

An In-Vehicle Data Recorder for Evaluation of Driving Behavior and Safety

Tsippy Lotan

OR YAROK,

38 Hashoftim St., Ramat Hasharon, Israel, 47210

Tel: (972) 3 760-6162 Fax: (972) 3 760-6160 Email: tsippy@oryarok.org.il

Tomer Toledo*

Department of Civil and Environmental Engineering,

Technion – Israel Institute of Technology, Haifa 32000, Israel

Tel: (972) 4 829-3080 Fax: (972) 4 829-5708 Email: toledo@technion.ac.il

* Corresponding author, Horev Fellow – supported by the Taub and Shalom Foundations

ABSTRACT

This paper describes the overall framework and components of an in-vehicle data recorder (IVDR) called DriveDiagnostics, and presents results from a study to validate its performance. The IVDR has been designed for the purpose of monitoring and analyzing driver behavior in normal driving situations and not only crash or pre-crash events. It records the movement of the vehicle and uses this information to indicate on the overall trip safety.

The validation study involved 33 drivers, whose vehicles were instrumented with the IVDR. The experiment first included a blind profiling stage in which drivers did not receive any feedback from the system, followed by a feedback stage in which drivers had access to personal web pages with the information recorded by the system. Data collected in the blind profiling stage was used to investigate the connection between drivers' safety indices as captured by the system, and historic crash data. The results show significant correlations between the two datasets. Thus, suggesting that the driving risk indices can be used as indicators to the risk of involvement in car crashes. This connection enabled us to also investigate the potential impact of the system on driving behavior and on safety. The results show that the initial exposure of drivers to the system has a significant positive impact on their behavior and on safety. Access to the feedback provided by the system has further impact on drivers' performance. However, if follow-up efforts are not made both these positive impact are not sustained over time.

INTRODUCTION

The human and cost implications of car crashes are staggering. Blincoe et al. (2002) estimated the direct cost of a car crash at \$14,000, out of which \$3,600 is the cost of damage to vehicles and other property. The total direct annual cost of car crashes in the US in 2000 was estimated at \$230.6 Billion, and the total cost to society at \$493.3 Billion. Thus, it is clear that the implications of a potential reduction in the risk of involvement in car crashes are large. There has been an increased interest in recent years in technology-based solutions that can assist drivers to reduce their risk of involvement in car crashes. One class of solutions that have been proposed are the installation of In-Vehicle Data Recorders (IVDR) that monitor and provide feedback on drivers' behavior.

In-vehicle data recorders (IVDR) are on-board devices that record information about the movement, control and performance of the vehicle (Correia et al. 2001). A number of IVDR systems have been developed in recent years. While their details and capabilities vary, the information they commonly collect may be classified in several categories (NHTSA 2001, Chidester et al. 2001):

1. Vehicle movement, which includes the longitudinal and lateral accelerations and the speed of the vehicle.
2. Driver control, which includes variables such as the engine throttle and brake application and wheel-angle.
3. Engine parameters, such as RPM.
4. State of the vehicle safety systems, such as air bags, seat belts, ABS and traction control.
5. Vehicle location, using GPS systems.
6. Time.
7. Visual documentation both inside and outside the vehicle.

Most of the applications of these systems have centered on the car crash event itself (e.g., crash investigations, emergency response, research and development of safety devices). However, the IVDR data may also be used in other avenues, such as prevention and training. The IVDR system described in this paper is specifically designed to collect driving behavior data that may be used to monitor and provide feedback to drivers for purposes of education and training. This direction has been adopted in several on-going recent studies, including the DriveAtlanta experiment (Georgia Tech, 2002) and the TripSense program (TripSense, 2005), which used IVDR data to determine insurance rates for participating vehicles. NHTSA (Neale et al., 2002) has recently conducted an ambitious study in which 100 vehicles were instrumented with IVDR as well as video cameras, radar sensors, GPS and lane trackers for a period of 13 months. Preliminary analysis of the huge data set collected in this study indicates great potential to enrich traffic safety research.

The limited empirical evidence reported in the literature indicates that installation of IVDR systems and the fact that drivers know their behavior on the road is monitored and documented affects driver behavior and safety. For example, Lehmann (1996) reports several case studies in which the installation of IVDR systems in various fleets resulted in reductions of 20-30% in crash rates, and even more significant reductions in the related costs. Similar reduction rates were reported in an experiment reported by Wouters and Bos (1997). While these results are promising, we are not aware of any studies that explain the causes of the safety improvements and therefore how they can be reproduced. For example, it is not clear to what extent these benefits are transferable to private vehicles, where the monitoring itself may not be an important deterrent of unsafe behavior. It is also important to investigate whether or not the safety effects stem from changes in drivers' perceptions and attitudes that

would affect driving behavior in the long-run and carry over to trips driven in vehicles that are not equipped with IVDR systems.

