STATE OF CALIFORNIA

DMV

DEPARTMENT OF MOTOR VEHICLES

CALIFORNIA'S THREE-TIER DRIVING-CENTERED ASSESSMENT SYSTEM

Outcome Analysis

November 2011

EDMUND G. BROWN JR.
Governor

TRACI STEVENS, Acting Secretary
Business, Transportation and Housing Agency

GEORGE VALVERDE
Director

RSS-11-234

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From - To)*
November 2011	Final Report	

4. TITLE AND SUBTITLE
California's Three-Tier Driving-Centered Assessment System – Outcome Analysis

5a. CONTRACT NUMBER
5b. GRANT NUMBER
5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Bayliss J. Camp, Ph.D.

5d. PROJECT NUMBER
5e. TASK NUMBER
5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
California Department of Motor Vehicles
Research and Development Branch
P.O. Box 932382
Sacramento, CA 94232-3820

8. PERFORMING ORGANIZATION REPORT NUMBER
CAL-DMV-RSS-11-234

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Office of Traffic Safety
2208 Kausen Drive, Suite 300
Elk Grove, CA 95758-7115

10. SPONSOR/MONITOR'S ACRONYM(S)
OTS

11. SPONSORING/MONITORING AGENCY REPORT NUMBER

12. DISTRIBUTION AVAILABILITY STATEMENT Unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT
This Outcome Analysis constitutes the second of two reports on the 3-Tier Assessment System, as piloted by California DMV in 2006-2007. It contains (a) a projection of the costs associated with the Area Driving Performance Evaluation, (b) a determination of the willingness of a participant to pay a fee for the Area Driving Performance Evaluation, (c) a determination of the percentage of drivers who were assessed to have a limitation, but who, upon completion of the assessment, were able to retain their driving privileges, (d) the utilization of certified driving rehabilitation specialists, and (e) the results regarding crash rates and retention of driving privileges. Together, these analyses examine the effectiveness of the 3-Tier Assessment System in identifying functional impairments, reducing crashes, and extending safe driving years for California drivers of all ages. These analyses are based upon 2 years of elapsed driving history for the 12,279 customers who participated in the Pilot, along with two control groups: 14,907 customers in the Baseline II cohort, and 10,551 customers in the Nearby cohort. Based on limited data, an estimation is provided of the costs of the anticipated increase in the use of the ADPE, as associated with the 3-Tier Assessment System. Very few customers were willing to pay a fee for the ADPE. The overwhelming majority of customers, even those with functional limitations, were able to retain their driving privilege. No customers reported using certified driving rehabilitation specialists. The analyses found no evidence for a reduction in crash risk subsequent to participation in the Pilot; however, the analyses found some evidence that the Pilot is associated with an increased amount of time to complete the renewal process, with an increase in the odds of failing to renew the driving privilege, and with an increase in the odds of receiving a restricted license. Recommendations regarding implementation and future research are included.

15. SUBJECT TERMS
3-Tier, Driving Wellness, Driving Fitness, Licensing Tests, Driving Assessment System, Functionally-limited Drivers, Driver Assessment

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THS PAGE			Douglas Rickard
Unclassified	Unclassified	Unclassified	None	146	19b. TELEPONE NUMBER (*Include area code*) (916) 657-5768

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI-Std Z39-18

PREFACE

This report is published by the California Department of Motor Vehicles Research and Development Branch as required by California Vehicle Code section 1659.9, and in fulfillment of California Assembly Bill 2541 (Daucher), calling for an evaluation to determine the effectiveness of the 3-Tier Assessment System in identifying functional impairments, reducing crashes, and prolonging safe driving years of all drivers regardless of age. The opinions, findings, and conclusions expressed in this report are those of the author and not necessarily those of the State of California.

ACKNOWLEDGEMENTS

This project was made possible with the support of the California Office of Traffic Safety and the National Highway Traffic Safety Administration (OTS Grant #TR1015). The present report is a product of the California DMV Research and Development (R&D) Branch, and was originally submitted to the California Legislature pursuant to the requirements of California Vehicle Code Section 1659.9.

This study was conducted under the supervision of David DeYoung, Chief of R&D. Robert Hagge, Manager of the Driver Competency and Safety Projects Section of R&D, also provided critical feedback on the manuscript at various stages of development. Scott Masten, Manager of the Alcohol and Drugs Projects Section of R&D, provided advice at several points, including assistance in extracting data from the Statewide Integrated Traffic Records System (SWITRS) for the purposes of identifying potential history confounds, advice regarding Cox proportional hazards regression modeling, and tutoring in how to calculate attributable/prevented fraction estimates. Len Marowitz, retired Manager of the Alcohol and Drugs Projects Section, provided valuable advice and guidance in the early stages of the evaluation of 3-Tier Pilot. Mike Gebers, Research Scientist III, provided advice regarding the use of ecological exposure variables. Debra Atkinson, Staff Services Manager I (retired annuitant), provided technical knowledge regarding DMV policies and procedures. R&D is part of DMV's Licensing Operations Division (LOD), under the leadership of Shamim (Mimi) Khan, Deputy Director.

The author would like to thank Loren Staplin, Chair of the Operator Education and Regulation Committee of the Transportation Research Board, for providing the author the opportunity to make a presentation of the initial findings of this report before an invited panel of traffic safety experts at the 2011 Transportation Research Board Annual Meeting. The author would also like to thank the individual members of that panel: John J. Joyce, former chief of Driver Safety Research at the Maryland Motor Vehicle Administration; Kim Snook, Director of the Iowa Office of Driver Services; and Elin Schold Davis, Project Coordinator of the Older Driver Initiative for the American Occupational Therapy Association. These individuals, as well as many members of the audience, provided critical feedback on certain technical and methodological questions.

The author would also like to thank Felix Elwert, Assistant Professor of Sociology at the University of Wisconsin – Madison, for answering certain technical questions regarding statistical methods.

Finally, the author would like to thank Doug Rickard, for his assistance in proofing this report as it was prepared for publication.

EXECUTIVE SUMMARY

In accordance with California Vehicle Code Section 1659.9, and with funding from the California Office of Traffic Safety (OTS Grant #TR1015), this report contains an outcome evaluation of the 3-Tier Assessment System (3TAS) as piloted by the California Department of Motor Vehicles (CA DMV) in 2006-2007. Per the authorizing legislation, this report covers the following topics:

(1) A projection of the costs associated with the area driving performance evaluation.

(2) A determination of the willingness of a participant to pay a fee for the area driving performance evaluation.

(3) A determination of the percentage of drivers who were assessed to have a limitation but who, upon completion of the assessment, were able to retain their driving privileges.

(4) The utilization of certified driving rehabilitation specialists.

(5) The results regarding crash rates and retention of driving privileges.

The main thrust of the analyses contained herein examine the effectiveness of the 3-Tier Assessment System, as piloted by CA DMV in 2006-2007, in identifying functional impairments, reducing crashes, and extending safe driving years for California drivers of all ages. These analyses are based on 2 years of elapsed driving history, and incorporate comparisons with two control groups processed according to standard CA DMV procedures in place at the time of the 3-Tier Pilot.

Statement of the Problem

The demographic composition of the driving population of California is shifting. Within the next two decades, the number of senior citizens (aged 65 and older) in the state is expected to double, while their share of the population will increase from just over 11% to almost 18%. This demographic change is expected to result in a major increase in the number of senior drivers, as well as a shift in their proportional share of drivers on California's roads. These changes have a number of implications in different areas of traffic safety. For CA DMV, these implications include (a) preserving and extending driving years for individuals who can safely operate a motor vehicle, and (b) identifying appropriate procedures for determining when individuals must restrict their driving privilege, or even cease driving altogether. The 3-Tier Pilot constituted a critical test of a new method by which the CA DMV may improve traffic safety through the licensing process.

Background

The 3-Tier Assessment System, as piloted by CA DMV in 2006-2007, built upon prior analytic work at CA DMV and elsewhere. In particular, the 3-Tier Pilot grew out of research conducted at CA DMV starting in 1993 with a cooperative agreement between the department and the National Highway Traffic Safety Administration (NHTSA). This research culminated in a series of reports (Hennessy & Janke, 2005; Hennessy & Janke, 2009), which form the main basis upon which the present 3-Tier Pilot was conducted.

Description of the 3-Tier Assessment System

The 3-Tier Assessment System consisted of a tiered series of screening and assessment tools, applied to drivers of any age. These screening tools were intended to identify customers in need of further assessment of the safety of their driving, and to identify customers who might benefit from education regarding how to safely compensate for identified limitations to driving-relevant functional abilities.

The Tier 1 screening tools consisted of a brief test of memory recall, two vision tests (visual acuity and contrast sensitivity), and a checklist of observed limitations to physical function in the upper and lower body. These tests were administered by front-line staff working in field offices as part of driver license renewal application processing. The Tier 1 tests were applied to all eligible customers, regardless of age, who were enrolled in the Pilot through the license renewal process.

The Tier 2 screening tools consisted of (a) the 18-question written knowledge test for renewal of a Class C (non-commercial) license, and (b) the Perceptual Response Test (PRT), a test of processing speed of the visual system, which also screens for potential driving-relevant limitations in perception and cognition. All eligible renewal customers enrolled in the 3-Tier Pilot took the written knowledge test. Those customers who failed one or more Tier 1 tests, or who failed the written knowledge test twice, were required to take the PRT.

Tier 3 included an on-road drive test, and educational intervention. The drive test consisted of either the Supplemental Driving Performance Evaluation (SDPE) or the Area Driving Performance Evaluation (ADPE), both of which are currently part of CA DMV's standard repertoire of driver license screening instruments. The educational intervention included drive-test preparation materials and a new curriculum developed for the 3-Tier Pilot which focused on compensation techniques for driving safely despite some limitation to functionality identified at Tier 1 or Tier 2. Customers identified at Tier 1 or Tier 2 with multiple driving-relevant limitations, or with a limitation that was serious in nature, were assessed at Tier 3. Customers enrolled in the 3-Tier Pilot as a result of a Driver Safety referral were assessed at Tier 3.

Methods

Data Collection

The 3-Tier Pilot was conducted in six CA DMV field offices between June 4 and October 31, 2007. The six Pilot field offices were: Carmichael, Fairfield, Folsom, Sacramento-Broadway, Sacramento-South, and Vacaville. For a variety of reasons, it was determined to be operationally infeasible to assign customers randomly to Pilot and control groups (as would occur in a randomized experimental design). Therefore, a quasi-experimental design was originally planned, whereby a comparison group of customers was identified who otherwise met all the eligibility requirements for participating in the Pilot, but who did not undergo Pilot processing. These customers were processed at the same six Pilot field offices between September 1, 2006, and January 31, 2007 (i.e., several months prior to implementation of the Pilot). This comparison group was termed the Baseline, as opposed to the Pilot, cohort.

In the Process Analysis (Camp, 2010b), initial descriptive statistics were calculated for the Baseline and Pilot customers. These initial descriptive statistics revealed that the two groups were significantly different in important ways, including age, gender, and prior violation record. Because these differences would have created substantial bias in any analysis of safety outcomes when comparing the two groups, two additional comparison cohorts were constructed, and the Baseline customers were omitted from most of the analyses discussed in the present manuscript.

The two additional comparison groups—termed Baseline II and Nearby, respectively—were selected based on their ability to yield the least-biased comparisons to the Pilot cohort, so that evaluation of the effects of the Pilot could be as objective and rigorous as possible. For the Baseline II and Nearby cohorts, the author consulted a separate database provided by CA DMV

Audits Branch. These data are collected in a routine, automated manner (i.e., not directly by Field Office or Driver Safety staff). The data extracted from this database contain certain limited information for customers conducting Class C (non-commercial) license renewal transactions in the six Pilot field offices, and a subset of seven other offices relatively near to Sacramento: Davis, Lodi, Napa, Placerville, Rocklin, Roseville, and Woodland. The Baseline II cohort consisted of customers processing their applications in the six Pilot offices between September 1, 2005 and January 31, 2006. The Nearby cohort consisted of customers processing their applications in the seven other offices between June 4, 2007 and October 31, 2007.

The customers sampled as part of the Baseline II cohort proved to be more similar in background (e.g., age, gender, and prior violation record) to Pilot customers than were customers in the original Baseline cohort. While the Baseline II cohort contains certain built-in controls for potential geographic biases—because it uses customers patronizing the same offices as in the Pilot cohort—it is unfortunately susceptible to historical biases. In particular, there has been a significant and substantial decline in motor vehicle crashes in California (and the Sacramento metropolitan area) during the past several years, including the years in between when the Baseline II and Pilot customers renewed their licenses. Therefore, in an effort to account for these potential historical influences, the Nearby cohort was also sampled. These customers were, like the Baseline II cohort, more similar in background (age, gender, and prior violation record) to Pilot customers than were customers in the original Baseline cohort. In order to account for potential geographical biases between the Nearby and Pilot customers (since they patronized different offices), some additional variables capturing geographic information were included to partly control for these differences. The analyses of the safety and mobility effects of the Pilot that are included in the present manuscript employ both the Baseline II and the Nearby cohorts as comparison groups. In general, outcomes are based upon data for the 2 years subsequent to enrollment in the Pilot or comparison groups. There are some limitations to, and potential biases associated with, the data available for the Baseline II and Nearby cohorts; these limitations are discussed at some length in the main body of the manuscript.

Statistical Analyses

The present report uses a variety of research design and statistical methods, depending on the specific sub-analyses. Descriptive and inferential statistics were calculated using IBM SPSS Statistics (ver. 19), and SAS (ver. 9.2) statistical software. Inferential statistics include chi-square analysis for comparisons involving categorical data, analysis of variance for comparisons involving differences among group means, ordinary least squares regression for modeling differences in linear outcome data (e.g., the number of days to complete a renewal application), logistic regression analysis for modeling differences in binary outcome data (e.g., crashed vs. crash-free in the 2 years subsequent to participation), and Cox proportional hazards analysis for determining whether groups differed in the number of days to their first subsequent crash.

Results

The results are divided into thirteen questions.

Question #1: What are the Costs Associated with the ADPE?

It appears that the 3-Tier Pilot was associated with some increase in the number of ADPE tests given, relative to the Nearby and Baseline II cohorts. This increase is difficult to quantify due to lack of definitive data on ADPE tests given to drivers in the two comparison cohorts. However, any increase appears to have been quite small in terms of absolute number—between 5 and 14 tests. Based on methods similar to those used in the Process Analysis (Camp, 2010b), and extrapolating for potential statewide implementation, it is anticipated that the additional ADPEs associated with 3TAS would require between 0.38 and 1.06 full-time equivalent (FTE) Licensing Registration Examiner (LRE) positions. It is not anticipated that the costs associated with conducting the ADPE would include anything other than LRE staff time. Therefore, the costs associated with the ADPE, specifically the anticipated increase in the use of this test if the 3-Tier Assessment System were implemented statewide, would be between $19,740 and $55,068 per year, on an ongoing basis.

Question #2: How Willing are Customers to Pay an Extra Fee for the ADPE?

While 3-Tier Pilot customers were surveyed as to their willingness to pay an extra fee for the ADPE, there is a great deal of missing data on customers' responses to this question. It is unknown what most drive-test customers (regardless of drive-test type) would be willing to pay if required to pay an additional fee to take an ADPE. It is also unknown what most of those customers who were specifically scheduled for an ADPE would be willing to pay, if required to pay an additional fee. At most it can be said that among those Pilot cohort customers for whom a survey response on this question was recorded, very few were willing to pay even as much as an additional $75 for such a test.

Question #3: How Many Customers Possessed Limitations, and What Percentage of Those With Limitations Retained Their Driving Privilege?

The driving-relevant functional limitations identified through the 3-Tier Assessment System came in three forms: physical, cognitive/perceptual, and visual. Each of these was identified through the use of multiple screening tools.

Very few 3-Tier Pilot customers lost their driving privilege through a formal licensing action such as a suspension or revocation. However, drivers can also informally remove themselves from the licensing system by delaying renewal of their license, or choosing not to renew their driving privilege and letting their license application lapse entirely. These informally unlicensed customers are included in the analyses, to capture more fully any potential effects of the 3-Tier Assessment System on retention of the driving privilege. For this question, retention of the driving privilege was measured two ways: (i) the number who remained unlicensed (lapsed, suspended, or revoked) for the entire 2-year follow-up period following enrollment in the Pilot, and (ii) the mean number of days between starting a license renewal application, and the issuance of a valid license.

