.502	0.212
v1	Minor veh.; lamps, noise	2954	1498	0.468	0.485	0.451
v2	Brakes, tires, unsafe vehicle	946	451	0.400	0.430	0.371
v3	Comm. veh.; size & weights	503	369	0.542	0.583	0.500
m1	Seat belt	12337	4858	0.376	0.384	0.368
m2	Speeding	173592	55211	0.319	0.321	0.317
m3	Careless driving	902	342	0.327	0.357	0.299
m4	Slow driving	45	11	0.119	0.237	0.055
m8	STOP sign, ROW violations	14024	3935	0.288	0.295	0.281
m9	PXO violations	1237	355	0.296	0.320	0.272
m10	Turn violations; right, left, U	18231	4942	0.283	0.289	0.277
m11	Unsafe move; open door	1649	542	0.334	0.355	0.312
m13	Disobey red light	13731	4270	0.313	0.321	0.306
m14	Amber light	3453	982	0.285	0.299	0.271
m15	Advance green	274	73	0.265	0.317	0.218
m16	Fail to share road	170	62	0.303	0.372	0.242
m17	Passing violations	1305	459	0.327	0.351	0.303
m18	Wrong-way one way street	1582	458	0.284	0.306	0.264
m19	Improper driving divided h'way	2599	900	0.361	0.379	0.344
m20	F.T.C.	934	337	0.344	0.374	0.316
m21	Emerg. veh., school X'ing	48	15	0.159	0.280	0.084
m22	R/R crossing violations	95	35	0.314	0.408	0.233
m24	Headlight beam not lowered	225	71	0.260	0.318	0.209
m25	Improper parking	145	77	0.407	0.484	0.334
m26	Fail to stop for school bus	604	133	0.281	0.317	0.249
m28	Disobey traffic signals	1650	529	0.322	0.344	0.301
m29	Fail to report accident	224	73	0.266	0.324	0.215
m30	Fail to remain at scene	236	75	0.315	0.373	0.261
m32	Dangerous driving C.C.C.	5	2	0.010	0.421	0.000
m33	Fail to remain at accident C.C.C.	66	39	0.377	0.491	0.274
m34	Dangerous driving C.C.C.	89	46	0.281	0.376	0.202
m35	Impaired driving C.C.C.	2381	1040	0.443	0.462	0.424
m36	Fail/refuse breath test C.C.C.	94	37	0.218	0.306	0.149
m37	Fail or ref. breath test C.C.C.	120	50	0.314	0.397	0.241
m38	Driving with >80 mgs. alcohol	3676	1502	0.386	0.401	0.371
m41	Crowding driver seat	120	36	0.280	0.362	0.211
m44	Radar device in vehicle	68	27	0.251	0.359	0.166
m45	No safe helmet, motorcycle	96	47	0.266	0.357	0.191
m46	Fail to signal to stop	18	6	0.080	0.277	0.019
m47	FTC, commercial vehicle	67	35	0.362	0.475	0.262
m48	Fail to stop for police officer	12	6	0.086	0.340	0.016

remaining 3 yr these drivers recorded 4,858 accidents, for an average of 0.394 accident per driver. To eliminate any bias due to differences in the age-gender distribution that might be associated with specific offenses, all averages were recalculated for a "standard population." The "standard population" used had an age-gender composition of those drivers who had exactly one conviction of any kind in the 4 yr. This is why in the "weighted average" column, the average number of accidents in the remaining 3 yr associated with this seat belt offense is listed as 0.376 rather than 0.394. Similarly, the 173,592 drivers who had only a single speeding conviction in some year have an adjusted average of 0.319 accident in 3 yr. The last two columns give 95 percent confidence limits for the weighted average.

Some of the 45 conviction categories were found to be associated with a similar weighted average and could be combined. The resulting 14 conviction groupings, the associated 3-yr average number of accidents, and 95 percent confidence limits are shown in Figure 1. Also shown are estimates of accident potential for conviction-free drivers, those who had no convictions of any type during 1 calendar yr.

In summary, the final 14 conviction categories to be used in analysis were established on the basis of the following considerations:

1. Conviction types within each category were similar in nature,

2. The accident potentials associated with each conviction within a category were similar, and

3. The numbers of drivers with offenses in each category was sufficient to provide a reliable estimate of accident potential for that category.

Inspection of Figure 1 leads to the question of why it is that convictions for, say, minor neglect of vehicular condition (No. 5) are found to be associated with more accidents than those offenses traditionally deemed very dangerous, such as speeding (No. 13) or running a red light (No. 12). Several reasons combine to explain this apparent paradox. First, the convictions associated with most accidents (Nos. 1 and 2 in Figure 1) are those characteristic of truck drivers. These drivers cover 10–20 times the distance of a passenger car driver. Therefore, it is to be expected that they will have, on the average, more accidents. Thus most of those who have a type 1 or 2 conviction are truck drivers who by virtue of exposure have a greater than average number of accidents. This, in turn, results in the average number of accidents associated with these conviction types being greater than those associated with other conviction types, simply because a greater percentage of drivers with this conviction type are truck drivers.