In this paper we describe a specialized IVDR called DriveDiagnostics. This system has been designed for the purpose of monitoring and analyzing driver behavior in normal driving situations and not only crash events. The rest of the paper is organized as follows: we first describe the overall framework and components of the IVDR system, the data it collects and analyzes and the information provided to users. Next, we describe an experiment designed to evaluate the relevance of the statistics calculated by the system to describe drivers' behavior and its impact on safety, and to evaluate the potential impact of the installation of the system and of the feedback it provides to drivers on their behavior. Within this experiment we use historic crash records to establish the connection between the data collected by the system and the actual involvement in car crashes at the level of the individual driver. Finally, we present on-going and potential applications of the IVDR data in driving behavior and safety research.

THE DRIVEDIAGNOSTICS SYSTEM

The overall framework of the DriveDiagnostics system is shown in Figure 1. The system incorporates four layers of data collection and analysis: measurement, identification, analysis and reporting.

The first layer in the system is the measurement module, which collects the two-dimensional acceleration and speed of the vehicle at a sampling rate of 40 measurements per second. The system also records the position of the vehicle using GPS. This raw information is analyzed in two information processing layers. The first is a detection and evaluation layer, which incorporates pattern recognition algorithms to identify and classify over 20 different maneuver types in the raw measurements. Examples of these maneuvers include lane changes with and without acceleration, sudden brakes, strong accelerations, excessive speed and so on. The quality of performance of the detected maneuvers is also evaluated. This evaluation is based on parameters of the detailed trajectory of the vehicle during the maneuver, such as its duration and smoothness and extent of sudden changes in the vehicle movement, and on the speed it is performed at. Unlike other similar systems, information transmission is done in real-time, continuously throughout the trip, and not only when a crash event occurs. The various information elements are transmitted through wireless networks to an application server, which maintains a database with vehicle-specific and driver-specific trip history and other relevant information, such as crash records, maintenance and fuel costs etc. The next layer, which resides in the application server synthesizes the specific maneuvers that were identified to evaluate an overall driving risk index at the level of the individual trip and of the vehicle overall performance, to characterize and classify the driver's profile and estimate the associated costs. In the current implementation drivers are classified in three categories (cautious, moderate and aggressive) based on the rate and severity of maneuvers they generate and on their speed profile.

The final layer is a reporting layer that provides feedback based on the information collected in the database. This may be done both off-line and in real-time. In an off-line application, various reports that summarize and compare information at the level of the driver, vehicle or an entire fleet are produced and viewed as printed reports or through a dedicated website. An example of a monthly driver report is shown in Figure 2. Each square in the summary chart corresponds to a trip, from engine start to turn off. The X axis indicates the day of the month and the Y axis indicates the number of trips performed during each day. Trips are color-coded

by their classification: green, yellow and red for trips in which the driver behavior was classified as cautious, moderate and aggressive, respectively. Real-time feedback, which typically includes warnings on aggressive behavior or on significant deviations from the normal driving patterns for the specific driver, can currently be provided in two ways: as a text message sent to the driver or to others (e.g. fleet managers, parents of a young driver) or using an in-vehicle display unit.