Separate analyses were conducted on the relationship between the level of functional limitation and retention of the driving privilege, for all individual screening tests and for the overall level of functional limitation, scored as "pass," "somewhat fail" (or SFail), and "extreme fail" (or XFail). The relationship between level of functional limitation and retention of the driving privilege was generally similar across individual screening tools. Customers screened as having a functional limitation, of whatever sort, were less likely to retain the driving privilege. For those who successfully retained the driving privilege, those screened as having a functional limitation, of whatever sort, generally took longer to complete their application processing.

This relationship between level of functional limitation and retention of the driving privilege that was observed among the specific individual screening tests also held for the overall 3-Tier Assessment System score. Among 3-Tier eligible customers coded as "pass," (8,887 customers) 98.8% retained the driving privilege. Among those coded as "SFail," (2,028 customers) 97.1% retained the driving privilege. Among those coded as "XFail," (1,241 customers) 85.7% retained the driving privilege. The mean days to licensure generally rises with the overall measure of limited functionality, from 7.9 days ("pass") to 17.7 days ("SFail") to 66.1 days ("XFail"). Thus, while there is a clear relationship between the level of limited functionality and license status, most customers ultimately retained their driving privilege.

Question #4: How Many Customers Used Certified Driving Rehabilitation Specialists?

Most 3-Tier eligible customers who were required to take a drive test reported that they had not used any type of behind-the-wheel training. The proportion who said they had used such training generally rises after each failed attempt at an SDPE. Among those customers who did report using such training, none reported using a certified driving rehabilitation specialist (CDRS) or occupational therapist (OT).

Question #5: What is the Relationship between the Screening Tests and Compensation Behaviors?

3-Tier Pilot customers were surveyed as to their driving habits. These data were then analyzed to determine their relationship to the overall level of functional limitation, as indicated by the overall 3-Tier score. The analyses were stratified by age.

Among Pilot customers aged 70+, those with a "pass" score reported driving 5.0 days per week. Those with an "SFail" score reported driving 4.8 days per week. Those with an "XFail" score reported driving 4.4 days per week. This pattern is similar to the pattern for younger drivers (<70), where those with a "pass" score reported driving 5.5 days per week, "SFails" reported driving 5.0 days per week, and "XFails" reported driving 4.6 days per week.

Among both older (70+) and younger (<70) drivers, the level of functional limitation (as indicated by the overall 3-Tier score) was associated with an increased likelihood that a driver would report "often" or "always" avoiding a risky or challenging driving situation. The association between level of functional limitation and compensation behaviors was especially strong for the following situations: avoiding driving at night, avoiding driving at sunrise or sunset, and avoiding freeways. Among older drivers, the association between overall 3-Tier score and compensation was also strong for avoiding driving along unfamiliar routes.

Question #6: What is the Inter-Correlation Among Outcomes on the Cognitive/Perceptual Screening Tests?

The 3-Tier Pilot included three tests that potentially screen for problems in cognition and perception: the memory recall test, the PRT, and the written renewal test. Analyses were conducted to determine the level of inter-correlation among outcomes on these tests. These analyses were stratified by age.

Among older customers (aged 70+), those who failed the memory recall test were significantly more likely to fail the written renewal test multiple times. Also among older drivers, those who failed the memory recall test were more likely to fail the PRT, though this association was not statistically significant. Finally, and again among older drivers, customers who failed the written renewal test multiple times were significantly more likely to fail the PRT. Among younger (<40 y.o.) and middle-age (40-69 y.o.) drivers, outcomes on any one cognitive/perceptual screening tests were not significantly associated with outcomes on any of the other cognitive/perceptual screening tests. This may in part have been due to small cell-size counts (i.e., very few people younger than 70 y.o. tended to fail the cognitive/perceptual screening tests, especially the memory recall test and the PRT). Thus, at least for older customers, the inter-correlation among the cognitive screening tests suggests that they all tap into a common domain of cognitive function.

Question #7: What is the Relationship between Screening Test Outcomes and Prior Crash Record?

The constituent elements of 3-Tier were chosen, on the basis of prior research (Hennessy and Janke, 2009) for, among other reasons, their validity for identifying drivers at risk of crashing. A series of confirmatory analyses were conducted to test the validity of the component screening tests, separately and in combination, at identifying drivers with an increased prior risk of crashing. These analyses were first stratified by age, without additional statistical controls. Logistic regressions were then constructed to model the relationship between screening test outcomes and 3-year prior crash record, incorporating controls for gender, possession of a suspension/revocation indicator, exposure (self-reported days of driving), and self-reported compensation behaviors.

THE VALIDITY OF THE INDIVIDUAL SCREENING TESTS AT PREDICTING PRIOR CRASHES

The number of observed physical limitations was not significantly associated with 3-year prior crash record for those younger than 40 y.o., or for those aged 70 and older. For customers aged 40-69, those with one observed limitation were more likely to have crashed than those with zero or two observed limitations. When statistical controls are introduced, customers with two or more observed physical limitations had substantially and significantly lower odds of having crashed than did those with no observed limitations, while those with one limitation were not significantly different from customers with no limitations.

Among 3-Tier Pilot customers, failure on the memory recall test was not significantly associated with 3-year prior crash record, for any age group. The same is true for failure on the visual acuity standard, and failure on the written renewal test. When statistical controls are introduced, customers who failed these tests were not significantly different from customers who passed these tests.

Failure on the Pelli-Robson contrast sensitivity chart was not significantly associated with 3-year prior crash record, for any age group. When statistical controls are introduced, customers who somewhat failed this test were significantly less likely to have crashed than were customers who passed this test. Customers who extreme failed were not significantly different from customers who passed this test.

Failure on the PRT was significantly associated with 3-year prior crash record for customers aged 70 and older, but not for customers younger than 70 y.o. Customers aged 70 and older who failed the PRT were somewhat more likely to have crashed than were customers of the same age who passed this test. When statistical controls are introduced, customers who failed this test had directionally greater odds of having crashed. However, this difference did not rise to conventional levels of statistical significance. This may be due in part to the fact that the test was not administered to all customers, which reduces the amount of statistical power.

THE VALIDITY OF THE SCREENING TESTS, IN COMBINATION, AT PREDICTING PRIOR CRASHES

A customer's level of limited functionality, as indicated by their overall 3-Tier score, was not significantly associated with 3-year prior crash record for any age group. When statistical controls are introduced, customers coded as "SFail" had significantly lower odds of having crashed as compared to customers coded as “pass.” Those coded as “XFail” were not significantly different from customers coded as “pass.”

Thus, whether considered separately or as a whole, the 3-Tier screening tests generally failed to consistently identify drivers with an elevated risk of having crashed in the 3 years prior to screening.

Question #8: What is the Relationship between Screening Test Outcomes and Drive Test Outcomes?

An analysis was conducted to examine the relationship between level of limited functionality, as indicated by overall 3-Tier score, and performance on an on-road test of driving skill. Three groups of customers were compared: Driver Safety referrals, limited-term license holders, and those required to take a drive test because of their overall 3-Tier score. The analysis was stratified by age.

Among drivers younger than 70 y.o, there were no statistically significant differences in fail rate on the first drive test across the three types of drive-test customers. Among drivers aged 70+, Driver Safety referrals were significantly more likely to fail the first drive test as compared to the other two types of customers. Also among those aged 70+, limited-term customers and 3-Tier XFails were not significantly different from each other in their fail rate on the first drive test.

The findings for this question are subject to severe methodological constraints. As noted in the Process Analysis (Camp, 2010b), several hundred customers coded as XFails under 3-Tier did not actually take a drive test as required under the Pilot. Many of these customers delayed renewing their licenses for some months and so avoided the drive-test requirement. It is not possible to speculate as to how these customers would have performed on the drive test had they actually taken one. The results discussed in this question therefore may have been quite different had these customers actually been tested as required under the Pilot.

Question #9: What is the Overall Program Effect of the 3-Tier Pilot on Reducing Crashes?

OVERALL PROGRAM EFFECTS ON SUBSEQUENT CRASHES

A series of analyses were conducted, comparing the 2-year subsequent crash rates of 3-Tier Pilot participants with the crash rates of customers in the Nearby and Baseline II cohorts, using the Cox proportional hazards technique. The outcome of interest consisted of a simple dummy variable (crashed versus crash-free) in the 2 years subsequent to the date of first contact. Crashes were included regardless of fault status. Each model controlled for age, sex, Driver Safety referral status, potential delays in licensure, annualized crash rate of the zip code of residence for the driver, and prior record (suspension/revocation indicators, number of convictions, and number of crashes in the 3 years before the date of first contact). Interaction terms were included to test for differences in program effects across age groups.

The comparison to the Baseline II cohort found that 3-Tier Pilot customers had substantially and significantly lower odds of crashing in the 2 years following enrollment. However, the crash rate among California drivers and drivers in the Sacramento area also declined between the period when the Baseline II cohort was enrolled (September 2005 through January 2006) and the period when the Pilot cohort was enrolled (June 2007 through October 2007). The apparent difference between the two cohorts in crash risk found in the statistical modeling was of approximately the same order as the secular decline in crash risk observed for other drivers between these two periods. There were no significant differences associated with the interaction terms of age and cohort status in this analysis.

The comparison of Pilot cohort participants to the Nearby cohort found that there was no significant difference in the odds of crashing in the 2 years subsequent to enrollment. There were no significant differences associated with the interaction terms of age and cohort status in this analysis.

In both cases, these analyses present severe methodological problems. The comparison between the Pilot and Baseline II cohort cannot be modeled in such a way so as to allow statistical separation of the effect of history from the potential effects of the Pilot. The comparison between the Pilot and Nearby cohorts include some statistical controls for differences in the driving environment (especially the ecological zip code crash variable). However, this may not entirely capture all differences in the driving environments, nor other unmeasured differences associated with the zip code of residence.

Despite these potential problems with the comparisons, it can be stated that there exists no evidence for an overall program effect of the 3-Tier Pilot on reducing subsequent crashes among participating customers. Similarly, there exists no evidence of an overall effect of the Pilot on the crash risk of older drivers, as indicated by the lack of significance in the interaction terms included in these analyses.

OVERALL PROGRAM EFFECTS ON SUBSEQUENT AT-FAULT CRASHES

A second set of analyses were conducted that focused on the fault status of subsequent crashes. Cox proportional hazards equations were constructed modeling the number of days to the first subsequent at-fault injury/fatal crash, as compared to customers who were in an injury/fatal crash for which they were found not to be at fault. The outcome of interest consisted of a dummy variable (at fault vs. not at fault) that was contingent upon being in a fatal/injury crash. Drivers who were not in a fatal/injury crash were excluded from the analysis. Interaction terms for age were not included in these models, due to lack of statistical power.

Compared to customers in the Baseline II cohort, Pilot customers had statistically significant lower odds of being found at fault for a subsequent injury/fatal crash. Compared to customers in the Nearby cohort, Pilot customers had directionally lower odds of being found at fault for a subsequent fatal/injury crash. However, the effect in this second equation does not meet conventional standards of statistical significance.

Together, these analyses suggest there exists weak, but inconsistent, evidence for a positive effect of the 3-Tier Pilot in reducing the likelihood of being found at fault for a subsequent injury/fatal crash. Due to the lack of statistical power, as well as severe methodological problems, this finding must be interpreted with a great deal of caution.

Question #10: What is the Overall Program Effect of the 3-Tier Pilot on Reducing Convictions?

A third set of analyses were conducted that focused on subsequent convictions, using a similar approach as that used to model crashes and fault status for injury/fatal crashes. The results showed that compared to drivers in the Baseline II cohort, there was no statistically significant difference in the odds of Pilot customers receiving a conviction in the follow-up period. Compared to drivers in the Nearby cohort, Pilot cohort customers had a small, but statistically significant, greater odds of receiving a conviction in the 2 years subsequent to enrollment in the

Pilot. Given the differences in driving environments among customers in these two cohorts—differences which may include the severity of law enforcement regimes across different jurisdictions—these differences may simply be an artifact of uncontrolled bias in the estimates. This finding should therefore be treated with a great deal of caution.

In sum, there is weak, but inconsistent, evidence that participation in the 3-Tier Pilot is associated with a small increase in the odds of receiving a conviction in the 2 years subsequent to participation.

Question #11: What Effect Did the Educational Material Distributed as Part of Tier 3 Have on Individual Risk of Crashing?

As noted in the 3-Tier Process Analysis (Camp, 2010b), it cannot be determined with any degree of confidence which individuals received the educational materials distributed as part of the 3-Tier Pilot, and which did not. It therefore cannot be determined whether the educational materials used as part of the 3-Tier Pilot had any effect on subsequent crash risk.

Question #12: What Effect Did the 3-Tier Pilot Have on Licensing/Mobility Options in the Post-Participation Period?

The overwhelming majority of 3-Tier customers retained their driving privilege upon completion of the Pilot. Even among the most severely functionally-limited group, over 85% of 3-Tier eligible customers successfully renewed their driving privilege. Nonetheless, customers who were found to possess driving-relevant limitations generally took longer to renew their license. Using the two comparison cohorts, the analysis was extended to determine if participation in the Pilot was associated with (i) delays in renewal of the driving privilege, (ii) the assignment of restrictions on the driving privilege, and (iii) retention of the driving privilege.

OVERALL PROGRAM EFFECTS ON TIME TO RENEWAL

In comparison to customers in the Nearby and Baseline II cohorts, 3-Tier Pilot renewal customers took approximately 3 to 5 days longer to renew their license. This extension of the time required to renew was associated with (a) a small, but nevertheless statistically significant, increase in the number of written renewal tests taken by Pilot customers, and (b) an increase in the number of drive tests required of Pilot customers.

OVERALL PROGRAM EFFECTS ON RESTRICTIONS

In comparison to customers in the Nearby and Baseline II cohorts, 3-Tier Pilot customers were significantly more likely to possess a license with the following restrictions: corrective lenses (01), no freeway driving (02), additional right-side mirrors (06), and area/route driving (13). The assignment of a restriction to driving only between sunrise and sunset (07) was significantly associated with Pilot participation only in one comparison (against the Baseline II cohort). For all other kinds of restrictions, including those normally associated with physical and mental (P&M) conditions, there was no statistically significant association between possession of a restricted license and participation in the 3-Tier Pilot.[1]

OVERALL PROGRAM EFFECTS ON LICENSURE

It was noted in the Process Analysis (Camp, 2010b) that surprisingly few suspension and revocation actions were taken specifically as a result of 3-Tier Pilot processing. It was therefore determined to expand the definition of the unlicensed to include all those who failed to renew their driving privilege for the entire 2-year follow-up period. This may include individuals who moved out of state, or whose license privilege was suspended or revoked. It also includes drivers who may have informally surrendered their driving privilege.

Among 3-Tier eligible renewal customers, participation in the Pilot is significantly and substantially associated with an increased odds of not possessing a valid license for the full 2-year period following the date of first contact. Depending on the comparison cohort, Pilot participation is associated with an increased odds of delicensure of approximately 50% (as against the Baseline II cohort) to 130% (as against the Nearby cohort). Using a population attributable fraction method, it is estimated that between 117 and 192 customers may have failed to renew their license as a result of participation in the 3-Tier Pilot.

Although it is normally assumed that delicensure is synonymous with cessation of driving, this may not actually be true. There exists evidence of driving among some drivers who lacked valid licenses during the follow-up period. Among customers younger than 70 y.o., approximately 30-32% of drivers accumulated a conviction during the period for which they did not possess a valid

[1]Other P&M-related restrictions include the following: adequate signaling devices (09), automatic transmission (10), adequate support to ensure proper driving position (11), hand-controlled brakes (16), knob attachment on steering wheel (17), full hand controls (22), and bioptic lens (44).

license. Among those aged 70 and older, approximately 4-6% of drivers accumulated a conviction during the period for which they did not possess a valid license.

Question #13: What Are the Cost Savings Associated with Projected Reductions in Crashes?