The second reason is easiest to explain through an example. Assume that 1,000 run-the-red offenses lead to 5 accidents and that 1,000 fail-to-signal-turn offenses lead to 1 accident. Thus running a red is a more dangerous offense than failure

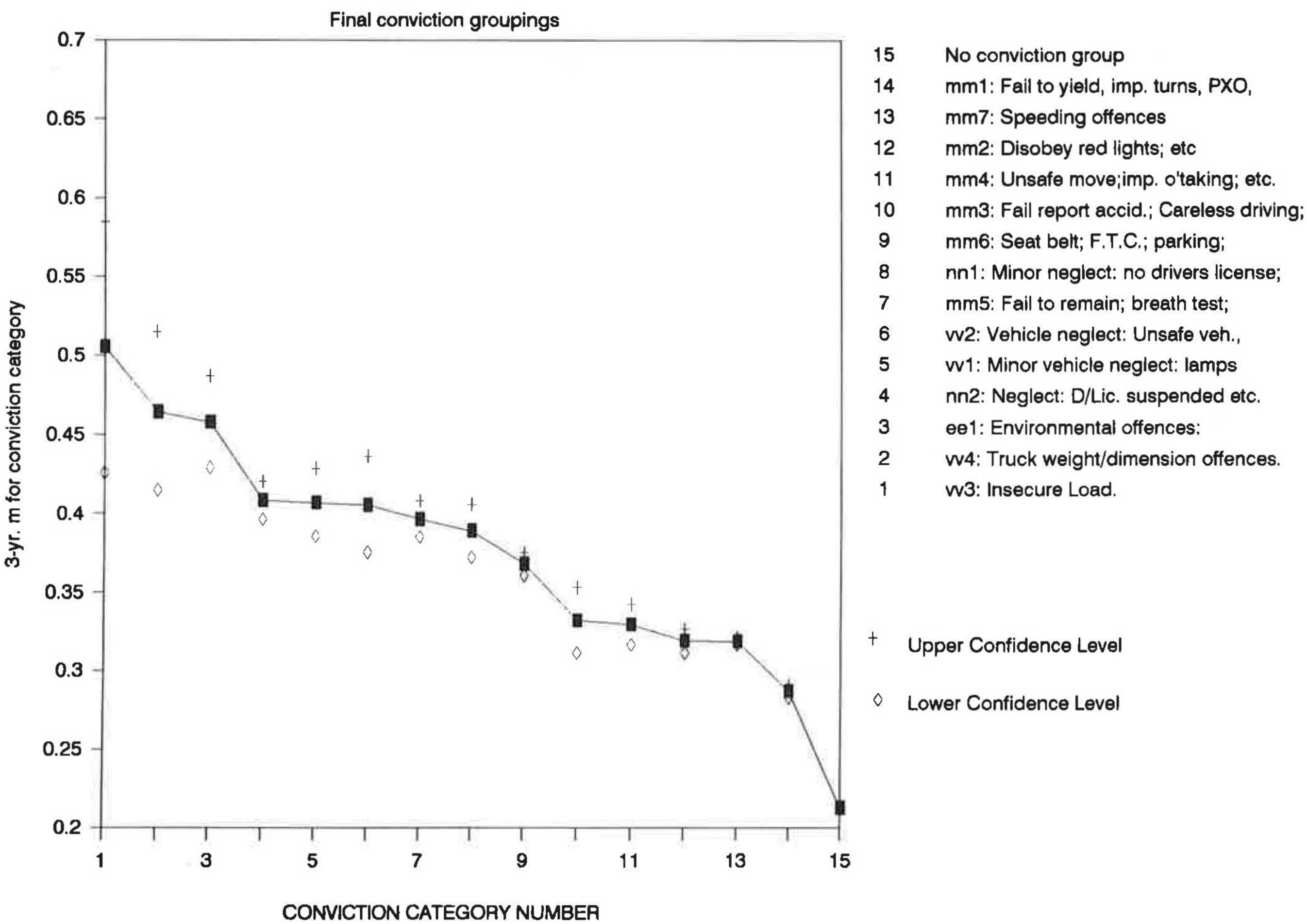

FIGURE 1 Expected numbers of accidents in a subsequent 3-yr period for drivers with one conviction in any year.

to indicate a turn. However, the enforcement for the offenses is unequal (perhaps because one is assumed to be more dangerous than the other). Assume that of the 1,000 run-the-red offenses, 10 lead to a conviction, whereas of the 1,000 fail-to-signal offenses 1 leads to a conviction. Thus, in a figure similar to Figure 1, we would see 5/10 = 0.5 accident per conviction for running the red and 1/1 = 1.0 accident per conviction for failure to indicate a turn. Even though, according to the starting assumption, running the red is five times more dangerous than failing to indicate a turn, because enforcement of the two offenses is unequal, the final result indicates the contrary. The problem is caused by the fact that the driver record contains information about convictions, not about the number of illegal actions committed by a driver.

A third reason might be related to the connection between different types of behavior and convictions. The incidents on a person's driving record, convictions and accidents, are indications of his or her overall driving behavior. As a result, the types of convictions committed by certain types of people may also provide insight into their potential for accidents. If a person engages in certain behaviors that lead to certain convictions, he or she may also engage in certain other behaviors that predispose the participant to accidents. To illustrate, drivers with environmental types of convictions (e.g., a noisy muffler) were found to have a higher weighted mean of accidents than drivers with most other types of convictions. Most drivers quickly have their noisy muffler fixed and are unlikely to receive this type of conviction. The attitude that results in drivers coming to the attention of the police and being charged with this offense may be related to a similar careless attitude toward behavior that results in accidents.

Two conclusions follow. First, one should not interpret the results in Figure 1 as providing information about the danger inherent in this or that offense. Second, one should not be surprised when, in the subsequent analysis, innocuous offenses prove to be strongly related to the driver's accident potential.