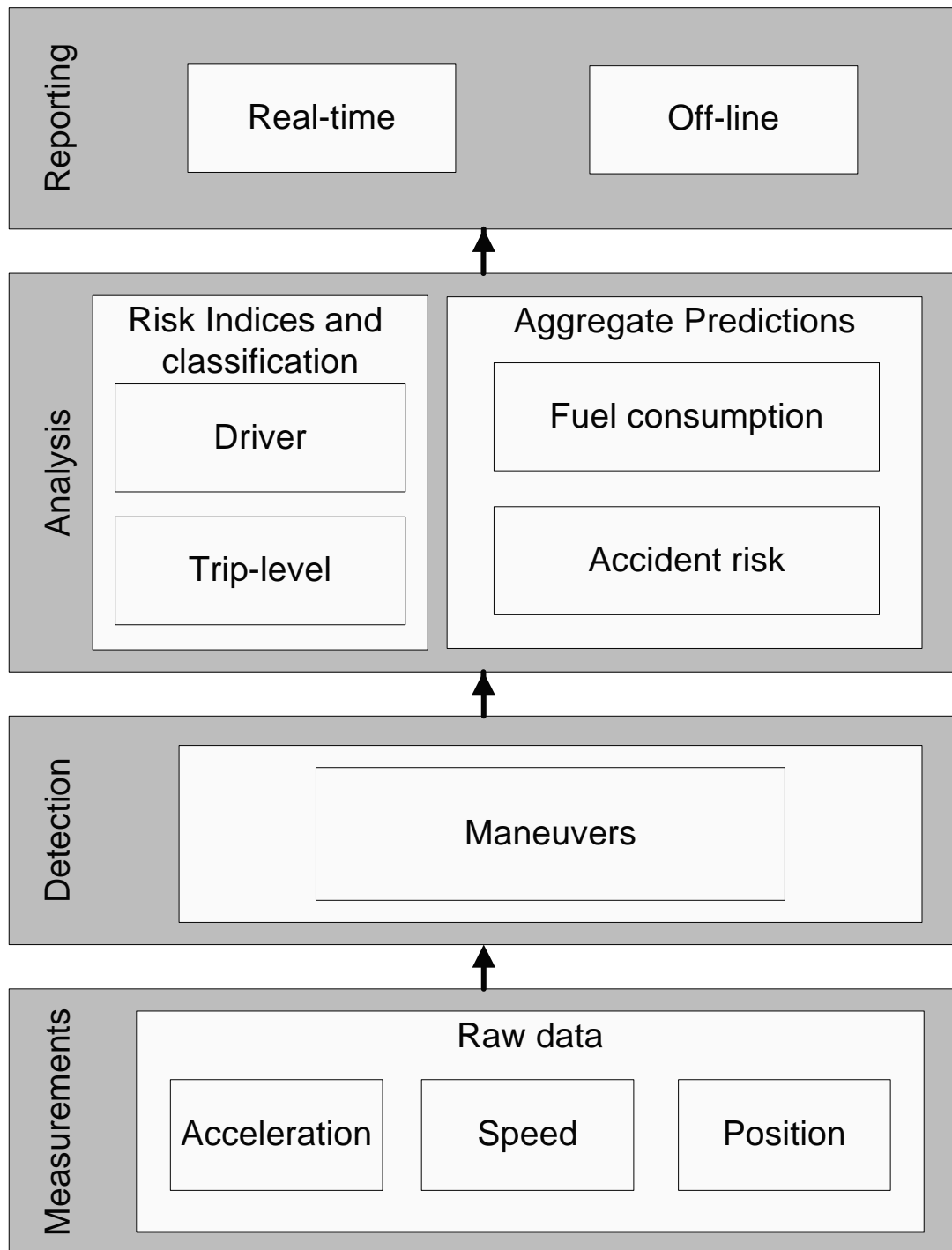


Figure 1 Overall framework of the DriveDiagnostics system

DriveDiagnostics Report generator

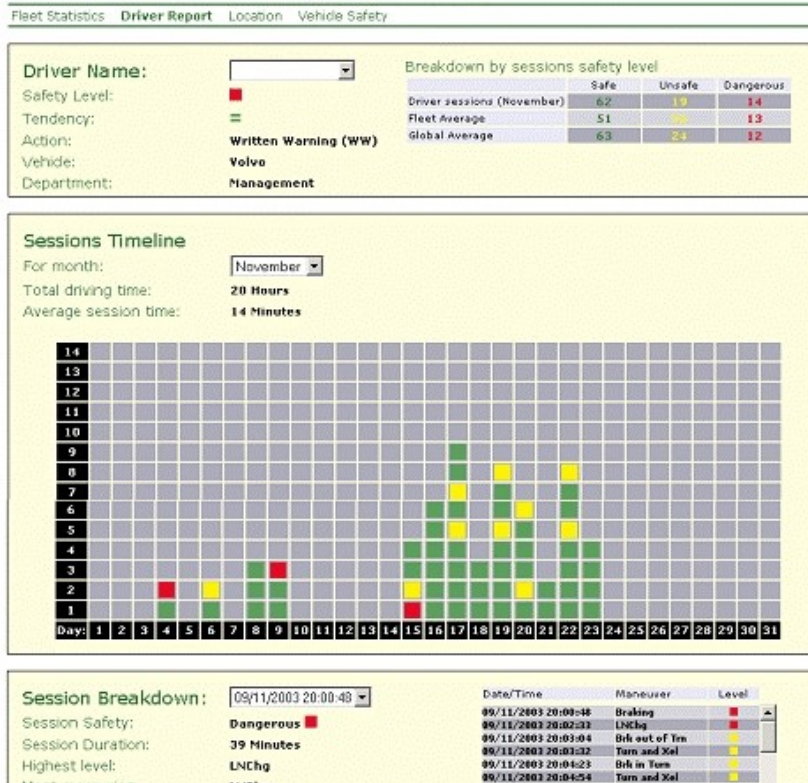


Figure 2 An example of a monthly driver report

The dimension of the sensor unit itself is about 11x6x3 centimeters. It is typically installed under the plastic panel underneath the handbrake, or in another hidden, flat location inside the vehicle. It requires a small amount of power (<250mA), and so is wired to the car battery.