It was anticipated that the 3-Tier Pilot would produce some demonstrable safety effect in terms of reductions in crash rates among participants, relative to customers who did not participate in the Pilot. Due to severe methodological constraints, it is somewhat difficult to test this idea rigorously. Nevertheless, using the best available data, with two comparison groups (the Baseline II and Nearby cohorts), there exists no evidence for an overall program effect of 3-Tier in reducing the crash propensity of participants.

That said, participation in the 3-Tier Pilot was associated with an increased likelihood of receiving a restricted license. However, the available data do not lend themselves to estimating whether possession of a restricted license is associated with reduced crash risk. Nor is it possible to determine whether the assignment of restrictions extended the valid license status of drivers who otherwise would have received a suspension or revocation.

It does appear that some small number of 3-Tier Pilot cohort participants (approximately 117-192 customers) surrendered their driving privilege in excess of what would have been expected, based on delicensure rates among the comparison cohorts. Delicensure may be associated with reduced exposure. This reduced exposure may take the form of driving cessation, especially among older customers, though this is not really possible to determine from the available data. To the extent that delicensure—or its logical extension, driving cessation—is associated with reduced exposure, it is also likely to be associated with reduced crash risk. However, because this effect is concentrated among such a small number of drivers, these potential crash savings are invisible within the overall program effect estimates.

For these reasons—the lack of an overall program effect, the impossibility of estimating crash-risk reduction among restricted drivers, the difficulty of estimating crash-risk reduction among the delicensed—it is not really possible to estimate cost savings that may be associated with the 3-Tier Pilot. As a logical corollary, it therefore remains possible that there exist unmeasured benefits associated with the 3-Tier Assessment System, as piloted by CA DMV in 2006-2007.

Discussion and Conclusions

The 3-Tier Pilot involved the first large-scale effort to test, in a real-time public-agency setting in California, the 3-Tier Assessment System. The original twin program goals of the 3-Tier Pilot included (a) preserving and extending the safe driving years of California drivers of all ages, and (b) reducing crashes and violations.

In regards to the first goal, it appears that the overall effect of the Pilot likely reduced the driving years of a small, but nonetheless non-trivial, number of drivers. This reduction occurred mostly through the mechanism of informal surrender of the driving privilege (failure to renew) rather than through formal licensing actions taken by CA DMV (e.g., suspensions or revocations).

In regards to the second goal, it must again be emphasized that the quasi-experimental nature of the pilot makes for quite strenuous methodological constraints on the estimation of any program effects. In particular, there remain substantial differences between the Pilot and the two comparison groups that cannot entirely be statistically eliminated. Given these differences, any conclusions regarding program effects on crash risk must be made with extreme caution.

That said, two points regarding crash reduction may be made. First, there exists no evidence for an overall program effect of the Pilot in reducing the likelihood of crashing. There exists weak, but inconsistent, evidence that the 3-Tier Assessment System may be associated with a reduction in the likelihood of being found at fault for a subsequent injury/fatal crash. However, these findings reached conventional standards of statistical significance in only one comparison, and then only barely.

Second, among a small minority of 3-Tier customers, the Pilot may have had a lasting impact on subsequent driving habits. Delicensure is presumably associated with a reduction in driving—though not, in all cases, with cessation of driving—and therefore the delicensure effect of the Pilot may have produced a small (but unmeasured) reduction in the number of crashes. How this effect actually played out among this small group of drivers cannot really be determined using these data. It may be the case that the Pilot encouraged some extremely functionally-limited drivers to cease driving altogether, when otherwise they would have crashed. It is also possible that the Pilot may have induced some drivers to cease driving prematurely. Not all those who failed to renew their license had been designated as extremely functionally limited. In any case, any potentially beneficial effect of delicensure on subsequent crash risk cannot really be estimated at the present time.

Recommendations

Recommendation #1: Whether to Implement the 3-Tier Assessment System, as a Whole, Statewide.

The analyses presented here found no evidence for an overall program effect of the 3-Tier Assessment System for reducing subsequent crashes, and weak evidence for an overall program effect in reducing subsequent at-fault injury/fatal crashes. These findings are subject to strenuous methodological constraints, and so must be treated with caution. There are three distinct issues regarding how these methodological constraints affect interpretation of the results of any overall safety effects of the Pilot.

First, it is possible that the 3-Tier Pilot produced beneficial crash savings that remain obscured or unmeasured because of various methodological problems and potential sources of bias discussed elsewhere in this report and in the Process Analysis (Camp, 2010a and 2010b).

Second, to a certain extent some of the methodological problems and potential sources of bias noted in this report and in the Process Analysis (Camp, 2010a and 2010b) constituted deviations from the protocols developed for implementation. It is possible that the analyses presented here found no evidence for an overall program effect because the Pilot was not implemented perfectly. That said, it must be emphasized that every reasonable effort was made to enact the 3-Tier Pilot according to the recommendations of the 3-Tier Task Force, to train staff appropriately in Pilot procedures, and to engage in ongoing quality-control efforts. Thus, while it may be the case that the Pilot was not implemented perfectly, it remains an open question as to whether it would be possible to implement the 3-Tier Assessment System statewide with as high a degree of integrity as was achieved during the Pilot.

Third, the 3-Tier Pilot was implemented somewhat differently from the model proposed originally by Hennessy and Janke (2009). These differences in implementation were due in part to certain organizational constraints attendant upon altering CA DMV field office procedures, in quite substantial ways, for a temporary period of time. The results presented here are not precisely a test of the impact of the 3-Tier Assessment System as proposed by Hennessy & Janke

(2009); rather, the results presented here more strictly relate to the 3-Tier Assessment System as piloted by CA DMV in 2006-2007.

The analyses presented here did show evidence of a reduction in licensure rates associated with the 3-Tier Assessment System. These findings are also subject to strenuous methodological constraints, although the evidence regarding delays in licensure, the assignment of license restrictions, and retention of the driving privilege are generally stronger, and more consistent, than the evidence regarding crash rates. However, it remains unclear whether those who failed to renew their driving privilege did so in a manner that preserved their safety (i.e., they otherwise might have crashed), or if they did so in a manner that was potentially premature (i.e., they otherwise would have continued to drive safely).

Given these findings regarding crash rates and retention of the driving privilege, it is not recommended to implement the 3-Tier Assessment System statewide at this point in time.

Recommendation #2: Whether to Implement Separately any of the Constituent Components of the 3-Tier Assessment System.

It is not possible to determine at this time, on the basis of data collected for the 3-Tier Pilot, whether the physical observation protocol, the memory recall test, the Pelli-Robson "fog chart," or the PRT, would operate effectively as stand-alone mechanisms for identifying drivers in need of additional education and assessment. Such a determination would require a quite different research design than was used for the present project. Such a determination would also run contrary to the original "ecological" perspective on safe driving that formed the theoretical basis for the development of the 3-Tier Assessment System, as explicated in Hennessy and Janke (2009).

Given these methodological and theoretical constraints, it is not recommended to implement at this time, separately and on a stand-alone basis, any of the new screening tests used in the 3-Tier Pilot (the physical observation protocol, memory recall test, Pelli-Robson "fog chart," and PRT).

Recommendation #3: Future Research: Cognitive Screening Tests.

It is recommended that additional research be conducted to identify, if possible, a Tier 1 screening test for the kinds of cognitive and perceptual deficits associated with dementia-type disorders, mild cognitive impairment, and early-stage Alzheimer's disease. Although the PRT appears to possess utility as a screening tool for (among other things) these purposes, the 3-Tier Pilot incorporated this test at Tier 2, rather than at Tier 1. Administering the PRT to all renewal customers would involve substantial disruptions to current office procedures; therefore it would not necessarily be appropriate as a Tier 1 screening tool.

Such a Tier 1 screening test for functional limitations in cognition or perception must, if at all possible, possess face-validity properties that connect the testing mechanism (and its outcomes) to safe driving. These face-validity properties must be obvious both to front-line office staff and CA DMV customers. Ideally, such a test would also be readily incorporated into whatever office procedures are at that time in place for processing driver license renewal customers.

Recommendation #4: Future Research: Educational Intervention.

It is desirable that any system of screening for driving-relevant functional limitations, as well as any system of assessment of drivers' abilities to safely compensate for identified functional limitations, incorporate educational materials. It is therefore recommended that additional research be conducted to identify appropriate materials that encourage safe driving, and to identify the appropriate mechanism(s) for delivering such materials to CA DMV customers.

Recommendation #5: Future Research: Use of Occupational Therapists/Certified Driving Rehabilitation Specialists

It is recommended that additional research be conducted regarding the relationship between different kinds of drive-test preparation—of which referral to an occupational therapist or certified driving rehabilitation specialist might be one form—and drive test outcomes, subsequent driving habits, and subsequent safe driving.

Recommendation #6: Future Research: Voluntary Delicensure.

There exists evidence that when delicensure occurred in the 3-Tier Pilot, it was informal, rather than formal. Instead of a suspension or revocation order, delicensure among Pilot customers appears to have more commonly taken the form of letting the renewal application lapse.

It is not clear what happened with these informally delicensed customers once they let their applications lapse. Some may still be driving in California, as indicated by the accumulation of later traffic convictions. Others may have ceased driving altogether. Still others may have moved to another jurisdiction. It is unknown in what proportion these forms of informal delicensure (unlicensed driving, driving cessation, and moving outside the state) occur. It is unknown to what extent different kinds of informal delicensure are associated with known demographic variables of interest, such as age and gender. It is unknown if delicensure is associated with organizational variables of interest (e.g., the office at which a renewal is conducted). It is unknown if informal delicensure is associated with changes to crash risk, for those who still drive. Finally, it is unknown if informal delicensure is associated with other kinds of social and health outcomes, especially in the case of those who cease driving.

For all of these reasons, it is recommended that additional research be conducted regarding informal delicensure, driving habits and exposure, individual demographic variables (e.g., age, gender, health status), organizational variables (office processing), and safety outcomes such as crash risk.

TABLE OF CONTENTS

PAGE

TABLE OF CONTENTS (continued)

PAGE

TABLE OF CONTENTS (continued)

PAGE

APPENDICES

NUMBER PAGE

LIST OF TABLES

TABLE OF CONTENTS (continued)

LIST OF TABLES (continued)

TABLE OF CONTENTS (continued)

LIST OF TABLES (continued)

TABLE OF CONTENTS (continued)

LIST OF TABLES (continued)

NUMBER PAGE

LIST OF FIGURES

INTRODUCTION

On September 14, 2006, California Governor Arnold Schwarzenegger signed Assembly Bill 2541 (Daucher) into law, adding Section 1659.9 to the California Vehicle Code, and calling for a pilot study by the California Department of Motor Vehicles (CA DMV) of the 3-Tier Assessment System (3TAS). This manuscript (the "outcome analysis") constitutes the second of two reports on 3TAS. Per the authorizing legislation, this report covers the following topics:

(1) A projection of the costs associated with the area driving performance evaluation.
(2) A determination of the willingness of a participant to pay a fee for the area driving performance evaluation.
(3) A determination of the percentage of drivers who were assessed to have a limitation but who, upon completion of the assessment, were able to retain their driving privileges.
(4) The utilization of certified driving rehabilitation specialists.
(5) The results regarding crash rates and retention of driving privileges.

The main thrust of the analyses contained herein examine the effectiveness of 3TAS as a means of identifying functional impairments, reducing crashes, and extending safe driving years for California drivers of all ages. These analyses are based on 2 years of elapsed driving history, and incorporate comparisons with two control groups processed according to standard CA DMV procedures in place at the time of the 3-Tier Pilot.

Statement of the Problem

The demographic composition of the driving population of California is shifting. California's Department of Finance (2007) estimates that the number of seniors (aged 65 and older) in the state is expected to double over the next two decades, rising from 4.4 million (in 2010) to 8.8 million (in 2030). Over that same period, seniors' proportional share of the total population will increase substantially, from 11.3% to almost 18%. It is expected that this shift in the population will result in a major increase in the number of senior drivers, as well as a substantial shift in their proportional share of drivers on California's roads. This is expected to have a number of implications in different areas of traffic safety. For CA DMV, these implications include (a) preserving and extending driving years for individuals who can safely operate a motor vehicle, and (b) identifying appropriate procedures for determining when individuals must restrict their driving privilege, or even cease driving altogether. The 3-Tier Pilot constituted a critical test of a new method by which the Department might fulfill one of its core mission goals, namely promoting traffic safety through the licensing process.

Background

3TAS builds upon prior analytic work at CA DMV and elsewhere. Indeed, research on the assessment of age-related functional limitations in driving is a growing field made richer by the multi-disciplinary nature of the object of study. Scholars and practitioners outside CA DMV have looked at a number of research questions related to the assessment of driving competency. Taken together, this body of research findings raises three key points that have shaped the development of 3TAS: (1) that an assessment system designed to identify driving-relevant functional limitations should apply to a broad spectrum of drivers, regardless of age, (2) that such a system must involve screening for potential driving-relevant limitations in multiple functional domains, including vision, cognition, and physical function, and (3) that such a system should incorporate a substantial educational and/or therapeutic component, so that drivers may retain the driving privilege as long as they can safely continue to operate a motor vehicle.

Prior Research at CA DMV

At CA DMV the research that led to 3TAS began as part of a cooperative agreement between the Department and the National Highway Traffic Safety Administration (NHTSA) starting in 1993. This arrangement led first to the publication of a substantial literature review on the relationship between age-related disabilities and safe driving (Janke, 1994). This latter monograph contained the first conceptual description of a proposed tiered system of assessment. An important component of that initial proposal was the involvement at "second-stage testing" (what later was termed Tier 2) of referral to a medical professional, or vision specialist, depending on the nature of the problem identified at the "first-stage testing" (later termed Tier 1). It was also suggested that some drivers might, at this second stage of testing, be identified as so functionally limited that they were too unsafe even to be given a road test.

This initial proposal then led to several pilot studies (Hennessy, 1995; Janke & Hersch, 1997). These pilot studies involved field-testing and validation of a number of screening tests thought potentially suitable for inclusion in the first stage (or Tier 1) of a tiered system of screening and assessment. Hennessy's (1995) study focused on vision testing in particular, and identified the Pelli-Robson low-contrast acuity test and the Perceptual Response Test (PRT), along with CA

DMV's current Snellen standard, as demonstrating predictive validity for identifying drivers with visual problems and potential crash risk. These findings regarding the PRT (and the larger UFOV test battery of which it forms a sub-test) have been validated in numerous other studies (Clay et al., 2005). The tests which form the UFOV battery have also been extended to use in identifying drivers with problems related to perception/cognition, especially as related to Parkinson's disease (Classen et al., 2009), dementias of the Alzheimer's type (Duchek, Hunt, Ball, Buckles, & Morris, 1998; Rizzo, Reinach, McGehee, & Dawson, 1997), and cognitive impairments stemming from other types of injuries and disorders (Fisk, Novack, Mennemeier, & Roenker, 2002; Mazer et al., 2003; Schultheis, Garay, & DeLuca, 2001). Janke & Hersch (1997), in a paired set of pilot studies, field-tested a broad range of non-driving tests for potential use in identifying drivers at risk of performing poorly on an on-road test of driving skill.

Together, these early pilot studies resulted in the publication of a call for further testing and refinement, as reported in Janke & Eberhard (1998) and Janke (2001), of a tiered system of assessment using an assessment battery covering a range of functions, and suitable for use in an agency context. The final selection of specific screening tests, associated screening cut-points, and combinatory assessment outcome categories (passing, somewhat functionally limited, and extremely functionally limited) was made on the basis of a lengthy (12-month) pilot in four Southern California field offices (Hennessy & Janke, 2005; Hennessy & Janke, 2009). On the basis of findings published in these latter reports, it was also suggested that some drivers—especially those identified as somewhat functionally limited—might derive particular benefit from education and training in techniques for safely compensating for conditions of which they may not have been previously aware. It was on the basis of proposals contained in these latter two reports that the 3-Tier Pilot, implemented in 2006-2007 in six field offices in Northern California, took shape as required by the passage of AB 2541 (Daucher) of 2006, modifying California Vehicle Code section 1659.9. Certain modifications to the 3-Tier Assessment System took place prior to implementation of the Pilot; these were made on the basis of input from members of the 3-Tier Task Force, an internal working group of CA DMV staff convened to oversee preparations for implementation. These modifications had no effect on the choice of screening tests, cut-points, or combinatory assessment outcomes; instead, they dealt largely with the circumstances under which customers would be given the option to take an Area Driving Performance Evaluation (ADPE)-type road test.

