Department of Transportation Office of the Chief Information Officer

Supporting Statement
Commercial Driver Individual Differences Study

Part B. Collections of Information Employing Statistical Methods

1. DESCRIBE POTENTIAL RESPONDENT UNIVERSE AND ANY SAMPLING SELECTION METHOD TO BE USED

The potential respondent universe will be interstate truck drivers with a Class A commercial driver's license. There are approximately 1.8 million of these drivers in the United States (American Trucking Associations, 2010) and the proposed project will sample 21,000 of these drivers. Table 1 reports basic demographic characteristics for drivers at the primary research site and a national sample of truck driver demographics reported in the 2009-2010 American Trucking Trends (American Trucking Associations, 2010). As shown in Table 1, the mean age, age distribution, gender, and percent of driver who are self-employed is very similar to the national population of truck drivers. Although table 1 suggests the study sample is likely to match the more general truck driver population along some important demographic characteristics, this study still makes use of a purposively selected sample. Thus, given the sampling approach, the study will not make inferences about the incidence of these factors in the general truck driver population. The intent of the study is to identify relationships between various personal factors and medical conditions and crash risk rather than trying to make population estimates for the truck driver populations.

Table 1. Demographic Characteristics of Drivers Attending J.B. Hunt's Driver Orientation Compared to the National Sample

Group	Mean Age	Age Distribution	Gender	Self-Employed
National Sample	43.1 years old	21 to 29: 10.1% 30 to 44: 41.8% 45 to 64: 44.5% > 65: 3.7%	4.6 % Female 95.4% Male	21% of drivers
Primary Research Site	44.3 years old	21 to 29: 7.2% 30 to 44: 41.5% 45 to 64: 48.8% >65: 2.5%	3.9% Female 96.1% Male	19.9% of drivers

The sampling methodology will include a prospective cohort approach and retrospective case-control approach to track driver crash/violation risk outputs and assess risk factors. In the prospective cohort approach, the 20,000 participants would complete the study instruments and then be followed for up to three years to track their driving records. In this retrospective case-control approach, all drivers in participating fleets will be monitored, and once a preventable

crash occurs, the driver (if not one of the targeted 20,000 driver participants) will be asked to participate in the study. The sample obtained will be a sample of convenience as participation will be voluntary.

A sample of 21,000 drivers is needed to obtain the minimum number of cases and controls (i.e., 3,000 of each). Table 2 illustrates the annual crash numbers in each severity level based on carrier-collected data from 21,000 truck drivers in the Hickman et al. (2009) study. As shown in Table 2, there should be a sufficient frequency of DOT recordable crashes over the three-year data collection period in the proposed project to obtain at least 3,000 cases (4,467 preventable and non-preventable DOT recordable crashes over three years). A DOT recordable crash is operationally defined as an occurrence involving a commercial motor vehicle on a public road in intrastate or interstate commerce, which results in: (i) a fatality; (ii) injury to a person requiring immediate treatment away from the scene of the accident; or (iii) disabling damage to a vehicle, requiring it to be towed.

Table 2. Annual Crash Severity Frequencies from Approximately 21,000 CMV drivers in the Onboard Safety Systems Effectiveness Evaluation study (Hickman et al., 2009)

Crash Severity	Frequency
Preventable DOT crashes	514
Non-preventable DOT crashes	975
Preventable Non-DOT crashes	6,627
Non-preventable Non-DOT crashes	7,034

2. DESCRIBE PROCEDURES FOR COLLECTING INFORMATION, INCLUDING STATISTICAL METHODOLOGY FOR STRATIFICATION AND SAMPLE SELECTION, ESTIMATION PROCEDURES, DEGREE OF ACCURACY NEEDED, AND LESS THAN ANNUAL PERIODIC DATA CYCLES.

The procedures for collecting information from participants will be divided into three phases:

- In **Phase 1**, the research team will survey 20,000 drivers. These drivers will be recruited using a prospective approach during their driving orientation at participating carriers. They will complete the Driver Survey packet and asked to make their 649-F medical forms available to researchers. Drivers participating in this study will be monitored for involvement in a DOT reportable crash for up to 36 months. The research team will collect additional data from the carriers employing the participating drivers.
- In **Phase 2,** the research team will contact participants involved in a DOT reportable crash and request that they complete a Follow-Up Survey packet that includes the Recent Life Experiences Measure, Job Descriptive Index, Epworth Sleepiness Scale, and Berlin Questionnaire. At least an additional 1,000 drivers who are involved in DOT reportable crashes that occurred in the 36-month monitoring period will be recruited into the study with a retrospective approach. They will complete both sets of surveys (Driver Survey packet and only the Job Descriptive Index from the Follow-Up questionnaire packet) and their 649-F medical forms and driving records will be made available to researchers.