Age and Gender Categories

It is well known that the average number of accidents for a driver depends on gender and age. To account for this fact, age and gender will be used in the analysis as "variables." It is relatively simple to account for gender because it comes in two natural categories. The relationship between age and number of accidents, however, is continuous in nature and distinctly nonlinear, as shown in Figure 2. To include age in the analysis, it was necessary to establish a number of age categories. After careful analysis, the boundaries between age groups were chosen so that the average number of accidents within each group remained nearly constant while sufficient numbers of drivers within each age category were still retained to maintain statistical reliability. The eight age categories chosen are indicated on Figure 2.

Exclusion of Drivers with Suspended Licenses

Some drivers in the sample had their licenses suspended during the study period. Many are drivers who had a number of convictions that carried points. The extent to which a suspended driver curtails his or her driving is unknown. This

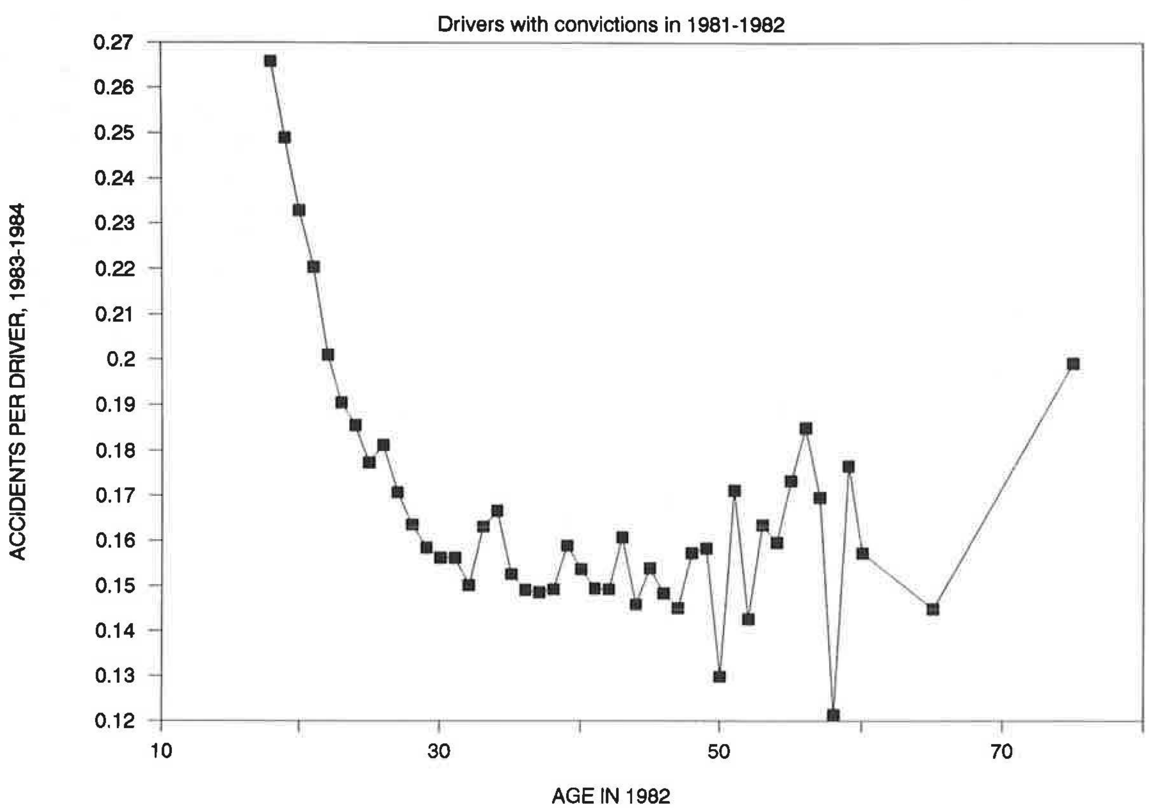

FIGURE 2 Relationship between age and accidents.

leads to serious difficulty in the statistical analysis. Consider a driver who in the first 2-yr period had many convictions and was suspended. That driver can be expected to drive less in the second 2-yr period and therefore to have proportionately fewer accidents. In the statistical analysis this would tend to create a negative correlation. That is, it would lead to the incorrect result that the larger the number of convictions in the first 2 yr, the fewer accidents a person is likely to have in a subsequent period. The net effect of this difficulty is to distort the results of analysis in some unpredictable way. In fact, in the initial statistical analyses, this distortion was so large that negative weights were produced for criminal code offenses for which drivers were likely to be suspended. As a result of this finding, it was necessary to remove from the data set and from subsequent analysis those drivers who were under suspension at any time in the period 1981–1984. The outcome of this decision is that whatever results are obtained in the course of the analysis apply directly only to those drivers whose licenses have not been suspended. The extension to drivers who were suspended under the present demerit point system is therefore an extrapolation.

ANALYSIS AND RESULTS

The various activities described so far (checking for representativeness, conducting consistency checks, selecting conviction categories, determining age groups, and removing suspended drivers) are all preliminary to the main activity, namely, the establishment of a relationship between information contained in a driver's record and his or her expected number of future accidents. The information used was a driver's gender, age, the count of accidents (at-fault, not-at-fault, or total), and the count of convictions in each of the 14 categories. This information from the first 2 yr is used to estimate "regression weights" that best fitted the accident record in the second 2-yr period. These regression weights are the relative number of points each conviction category should be assigned for the best prediction of the likelihood of an accident in the second 2-yr period. Only the records of those drivers who had at least one conviction in the first 2-yr period were used (n'' 170,000). The tool of analysis was GLIM (*1*), which yields maximum likelihood estimates of the regression weights and facilitates estimation using the negative binomial error structure, which is appropriate in this case.