Research by Scholars and Practitioners Outside CA DMV

Any assessment of driving competency cannot be based solely upon age. This is true for empirical as much as for conceptual or political reasons. Empirically speaking, the elderly are among the safest drivers on the road, especially when measured in terms of per-driver crash rates (Evans, 1988a; US Department of Transportation, 1993). However, there exists a good deal of variation within age groups, and crash rates as measured on a per-driver (or a per-able-driver) basis are higher among the oldest seniors (those aged 80+) than they are among younger seniors (e.g., those aged 65-74) (Eberhard & Mitchell, 2009; Lyman, Ferguson, Braver, & Williams, 2002; US Department of Transportation, 2003). When measured on a per-mile basis, crash rates are generally higher for older drivers than for all others except novice younger drivers (Janke, Masten, McKenzie, Gebers, & Kelsey, 2003; National Highway Traffic Safety Administration, 2005; Williams & Carsten, 1989). That said, when it comes to traffic citations, violation rates decline with age. Seniors are always among the most law-abiding drivers on the road, whether measured on a per-driver or per-mile basis; this is especially true for violations involving alcohol (Janke et al., 2003; National Highway Traffic Safety Administration, 2005).

Not only are (some) seniors at risk for crashing, but they are also at elevated risk for crashes that result in injuries and fatalities (Barancik et al., 1986; Barr, 1991; Bédard, Guyatt, Stones, & Hirdes, 2002; Evans, 1991; Hanrahan, Layde, Zhu, Guse, & Hargarten, 2009; Massie & Campbell, 1993; McCoy, Johnston, & Duthie, 1989; Newgard, 2008; Retchin, Cox, Fox, & Irwin, 1988; Reuben, Siliman, & Traines, 1988; Ryan, Legge, & Rosman, 1998). Fatalities and injuries among senior drivers are largely a product of the physical frailty typically associated with age (Evans, 1988b; Li, Braver, & Chen, 2003; Meuleners, Harding, Lee, & Legge, 2006; Viano, Culver, Evans, Frick, & Scott, 1990). Thus, while crashes among seniors are not frequent, when they happen they tend to have serious consequences. However, this is really only true for senior drivers themselves (and perhaps their passengers), and not for other road users (Braver & Trempel, 2004; Dellinger, Kresnow, White, & Sehgal, 2004; Dulisse, 1997; Evans, 2000; Langford, Bohensky, Koppel, & Newstead, 2008b). Furthermore, crash rates for all age groups, including seniors, have been on the decline over the past decade (Eberhard, 2008), likely due to improvements to road safety, vehicle crashworthiness, and increased use of safety devices such as seat belts and airbags. There is even some evidence that fatality risk among seniors is declining, and declining faster than the fatality rate among middle-age drivers (Cheung & McCartt, 2011).

Crash risk among older drivers is directly related to the frequency and type of driving done by individual seniors and this in turn complicates the way in which this category of crash risk ought to be conceptualized (Hakamies-Blomqvist, Raitanen, & O'Neil, 2002; Langford, Methorst, & Hakamies-Blomqvist, 2006). Many seniors reduce their amount of driving by, for instance, making fewer and shorter trips (Adler & Kuskowski, 2003; Benekohal, Michaels, Shim, & Resende, 1994; Blanchard & Myers, 2010; Charlton et al., 2006; Gallo, Rebok, & Lesiker, 1999; Klavora & Heslegrave, 2002; Kostyniuk, Shope, & Molnar, 2000; Marottoli et al., 1993; Raitanen, Tormakangas, Mollenkopf, & Marcellini, 2003; Ruechel & Mann, 2005; Stutts, 1998). This self-restricting behavior limits their exposure. In other words, the less often a person drives, the fewer opportunities they have of crashing. Many seniors also alter their driving behavior in particular ways to limit the incidence of risky or demanding traffic situations; this may include restricting their driving to routes with lower speed limits, to routes that involve fewer turns (especially unprotected left turns), to certain times of day (i.e., avoiding driving at night or during rush hour), to days when the weather is clear, etc. (Adler, Rottunda, & Kuskowski, 1999; Baldock, Mathias, McLean, & Berndt, 2006; Hakamies-Blomqvist & Wohlstrom, 1998; Kostyniuk & Molnar, 2005; Preusser, Williams, Ferguson, Ulmer, & Weinstein, 1998; Stalvey & Owsley, 2000; for a discussion of self-restricting behavior among younger drivers, see Naumann, Dellinger, & Kresnow, 2011). As seniors adopt restrictions to their driving, this generally involves reducing the proportion of driving done on freeways, and consequently increasing the proportion of driving on city streets, where crashes are more likely on a per-mile basis regardless of the age of the person doing the driving (Janke, 1991; Eberhard, 2008). However, the type of alterations that drivers make to their driving depends in part on what sort of driving environment they face—and this varies depending on if a driver lives in an urban or rural environment (Hanson & Hildebrand, 2011a, 2011b). It should be noted, however, that self-restricting behavior occurs at such widely varying rates across studies (Molnar & Eby, 2008) as to invite some skepticism as to the reliability of self-reported self-restricting driving behavior. There appears to be basic problems with self-reporting of driving exposure when compared to objective measures of driving, such as provided by electronic trip meters (Blanchard, Myers, & Porter, 2010; Huebner, Porter, & Marshall, 2006).

Self-restriction, in turn, derives in large part from drivers' knowledge of their own limitations in vision, cognitive health, or physical frailty (Alvarez & Fierro, 2008; Ball et al., 1998; Dellinger, Sehgal, Sleet, & Barrett-Connor, 2001; Dobbs & Carr, 2005; Foley, Masaki, Ross, & White, 2000; Freund, Colgrove, Burke, & McLeod, 2005; Johnson, 1998; Kington, Reuben, Rogowski, & Lillard, 1994; Marshall, 2008; Molnar, Eby, Kartje, & St. Louis, 2010; Owsley, Stalvey, & Phillips, 2003; Stalvey & Owsley, 2000; Vance et al., 2006; West, et al., 2003). This knowledge of one's own limitations is, ideally, then related to driver's self-understanding of their own competence behind the wheel (Sundström, 2011). Of course, drivers who are unaware of their limitations (or deny the relationship between a particular limitation and safe driving) are unlikely to take precautions when driving (Holland & Rabbitt, 1990). This lack of self-awareness is especially troubling for patients with dementia or other cognitive limitations (Friedland et al., 1988; Kazniak, Keyl, & Albert, 1991). Self-regulation of driving is also tied to household composition (i.e., living alone versus with a spouse or other family member) and psychological factors such as feelings of independence and self-worth (Donorfio, D'Ambrosio, Coughlin, & Mohyde, 2009; Johnson, 2002).

Functional limitations in vision, cognition, and physical strength or coordination are, of course, related to safe driving. The literature on this last point is quite large (but see Charlton et al., 2004; Dobbs, 2005; Janke, 1994; Marottoli, Cooney, Wagner, Doucette, & Tinetti, 1994; and Vaa, 2004 for recent reviews of the literature and synthetic meta-analyses). In sum, crash risk among seniors is a product of the combination of at least two factors: (a) functional limitations in physical capabilities, visual health, and cognitive capacity that may impair driving skill but also (b) the degree to which drivers are aware of these limitations and compensate for them by changing where, when, and how they drive.

It is undeniably the case that the frequency of driving-related impairments in physical function, vision, and cognitive health increases with age (Attebo, Mitchell, & Smith, 1996; Hakamies-Blomqvist, 1993; Lyman, McGwin, & Sims, 2001; Marottoli & Drickamer, 1993; Molnar & Eby, 2008; Ni, Kang, & Andersen, 2010; Stelmach & Nahom, 1992; Stutts, Stewart, & Martell, 1998). However, it does not follow that assessment of driving skill ought to be based upon age (Organization for Economic Cooperation and Development, 2001). Leaving aside variation between younger and older drivers, there is simply too much variation even within age groups (by gender as well as by health status) for a strictly age-based assessment system to be of any real utility (Baker, Falb, Voas, & Lacey, 2003; Hu, Jones, Reuscher, Schmoyer, & Truett, 2000; Hu, Trumble, Foley, Eberhard, & Wallace, 1998; Oxley, Charlton, Koppel, Scully, & Fildes, 2005; Rabbit, 1993; Waller, 1992). Stated more plainly: some younger drivers have functional limitations that impair driving, while many older drivers do not possess substantial functional limitations that impair their driving. Even among those older drivers who do possess driving-relevant functional limitations there is a good deal of evidence that many drivers can compensate (by, for instance, self-restricting) for their limitations (Ball et al., 1998; Charlton, Oxley, Fildes, Oxley, & Newstead, 2003; Owsley, Stalvey, Wells, & Sloane, 1999). Thus, an age-based system of assessment would likely miss some drivers who ought to be assessed, subject substantial numbers of drivers to potentially fruitless scrutiny, and at the same time use scarce public resources inefficiently. Furthermore, there is little evidence that jurisdictions which have imposed age-based assessment of driving skill have seen significant reductions of motor vehicle crashes (Grabowski, Campbell, & Morrisey, 2004; Hakamies-Blomqvist, Johansson, & Lundberg, 1996; Lange & McKnight, 1996; Langford, Bohensky, Koppel, & Newstead, 2008a; Langford, Fitzharris, Koppel, & Newstead, 2004; Langford, Fitzharris, Newstead, & Koppel, 2004; Levy, Vernick, & Howard, 1995; Rock, 1998; Torpey, 1986). The only age-based licensing policy that shows specific results in reducing crashes and deaths among older drivers is a requirement to renew a license in a DMV office in person, rather than by mail (Grabowski et al., 2004).

This then argues for exploring the utility of a function-based system of driving assessment that would identify drivers with limitations that may affect their driving, regardless of age (Fildes et al., 2000; Janke & Eberhard, 1998). Driving involves a multitude of competencies, however, and no single functional domain—not even vision—exerts a predominant effect on overall driving skill (Anstey, Wood, Lord, & Walker, 2005; Bédard, Weaver, Dārzinš, & Porter, 2008; Galski, Bruno, & Ehle, 1992; George & Smiley, 1999; Simms, 1985; Stav, Justiss, McCarthy, Mann, & Lanford, 2008). Thus, jurisdictions that have adopted systems of post-licensing driving competency assessment often use assessment batteries that cover several domains (Ball et al., 2006; Diller et al., 2001; Fildes et al., 2008; Kantor, Mauger, Richardson, & Tschantz-Unroe, 2004; Langford, 2008; Vernon et al., 2002; Staplin, Gish, & Wagner, 2003; Staplin, Lococo, Gish, & Decina, 2003a; Wheatley & DiStefano, 2008). Usually, these assessment systems incorporate some form of the on-road drive test, which has been shown both to have utility for demonstrating one's overall driving competency, as well as one's ability to compensate for a given limitation (Brook, Qustad, Patterson, & Valois, 1992; DiStefano & McDonald, 2003; Fox, Bowden, & Smith, 1998; Hunt et al., 1997; Justiss, Mann, Stav, & Velozo, 2006; Molnar, Patel, Marshall, Man-Son-Hing, & Wilson, 2006; Withaar, Brouwer, & Van Zomeren, 2000; Wood & Mallon, 2001).

Among the potential outcomes of a function-based system of driving assessment are various kinds of restrictions on the driving privilege. License restrictions are currently used by CA DMV as required to preserve the safety of the driver and the motoring public. Such restrictions include cases where drivers' functional abilities may require changes to the situations in which they drive (for instance, at night or on the freeway, or only when accompanied by another licensed adult driver). Other kinds of restrictions may require the use of adaptive mechanisms to improve the facility with which a given driver can maneuver their vehicle (for instance, when wearing corrective lenses, using hand controls, installing additional mirrors, etc.). In some cases, drivers may even be restricted to using certain defined routes or areas. Especially in the case of older drivers, license restrictions may be associated with reductions in crash risk (Nasvadi & Wister, 2009) and avoiding premature driving cessation (Braitman, Chaudhary, & McCartt, 2010; Langford & Koppel, 2011).

Aside from formal licensing actions, research suggests that any system of driving assessment can usefully incorporate therapeutic and educational components (Korner-Bitensky, Kua, von Zweck, & Van Benthem, 2009). These are expected to result in positive safety and mobility outcomes by (a) potentially reducing an individual's risk of crashing, and (b) potentially extending an individual's safe driving years and so preserving their options regarding personal mobility. Referral to a health professional is, in California and some other jurisdictions, sometimes called for in those cases where conditions known to affect driving are medically indicated (Adler & Silverstein, 2008; Fildes et al., 2008; Lococo & Staplin, 2005; Soderstrom & Joyce, 2008; Wheatley & Di Stefano, 2008). In many instances conditions that affect driving may be treatable, or at least to some extent managed, though medical means (Dobbs, Carr, & Morris, 2002; Marottoli et al., 2007). Alternatively, driving skill can in some cases be corrected or improved through driving rehabilitation and training—as, for instance, under the care of an occupational therapist or licensed driving instructor, or even in consultation with family members (Bédard, Isherwood, Moore, Gibbons, & Lindstrom, 2004; Eby, Molnar, Shope, Vivoda, & Fordyce, 2003; Kua, Korner-Bitensky, Desrosiers, Man-Son-Hing, & Marshall, 2007; Lyman, McGwin, & Sims, 2001; Man-Son-Hing, Marshall, Molnar, & Wilson, 2007; Ranney & Hunt, 1997; Sommer, Falkmer, Bekiaris, & Panou, 2004; Wheatley & Di Stefano, 2008). Certainly referrals to health professionals and other kinds of community resources may be necessary in situations where a driver is required to cease driving and transition to other transportation options (Adler & Rottunda, 2006; Adler & Silverstein, 2008; Bryanton, Weeks, & Lees, 2010; Carr et al., 1991; Edwards, Bart, O'Connor, and Cissell, 2009; Gilley et al., 1991; Jett, Tappan, & Rosselli, 2005; Johnson, 2008; O'Neill, 1997; Taylor & Tripodes, 2001). This last point cannot be understated; as noted by many, many authors, access to personal transportation (usually in the form of a vehicle owned and operated by a member of the household) is an important factor in the maintenance of overall health and well-being (Banister & Bowling, 2004; Edwards, Lunsman, Perkins, Rebok, & Roth, 2009; Eisenhandler, 1993; Freeman, Gange, Muñoz, & West, 2006; Fonda, Wallace, & Herzog, 2001; Harper & Schatz, 1998; Harris, 2000; Harrison & Ragland, 2003; Johnson, 1998; Kostiniuk, Shope, & Molnar, 2000; Marottoli et al., 1997; Marottoli et al., 2000; Oxley & Whelan, 2008; Persson, 1993; Rabbitt, Carmichael, Shilling, & Sutcliffe, 2002; Ragland, Satariano, & MacLeon, 2005; Stutts, Wilkins, Reinfurt, Rodgman, & Van Heusen-Causey, 2001; Taylor & Tripodes, 2001; Windsor, Anstey, Butterworth, Luszcz, & Andrews, 2007; Yasuda, Wilson, & von Mering, 1997).

CURRENT PROJECT

CA DMV's 3-Tier Assessment System Pilot

The 3-Tier Assessment System, as piloted by CA DMV in 2006-2007, consisted of a tiered series of screening and assessment tools by which the department sought to identify customers potentially in need of further assessment and/or education regarding their driving.