- In **Phase 3**, matched controls for participants involved in DOT reportable crashes will be identified from the original prospective approach and will complete the Follow-Up Survey packet.

The study's sampling methodology is based on a pre-cursor study to the CDIDS regarding participant recruitment, titled *Commercial Motor Vehicle Driver Risk Factor Study* (IRB# 06-389). Researchers found that the *only* feasible way the proposed study could be accomplished was by recruiting CMV driver participants at the medical certification stage (i.e., recruiting drivers during completion of their Medical Examination Report for Commercial Driver Fitness Determination, Form 649-F) and by working directly with large fleets.

Phase 1: Driver Survey

After establishment of the participating carriers, VTTI researchers will train fleet staff on IRB human subject protection. After fleet staff is fully IRB trained, they will introduce the CDIDS study to potential participants using a script provided by the VTTI research team. Fleet staff will introduce and recruit drivers during driver orientation according to the recruitment protocol and training that they received from the VTTI research team. Drivers that express interest in the study will obtain a study packet of materials which will include an explanation of the purpose of the study and the requirements (Appendix C), as well as an Informed Consent Form (ICF (Appendix D) and a Driver Survey (Appendix F). Drivers will be instructed to review and complete the materials on their own time, away from the fleet terminal, and mail it to VTTI in the provided sealed, tamper-proof envelope or complete the online version.

The 649-F medical form (Appendix G) will also be made available to researchers once a driver agrees to participate in the study. Drivers participating in this study will be monitored for involvement in a DOT reportable crash for up to 36 months after entering the study. Twenty thousand of the at least 21,000 drivers will be recruited and participate in this fashion.

Fleet managers will also participate by completing a survey regarding fleet characteristics (e.g., haul type) and protocols (e.g., pay, training; Appendix L).

Phase 2: DOT Reportable Crashes

Fleet staff participating in the CDIDS study will send VTTI researchers a monthly file of all DOT reportable crashes which occur in their fleet. If VTTI researchers find that a participant in the study is involved in a crash, they will contact the participant following a phone script (Appendix I) and request that they complete a Follow-Up Survey packet (Appendix J). At least 2,000 drivers are expected to complete the Follow-Up Survey packet due to a DOT reportable crash. Due to the time lag in receiving crash reports from the participating carriers, there could be up to a month between a crash and a driver receiving a follow-up questionnaire (though it is likely that most driver will be contacted in a week or two after the crash).

If the drivers involved in the crashes are not already participants in the study, VTTI researchers will contact them following a phone script (Appendix H), describe the study, and request participation. If the driver agrees to participate, VTTI will send them a set of surveys to complete

and return by mail, or give them access to a secure online version. At least 1,000 drivers who are involved in a DOT reportable crash and are not already participants of the study will be recruited into the study after the crash by using this retrospective approach. They will complete both sets of surveys and their 649-F medical form and driving record will be made available to researchers.

Phase 3: Matched Controls

Three years after the commencement of recruiting driver participants, VTTI researchers will distribute the Follow-Up Survey to matched controls for those participants who were involved in a DOT reportable crash. These controls will be participants who previously completed the Driver Survey packet and were not involved in a DOT reportable crash over the course of the study.

The case-control methodology involves comparisons of predictor factors (driver characteristics) between or among groups defined by one or more outcome criteria (in this case DOT reportable crashes). Good research design requires that criterion outcome measures be defined prior to data gathering and that these criteria be explicit and reliable. Although outcome criteria must be explicitly defined prior to data collection, the methodology allows for flexibility in the various *post hoc* group comparisons among various outcomes. Thus, even though the method is called "case-control," it is really a group comparison methodology where the characteristics of different outcome groups, however defined, may be explored and compared *post hoc* to the data collection.

In this study, researchers will classify drivers as having been involved in a DOT reportable crash or not, and then compare results with controls. Matched controls will be based on age, gender, and fleet carrier. It is anticipated that at least 3,000 case control drivers will participate in this phase.