Schemes and Variants Examined

The Ontario Ministry of Transportation was interested in a number of variants, each using different sets of variables. These fall into three categories:

1. Models that made use of age and gender information and models that did not;
2. Models that assigned points for accidents (with the further distinction between at-fault, not-at-fault, and total) and models that did not assign points for accidents; and
3. Models that assigned different numbers of points for each conviction category and models in which all convictions carried the same weight.

In total, 16 different combinations of variables from the first 2-yr period were used to estimate "weights" to calculate the expected number of accidents in the second 2-yr period. Each of these 16 combinations results in a prediction equation that is termed a "model." Table 2 shows which variables were used in each "model."

Details about the models are given elsewhere (*2*). Here the essential nature of the models is illustrated, using Model A2. Consider a female driver, 24 yr of age, who in the first period had two speeding convictions: one conviction for failing to yield the right of way and one for an at-fault accident. Regression weights for Model A2 are shown in Table 3. The base driver for all models is a 17–20-yr-old male who is conviction and accident free in the first period and who is expected to have 0.176 accident in the second period. From this value one has to subtract 0.061 for being female and 0.039 for being 24 yr old. One has to add 0.027 for each speeding conviction, 0.027 for the right-of-way conviction, and 0.058 for the accident. On the basis of Model A2, this driver is expected to have $0.176 - 0.061 - 0.039 + 2(0.027) + 0.027 + 0.058 = 0.215$ accident in the second 2-yr period.

The Distribution of "Accident Potential"

The models estimate for each driver the number of accidents he or she is expected to have per year in the second 2-yr period. For brevity, this number is called a driver's accident potential. Of course, not all drivers have the same accident potential: some drive more, some drive less; some take risks, others are more cautious. Before examining results for each model, let us examine the diversity of accident potential in the population of Ontario drivers. This will reveal how many drivers there are in the population who have a high accident potential. How many of these "high accident potential" drivers will indeed be identified for postlicensing control under the current demerit point scheme and the new models is examined later.

The number of accidents in the second 2-yr period was used to estimate the mean accident potential (0.055 accident/year) and its standard deviation (0.055 accident/year) in the total driver record sample. Details of the method are given elsewhere (*1*). This information was then used to plot the distribution of accident potential in a population of 5 million Ontario drivers shown in Figure 3. Using Figure 3 it can be established how many drivers in the population have an accident potential between any two chosen levels. Thus, for example, almost 90,000 drivers are estimated to have an accident potential of 0.22 accident/year or higher.

Performance of the Current Demerit Point System and of the New Models

Because the 16 new models were derived using appropriate statistical methods rather than by subjectively weighting each offense according to its perceived seriousness, they should perform better than the current demerit point system. However, all models face the same difficulties as the current demerit point system. Namely, because of the randomness inherent in the process of accident occurrence and the randomness inherent in the process by which drivers acquire convictions,

TABLE 2 VARIABLES USED FOR REGRESSION RUNS (*x* INDICATES VARIABLE USED IN RUN)

Run	Age & Sex Dummy Variables	Variables For Each Conviction Group	Variable for Total Convictions	Accident Variables		
				Total	Fault	At Fault
A1	x	x				
A2	x	x		x		
A3	x	x			x	
A4	x	x			x	x
B1		x				
B2		x		x		
B3		x			x	
B4		x			x	x
C1	x		x			
C2	x		x	x		
C3	x		x		x	
C4	x		x		x	x
D1			x			
D2			x	x		
D3			x		x	
D4			x		x	x

Note: x = indicates variable used in run.

a 2-yr record is just too short for an accurate estimate of a driver's accident potential. As will be seen, the new models are an improvement on the current system but, like the current system, still fail to detect many of the high accident potential drivers. In addition, many drivers identified by the models do not have a high accident potential.

Two measures of performance will be used to judge the quality of a model. The first measure of performance is straightforward. Consider, for example, the 10,000 drivers who in the first 2-yr period had the most demerit points (last row in Table 4). Checking the accident records of the same drivers, we find that during the second 2-yr period they had 1,452 accidents per year (see sum of first three columns in Table 4). Consider now another group of 10,000, this time those who in the first 2-yr period had the most accidents (the second from last row in Table 4). This second group recorded 1,828 accidents per year in the subsequent 2-yr period. Evidently, it is better to identify drivers by their previous accident record than by previous demerit points. Imagine now that a third group of 10,000 drivers is identified, this time those for whom Model A4 estimates the highest accident potential on the basis of their age and gender, as well as of convictions and accidents in the first 2-yr period. This group has 2,084 accidents per year in the second period. Thus, selection by Model A4 gives a richer catch than selection either by previous accidents or by the current demerit point system. In interpreting