The DriveDiagnostics system has so far been installed in almost 100 vehicles in a series of pilots validating its measurements and algorithms. Approximately 15000 trips have been analyzed so far. Preliminary results show very promising potential for the technology to positively affect the behavior of drivers. We next report findings of these pilots related to the validation of the system through the connection between the statistics collected and analyzed by the system and traffic safety, and to the potential impact of the installation and the feedback from the system on drivers' behavior.

VALIDATION STUDY

Experiment setup

In order to evaluate the usefulness of the information provided by the system and its impact on drivers' behavior, an experiment involving 33 drivers was conducted. All the drivers that participated in this experiment were employed by two companies, which provide their employees with company-owned medium-sized family cars as part of their employment benefits. This practice is quite common in Israel. The vehicles of these drivers were instrumented with the DriveDiagnostics system. The experiment itself included two major stages:

1. **Blind-profiling stage:** in the initial stage, the vehicles were instrumented. Privacy protection laws dictated that the drivers had to be informed about the installation of the systems. However they received no explanation about the nature of these devices and their purpose, or any feedback from them. It was therefore expected that during this blind profiling period, the installation had minimal effect on their behavior. This stage typically lasted for one or two months.
2. **Feedback stage:** at the end of the initial stage, the drivers were invited to a group meeting with the company's safety officer. In this meeting they learned about the true character of the system. In addition, personal meetings were held with all drivers. In these meetings the information about their driving behavior was discussed. Following these meetings the drivers received access codes to the personal web page, which present the information recorded relating to all the trips they have made (as shown in Figure 2). Drivers could access the information about their own trips only, but also received information about the fleet averages in order to put their own figures in context. These web pages are continuously updated in real-time with new information as new trips are made.

In addition to the data collected by the DriveDiagnostics system, two additional items of information were collected:

1. Historic crash data for the drivers that participated in the experiment. The data was obtained from the records of the two companies. The data included the numbers of crashes and crashes at fault, and the associated repair costs for each driver in the last five years. It is important to note that the companies are responsible for all expenses related to maintenance and service of all the vehicles that participated in the experiment. Drivers do not contribute towards these expenses even in cases of crashes at fault. Thus, they do not have any incentive not to report car crashes.
2. Records of all the log-ins made by all drivers to their personal web pages were collected from the server managing the drivers' web pages,

We next use these data to establish the connection between the information obtained by the IVDR system and drivers' risk of involvement in car crashes and to evaluate the potential of the installation of the system and the feedback it provides to affect drivers' behavior.

Connection between driving profiles and crash rates

The classification of trips and drivers as cautious, moderate and aggressive based on the maneuvers they make and the way they make them may be intuitive, but it must be shown that the measurements and the algorithms applied in the analysis can indeed be used as indicators for the risk of involvement in car crashes at the level of the individual driver. The crash data was available for 30 of the drivers that participated in the experiment. The data included records of 57 car crashes, with average repair costs of about 2000 New Israeli Sheqels (4.5 NIS \approx \$1, approximately \$450) per crash.

The data collected during the blind profiling, the initial stage before drivers received any feedback from the system, was used to characterize the habitual driving behavior of these drivers and to study the connection with their crash records using regression analysis. The explanatory variable used in these regression models is the risk index the system calculates for each driver. This risk index is the basis for the classification of drivers as cautious, moderate or aggressive. It depends on the quantities, types and severity of the maneuvers performed by the driver. These indices are typically in the range of 0 to 10 (with 10 being the most aggressive). The average and standard deviation for the 30 drivers in this experiment

were 3.03 and 2.41, respectively. Several functional forms were tested for the regression models. The functional form that best fit the data was:

$$y_i = \beta_0 + \beta_1 e^{x_i} + \varepsilon_i \quad (1)$$

where y_i is the car crash statistic for driver i . x_i is the risk index assigned to that driver, β_0 and β_1 are parameters, and ε_i is an error term.