Response rates are expected to be 76%. Procedures taken to maximize response are as follows:

- Distribution of study summary. Summary will describe the procedures and goals of the study in simple terms.
- Descriptions of the survey will de-emphasize the government's role in the survey and emphasize University sponsorship and assurances of confidentiality. The rationale for this is that given the authoritative role of the government in imposing regulations on truck drivers, these drivers may feel more comfortable discussing their medical, psychological, and personal history with third-party researchers who do not have authority over them.
- Driver Survey packets will contain business cards with the name of the Principal Investigator from VTTI.
- Measures have been designed to provide the most essential and important information relating to potential risk factors while minimizing items that may be considered offensive or intrusive.
- Appropriate incentive levels will be included for each phase of the study.

Commercial truck drivers are an independent, hard-to-reach, and somewhat autonomous group. The approach of primarily recruiting drivers through their employing carriers is intended to maximize subject participation, as are other methods to be used in the study. But, given the nature of commercial drivers and their work schedules, response rates are admittedly likely to be somewhat lower than they would be for other groups.

The principal statistical methods for analyzing antecedent and outcome variable relationships are expected to include t-tests (to compare group means) and odds ratios. Odds ratios are used to approximate relative risks of crashes (or other safety outcomes) based on independent variables of interest. For example, if one has crash involvement data for subjects classified as having or not having sleep apnea, one can calculate the odds of crash involvement given sleep apnea versus the odds without sleep apnea. Odds ratios of "1" indicate the outcome is equally likely to occur given the condition. An odds ratio greater than "1" indicates the outcome is more likely to occur given the condition. Odds ratios of less than "1" indicate the outcome is less likely to occur (Pedhazur, 1997). To determine if an odds ratio is significantly different from "1," a 95 percent confidence interval can be constructed, if "1" is not contained within the confidence interval then it is implied that there is a significant difference. An advantage of odds ratios is that they provide an easily understandable quantitative measure of risk association. The study will also employ logistic regression analysis. Logistic regression is a statistical technique used to model dichotomous outcomes based on multiple predictor variables. These predictor variables may be of any type (e.g., categorical, ordinal, continuous). Logistic regression identifies predictors that affect the probability of an outcome. The principal outcome in the current study is a case event; therefore, the logistic regression will be able to predict the probability of a case event based on the variables of interest. The logistic regression framework is flexible to accommodate for high-level effects, such as the effects of fleet and regions. Therefore, logistic regression is a way of assessing the individual and joint effects of a number of risk factors and is the primary analysis tool in this study.

Since the study design incorporates a "matching" component for some variables (e.g., driver age, gender), the analysis will require a focus on the matched pair rather than the individual subject. In this type of analysis, only the pairs in which the members differ in the risk factor under study contribute to the test statistic. Hence, we would use McNemar's test to test for the statistical significance of the difference between the two proportions. The proper method for estimating the odds ratio for matched pairs is to treat each pair as a stratum (Fleiss, Levin & Paik, 2003). The analysis team will also use a conditional logistic regression technique with match variables as strata to produce appropriate measures of relative risk (e.g., PHREG in SAS for a 1:2 case-control ratio).

The proposed study will not include less than annual periodic data cycles**.**

3. DESCRIBE METHODS TO MAXIMIZE RESPONSES RATE AND TO DEAL WITH THE ISSUES OF NON-RESPONSE.

Response rates are expected to be a minimum of 76%. A previous study by Hickman and Hanowski (2011) demonstrates a 76% driver response rate for a study more intrusive than CDIDS, including video of drivers as well as questionnaires. Therefore, VTTI researchers

believe that CDIDS, which is requesting drivers to only complete questionnaires and have their medical forms released, will have at least a 76% response rate if not higher.

To maximize the response rate, participants will be compensated with a $20 debit card after completing the initial Driver Survey packet which takes an hour or less to complete. In addition, participants requested to complete the Follow-Up Survey will be compensated with another $12 debit card. In addition, the following procedures will also be taken:

- Distribution of study summary. Summary will describe the procedures and goals of the study in simple terms.
- Participants will have the option of completing the questionnaires via a paper version or through a secure internet site.
- Descriptions of the survey will de-emphasize the government's role in the survey and emphasize University sponsorship. The rationale for this is that given the authoritative role of the government in imposing regulations on truck drivers, these drivers may feel more comfortable discussing their medical, psychological, and personal history with third-party researchers who do not have authority over them.
- Driver Survey packets will contain business cards with the name of the Principal Investigator from VTTI.
- Measures have been designed to provide the most essential and important information relating to potential risk factors while minimizing items that may be considered offensive or intrusive.
- Follow-up with respondents who agree to complete the Follow-Up questionnaires by calling them a month after sending out their questionnaires if they have not yet returned them completed.
- If we find missing items in a completed survey packet, a VTTI researcher will contact the participant by phone and ask them the questions to complete the missing data.
- Use pre-paid tamper-proof envelopes to simplify mailing process for respondents.