TABLE 3 REGRESSION COEFFICIENTS FOR MODEL A2

Description of Variable	Estimated Coefficient	Standard Error
Intercept (Male <21)	0.1763	0.002876
Dummy variable for age 21-25	-0.03856	0.002852
Dummy variable for age 26-30	-0.0565	0.002913
Dummy variable for age 31-35	-0.06006	0.003044
Dummy variable for age 36-40	-0.05202	0.003516
Dummy variable for age 41-50	-0.06013	0.00351
Dummy variable for age 51-60	-0.05774	0.004832
Dummy variable for age > 60	-0.06156	0.007068
Dummy variable for female	-0.06122	0.001735
mm1: Fail to yield, imp. turns, PXO, amber violations, etc.	0.02696	0.001753
mm2: Disobey red lights; rail crossing violations	0.042125	0.002916
mm3: Fail report acc.; careless driving; dang. driving; crim. neg. caus. death	0.023304	0.008613
mm4: Unsafe move; imp. o'taking; disobey signs	0.06359	0.004959
mm5: Fail to remain; breath test; alcohol; impairment	0.2444	0.054575
mm6: Seat belt; F.T.C.; parking; divided h'way offences	0.03221	0.00213
nn1: Minor neglect; no drivers license; permits; insurance, address change	0.02624	0.010463
nn2: Neglect: D/Lic. suspended or not produced; plates, insurance	0.03366	0.003708
vv1: Minor vehicle neglect: lamps windows obstructed, etc.	0.04954	0.006561
vv2: Vehicle neglect: unsafe veh., brakes, tires	0.08673	0.009296
vv3: Insecure load	0.159215	0.02721
vv4: Weights and dimension offences	0.11074	0.011895
ee1: Environmental offences: noise, fumes	0.08748	0.008334
mm7: Speeding offences	0.026515	0.000866
Total accidents in period 1	0.05831	0.001580

these results one has to keep in mind that the count of (Period 2) accidents is always subject to random fluctuations.

Several conclusions emerge. First, one can do better than to use the current demerit point weights. Second, it is important to make use of the driver's record of accidents. This is already evident from the comparisons of the last two rows. It also emerges, however, from the poor performance of Models A1, B1, C1, and D1, which do not make use of accident data. In fact, the top 1,000 drivers can be well identified by their previous accident record alone. Third, the more drivers are identified, the lesser the "yield." Thus, the top 1,000 drivers have a Period 2 accident rate of "0.3 accident/year, which is approximately six times the population average; for the first 10,000 drivers, the average accident rate is "0.2, and so on.

The measure of performance examined so far leaves the impression that the drivers who are identified indeed have an accident potential that is substantially larger than that for the population of all drivers. Although this is true for the group "on the average," this group itself may not be a homogeneous one. The second measure of performance by which the quality of the alternative models is to be judged relates to the diversity of accident potential within the group of drivers that these models identify.

A weighting scheme is like a net with which an attempt is made to catch drivers who, based on their 2-yr record, are likely to have an unusually high number of accidents in the next 2 yr. For illustration here, consider "unusually high" to be 3 standard deviations above the mean. Because the mean for an Ontario driver is 0.055 accident/year and the standard deviation happens also to be 0.055 accident/year, it is hoped that drivers whose accident potential is larger than 0.22 accident per year will be identified. If we manage to identify such a driver, we will call this a "hit." Conversely, if based on the 2-yr record we identify, and call in for treatment, a driver

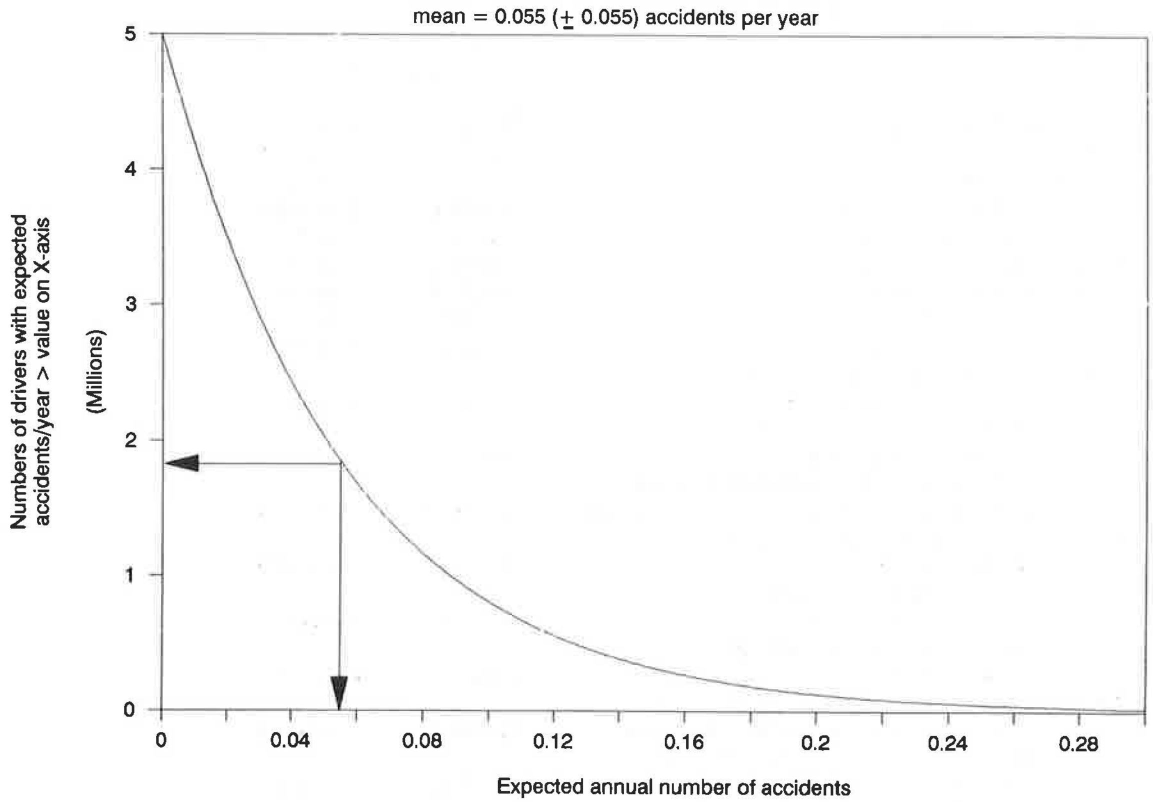

FIGURE 3 Distribution of accident potential in Ontario driver population.

whose accident potential is lower than the average accident potential in the population (that is, 0.05 accident per year), we will call this a "false alarm."