The 3-Tier Pilot incorporated two groups of customers, referred to respectively as "renewals" and "referrals." 3-Tier eligible renewal customers included all those required to renew their application in a field office (as opposed to renewing by mail), and who were required to take the 18-question written knowledge test as part of their renewal application, according to CA DMV policy in place at the time of the Pilot. As a general matter, this included customers of any age with a limited-term license, who had negligent operator "points" on their record, or who had recent at-fault crashes. This also included all customers, regardless of prior record, aged 70 and older. Customers who elected to take the written renewal test in a language other than English, or who were applying for licenses other than Class C (non-commercial vehicle), were excluded from the 3-Tier Pilot. 3-Tier eligible referral customers included certain individuals referred to CA DMV's Driver Safety Branch, typically by law enforcement or a physician, for examination of their ability to safely operate a motor vehicle.[2]

The Tier 1 screening tools included four components: a simple memory recall test, a checklist for the structured observation of physical limitations (both upper- and lower-body) that might potentially affect driving, and two vision tests. The vision tests measured, respectively, distance acuity and contrast sensitivity. The acuity test used CA DMV's current Snellen chart, and associated vision standard. The test of contrast sensitivity used the Pelli-Robson chart. These screening tests were administered by a front-line staff-person working in a CA DMV field office as part of a customer's renewal application processing. The Tier 1 screening tests were applied

[2]CA DMV accepts referrals requesting the evaluation of individuals' safe driving from other sources besides law enforcement and physicians. These include DMV staff (e.g., a licensing registration examiner who observes dangerous maneuvers during the course of a drive test), family members of a driver, drivers themselves (i.e., self-report), or concerned members of the community. As a practical matter, a plurality of Driver Safety referrals in the 3-Tier Pilot came from law enforcement, followed by referrals from physicians and field office staff. There were a handful of referrals from family members and concerned members of the community. There were no self-reported Driver Safety referrals in the Pilot.

to all 3-Tier eligible renewal customers. Referrals enrolled in the Pilot were not screened using the Tier 1 tests; instead, they were assessed only at Tier 3.

The Tier 2 screening tests consisted of CA DMV's current 18-question written knowledge test for renewal of a Class C license, and a computer-based test of visual function, the Perceptual Response Test (PRT). The PRT measures the processing speed of the visual system, and has been shown to be reliably associated with dementia-type cognitive disorders, as well as crash risk. All 3-Tier eligible renewal customers enrolled in the Pilot took the written knowledge test. However, only those customers flagged at Tier 1 as having one or more driving-relevant limitations, or who failed the written knowledge test twice, were required to take the PRT. These tests were typically administered by front-line field office staff.

Tier 3 consisted of two components: an on-road drive test, and the distribution of written- and video-based educational materials. The drive test consisted either of the Supplemental Driving Performance Evaluation (SDPE), or the Area Driving Performance Evaluation (ADPE), both of which have long been components of CA DMV's system of assessing driver competency (for more details on the SDPE drive test, see Masten, 1998b and 1998c). The educational materials were tailored to a given customer's needs and circumstances, depending on their Tier 1 and Tier 2 screening test outcomes. These educational materials were intended (a) to prepare them for an on-road driving test (if that was required) and/or (b) to educate them about potential means by which they might compensate for any identified limitations to their contrast sensitivity or perceptual speed. The educational materials that focused on drive-test preparation were given to all customers required to take a drive test as part of the Pilot. The educational materials that focused on methods of compensating for identified limitations to contrast sensitivity or perceptual speed were intended to be given out according to a randomizing protocol that would allow for later analysis of the effectiveness of these materials in improving safe driving. If a renewal customer was identified at Tier 1 and/or Tier 2 as possessing multiple potential driving-relevant limitations, or with a single limitation that was serious in nature, they were assessed at Tier 3. All Driver Safety referral customers were assessed at Tier 3. The Tier 3 assessments were typically administered by a Licensing Registration Examiner (LRE), or a field office manager.

For more details on how the 3-Tier Assessment System worked as an integrated system during the Pilot, see the Process Analysis (Camp, 2010b), especially pp. 17-31. For more details on the development of the 3-Tier Assessment System, including the choice of component screening tests, the determination of cutoff scores for each test, and the theoretical background for this type of ecological perspective on driver assessment, see the 3-Tier Technical Report by Hennessy and Janke (2009).

METHODS

Collection of Data for the Pilot, Nearby, and Baseline I and II Cohorts

The collection of data involved three phases, corresponding to (i) the Baseline I cohort, (ii) the Pilot cohort, and (iii) the Baseline II and Nearby cohorts. Each of these three data-collection phases involved substantially different methodologies.

For the Baseline I cohort, the license renewal application forms (DL1 RN and DL44) were retained for all 3-Tier eligible renewal customers processing their applications at the six Pilot field offices (Carmichael, Fairfield, Folsom, Sacramento-Broadway, Sacramento-South, and Vacaville). The starting point of the Baseline I cohort was 9 months prior to the starting point of the 3-Tier Pilot cohort–so, September 2006. The data collection for Baseline I continued for 5 months (through January 2007), a similar amount of time as was used for the actual Pilot. No additional processing was given to these customers other than what was normally conducted for a license renewal application. For those customers who were required to take a drive test, the drive test score sheet was also retained. In addition, the reason for the required drive test was recorded on a separate sheet of paper (e.g., a Driver Safety referral, failure of the visual acuity standard, or possession of a limited-term license). For Driver Safety referral customers, the DL11D was also retained; this document was used to transmit information between the Driver Safety office and the field offices.

For the Pilot cohort, the primary form of data collection involved the recording of information for each individual customer by field office employees during the course of their license renewal (or referral) transaction. The physical paperwork used to collect and record data consisted of several types of forms, depending on the circumstances of a given customer's renewal or referral transaction process. For renewal customers, these forms included the Driving Information Survey, the Tier 1 Score Sheet, and the 3-Tier Tracking Sheet. For Driver Safety referral customers, these forms included the 3-Tier Tracking Sheet and the DL11D used to communicate information between the Driver Safety office and the field offices. For customers required to take an on-road test of driving skill, a copy of their drive-test score sheet was also retained. For customers who were required to undergo a vision examination, a copy of the DL44 used to record the results of that examination was retained. In a few instances, additional information was recorded on a form titled DL11, which is normally used to transmit information from a field office to DMV Headquarters. These latter instances usually had to do with the addition, continuation, or removal of a limited-term status for a license, or the addition, continuation, or removal of restrictions on the driving privilege.

All of the forms collected during the Pilot were scanned and converted to a PDF document at each field office. The PDF version was then stored on a secure electronic server for access by DMV Research and Development (R&D); the original paper documents were subsequently destroyed to protect customers' personal information. Upon completion of the 3-Tier Pilot, the data contained in these PDF forms were entered into an electronic database suitable for statistical analysis via SPSS/PASW (ver. 19) or SAS (ver. 9.2), depending on the statistical technique required.

During this phase of the project, a certain number of customer files were identified as erroneously processed. These errors took various forms, as detailed in the Process Analysis (Camp, 2010b). In addition, a certain number of customers failed to complete their applications within the time frame allotted for the Pilot (i.e., by 12/31/07). For the purposes of the Process Analysis, these customers were grouped separately from those who had been correctly processed. This was done to enable the various estimations of customer flows, time added to processing, costing, etc., of the 3-Tier Assessment System when functioning according to the original assumptions and design. For the purposes of the present Outcome Analysis, these customers' data were retained for several reasons, not the least of which was to ensure an accurate estimation of the potential effects of the 3-Tier Assessment System on the retention of individuals' driving privileges. However, the data were retained only for those customers who were, as far as could be determined, eligible for inclusion in the 3-Tier Pilot. Customers who began their license renewal application prior to the start of the Pilot (6/4/07), who possessed a motorcycle or commercial license (or whose immediately previous license had included a motorcycle or commercial class), or who took the written renewal test in a language other than English, were deemed ineligible for inclusion, and their data were excluded from the present analysis. Customers who subsequently added a motorcycle or commercial class to their license were retained, along with a variable indicating the date on which they changed license class. Customers who subsequently died were also retained, along with a variable indicating the date on which their passing was reported to CA DMV.

For the Baseline II and Nearby cohorts, the author consulted a separate database provided by CA DMV Audits Branch. These data were provided by Audits Branch with the original intention that they would help the author determine whether (and to what degree) potential 3-Tier customers avoided enrollment in the Pilot by conducting their license renewal transactions at non 3-Tier offices (in the event, this probably occurred relatively rarely, and was thus probably not a major source of potential bias to the results; see Camp, 2010b for more details). This database included certain limited data for all customers conducting Class C (non-commercial) license renewal transactions in both the 3-Tier Pilot offices and a subset of other offices in the greater Sacramento metropolitan area. These "nearby" offices included: Placerville, Roseville, Rocklin, Woodland, Davis, Napa, and Lodi. The Baseline II customers were processed at the Pilot offices but were culled from that portion of the database covering the period one year prior to when the Baseline I cohort had been processed. For both cohorts, customers who began their license application during the appropriate period (6/4/07-10/31/07 for the Nearby cohort, 9/1/05-1/31/06 for the Baseline II cohort) were retained. Customers whose record indicated previous possession of a motorcycle or commercial license, or who showed evidence of having taken the written knowledge test in a language other than English, were excluded from the sample. Customers who failed to complete their application within the appropriate time period were retained for comparison with the lagging/erroneously processed customers in the Pilot cohort.

Exposure Measures

Six variables were constructed to capture individual differences in exposure. Two of these were based on survey data collected among Pilot cohort customers. The remaining four are based upon application and licensing data available for customers in all three study cohorts.

At the time of their enrollment in the Pilot cohort, customers were asked to fill out a "Driving Information Survey." This survey is similar to one used in prior studies that lead to the 3-Tier Pilot (Hennessey, 1995; Janke & Hersch, 1997), which was in turn based upon an instrument developed by Owsley and colleagues (Ball, Owsley, Sloane, Roenker, & Bruni, 1993; Ball et al., 1998; Owsley, Ball, Sloane, Roenker, & Bruni, 1991). The survey, as used in the 3-Tier pilot, consisted of nine questions. Of these, eight asked about avoidance behaviors—whether the respondent avoided various kinds of risky and challenging situations: driving at night, in rainy or foggy weather, at dusk/dawn, without passengers, making left turns across oncoming traffic, in heavy traffic, on the freeway, and on unfamiliar routes. The answers ("never," "sometimes," "often," and "always") are easily transformed to numeric values for use as ordinal variables. The ninth question asked simply how many days per week the respondent normally drove. This yielded a second exposure measure with values ranging from zero to seven; customers indicating that they did not drive at all most weeks had values on this measure set to zero. Missing answers were not imputed for any questions taken from the Driving Information Survey. For the precise question wording, see Appendix A.

Third, a variable was constructed to capture the effects of any potential changes in driving associated with delays in the renewal of the driving privilege. This variable measured the number of days between the issue date of the renewed license (or the date of the ending of a Driver Safety action) and 2 years from the application (or Driver Safety action) start date. For those customers who failed to renew their driving privilege during the two year follow-up period, this measure was set to zero.

Fourth, a variable was constructed to reference the office at which a customer started their license renewal transaction, or (for Driver Safety referral customers) the office at which they conducted their first drive test. This was intended to capture potential differences in office processing. It may also capture otherwise unknown demographic information (such as socioeconomic status). Finally, it may also serve as a potential proxy for the environment in which a given customer conducts their driving. This measure is necessarily imperfect: customers are by no means required to use the office nearest their home or place of business. Nor should we assume that drivers do most of their driving in the immediate vicinity of a given office.

Fifth, a variable was used that consisted of an ecological measure of the number of crashes per licensed driver per year, for the zip code of recorded residence for each customer. This was multiplied by a factor of 100 to increase the range of variation. Customers with an out-of-state zip code had values on this measure set to missing. Customers with confidential addresses, or who lived in zip codes with fewer than 100 residents, had values on this measure set to the statewide mean. This measure is somewhat imperfect, as it necessarily captures only a portion of any given customer's driving environment.[3] In a manner similar to the office variable, the zip code crash index may capture certain otherwise-unmeasured sociodemographic differences associated with a given driver's neighborhood of residence. Especially for comparisons involving the Pilot and Nearby cohorts, this measure was deemed preferable to the office measure (discussed in the previous paragraph) precisely because these two cohorts were not patronizing the same field offices. However, using zip codes does not capture those potential differences that may arise from variation in office processing. For examples of other studies that have used this index, see Peck and Gebers (1991), Oulad Daoud and Tashima (2011), and Gebers (forthcoming).

[3]It also assumes that individuals do at least some driving within their stated zip code of residence.

Finally, for the analyses of conviction outcomes, a variable was used that consisted of an ecological measure of the number of convictions per licensed driver per year, for the zip code of recorded residence for each customer. This measure involves similar methodological limitations as the ecological crash variable discussed above.

Descriptive Statistics for Pilot, Baseline I, Baseline II, and Nearby Cohorts

Table 1 contains basic demographic data for the key control variables included in most of the analyses conducted for the present study.

Table 1

Descriptive Statistics for the Four Study Cohorts

	Cohort			
	Pilot	Baseline I	Nearby	Baseline II
Cohort start/end dates	June 2007 Oct. 2007	Sept. 2006 Jan. 2007	June 2007 Oct. 2007	Sept. 2005 Jan. 2006
Customer count	12,279	4,844	10,551	14,907
Mean age (*SD*)	55.7 (20.7)	42.1 (20.44)	55.8 (20.6)	52.5 (22.0)
Percent aged 70+	44.3	17.1	43.4	37.4
Percent male	52.1	58.3	51.5	53.2
Mean number of crashes, prior 3 years (*SD*)	0.21 (0.54)	0.27 (0.60)	0.18 (0.48)	0.26 (0.59)
Mean number of convictions, prior 3 years (*SD*)	0.62 (1.22)	1.09 (1.63)	0.54 (1.09)	0.73 (1.31)
Self-reported days of driving per week	5.28 (1.93)	N.A.	N.A.	N.A.
Mean number of valid days of licensure during two-year follow-up (*SD*)	695.25 (131.48)	717.51 (72.29)	714.34 (86.38)	705.92 (108.16)
Mean number of crashes (x100) per licensed driver per year in zip code of residence (*SD*)	3.95 (0.42)	3.95 (0.43)	3.56 (0.46)	3.94 (0.45)
Mean number of convictions (x100) per licensed driver per year in zip code of residence (*SD*)	16.46 (2.22)	16.71 (2.16)	15.95 (3.00)	16.52 (2.24)

Left out of Table 1 are a number of customers (n=1855) processed at the Pilot offices, during the months that the 3-Tier Pilot occurred (June-October 2007), who were not enrolled in the Pilot. These "Audits-only" customers were identified through the Audits database. They were, on average, substantially younger than Pilot customers (45.2 years), and somewhat more likely to be male (56.0%). Because this group likely includes customers who elected to take the written renewal test in a language other than English, or who were otherwise ineligible for enrollment in the Pilot, they were not included in any of the analyses discussed in the remainder of this monograph.

Methodological Implications of the Differences in Data Collection Methods

As noted in Camp (2010b), and replicated in Table 1, those customers comprising the Baseline I cohort were substantially and significantly different from customers in the Pilot cohort on several key demographic and behavioral variables. These differences appear to be associated with a quinquennial shift in the number of customers renewing their driver license in the field offices. This shift involves a substantial drop in customer volumes every fifth year subsequent to 1997, when CA DMV extended the standard license renewal term from 4 years to 5. Although Driver Safety referral customers and limited-term license holders were statistically similar across the two cohorts, this was not true for renewal customers with full-term licenses. For the latter type of customer, those in the Baseline I cohort tended to be younger (by about 14 years, on average), were more likely to be male, and had on average almost three times as many negligent operator "points" on their record. These three variables (gender, age, and prior violation record) constitute key control variables for the safety outcomes of interest for the current analysis. The measured differences between the two cohorts on these variables are quite large, and it is possible that these differences may be associated with other, unmeasured, differences between the two cohorts.[4] It was therefore determined at the time of the writing of the Process Analysis that it would be difficult to introduce appropriate statistical controls to eliminate likely sources of bias. The Baseline I cohort data were thus set aside, and will not form any part of the analyses conducted for the present report. Instead, it was deemed necessary to engage in post-hoc construction of additional, more statistically comparable, control groups (the Baseline II and Nearby cohorts).

[4]These unmeasured differences might, for instance, involve a greater proportion among the Baseline I cohort of customers who had recently moved to the area from out of state.

The three cohorts used for the present analysis (Pilot, Nearby, and Baseline II) were constructed using two very different methods. Pilot customers were identified during field office transactions, by field office and Driver Safety staff. Nearby and Baseline II customers were sampled from a source (the Audits database) that collects transaction information in a routine, automated manner based on computer data input by staff members as part of license renewal transactions. These two different data collection regimens involved certain implications for (a) available data used to construct variables, and (b) potential sources of bias.