There will be several strategies to deal with nonresponse bias. These include:

1. As the research team will have access to all 649-F medical forms from all drivers in each orientation session, the research team will have access to general demographic information (age, gender, etc.) that will allow the research team to assess the make-up of non-responders. Thus, the demographic information from non-responders can be compared to responders to assess for potential non-response bias.
2. Generalize to the respondents only. This strategy avoids making erroneous inferences about the larger population.
3. Call back nonrespondents. Nonrespondents will be contacted to assess why they did not respond to help determine the extent of response bias. Nonrespondents who cannot be contacted can continue to bias the sample estimates.
4. Compare data in hand on respondents and nonrespondents. If data (e.g., gender, age, race, is available) the composition of respondents will be compared with that of

nonrespondents to see if there are any differences. The presence of differences indicates response bias and that caution is necessary in making inferences.

4. DESCRIBE TESTS OF PROCEDURES OR METHODS TO BE UNDERTAKEN.

A test of the CDIDS research procedures will be conducted with nine individuals at a participating fleet terminal. The VTTI research team will follow the Phase 1 procedures as outlined in this statement and refine the process and instrument(s) if necessary. Doing so will assure we are minimizing burden and maximizing the utility of the information collection.

The principal statistical methods for analyzing antecedent and outcome variable relationships are expected to include t-tests (to compare group means) and odds ratios. Odds ratios are used to approximate relative risks of crashes (or other safety outcomes) based on independent variables of interest. For example, if one has crash involvement data for subjects classified as having or not having sleep apnea, one can calculate the odds of crash involvement given sleep apnea versus the odds without sleep apnea. Odds ratios of "1" indicate the outcome is equally likely to occur given the condition. An odds ratio greater than "1" indicates the outcome is more likely to occur given the condition. Odds ratios of less than "1" indicate the outcome is less likely to occur (Pedhazur, 1997). To determine if an odds ratio is significantly different from "1," a 95 percent confidence interval can be constructed, if "1" is not contained within the confidence interval then it is implied that there is a significant difference. An advantage of odds ratios is that they provide an easily understandable quantitative measure of risk association. The study will also employ logistic regression analysis. Logistic regression is a statistical technique used to model dichotomous outcomes based on multiple predictor variables. These predictor variables may be of any type (e.g., categorical, ordinal, continuous). Logistic regression identifies predictors that affect the probability of an outcome. The principal outcome in the current study is a case event; therefore, the logistic regression will be able to predict the probability of a case event based on the variables of interest. The logistic regression framework is flexible to accommodate for high-level effects, such as the effects of fleet and regions. Therefore, logistic regression is a way of assessing the individual and joint effects of a number of risk factors and is the primary analysis tool in this study.

Since the study design incorporates a "matching" component for some variables (e.g., driver age, gender), the analysis will require a focus on the matched pair rather than the individual subject. In this type of analysis, only the pairs in which the members differ in the risk factor under study contribute to the test statistic. Hence, we would use McNemar's test to test for the statistical significance of the difference between the two proportions. The proper method for estimating the odds ratio for matched pairs is to treat each pair as a stratum (Fleiss, Levin & Paik, 2003). The analysis team will also use a conditional logistic regression technique with match variables as strata to produce appropriate measures of relative risk (e.g., PHREG in SAS for a 1:2 case-control ratio).

5. PROVIDE NAME AND TELEPHONE NUMBER OF INDIVIDUALDS WHO WERE CONSULTED ON STATICTICAL ASPECTS OF THE INFORMATION COLLECTION AND WHO WILL ACTUALLY COLLECT AND/OR ANALYZE THE INFORMATION

Project Leads for this information collection request:

Theresa Hallquist
DOT Project Manager
Federal Motor Carrier Safety Administration
1200 New Jersey Avenue, SE
Washington, DC 20590-0001
(202) 366-1064
theresa.hallquist@dot.gov

Jeff Hickman
Group Leader
VTTI (0536)
3500 Transportation Research Plaza
Blacksburg, VA 24060
(540) 231-1542
jhickman@vtti.vt.edu

REFERENCES

American Trucking Associations. (2009). *American Trucking Trends 2009-2010*. American Trucking Associations: Arlington, VA.