To illustrate, Models A4 and B4 are used. (The variables that these models use to estimate accident potential are shown in Table 2.) Drivers were ranked in terms of accident potential as estimated by each model based on their record during the first 2-yr period. Table 5 shows the hits (drivers correctly determined to have accident potential larger than 0.22 accident/year) and also the false alarms (drivers whose accident potential is below the population average of accidents/year) for consecutive groups of drivers identified by the two models as ranking highest on accident potential from a population of 5 million drivers.

Looking again at Figure 3, it can be seen that out of 5 million drivers, 90,000 drivers are expected to have an accident potential larger than 0.22 accident/year. As Table 5 shows, using Model A4 to select the 10,000 drivers with the worst records will catch 3,697 of the high accident potential drivers; calling in the next 10,000 will identify 2,679 more hits. Calling in the next 100,000 will yield another 15,987 hits. Thus, even after those 120,000 drivers of 5 million who have the highest estimated accident potential according to Model A4 have been selected for treatment, only 22,363 hits can be expected. Of the 90,000 drivers in the population who have an unusually high accident potential (>0.22) 67,637 remain still unidentified. The whole driver population would have to be called in before all the hits would be identified.

Table 6 compares performances among the 16 models, and the current demerit point system, in terms of hits and false alarms for the worst 10,000 drivers identified by each model. It should be noted that, although there always will be considerable overlap between groups of drivers identified by different models, there also will be systematic differences. Thus, for example, the use of Model Series *A* and *B* will lead to groups that contain more truck drivers than the current system, simply because the current system does not assign any points for truck weight or truck dimension offenses, whereas Models A and B weight these heavily.

Although comparison in terms of hits and false alarms is good for purposes of illustration, it depends on a rather arbitrary definition of what is to be considered an "unusually high" accident potential. A more comprehensive way to characterize the performance of different models is by continuous curves, as shown in Figure 4.

In Figure 4, accident potential is measured on the abscissa. The ordinate gives the number of drivers out of 10,000 whose true accident potential exceeds the value on the abscissa. The lowest curve represents the current demerit point scheme. In a group of 10,000 drivers who, in a population of 5 million, have the most demerit points, one can expect to find 2,800 who have an accident potential above 0.2 accident per year. The highest curves represent Models A4 and B4. In a group of 10,000 drivers who, in a population of 5 million, have the highest estimated accident potential by Model A4, one can expect to find some 4,200 drivers who have an accident potential above 0.2 accident per year. Thus, the higher the curve, the better the "net."

TABLE 4 ACCIDENTS PER YEAR RECORDED BY DRIVERS SELECTED BY VARIOUS MODELS

Model	Drivers estimated by model to be in:					
	Top 1,000	Next 4,000	Next 5,000	Next 10,000	Next 100,000	Total
A1	188	712	904	1660	13308	16772
A2	324	856	936	1798	14192	18016
A3	276	736	860	1712	13700	17284
A4	320	868	896	1736	14060	17880
B1	212	704	824	1548	12888	16176
B2	320	832	980	1628	13688	17448
B3	272	748	804	1712	13336	16872
B4	304	876	956	1636	13684	17456
C1	208	744	760	1592	13016	16320
C2	356	808	900	1672	13732	17468
C3	276	780	780	1424	13684	16944
C4	356	804	928	1616	13748	17452
D1	176	748	688	1432	12260	15304
D2	352	824	784	1608	13268	16836
D3	244	756	852	1360	13084	16296
D4	364	840	788	1576	13216	16784
Accs.*	312	756	760	1432	10780	14040
DP**	180	640	632	Not	Available	

* - Drivers with the highest accident counts in period 1 were selected

** - Drivers with the highest demerit points acquired in period 1

TABLE 5 FIGURES FOR MODELS A4 AND B4

Drivers estimated by model to be in:		Number of drivers expected to have:			
		m > 0.22		m > 0.05	
		Model A4	Model B4	Model A4	Model B4
the top	1,000	528	541	39	45
the next	4,000	1568	1595	246	289
the next	5,000	1601	1620	390	458
the next	10,000	2679	2657	933	1110
the next	100,000	15987	15291	14198	16928
TOTALS:	120,000	22363	21704	15806	18830

TABLE 6 FIGURES OF MERIT FOR 10,000 DRIVERS WITH HIGHEST *ms* (FOR EACH MODEL)

Model	Number of drivers expected to have: $m > 0.22$	$m < 0.05$
A1	3258	908
A2	3691	676
A3	3449	817
A4	3698	674
B1	3331	1062
B2	3750	806
B3	3516	923
B4	3757	792
C1	2911	1024
C2	3411	756
C3	3147	922
C4	3429	752
D1	2978	1211
D2	3441	909
D3	3155	1101
D4	3451	906
CURRENT DP	2231	1251