The Audits database includes data used to construct four important variables for the Nearby and Baseline II cohorts: the number of written test attempts, the number of drive test attempts, the current license term (i.e., limited vs. full-term), and current license restrictions. By comparison with the Pilot cohort, certain key information is unavailable, or limited. In particular, it is not possible to verify in all cases the reason why a drive test may have been required of a customer. Also, the Audits database does not necessarily accurately record whether a given customer passed the written renewal test on the first or second attempt. For third (and subsequent) attempts, the accuracy of the data improves, largely because three consecutive failures of the written renewal test means that a customer is required to submit a new license application (and pay a new application fee). Finally, it is not really possible to accurately track the vision referral process via the Audits database for those customers who fail the visual acuity standard.

These differences aside, it bears emphasizing that the traffic safety and licensure outcomes of interest for the present analysis—crashes, convictions, retention of the driving privilege, and restrictions on the driving privilege—are based upon the driver record master file, and so are methodologically comparable across all study cohorts.

Apart from these differences in the amount, and quality, of data collected through these different methodologies, there are two potential sources of bias. In the first instance, the Audits database only imperfectly tracks the processing of Driver Safety referrals. By comparison with the Pilot cohort, both the Nearby and Baseline II cohorts have very few representatives of this type of customer. This prevents any direct estimation of the potential effects of the 3-Tier Pilot on the safety outcomes of Driver Safety customers specifically. However, given that Driver Safety customers constituted approximately 1% of the Pilot cohort, it is not anticipated that the absence of this group in the other two cohorts would substantially bias the size and direction of any estimated effects of the 3-Tier Pilot on safety outcomes overall.

Secondly, but more problematically, the Audits database only imperfectly records the language in which a given customer chose to take the written renewal test. Although the database contains several variables that may indicate the use of languages other than English when taking the written test, data-entry of this information (which occurs by field office staff during the course of the renewal transaction) is not required for the issuance of a license. Furthermore, it appears that this data field is not used as often as one might expect given other sources of information the department collects on language use in the field offices (unfortunately, these latter sources do not lend themselves to direct use in the present analysis, as they are reported in an aggregated format at the office level). It therefore appears that some unknown, but substantial, number of customers are included in the Audits-derived cohorts who may have taken the written renewal test in a language other than English. This appears most obviously in the fact that the Baseline II cohort includes about 20% more customers than the Pilot cohort, despite the fact that these samples were drawn from the same field offices, for a similar length of time (5 months). These differences may bias known demographic and behavioral variables used as statistical controls. In addition, they may bias the outcome variables of interest for this study. It is possible that this source of bias may affect the Baseline II cohort differently from the Nearby cohort; the offices serving the Baseline II cohort lie in neighborhoods of the greater Sacramento metropolitan area which, according to the 2010 US Census, contain somewhat higher concentrations of immigrant populations than is true for the Nearby cohort offices. The size and direction of this potential source of bias is not precisely known, but it does appear that the Baseline II cohort is slightly younger than the Pilot cohort (by about 3 years, on average), marginally more likely to be male, and somewhat more likely to have crashes and convictions on the driver record in the 3 years prior to enrollment in the study. These marginal differences across cohorts indicate the necessity of including all appropriate controls, to reduce the potential for bias in the analyses.

Other Methodological Threats and Sources of Potential Bias

In the Process Analysis, several potential threats to methodological validity were discussed. These included:

(1) The generalizability of results based upon a non-random sample of the population;
(2) Customer migration from Pilot to non-Pilot offices;
(3) The various errors in processing that occurred during the Pilot;
(4) The existence of customers who failed to renew their licenses by the end of the Pilot period;

(5) The comparability of the Pilot and Baseline I cohorts;

(6) The limitation of the Pilot to customers taking the written renewal test in English;

(7) The probable variation in the administration of the memory recall test; and

(8) The variation in the administration of the Pelli-Robson contrast sensitivity test.

In addition, four potential sources of bias were introduced by the creation of the Nearby and Baseline II cohorts:

(1) Non-concurrence in the observation period of post-participation driving (Baseline II vs. Pilot);

(2) Potential differences in driving environments (Nearby vs. Pilot);

(3) Potential remaining unmeasured differences in the cohort groups (Baseline II vs. Nearby vs. Pilot); and

(4) Potential overlap in customers among the three study cohorts (Baseline II vs. Nearby vs. Pilot).

Generalizability of the Results

As noted in the Process Analysis (Camp, 2010b), the 3-Tier Pilot incorporated a quasi-experimental method, with a purposive sampling frame. Because the population enrolled in various cohorts—Pilot, Baseline II, and Nearby—did not comprise a random sample of the population, any generalization of the results of the analyses in the present report must be severely qualified. This qualification certainly implicates the Process Analysis, and may implicate the present Outcome Analysis. The drivers enrolled in the various cohorts may differ in various ways from other drivers in California, in ways that cannot be specified using data collected as part of the Pilot. That said, there is no *a priori* reason to believe that the three study cohorts differ from otherwise 3-Tier eligible customers in other parts of the state in terms of the relationships among safe driving outcomes (crashes and violations), licensing outcomes (retention of, and restrictions on, the driving privilege), and those variables that predict these outcomes. In other words, there is no *a priori* reason to believe that the causal or correlative relationships among the variables of concern are of a different order or magnitude than would be obtained through a more rigorous method of random sampling. That said, this logical possibility cannot be excluded from consideration. Also, it is true that the customers enrolled in the study cohorts are, because of current CA DMV policies governing in-office license renewals, somewhat older, and somewhat more likely to have crashes and violations on their record, as compared to the general population of California drivers (Gebers & Peck, 1994). From a methodological perspective, this constrains some of the available variance in certain variables, especially age, prior crashes, and prior violations. Simply put, there are relatively few younger drivers with clean records in any of the study cohorts. We therefore cannot speculate on the effects the 3-Tier Assessment System might have on this type of driver.

Customer Migration

Several analyses were included in the Process Analysis to estimate the degree to which customers attempted to avoid participation in the Pilot, by conducting their transactions at non 3-Tier offices. It was found on the basis of these analyses that customer migration was quite uncommon. To the degree that customers were specifically identified as migrating from a Pilot to a non-Pilot office (i.e., they started their application at one office, but completed it at a different location), these were retained for the present analysis and form a subgroup of the 3-Tier eligible, erroneously-processed customers. While migrating customers appear to be somewhat different from other types of study participants (they tend to be younger, and more likely to have violations on their prior record), they form a relatively small proportion (<1%) of the overall Pilot population. In addition to this small group of customers specifically identified as having switched, it is possible that some unknown number of customers began their applications at an office to which they otherwise would not have gone, absent the Pilot. It is essentially impossible to determine the number of such customers (much less their demographic characteristics), and so impossible to determine whether or not their existence biases the results of the present analysis. If a customer began their license application process at a non 3-Tier office (including any Nearby cohort office), and then migrated to a Pilot office, they were excluded from the analysis as ineligible for inclusion in the Pilot. There were only a handful of such cases.

Errors in Processing

In the Process Analysis, three sub-groups of erroneously-processed customers were discussed in terms of the potential for methodological invalidity they represent. These three groups included: non 3-Tier eligible customers, customers with missing documents or data, and customers whose license renewal application was improperly processed in some way.

Customers who were ineligible for inclusion in the Pilot included those who began their applications prior to the start of the Pilot (6/1/07), those whose most recent driver license included a commercial or motorcycle class, those who took the written renewal test in a language other than English, and those who began their applications at a non 3-Tier office. All of these Pilot cohort customers have been excluded from the present analyses.

Customers in the Nearby and Baseline II cohorts who possessed similar indications of ineligibility for inclusion in the Pilot have also been excluded from the present analyses.

Secondly, several hundred customers were identified in the Process Analysis with missing documents or data. At that time, it was assumed that this group would need to be excluded from the present analyses. However, in the course of the construction of the Audits-derived databases (Nearby and Baseline II), it was determined that some of these missing data values could, in fact, be reasonably reconstructed for this group of Pilot customers. This was particularly true for customers with missing results for the written renewal test, and customers with missing results for an on-road drive test. In both cases, the Audits database contains information usable for the imputation of these variable values. The remaining customers with unimputable missing data were retained in the analysis as 3-Tier eligible, but an additional set of variables was created to flag the error(s) in processing.

Among the Nearby and Baseline II cohorts, there is no parallel group of customers with missing data, though there do exist customers with incomplete applications (discussed in the next sub-section, on lagging applicants).

Thirdly, there remained those customers who, in some manner, were not processed according to the training protocols developed for the Pilot. There were three main types of such customers: those who received (or failed to receive) the educational intervention in contravention of the protocols governing its distribution, customers who were not given the PRT, and customers who were re-licensed without an on-road drive test that was otherwise required according to 3-TAS protocols.

Because the data regarding the distribution of the educational materials is too corrupt to be reliable, it simply cannot be known who received the educational intervention and who did not. Therefore, this information was left unmarked in the present analysis, and consequently constitutes a potential source of unestimated and uncontrollable bias to the findings. It is entirely possible that the widespread distribution of these materials to somewhat functionally limited drivers improved their driving competency. However, there is no way to way to measure the size of this potential effect, if it in fact exists.

Customers who did not take the PRT, even though required to do so according to Pilot protocols, were retained for the present analysis and included in the larger group of 3-Tier eligible, but erroneously-processed, customers.

Customers who were re-licensed without a required on-road test of driving skill were also retained for the present analysis and included in the larger group of 3-Tier eligible, but erroneously-processed, customers.

Customers in the Nearby and Baseline II cohorts were, by definition, processed correctly according to current (non 3-Tier) policies and procedures. If there exist errors in processing (such as, for instance, issuance of a license despite failure on one or more screening tests), these are not visible using available variables in the Audits database.

Customers with Lagging Applications

Several hundred customers enrolled in the Pilot failed to complete their license renewal applications by the end of the Pilot period (10/31/07) or by the end of the 2 months allowed for follow-up processing (so, by 12/31/07). As noted in the Process Analysis, these customers were somewhat different from other Pilot customers with completed applications. They were retained for the present analysis in part to control for potential sources of bias. They were also retained in order to identify more accurately the potential effects of the Pilot on licensing outcomes, including delays to licensure and voluntary (informal) delicensure.

A similar set of customers in the Nearby and Baseline II cohorts were identified for comparison. These included drivers who started a license application within the cohort period (6/4/07-10/31/07 for the Nearby cohort, 9/1/05-1/31/06 for the Baseline II cohort), but who failed to complete their application within the 2 months allowed for follow-up processing (so, by 12/31/07 for the Nearby cohort, and 3/31/06 for the Baseline II cohort).

Comparability of the Baseline I and Pilot Cohorts

At the time of the writing of the Process Analysis, it was anticipated that it would be necessary to introduce statistical controls to account for existing demographic and behavioral differences between the Baseline I and Pilot cohorts. After the publication of that report, it was determined that it was possible to construct additional comparison groups—the Baseline II and Nearby cohorts—with more narrowly similar demographic and behavioral profiles. The post-hoc construction of these samples eliminated the necessity of relying upon the Baseline I cohort for the estimation of any potential safety effects associated with the Pilot.

The Limitation of the Pilot to Customers Taking the Written Renewal Test in English

As noted in the Process Analysis, it is difficult to determine whether, to what degree, and in what direction, the limitation of the Pilot to English-speaking customers may bias the present analysis. Although there is some research on the relationship between race/ethnicity and the kinds of driving-relevant conditions flagged by the contrast sensitivity test (diabetes, macular degeneration, cataracts, etc.) or between race/ethnicity and the kinds of conditions flagged by the various cognitive screening tests (e.g., dementia disorders), this research has tended to focus on black/white racial differences rather than Hispanic/non-Hispanic or white/Asian ethnic and racial differences. The author knows of no studies that incorporate immigration status as a variable in examining cross-racial/ethnic differences in these kinds of conditions. Data were not collected on customers' race/ethnicity, or their immigration status, for this project. Therefore any potential effects of these variables on the findings contained herein must remain unestimated.

It is possible, due to the data-collection methods used to construct the Nearby and Baseline II cohorts, that these groups contain a higher proportion than the Pilot cohort of customers who took the written renewal test in a language other than English. The Audits database contains certain information noting the language in which a customer took their written renewal test. However, there is reason to believe that these data-fields are imperfectly populated. Specifically, there is reason to believe that persons who take the written renewal test in a non-English format are not always flagged in the database as having done so. However, it does not appear that the reverse occurs; in other words, it does not appear that English-language test takers are flagged as having taking a foreign-language test. To the extent that the available evidence suggests that a given customer took the written renewal test in a non-English format, they were excluded from the Nearby and Baseline II cohorts.

The Probable Variation in the Administration of the Memory Recall Test

On the basis of staff interviews reported in the Process Analysis Appendix (Camp, 2010a), there exists evidence of some degree of variation in the administration of the memory recall test. This variation may have resulted in a net lower fail rate on this screening test than would have occurred otherwise. In other words, it may be that relatively fewer people were assessed at Tiers

2 and 3 than might have occurred under other circumstances. Statistically, this variation may have weakened the relationship between outcomes on this screening test and safety outcomes such as crashes. By extension, this variation may have weakened the relationship between the combined 3TAS score (of which this screening test formed a part) and measured safety outcomes. Furthermore, this variation might have weakened any intervention effect of the pilot on safety outcomes, since customers who did not fail this test (but should have) were consequently less likely to be given education about safe driving, or be subject to a licensing action. However, there exists no baseline parameter against which to compare the fail rate for this screening test, as found in the 3-Tier Pilot. It is therefore impossible to estimate the size or impact of this potential source of bias. There exists no way to identify a parallel group of customers in either the Nearby or Baseline II cohorts.

The Variation in the Administration of the Pelli-Robson Contrast Sensitivity Test

On the basis both of staff interviews and statistical modeling reported in the Process Analysis Appendix, there exists evidence of variation in the administration of the Pelli-Robson contrast sensitivity test. The net effect of this variation appears to have been a lower fail-rate on this screening test than would have occurred otherwise. This in turn suggests that some unknown number of customers were not assessed at Tiers 2 and 3, despite the existence of driving-relevant potential limitations in their vision. If there exists a group of such customers, they were by definition not given the educational materials on this subject; hence they may not have benefitted from whatever safety effects derive from participation in the more advanced assessment tiers of the 3-Tier Pilot. As with the variation in the implementation of the memory recall test, this may have statistically weakened the association between outcomes on this screening test and measured crashes, statistically weakened the association between the combined 3TAS score and measured crashes, and weakened any intervention effect of the Pilot on safety outcomes. However, it is not possible to identify retroactively those customers who should have been assessed at Tiers 2 and 3, due to improperly passing the Pelli-Robson contrast sensitivity test. Furthermore, it is not possible to identify a parallel group of customers in either the Nearby or Baseline II cohorts. As a result, this potential source of methodological bias must remain unestimated and uncontrolled.

Non-Concurrence in the Post-Participation Observation Period (Baseline II vs. Pilot)

The Baseline II cohort was drawn from a group of customers processing their renewal applications approximately 18 months prior to the Pilot cohort. For both cohorts, the post-participation follow-up period consisted of an identical amount of time (2 years, though in some individual cases this may have been somewhat less due to the death of the driver, or a change in license class). However, these follow-up periods only partially overlap. It is therefore possible that non-Pilot related changes to road safety may affect crash rates in ways that would appear as Pilot-associated safety effects. Unfortunately, these effects cannot be statistically disentangled from any estimated differences between the Pilot and Baseline II cohorts precisely because of the almost total lack of overlap between the two observation periods. Simply reducing the observation period to that short period of overlap between the two cohorts would introduce the additional difficulty of comparing one cohort who had recently renewed their licenses (Pilot) with another cohort whose renewal experiences were almost 2 years in the past (Baseline II). In sum, it must be noted that if the analysis estimates suggest any safety savings associated with comparing the Baseline II and Pilot cohorts, these safety savings may be attributable to unmeasured non-Pilot dynamic changes to the driving environment.