Regression results showing the connection between the driving risk indices and the various car crash rates and costs are presented in Table 1. The fit of the various models, shown by the R^2 statistics are reasonable. The correlations between drivers' risk indices and the crash involvement data, $r(e^{x_i}, y_i)$, are in the range of 0.632 to 0.873. Furthermore, in all cases, the t-statistics (shown in parentheses) of all coefficients are highly significant. Thus, strengthening the conclusion that the drivers' risk indices computed by the DriveDiagnostics system can be used as indicators to the risk of involvement in car crashes. The identity of the company in which the driver is employed did not have a significant impact on the regression results.

Table 1 – Regression results linking driving risk indices to crash rates and costs

y_i	R^2	$r(e^{x_i}, y_i)$	β_0	β_1
Number of crashes per year	0.460	0.678	0.424 (4.7)	$1.551 \cdot 10^{-4}$ (4.9)
Number of crashes at fault per year	0.763	0.873	0.131 (3.1)	$1.401 \cdot 10^{-4}$ (9.5)
Cost of crashes per year (NIS)	0.524	0.724	531.0 (2.8)	0.368 (5.6)
Cost of crashes at fault per year (NIS)	0.400	0.632	297.0 (1.7)	0.268 (4.3)

4.5 NIS \approx \$1

Feedback usage

IVDR systems can affect drivers' behavior in two ways. First, the instrumentation of the vehicles and the knowledge that their actions are being monitored can by itself be a moderating factor. Second, the feedback drivers receive about their behavior may enable them to improve their performance. In this experiment, the feedback drivers receive includes not only the records of their own behavior, but also a comparison to the performance of the entire fleet. The information and feedback generated by the IVDR system is provided to drivers and to the companies' safety officers only through the dedicated web server. Therefore, the number of times drivers access the feedback on the website is a useful indication to the level of interest and usage of the information.

The average number of times drivers accessed the web page for each month after they were first introduced to it is shown in Figure 3. In the first month the system drew considerable attention with an average of 14.78 log-ins per driver. However, in subsequent months, interest in the feedback was continuously reduced to a level of 2.33 log-ins in the 5th month. It should be noted that in the experiment, there were no follow-up activities beyond the initial meetings in which the system was introduced. The results suggest that it is not enough to simply provide the information, and that routine follow-up activities may be necessary in order to maintain a high level of interest in the feedback.

We also examined the question whether the habitual driving profiles that were captured during the blind profiling stage are useful in explaining the frequency of access to the feedback. However, the correlation between the blind profiling driving risk indices and the number of log-ins was very low (0.16 for the first month log-ins, and even lower if additional months are considered).

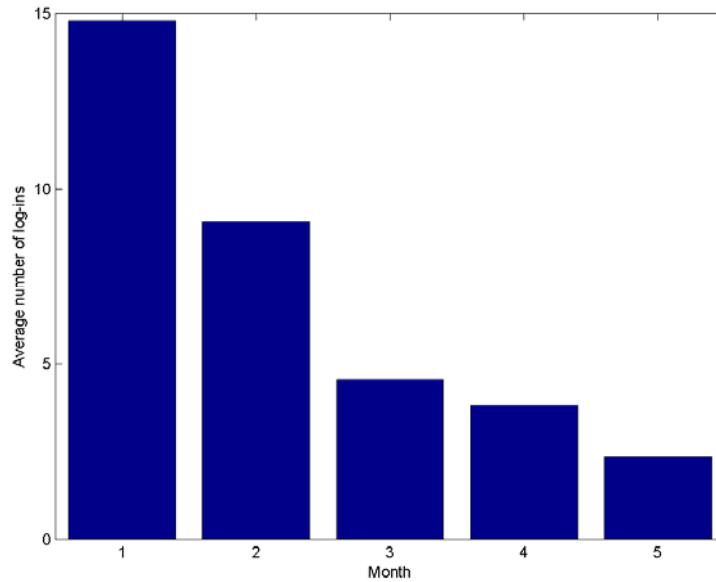


Figure 3 Average number of log-ins as a function of time

Impact of the feedback on drivers' behavior

The ultimate goal of the IVDR system is to positively affect driving behavior. In order to evaluate the impact the system has on driving behavior, it is useful to investigate how drivers' performance changes in the presence of the system. The results presented in this section are based on the records of 27 drivers for whom the data included records of at least four months of exposure to the feedback. Figure 4 shows the average driving risk indices for the months before and after drivers were informed about the system. The results indicate that the initial exposure of drivers to the system and the feedback it provides has a significant impact on driving behavior. The average driving risk indices dropped from 2.50 prior to the exposure to the system to 1.55 in the first month that feedback was provided. This moderating effect remained roughly constant for three months. However, similar to the case of access to the feedback, the impact of the system on driving risk indices was diminished in the next months. By the 5th month driving risk indices were back to the initial values and even slightly higher (average of 2.72). This result again suggests that while the initial impact of the system can be significant, it decreases over time without routine follow-up or maintenance efforts.