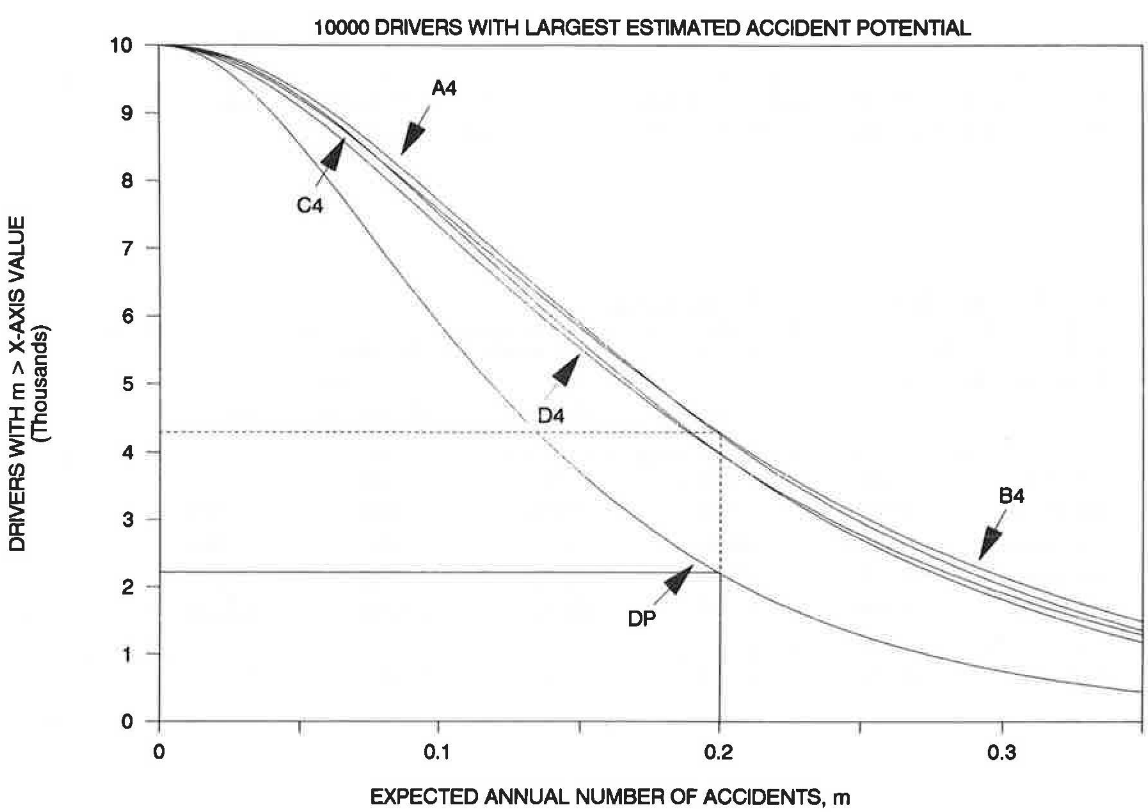

FIGURE 4 Current DP versus Schemes A4, B4, C4, D4.

SUMMARY AND CONCLUSIONS

The purpose of this work was to use a driver's record of convictions and accidents to predict, as well as possible, which drivers, based on their past record, are most likely to have an accident in the near future.

A sample of 827,955 records of drivers licensed to drive in Ontario during 1981–1985 has been examined. Each driver record contained information about the driver's gender, age, and details of accidents, convictions, demerit points, and suspensions.

In preparation for analysis, the many hundreds of offense types had to be grouped into a manageable number of categories. First, offenses that were similar in nature were put in the same group; then those offenses that were associated with a similar average number of accidents were consolidated. It turned out, for example, that drivers who in 1 yr had a single speeding conviction had fewer accidents in the remaining 3 yr than other drivers who had a single conviction in that year for a relatively minor offense, such as a missing lamp. This finding may initially be puzzling but, on reflection, aids the correct interpretation of later results. It arises partly because not all illegal behaviors lead to convictions at the same rate, partly because some offenses are specific to truck drivers who drive 10 to 20 times as much as car drivers and tend to have proportionately more accidents, and partly because behavior that results in a fairly innocuous offense, such as a noisy muffler, may be of the type that also leads to accidents. Therefore neither the ratio of accidents to convictions nor the "weights" that are later attached to particular offenses are an indication of the gravity of those offenses.

Drivers with a suspended license will curtail their driving to some extent. This is why, during the suspension period, one would expect some reduction in the number of accidents in which they are involved. However, the extent to which driving is curtailed is unknown. To assume that suspended drivers stop driving would introduce a bias into the analysis; assuming that they continue to drive would cause another bias. To protect the integrity of the results, drivers who were suspended had to be removed from the analysis. Therefore, the conclusions of this study apply directly only to drivers who have not been suspended. The extension to drivers who were suspended under the present demerit point system is therefore an extrapolation.

Sixteen models have been examined to estimate the expected number of future accidents for a driver based on age, gender, convictions, and accidents. The models differ from one another in the information they use. Some make use of age and gender, others do not. In some, each of 14 types of convictions is given a different weight; in others, all convictions have the same weight. In some, at-fault accidents are counted separately from not-at-fault accidents; in others, they are lumped together. All 16 models have a common structure: a "weighted sum" of convictions and, in some models, of accidents.