Potential Differences in the Driving Environment (Nearby vs. Pilot)

The Nearby cohort consisted of customers processing their renewal applications at offices different from those used by Pilot cohort customers. In some cases these offices were approximately 100 miles distant from each other. Customers using offices at such distances from each other are likely to live and work in very different areas, and hence are likely to encounter quite different driving environments as part of their normal road usage. This may in turn have an effect on their risk of crashing, as well as other sorts of safety outcomes discussed in the present analysis. To the extent possible, this is controlled for in the present analysis through the use of the ecological crash- and conviction-exposure variables discussed above (p. 18). However, this may not completely capture differences in exposure that exist between the two cohorts. Therefore, any findings related to potential safety savings associated with comparing the Nearby and Pilot cohorts may be biased by unmeasured non-Pilot static differences in the driving environment.

Potential Remaining Unmeasured Group Differences (Baseline II vs. Nearby vs. Pilot)

Every effort was made to include available variables that might control for between-group differences among the three cohorts. However, due to the non-randomized nature of group assignment, there likely remain unmeasured differences across groups. This is especially true for comparisons between the Pilot and Nearby cohorts, which include drivers living in very different neighborhoods (and cities) across a relatively wide swath of Northern California. This is also perhaps true for comparisons between the Pilot and Baseline II cohorts; although these customers processed their applications at the same offices, there exists evidence of remaining unmeasured between-group differences (see Table 1, p. 19). Because these differences are unmeasured, they cannot be controlled for, and so remain potential sources of bias to any inter-cohort comparisons contained in this monograph.

Potential Overlap in Customers Among the Three Study Cohorts (Baseline II vs. Nearby vs. Pilot)

The sampling of the three cohorts was done in such a way as to minimize the number of times individual customers appeared twice in different groups. Still, a handful of customers initially appeared as duplicates. In general, if a customer appears in the Pilot cohort, they were counted as a Pilot customer and deleted from the other datasets if they appeared there as well. If a customer appeared in both the Nearby and Baseline II datasets (but not in the Pilot cohort), they were deleted from the Nearby cohort and retained as a Baseline II customer.

RESULTS

Question #1: What are the Costs Associated with the ADPE?

In the Process Analysis (Camp, 2010b) it was noted that, for at least two reasons, it was difficult to determine if the 3-Tier Pilot resulted in an increase in the number of administered Area Driving Performance Evaluations (ADPEs). Eighteen such tests were conducted during the Pilot versus one ADPE test during Baseline I. Although this appears to be a substantial difference, there also exists a substantial difference in the mean age between Baseline I cohort and Pilot cohort customers. The ADPE is typically administered to customers of relatively advanced age. Therefore, the simple fact that Pilot cohort customers were, on average, almost 14 years older than Baseline I cohort customers could have accounted for some unknown portion of the difference between the two cohorts in the number of administered ADPEs. Second, data collected as part of the staff interviews suggested that at least some of the Pilot cohort customers who took an ADPE drive test might have taken such a test even if they hadn't been participating in the Pilot. In a related problem, it was also difficult to determine if there existed a reliable difference in the length of the ADPE tests administered during the Baseline I as compared to those administered during the Pilot. For these reasons, it was assumed in the Process Analysis that the 3-Tier Pilot did not result in any excess costs associated with the administration of additional ADPE tests.

It is possible to construct two additional estimates of the number of potential additional ADPEs associated with the 3-Tier Assessment System, using data drawn from the Audits database for the Baseline II and Nearby cohorts. Both the Baseline II and Nearby cohorts have mean population demographics, including age, that are relatively similar to those of the Pilot cohort (see Table 1, p. 19). Table 2 displays the number of customers within each cohort (Pilot, Baseline II, and Nearby) who were issued a license with an area restriction, and the number of drive tests administered to these customers. For each cohort, the number of restrictions, and the number of drive tests, is then converted to a rate to adjust for differences in the size of each cohort.

These data should be interpreted with a great deal of caution. First, the estimation of the number of customers in the Pilot cohort with an area restriction according to Audits differs slightly from the estimation provided in the Process Analysis, which was based upon physical paperwork submitted during the course of the Pilot. Specifically, in one instance a customer took (and passed) an area drive test; a temporary license was then issued with three restrictions: requiring corrective lenses, prohibiting freeway driving, and prohibiting driving at night. An area restriction was also requested on the advice of the drive-test examiner. Because area restrictions are customized, they require non-routine processing. As a result, these requests are handled at CA DMV headquarters rather than at the field office. They therefore may take some days or weeks to be added to a customer's permanent driver record. Thus, although 14 customers exited the Pilot with an area restriction, the information provided by the Audits database indicates that only 13 did so.

Table 2

Area Restrictions Assigned and Associated Drive Tests (Pilot, Baseline II, and Nearby Cohorts)

	Pilot	Baseline II	Nearby
Total *N* of 3-Tier eligible customers.	12,279	14,907	10,551
Total *N* of customers with an area restriction, according to the Audits database.	13	2	1
Area restriction rate, per 10,000 customers.	10.6	1.3	1.0
Total *N* of drive tests administered, according to the Audits database.	859	341	289
Total *N* of drive tests given to customers with area restrictions, according to the Audits database.	17	2[1]	0[2]
Imputed drive test rate for customers with area restrictions, per 1000 drive tests.	19.8	5.9	0.0

[1]Both of these tests were administered to a single customer, who was undergoing a Driver Safety re-examination as part of a reinstatement proceeding.
[2]This person was not required to take a drive test when they renewed their license.

This same problem represents particular difficulties for the Baseline II and Nearby cohorts, where no physical paperwork exists independent of the Audits database. Although it is possible to obtain the current restriction status for drivers in these cohorts, it is simply not possible to determine when restrictions were applied to the license. An additional search was conducted to examine those customers who, according to the Audits database, possessed license restrictions including code 02 (no freeway driving). This code is required to be added whenever an area restriction is given. Code 02 can be added at the time of the issuance of the temporary license (i.e., in the field office), and therefore may be a precursor to the addition of an area restriction. No additional customers with area restrictions in the Nearby cohort were identified by this method. There are two additional customers in the Baseline II cohort who were identified by this method. Both of these customers took drive tests at the time they renewed their application; in both cases these drive tests appear to be associated with the possession of a limited-term license. However, in neither case can it be determined with full confidence when the area restriction was placed on their record, especially as there were at least two license renewal cycles for these customers between when the Audits database was constructed (2005/6) and when their records were checked for an area restriction (2011).

Second, the total number of drive tests reported in the Audits database is somewhat lower than the total number of drive tests administered, according to the paperwork submitted during the 3-Tier Pilot. If one includes all 3-Tier eligible customers—such as those erroneously processed, or with lagging applications—and uses information gathered in paper format, then 912 drive tests were administered to 3-Tier Pilot cohort customers. As noted in Table 2, for these same customers the Audits database records 859 drive tests. This represents a potential undercount of 53 drive tests, or approximately 6%. To some degree (but not entirely) this discrepancy is explained by Driver Safety referral customers, whose drive tests were not typically captured in the Audits database used for this project. The Audits database was constructed to capture driver license renewal transactions; most Driver Safety referral customers were not renewing their license. In three cases, customers who took area drive tests in the Pilot were enrolled because of a Driver Safety procedure; in two of these, the customers' drive tests were not recorded in the Audits database. Those customers in the Baseline II and Nearby periods who were assigned an area restriction as part of a Driver Safety procedure may not be captured in the Audits database, and therefore may not be included in the calculations presented in Table 2.

Third, because of the way the Audits database records data, it is not technically possible to determine what form of drive test was taken (DPE, SDPE, or ADPE)—only that a drive test was administered, whether it was for a commercial or a non-commercial license, and the eventual outcome of the test (pass versus fail). During the Pilot, it is known on the basis of paperwork that at least two customers took (and passed) an ADPE test, subsequent to which they decided to take an SDPE test. Because these customers exited the program without an area restriction, absent additional information it would appear that they did not take an ADPE test, according to the Audits database. Thus, although 19 total ADPE drive tests were conducted according to Pilot paperwork, according to Audits it would appear that 17 drive tests were given to customers who exited the program with an area restriction. This includes six cases where a customer first failed an SDPE test, then took an ADPE.

Taken together, these methodological limitations suggest that the Audits database, as used for this analysis, only imperfectly records the issuance of area restrictions, and only by a very flawed method of imputation can the number of ADPE drive tests be estimated. That said, it is not clear that the methodological limitations of estimating ADPE volumes from the Audits database would affect the three cohorts in substantially different ways. In other words, while it is probably the case that using the Audits database produces an imprecise count of ADPE drive tests, there is no obvious reason why this imprecision should affect one cohort but not another.

In sum, while using the Audits database to estimate the number of ADPE tests given during the Pilot is imperfect, nevertheless it appears that the Pilot was associated with some increase in the number of ADPE tests given, relative to the number of such tests administered to customers in the Nearby and Baseline II cohorts. This increase was quite small in terms of absolute number. Given the rarity of such tests, it is difficult to construct a reliable estimate of the expected increase in the number of ADPE drive tests if the 3-Tier Pilot is adopted for statewide implementation. Nevertheless, it is possible to construct an estimate of the costs associated with the ADPE, using similar methods as contained in the Process Analysis (Camp 2010b) to estimate the costs associated with other components of 3TAS.

As reported in the Process Analysis, of the 19 ADPE tests conducted according to Pilot paperwork, 14 were administered to renewal customers and 5 were administered to Driver Safety referrals. It is not anticipated that the implementation of 3TAS would substantially affect the number of ADPE tests associated with Driver Safety referral processing. Therefore, the estimation of costs associated with the ADPE will be based on the number of tests taken by renewal customers. Of the 14 ADPE tests administered to renewal customers, half were given to customers who already possessed a limited-term license (and so were required to take a drive test in any case), and half were given to customers whose drive test was specifically a result of 3-Tier processing. It is possible that some of the limited-term customers may have chosen an ADPE, rather than an SPDE, in part because of their participation in the Pilot. However, it is also possible that some of these limited-term customers would have taken an ADPE regardless of their participation in the Pilot. This necessitates the use of a range in the estimation of additional ADPEs resulting from the 3-Tier Pilot, and their associated costs. This range will assume, at the low end, that all limited-term customers who took an ADPE in the 3-Tier Pilot would have taken such a test regardless of their participation in the Pilot. At the high end, this range will assume that these customers would have taken an SDPE absent their participation in the Pilot, and therefore all limited-term ADPE tests in the 3-Tier Pilot were in excess of what would have occurred had the Pilot not been implemented. This range is also affected by the cohort comparison; according to the Audits database, while two ADPE tests were affirmatively conducted among the Baseline II cohort, no ADPE tests were conducted among the Nearby cohort. For the purposes of costing, it will be assumed that an additional two customers in the Baseline II cohort whose area restriction appeared sometime subsequent to their license renewal in 2005/6 in fact took SDPE tests at the time of their participation in Baseline II; this has the effect of slightly inflating the assumed Pilot-associated increase in ADPE tests. Taken together, these assumptions suggest that the 3-Tier Pilot was associated with somewhere between 5 and 14 excess ADPE drive tests.

During the Pilot, Licensing Registration Examiners (LREs) were asked to record the amount of time it took to conduct an area drive test. These estimates included the time the LRE spent traveling to and from a customer's house, the time spent greeting a customer and discussing the area routes (or boundaries), and the time spent on the actual drive test itself. The average amount of time required to conduct an ADPE test varied somewhat according to customer type (limited-term versus regular-term renewal), but these differences are likely a product of (a) the small number of tests given, and (b) the inherently large variation in the length of an ADPE test. For the purposes of costing, the mean length of an ADPE drive test is assumed to be 110 minutes, the average of those tests given to regular-term renewals in the Pilot.[5]

Because the Pilot occurred over a 5-month period, it will be assumed that between 1 (5/5) and 2.8 (14/5) ADPE tests were given per month. If each test took an average of 110 minutes, this would yield between 110 and 308 minutes—or between 1.83 and 5.13 hours per month—of additional LRE time required to conducted 3TAS-associated ADPE tests during the Pilot. Dividing by the number of work hours in an average month for an employee of the State of California (173.33) yields the following estimate of the number of fulltime equivalent (FTE) positions required for the six field offices that participated in the 3-Tier Pilot: 0.01 to 0.03 FTEs.

In order to extrapolate these estimates statewide, two potential methods were used in the Process Analysis. The first method assumes that the six Pilot field offices were roughly representative of all 169 CA DMV offices; therefore, if we multiple the number of FTEs by a ratio of 169/6 (or 28.2), this would suggest that approximately 0.30 to 0.83 FTE LRE positions would be necessary statewide to handle the additional time associated with 3TAS-generated ADPE tests. The second method used a ratio of the total number of DL transactions processed statewide over the total

[5]The average length of a limited-term renewal ADPE drive test was 99 minutes.

number of DL transactions processed in the six Pilot offices. This ratio (24.5) would suggest that approximately 0.26 to 0.72 FTE LRE positions would be necessary statewide to handle the additional time associated with 3TAS-generated ADPE tests. The estimates yielded by these two methods are very close; therefore, the first method (based on the ratio of Pilot offices to the number of offices statewide) will be used to estimate costs.

It is also necessary to adjust these estimates for the increase in customer load expected to result from the extension of 3TAS to all languages. This factor, as detailed in the Process Analysis, was 1.27. This increases the estimated staffing to between 0.38 and 1.06 FTE.

To convert these estimates of staff positions to dollars, the mean salary (and associated benefits) for an LRE was taken from the 2009/2010 fiscal year: $4,329/month.[6] This yields an expected monthly cost of additional LREs necessary to conduct 3TAS-associated ADPE tests of $1,645-$4,589. Converted to an annualized basis, this suggests that the ongoing costs associated with 3TAS-generated ADPE tests would be between $19,740 and $55,068 statewide per year. It is not anticipated that the costs associated with conducting the ADPE would include anything other than LRE staff time.

<u>Question #2: How Willing are Customers to Pay an Extra Fee for the ADPE?</u>

During the 3-Tier Pilot, customers who were required to undertake a road test had an extensive interview with a manager or drive-test examiner prior to the test. This interview covered a number of topics, the most important of which involved (a) explaining the reasons why the road test was required of that customer, and (b) sharing with the customer those educational materials that had been developed to help the customer improve their driving skills and prepare for the test.

As part of the interview, the manager or drive-test examiner was trained to ask the customer a series of questions about the customer's driving habits in order to determine which type of drive-test was most appropriate for the customer's circumstances. If the customer was scheduled for an ADPE, the manager was trained to ask the customer the following question:

> "You are being considered (or scheduled) for an Area driving test. Currently there is no additional cost to you to take an area drive test. This drive test requires 4 times as much of an examiner's time as does the SDPE, i.e. 2 hours versus 30 minutes. Before you

[6]This does not include reduction for furloughs that were specific to that year's fiscal emergency. The benefit rate was set at 46.762%.

answer, keep in mind that taking an Area drive test does not guarantee that you will pass it or be licensed. If you were required to pay more for an area drive test, would you pay $125, $100, $75 or No More?"

As a general matter, there is a great deal of missing data in the customer responses to this question. Across all drive test attempts, in only one third (6 out of 19) of ADPE drive test attempts is there an answer recorded for the customer. On the other hand, there are an additional 39 answers recorded among customers who took an SDPE. These data are displayed in Table 3.

Table 3

Willingness of Pilot Customers to Pay Additional Fees for an Area Drive Test

Drive test type (total *N*)	No answer recorded (% of total)	Pay no more ($0) for an area test	Pay $75 more (%)	Pay $100 more (%)	Pay $125 more (%)
ADPE (19)	13 (68.4)	5 (26.3)	1 (5.3)	0 (0.0)	0 (0.0)
SDPE (948)	909 (95.9)	37 (3.9)	2 (<0.1)	0 (0.0)	0 (0.0)

Note. Includes all drive-test attempts (first, second, and third). Individual customers may have been asked this question more than once; where such instances occurred, the answers are counted separately.

Given the amount of missing data, as well as the fact that the majority of people who answered the question were not, in fact, being scheduled for an ADPE, these data should be interpreted with a great deal of caution. It is simply unknown what most drive-test customers, even those customers scheduled for an ADPE, would be willing to pay if required to pay an additional fee to take this type of test. At most, these data suggest that very few Pilot cohort customers, when asked whether they would be willing to pay more for an ADPE, signaled a willingness to pay even as much as $75 more for such a test.