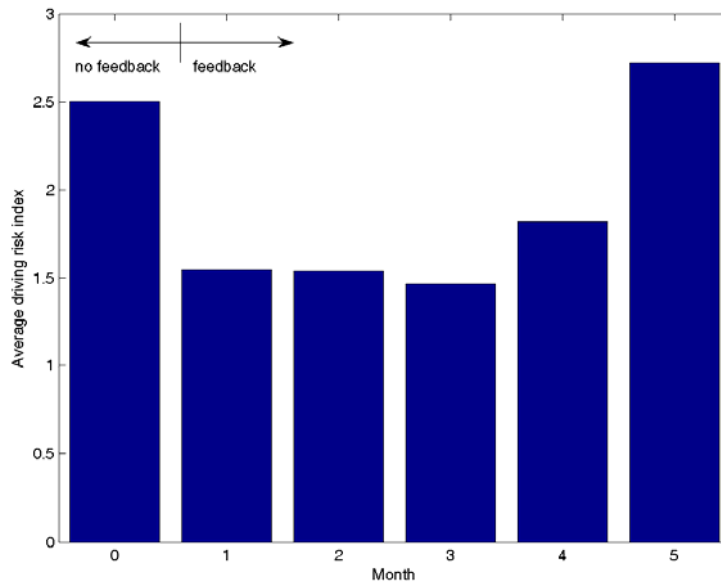


Figure 4 Average driving risk indices as a function of time

The potential of the system to change drivers' behavior in the long-term is through the feedback it provides. We next develop a model to examine the impact of the initial exposure of drivers' to the system and the extent of their usage of the feedback they receive (as measured by the number of times they access the web page). The data used for estimation included 123 observations of the 27 drivers who have used the system for four or five months after the initial blind profiling stage. The data includes one observation for each driver for every month. In order to account for the correlations among the observations of the same driver due to the drivers' unobserved characteristics, we used a fixed effects specification (see for example, Pindyck and Rubinfeld, 1997). This specification is given by:

$$y_{it} = \beta X_{it} + \gamma_i W_i + \varepsilon_{it} \quad (2)$$

where y_{it} is the risk index for driver i in month t . X_{it} are vectors of explanatory variables. β are the corresponding parameters. ε_{it} is a generic error term. γ_i are the parameters of the individual-specific effects, W_i , which are defined by:

$$W_i = \begin{cases} 1 & \text{for individual } i \\ 0 & \text{otherwise} \end{cases}$$

Estimation results for this model are presented in Table 2. The table does not show the values of the coefficients of the individual-specific effects (25 coefficients) and the model constant. These value are omitted because they depend on the alternative that is chosen as the base. The variable $risk_index(0)$, which captures drivers' risk indices in the blind profiling stage also depends on this choice. However it is presented in order to provide the full specification of the model. $\Delta risk_index(0, t-1)$ is the difference between the initial risk index for the

driver and the risk index in the previous month for each observation. *log ins* is the number of times the driver accessed the feedback in the month. The fixed effects model was superior to a pooled model that ignores the panel nature of the data. The F-Statistics for the test of the null hypothesis that all individual-specific effects are jointly equal to zero is 1.67 with 25 and 92 degrees of freedom. Thus, the hypothesis is rejected at the 5% level.

Table 2 – Regression results for monthly driving risk indices

x	β	t-statistic
$risk_index(0)$	1.156	2.8
$\Delta risk_index(0,t-1)$	-0.317	-3.2
$log\ ins$	-0.069	-4.1
$(log\ ins)^2$	0.00062	2.6

The initial risk indices that were recorded for the various drivers represent their habitual driving. These variables have a significant positive impact on the risk indices in subsequent months. Coupled with the individual-specific constants, these variables capture differences in the behavior of different drivers due to differences in personal characteristics. The coefficient of this variable is roughly a unit, which indicates these risk indices can be viewed as a basis that risk indices in subsequent months deviate from. $\Delta risk_index(0,t-1)$ captures the deviation of the risk index in the previous month from the habitual driving profile. Positive values of this variable are obtained when the risk index in the previous month was lower compared to the initial risk index. The results show that lower than habitual risk indices in a given month indicate lower risk indices in the next month as well. Thus, suggesting that the risk indices of a given driver are correlated over time.