Two measures of performance were used to judge the quality of a model. The first measure of performance is the number of second-period accidents in a group of drivers (of fixed size) identified on the basis of their first-period record. Thus, those 10,000 drivers who in the first 2-yr period had the most demerit points recorded 1,452 accidents per year during the second 2-yr period. Those 10,000 drivers who in the first 2-yr period had the most accidents recorded 1,828 accidents per year in the subsequent 2-yr period. A third group of 10,000 drivers, those who by Model A2 have the highest accident potential when calculated on the basis of the first period data, recorded 2,116 accidents per year in the second period. Thus, selection by Model A2 gives a richer catch than selection either by previous accidents or by current demerit points. On this score, Model A2 performs best.

Several conclusions emerge. First, one can do better than to use the current demerit point weights. Second, it is important to make use of the driver's record of accidents. Third, the more drivers are identified, the lesser the "yield." Thus, the top 1,000 drivers have an accident rate of ″0.3 accident/year, which is approximately six times the population average; for the first 10,000 drivers, the average accident rate is ″0.2, and so on.

Not all drivers have the same expected number of accidents per year. Some drive more, some less; many drivers are prudent, some take unwise risks. On the basis of the accident data, it is shown how many drivers in Ontario have what expected number of accidents. Thus, for example, of 5 million drivers, some 90,000 have an expected number of accidents that is 3 standard deviations above the average for the population. It is these "high accident potential" drivers whom a demerit point scheme aims to identify.

A 2-yr record of convictions and accidents is just too short for estimating a driver's expected number of accidents with accuracy. This is why some of those identified by the model as having the highest expected number of accidents turn out, in reality, to be just average drivers. Conversely, this is why most high accident potential drivers may not have, in 2 yr, a record that identifies them as such. It was shown, for example, that of the 10,000 drivers who, by the "richest" model (A4) were estimated to have the highest expected number of accidents, 3,698 have an accident potential in excess of 3 standard deviations above the mean for the population. At the same time, 674 of those 10,000 were average drivers or better.

It may be of interest to note that little is gained by giving different numbers of points to different offenses. Model D1 uses simply 1 point per conviction and no accident data; Model D2 uses 1 point per conviction and 1.88 points per accident. Model D2 is only slightly worse than Model B2, which assigns different numbers of points to each of 14 offense classes. Table 7 compares hits and false alarms among the worst 10,000 drivers for those models. The more drivers selected, however, the more separate weights improve performance (in terms of predicting the number of future accidents). For the 100,000 worst drivers, separate weights help to increase the number of hits and to reduce the number of false alarms by about 10 percent.

In summary, if the purpose of a demerit point system is to identify drivers who are most likely to have an accident, the scheme used now is not as efficient as alternative schemes would be. Even by giving equal weights to all convictions and a weight of ″2 to an accident (D2), one can do much better. It is important to use data about accident involvement, but it does not pay to differentiate between at-fault and not-at-fault accidents.

For the top 5,000 or so drivers, the inclusion of age and gender information appears to be unimportant. For the next 100,000 drivers, consideration of age and gender improves performance (in terms of predicting future accidents) by a few percentage points. Consideration of age and gender does

TABLE 7 COMPARISON OF HITS AND FALSE ALARMS AMONG WORST 10,000 DRIVERS

	Hits >0.22	False Alarms <0.055 acc./yr.
Current Demerit Points	2231	1251
D1	2973	1211
D2	3441	909
B2	3750	806

not seem to increase the number of hits, but it helps in reducing the number of false alarms by some 10 percent.

With all this, one has to keep in mind that if only a few drivers are identified (say about 10,000), 30–40 percent of those will be genuine high accident potential drivers and 6–10 percent will be falsely identified better-than-average drivers. However, only 3 percent of all high accident potential drivers in the population will be in this group of 10,000 drivers. It does not help much to increase the size of the group because performance deteriorates with size. Thus, in a group of 120,000 drivers, only "19 percent genuinely have a high accident potential, whereas 13–16 percent are falsely identified. Even when as many as 120,000 drivers are identified by the richest model, only 22,363 of the 90,000 "high accident potential" drivers are caught in the net.

The performance of models for the estimation of a person's accident potential can be further improved. Consideration should be given to a system that tracks a person's accident potential nearly continuously. If during a certain period of time (measured in weeks) the driver did not acquire a conviction and was not involved in an accident, his or her estimated accident potential would be revised downward. If during that period of time, convictions or accidents were recorded, the estimated accident potential would correspondingly be revised upward. A person's aging, the general accident trend, and seasonal variation would also be reflected in these revisions. In this manner, a person's current estimated accident potential could be made a reflection of his or her entire driving history. In such a scheme, there is no need to specify an arbitrary period of time after which points are forgiven.

In the models developed so far, involvement in an accident adds a fixed amount to a driver's accident potential. Under the newly suggested scheme, an accident by a person with an already high accident potential would be weighed more heavily. In general, a "revision" scheme of this nature relies on solid mathematical logic and is expected to perform better than other possible schemes.

REFERENCES

1. E. Hauer, A. Smiley, and B. Persaud. *Accidents, Convictions and Demerit Points: An Ontario Driver Study—Technical Report*. Prepared for the Ontario Ministry of Transportation, Downsview, 1988.
2. E. Hauer, A. Smiley, and B. Persaud. *Accidents, Convictions and Demerit Points: An Ontario Driver Study—Summary Report*. Prepared for the Ontario Ministry of Transportation, Downsview, 1988.

Publication of this paper sponsored by Committee on Traffic Records and Accident Analysis.