Bureau of Labor Statistics. (2009). *Occupational Employment and Wages: 53-3032 First-Line Supervisors/Managers of Transportation and Material-Moving Machine and Vehicle Operators*. Retrieved on March 10, 2011 from: http://www.bls.gov/oes/current/oes531031.htm.

Bureau of Labor Statistics. (2009). *Occupational Employment and Wages: 53-3032 Truck Drivers, Heavy and Tractor-Trailer*. Retrieved on March 10, 2011 from: http://www.bls.gov/oes/current/oes533032.htm.

Craft, R., & Blower, D. (2004). *The Large Truck Crash Causation Study*. Paper presented and distributed at the November 17, 2004 FMCSA R&T Stakeholder Forum. Arlington, VA.

Fleiss, J.L., Levin, B. & Paik, M. C. (2003). *Statistical Methods for Rates and Proportions (3rd Ed)*. New York, NY: John Wiley & Sons.

Hanowski, R. J., Blanco, M., Nakata, A., Hickman, J. S., Schaudt, W. A., Fumero, M. C., Olson, R. L., Jermeland, J., Greening, M., Holbrook, G. T., Knipling, R. R.., & Madison, P. (2005). *The drowsy driver warning system field operational test, data collection final report*. Contract No. DTNH22-00-C-07007, Task Order 14. Blacksburg, VA: Virginia Tech Transportation Institute.

Hickman, J. S. & Hanowski, R. J. (2011). Use of a video monitoring approach to reduce at-risk driving behaviors in commercial vehicle operations. *Transportation Research, 14*, 189-198.

Hickman, J.S., Campbell, J.L., Guo, F., Brown, J., Richard, C., Salyer, D., & Hanowski, R.J. (2009). Onboard Safety Systems Effectiveness Evaluation: Task 2.4 Final Research Design Plan (DTMC75-09-C-00022). Federal Motor Carrier Safety Administration: Washington, D.C.

Knipling, R.R. (2005a). Evidence and Dimensions of Commercial Driver Differential Crash Risk. *Proceedings of the Driving Assessment 2005 Conference*. Rockport, ME.

Knipling, R.R. (2005b). *Individual Differences in Commercial Driver Fatigue Susceptibility: Evidence and Implications*. Paper and presentation for the Fatigue Management in Transportation Operations International Conference, Seattle.

Knipling, R. R., Hickman, J. S., Hanowski, R. J. & Blanco, M. (2005). *Phase I- preliminary analysis of data collected in the drowsy driver warning system field operational test: Task 4, Phase II research plan*. Contract No. DTNH22-00-C-07007 (Task Order No. 21). Blacksburg, VA: Virginia Tech Transportation Institute.

Knipling, R.R., Boyle, L.N., Hickman, J.S., York, J. S., Daecher, C., Olsen, E.C.B., & Prailey, T. D. (2004). *Synthesis Report #4: Individual Differences and the High-Risk Commercial Driver*. Project Final Report, Transportation Research Board Commercial Truck and Bus Synthesis Program. ISSN 1544-6808, ISSN 0-309-08810-0, available at http://trb.org/news/blurb_browse.asp?id=11.

Lancaster, R., & Ward, R. (2002) *The contribution of individual factors to driving behaviour: Implications for managing work-related road safety.* Entec UK Limited, Health and Safety Executive, Research Report 020, United Kingdom.

Pedhazur, E. J. (1997). *Multiple regression in behavioral research: explanation and prediction* (3rd ed.). Forth Worth: Harcourt Brace College Publishers.

Treat, J.R., Tumbas, N. S., McDonald, S. T., Shinar, D., Hume, R. D., Mayer, R. E., Stansifer, R. L., & Catellan, N. J. (1979). *Tri-Level Study of the Causes of Traffic Crashes: Final Report Volume I: Causal Factor Tabulations and Assessments.* Institute for Research in Public Safety, Indiana University, DOT HS-805 085.

Tsuang, M. T., Boor, M., & Fleming, J. A. (1985). Psychiatric aspects of traffic accidents. *American Journal of Psychiatry, 142,* 538-546.