The results show that the temporal variability in the risk indices of a given driver over time can be explained by the access to the feedback from the system, as measured by the number of log-ins to the web site. Higher levels of access to the feedback are related to lower drivers' risk indices, which imply safer driving. Thus, suggesting that the feedback the system provides can be useful in moderating driving behavior. Figure 5 and Figure 6 further illustrate the connection between access to the feedback and driving risk indices. Figure 5 shows the marginal impact of the access to the feedback on risk indices. This impact is negative, which implies that risk indices decrease with every additional access to the feedback. This negative impact occurs at a diminishing rate, i.e. the marginal impact on driving risk indices of additional log-ins to the web site is lower for drivers that access the feedback more frequently compared to drivers that make infrequent visits to the web site. Figure 6 shows the log-ins elasticity of risk indices predicted by the model for a base risk index of 2.5. The elasticity captures the ratio of the rate of change in driving risk indices to the rate of change in the number of log-ins. It is negative, which again reflects the negative correlation between the number of log-ins and the driving risk indices. The value of the elasticity increases in absolute value as the numbers of log-ins increase, but at a diminishing rate because of the diminishing marginal impact of log-ins.

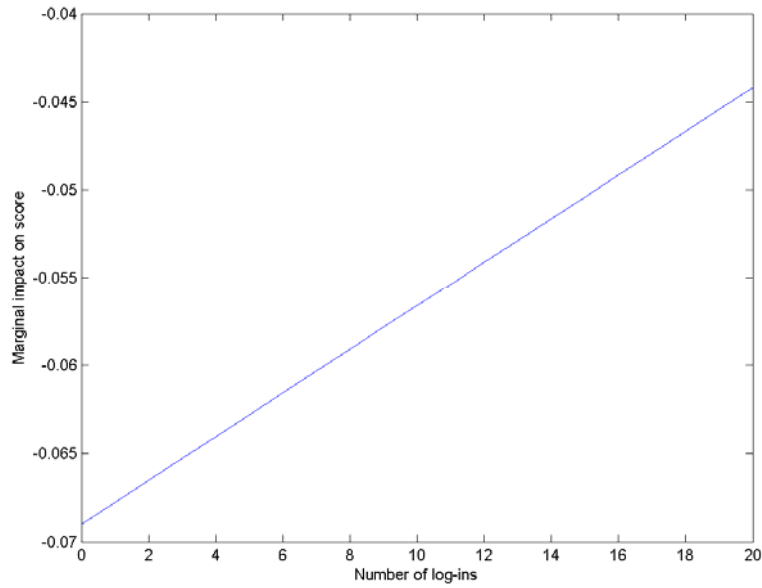


Figure 5 Marginal impact of log-ins on the risk index

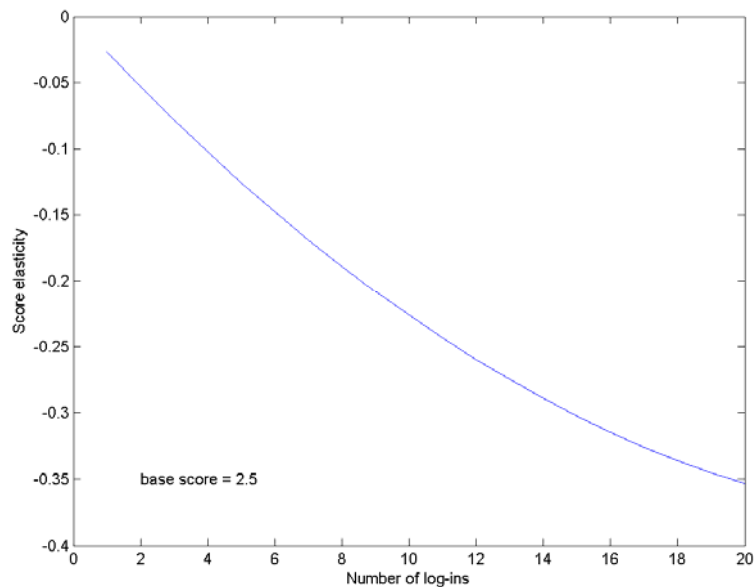


Figure 6 risk index - log-ins elasticity

CONCLUSION

This paper describes the overall framework and the components of an IVDR system called DriveDiagnostics, and presents results from a study to validate its performance and algorithms. This system has been designed for the purpose of monitoring and analyzing drivers' behavior in normal driving situations and not only crash or pre-crash events. The system records the movement of the vehicle and uses this information to identify and classify over 20 different maneuver types. These maneuvers are then used to calculate an overall driving risk index at the level of a single trip and for individual drivers.

For the validation, we used data collected by the system in the blind profiling stage, before drivers were exposed to the system to investigate the connection between drivers' profiles as captured by the system, and historic crash data. The results show significant correlations between the two data sets. Thus, suggesting that the driving risk indices calculated by the system can be used as indicators to the risk of involvement in car crashes at the level of the individual driver. The connection between driving risk indices and crash rates and costs allowed us to investigate the potential impact of the system on driving behavior and on safety. The results show that the initial exposure of drivers to the system has a significant positive impact on their behavior and on safety. Furthermore, access to the feedback provided by the system can further impact drivers performance in the desired direction. However, if follow-up efforts are not made both these positive impact are not sustained over time. In this experiment, the initial impact of the system diminished with time and disappeared within five months. Similarly, drivers initially made extensive use of the feedback from the system, but they accessed it less and less frequently as time passed.

An IVDR system that can monitor drivers' behavior and produce statistics that indicate on safety performance may be a useful tool in many studies related to driving behavior and safety. The DriveDiagnostics system is currently used in several research studies and is planned to be used in others. Examples of these studies include:

1. A study of the driving behavior of novice young drivers and their families during the period of accompanied driving, which is mandated for young drivers in Israel immediately after licensure. This study aims to evaluate the effectiveness of a program designed to increase awareness and promote the accompanied driving practice. The study looks at the impact of the extent of accompanied driving on the performance of the young drivers and the other members of the family as well as at issues of inter-generation transfer of behaviors. This study is described in further detail in Lotan and Toledo (2005).
2. A study of differences between the behaviors of professional and non-professional drivers. The purpose of this study is to identify problem areas and training needs for these groups so that better programs can be designed.
3. The behaviors of drivers from several fleets are compared in order to investigate the impact of the safety policies and practices of the various companies on their performance.
4. Drivers that use multiple vehicles will be studied to learn about the impact of the vehicle type and of the circumstances of the various trips on their behaviors.

ACKNOWLEDGEMENT

The authors thank the anonymous referees for their comments that helped improve the quality of this paper.

REFERENCES

- Blincoe L., Seay A., Zaloshnja E., Miller T., Romano E., Lucher S. and Spicer R. (2002) Economic cost of motor vehicle crashes 2000, NHTSA Technical Report DOT-HS-809-446, Washington DC.
- Chidester A.C.D, Hinch J. and Roston T.A. (2001) Real world experiences with event data recorders, Proceedings of the 17th International Technical Conference on the Enhanced Safety of Vehicles (ESV), paper 247.

- Correia J.T., Iliadis K.A., McCarron E.S. and Smolej M.A. (2001) Utilizing data from automotive event data recorders, Proceedings of the Canadian Multidisciplinary Road Safety Conference XII, London Ontario.
- Georgia Tech (2002) Drive Atlanta: study will monitor driver behavior to determine role of speed in highway crashes, Georgia Tech Research News, Web-page: <http://www.gtresearchnews.gatech.edu/newsrelease/DRIVEATL.htm>.
- Lehmann G. (1996) The Features of the accident data recorder and its contribution to road safety, Proceedings of the 15th International Technical Conference on Enhanced Safety of Vehicles (ESV), Volume 2, pp. 1565-1568.
- Lotan T. and Toledo T. (2005) Evaluating the safety implications and benefits of an in-vehicle data recorder to young drivers, Proceedings of the 3rd International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design, Rockport, ME, pp. 440-447.
- Neale V.L., Klauer S.G., Knipling R.R., Dingus T.A., Holbrook G.T. and Petersen A. (2002) The 100 car naturalistic driving study phase I: experimental design, Report DOT-HS-808-536, Department of Transportation, Washington DC.
- NHTSA (2001) Event data recorders: summary of findings by the NHTSA EDR working group, Report NHTSA-99-5218, Department of Transportation, Washington DC.
- Pindyck R.S. and Rubinfeld D.L. (1997), Econometric models and economic forecasts, 4th edition. Irwin McGraw-Hill, Boston MA.
- TripSense (2005) Web-Page: <https://tripsense.progressive.com/home.aspx>, Progressive Insurance Company.
- Wouters P.I.J. and Bos J.M.J. (1997) The impact of driving monitoring with vehicle data recorders on accident occurrence: methodology and results of a field trial in Belgium and the Netherlands. SWOV, Institute for Road Safety Research, Report R-97-8.