Question #3: How Many Customers Possessed Limitations, and What Percentage of Those With Limitations Retained Their Driving Privilege?

The driving-relevant limitations identified through 3TAS came in three forms: physical, cognitive, and visual. Each of these forms of limitation was identified through the use of multiple screening tools. Physical limitations, for instance, were identified through a two-stage observation process. The first stage focused on the identification of limitations in the upper body (arms, trunk, and hands), while the second focused on the identification of limitations in the lower body (legs and feet). Cognitive limitations were identified through the use of three screening tools: the memory recall test, the written knowledge test, and the PRT. Visual limitations were identified through the department's regular visual acuity standard (Snellen chart) and the Pelli-Robson contrast sensitivity chart.

The tables in this section display information regarding the proportion of customers within the Pilot cohort who were identified as having a given form of driving-relevant limitation. For each form of limitation, the table also includes what proportion of customers retained their driving privilege. The retention of the driving privilege is measured two ways: (a) the proportion possessing a valid license as of the final date of the follow-up period (2 years from the date that the renewal application was begun, or from the first drive test required by the Driver Safety Office, as appropriate to the individual customer), and (b) the mean number of days between the license renewal date and the date of licensure. If a customer failed to renew their driving privilege during the follow-up period, their data was left missing on this second measure, and so not included in the calculation of the mean time to renew.

As discussed in the Process Analysis (Camp, 2010b), a substantial number of customers were erroneously processed in some manner. Typically, this took the form of missing documents which contained key information not imputable from other sources. In addition, a substantial number of customers did not complete processing before the termination of the Pilot. In order to capture the full magnitude of any potential effects of the Pilot on license retention, it is necessary to include these customers in the analysis. For the purposes of comparison, the cells titled "correctly processed" *exclude* erroneously-processed and lagging customers, while the cells titled "3-Tier eligible" *include* both erroneously-processed and lagging customers.

It must be noted here that the figures reported in both the "correctly processed" columns and the "3-Tier eligible" columns differ slightly from those previously published in the Process Analysis (Camp, 2010b). Through the construction of the Audits database—which occurred after the publication of the Process Analysis—additional data were made available that altered the designation of customers as either correctly processed, incorrectly processed, or ineligible for inclusion in the Pilot. For instance, some customers that were thought to be correctly processed were, on the basis of information contained in the Audits database, discovered to have started their renewal applications prior to the beginning of the Pilot. These customers were therefore (newly) designated as ineligible for the 3-Tier Pilot. These customers are not included in the tables below. On the other hand, some customers who had been designated as incorrectly processed had missing data that was, on the basis of information contained in the Audits database, amenable to imputation. This occurred most often with the recording of written test results. These customers are included in the tables below. The reclassification of customers into different categories occurred in only a small number of cases, and therefore does not materially change the substantive findings of the Process Analysis.

Licensure Outcomes and Screening for Driving-Relevant Physical Limitations

Table 4 provides a cross-tabulation of the proportion of customers remaining unlicensed, and the mean number of days between the start of the application and the renewal of the driving privilege, according to the number of physical limitations observed at Tier 1. Due to the relatively small number of such observed limitations, it is not meaningful to provide a breakdown according to specific type (e.g., cannot walk unaided vs. uncontrollable shaking in the hands). Because Driver Safety referral customers were not given the Tier 1 tests, they are not included in the results of Table 4.

Among correctly-processed customers, fewer than a dozen remained unlicensed at the end of the follow-up period. The proportion unlicensed rises with the number of observed physical limitations, though this proportion never exceeds 5%. That said, the amount of time required to renew the driving privilege rises substantially with the number of observed physical limitations. This rise is largely associated with the drive test requirement, and the scheduling of such tests, sometimes days (or even weeks) in the future. Because so few correctly-processed customers failed to renew their driving privilege—by definition, one of the main criteria for "incorrect processing" was failure to complete the renewal application—the remainder of this section will focus on the licensure outcomes among the 3-Tier eligible pool of customers.

Among all 3-Tier eligible customers—including those whose applications were erroneously processed, or who failed to complete their application by the end of the Pilot period—the proportion remaining unlicensed rises with the number of observed physical limitations. Slightly more than one out of every five customers who had at least two observed physical limitations remained unlicensed at the end of the follow-up period. In a parallel manner, the number of days required to renew the driving privilege rises with the number of observed physical limitations.

Table 4

Observed Physical Limitations Among Pilot Renewal Customers, and Proportion Retaining Their Driving Privilege

	Observed physical limitations (*N* of customers)	Number unlicensed[1] (%)	Mean days to licensure[2] (*SD*)
Correctly-processed	0 limitations (10,389)	6 (0.1)	4.9 (16.6)
	1 limitation (395)	3 (0.8)	17.2 (36.7)
	2+ limitations (40)	2 (5.0)	58.1 (115.5)
	Total unlicensed/overall mean	11 (0.1)	5.5 (19.3)
	Effective *N*	10,824	10,813
3-Tier eligible	0 limitations (11,585)	283 (2.4)	14.1 (53.2)
	1 limitation (498)	40 (8.0)	24.9 (57.8)
	2+ limitations (73)	16 (21.9)	74.3 (128.8)
	Total unlicensed/overall mean	339 (2.8)	14.8 (54.2)
	Effective *N*	12,156	11,817

[1]The number unlicensed includes any customer whose driving privilege was expired, suspended, or revoked for the entire two-year follow-up period following their date of first contact.
[2]The mean days to licensure only includes those customers who eventually renewed their driving privilege.

Licensure Outcomes and Screening for Driving-Relevant Cognitive/Perceptual Limitations

Table 5 provides the same information as Table 4, but for outcomes on the memory recall test. Again, because Driver Safety referral customers were not given the Tier 1 tests, they are not included in the results of Table 5.

As noted in the Process Analysis (Camp, 2010b), relatively few customers—approximately 1%, among correctly processed customers—failed the memory recall test. This low fail rate may, in part, have derived from skepticism on the part of staff regarding the utility, and traffic safety relevance, of this test. Although this low fail rate introduces some doubt regarding the uniformity with which this test was applied, a similar pattern emerges in Table 5 as in Table 4: those who failed this test were somewhat less likely to retain their driving privilege, and took somewhat longer to renew their licenses. That said, among those who failed the memory recall test, nearly 90% of 3-Tier eligible customers eventually renewed their driving privilege.

Table 5

Memory Recall Test Outcomes Among Pilot Renewal Customers, and Proportion Retaining Their Driving Privilege

	Memory recall test (*N* of customers)	Number unlicensed[1] (%)	Mean days to licensure[2] (*SD*)
Correctly-processed	Pass (10,706)	8 (0.1)	5.4 (18.8)
	Fail (118)	3 (2.5)	16.1 (46.5)
	Total unlicensed/ overall mean	11 (0.1)	5.5 (19.3)
	Effective *N*	10,824	10,813
3-Tier eligible	Pass (11,981)	320 (2.7)	14.5 (53.8)
	Fail (175)	19 (10.9)	35.1 (77.8)
	Total unlicensed/ overall mean	339 (2.8)	14.8 (54.2)
	Effective *N*	12,156	11,817

[1]The number unlicensed includes any customer whose driving privilege was expired, suspended, or revoked for the entire two-year follow-up period following their date of first contact.
[2]The mean days to licensure only includes those customers who eventually renewed their driving privilege.

The second component of the cognitive screening consisted of the 18-question written renewal test. Based in part on prior research (Janke & Hersch, 1997), it was conjectured that multiple failures on the written test may constitute a potential sign of cognitive limitations. The relationship between outcomes on this test, and retention of the driving privilege, are displayed in Table 6. Although Driver Safety referral customers were usually required to take the written test as part of their processing, the results of these tests were not routinely recorded as part of their 3-Tier records. They are therefore not included in Table 6. However, a handful of 3-Tier eligible customers whose Tier 1 scores were unimputable (and so were excluded from Tables 4 and 5) had written test results that were imputable through the Audits database. They are therefore included in Table 6, below.

Table 6

Written Renewal Test Attempts Among Pilot Renewal Customers, and Proportion Retaining Their Driving Privilege

	Written test attempt (*N* of customers)	Number unlicensed[1] (%)	Mean days to licensure[2] (*SD*)
Correctly-processed	Passed on 1st or 2nd attempt (10,376)	10 (0.1)	4.5 (17.7)
	Passed on 3rd attempt (345)	1 (0.3)	23.7 (29.4)
	Required at least 4 attempts (103)	0 (0.0)	50.0 (37.5)
	Total unlicensed/overall mean	11 (0.1)	5.5 (19.3)
	Effective *N*	10,824	10,813
3-Tier eligible	Passed on 1st or 2nd attempt (11,226)	211 (1.9)	10.7 (47.0)
	Passed on 3rd attempt (598)	82 (13.7)	60.8 (94.6)
	Required at least 4 attempts (336)	44 (13.1)	90.0 (101.6)
	Total unlicensed[3]/ overall mean	339 (2.8)	14.8 (54.3)
	Effective *N*	12,162	11,823

[1]The number unlicensed includes any customer whose driving privilege was expired, suspended, or revoked for the entire two-year follow-up period following their date of first contact.
[2]The mean days to licensure only includes those customers who eventually renewed their driving privilege.
[3]This includes two customers who did not attempt the written test at all; both failed to renew their driving privilege in the follow-up period.

As with previous screening tests, it is clear that multiple attempts on the written test are associated with a higher likelihood of delicensure, as well as with a lengthened time necessary to renew the driving privilege. The relationship between failure on the written test and days to renew is especially strong: 3-Tier eligible customers who had to make three attempts to pass the test took at least 2 months to renew, while those who had to make at least four attempts took, on average, 3 months to renew their driving privilege. This is probably in part a product of customers taking the time to study the *California Driver Handbook*, in preparation for taking the test a second, third, or even fourth time. Among those who had to attempt the written test four or more times, the length of time to renew the driving privilege was also likely in part a product of the drive-test requirement. That said, even among those 3-Tier eligible customers who failed the written test multiple times, six out of every seven customers eventually renewed their driving privilege.

The final component of the cognitive screening tests involved the PRT. These results are displayed in Table 7. Driver Safety referral customers were not required to take the PRT, and so are excluded from this analysis. For the purposes of these calculations, a "somewhat fail" (SFail, or a score of 24-40), an "extreme fail" (XFail, or a score between 41 and 500), and an "abort" (a score greater than 500) were combined. The effective N of this table differs slightly from those reported in the tables above, due to missing data.

Consistent with prior findings, customers who failed on the PRT were somewhat less likely to renew their driving privilege. Because failure on the PRT resulted in the customer having to take an on-road test, it is not surprising that the number of days to renew increases substantially among those who could not pass this screening test. Even among those who failed this test, however, about 85% of 3-Tier eligible customers eventually renewed their driving privilege during the follow-up period.

Table 7

PRT Outcomes Among Pilot Renewal Customers, and
Proportion Retaining Their Driving Privilege

	PRT outcome (*N* of customers)	Number unlicensed[1] (%)	Mean days to licensure[2] (*SD*)
Correctly-processed	Not required (8,541)	0 (0.0)	3.4 (13.5)
	Pass (2,009)	4 (0.2)	10.1 (26.8)
	Fail (276)	7 (2.5)	38.3 (54.4)
	Total unlicensed/overall mean	11 (0.1)	5.5 (19.6)
	Effective *N*	10,826	10,815
3-Tier eligible	Not required (9,487)	218 (2.3)	12.7 (53.2)
	Pass (2,234)	54 (2.4)	18.1 (53.2)
	Fail (443)	67 (15.1)	49.2 (74.5)
	Total unlicensed/overall mean	339 (2.8)	14.9 (54.4)
	Effective *N*	12,164	11,825

Note. PRT "Fail" includes SFail (score of 24-40), XFail (score of 41-500), and abort (timed out).
[1]The number unlicensed includes any customer whose driving privilege was expired, suspended, or revoked for the entire two-year follow-up period following their date of first contact.
[2]The mean days to licensure only includes those customers who eventually renewed their driving privilege.

Licensure Outcomes and Screening for Driving-Relevant Visual Limitations

Table 8 displays information for outcomes on the visual acuity test. Although Driver Safety referral customers generally must pass the visual acuity standard prior to taking an on-road drive test, these data were not consistently collected as part of the 3-Tier Pilot. Therefore, Driver Safety referral customers are excluded from Table 8. The proportion noted as passing includes customers who may have failed on their first attempt, but passed on a second attempt (for instance, after seeing a vision specialist and acquiring a new prescription for corrective lenses).

Table 8 provides, among other information, a partial comparison of the various 3-Tier screening tests to current (non-3-Tier) CA DMV procedures in terms of their effects on the likelihood of retaining the driving privilege. The 3-Tier Assessment System involved no changes to the current CA DMV visual acuity standard. Among 3-Tier eligible customers, almost 17% of customers who fail the visual acuity standard failed to renew their driving privilege in the succeeding 24 months. The relationship between failure on this screening test and licensure rates is of approximately the same magnitude as that found with other, 3-Tier-specific, screening tests for physical and cognitive limitations.

Table 8

Visual Acuity Test Outcomes Among Pilot Renewal Customers, and Proportion Retaining Their Driving Privilege

	Visual acuity test (*N* of customers)	Number unlicensed[1] (%)	Mean days to licensure[2] (*SD*)
Correctly-processed	Pass (10,639)	8 (0.1)	5.0 (18.6)
	Fail (185)	3 (1.6)	36.0 (30.6)
	Total unlicensed/overall mean	11 (0.1)	5.5 (19.3)
	Effective *N*	10,824	10,813
3-Tier eligible	Pass (11,744)	270 (2.3)	12.8 (50.6)
	Fail (412)	69 (16.7)	81.0 (105.4)
	Total unlicensed/overall mean	339 (2.8)	14.8 (54.2)
	Effective *N*	12,156	11,817

[1]The number unlicensed includes any customer whose driving privilege was expired, suspended, or revoked for the entire two-year follow-up period following their date of first contact.
[2]The mean days to licensure only includes those customers who eventually renewed their driving privilege.

Table 9 displays the relationship between outcomes on the Pelli-Robson contrast sensitivity test and licensing outcomes. As with the other screening tests used as part of 3TAS, failure on the contrast sensitivity test is associated with an increased likelihood of failing to renew the driving privilege. The largest difference occurs between "extreme fails" (those who were required to visit a specialist for professional examination of their vision health) and customers who passed or somewhat failed on this assessment.

Table 9

Pelli-Robson Chart Outcomes Among Pilot Renewal Customers, and Proportion Retaining Their Driving Privilege

	Pelli-Robson Chart (*N* of customers)	Number unlicensed[1] (%)	Mean days to licensure[2] (*SD*)
Correctly-processed	Pass (9,153)	0 (0.0)	4.3 (17.3)
	Somewhat fail (1,487)	9 (0.6)	9.3 (24.2)
	Extreme fail (77)	2 (2.6)	35.6 (35.1)
	Not tested (visual acuity fail) (107)	0 (0.0)	39.3 (31.5)
	Total unlicensed/overall mean	11 (0.1)	5.5 (19.3)
	Effective *N*	10,824	10,813
3-Tier eligible	Pass (10,006)	182 (1.8)	12.3 (51.7)
	Somewhat fail (1,717)	81 (4.7)	15.4 (43.2)
	Extreme fail (124)	18 (14.5)	53.6 (81.1)
	Not tested (visual acuity fail) (309)	58 (18.8)	90.7 (110.2)
	Total unlicensed/overall mean	339 (2.8)	14.8 (54.2)
	Effective *N*	12,156	11,817

[1]The number unlicensed includes any customer whose driving privilege was expired, suspended, or revoked for the entire two-year follow-up period following their date of first contact.
[2]The mean days to licensure only includes those customers who eventually renewed their driving privilege.

Licensure Outcomes and Screening for All Driving-Relevant Limitations (Combined)

Finally, Table 10 presents the association between licensure rates (and mean time to licensure) with a combined scale of 3-Tier Assessment System outcomes.

Table 10

Combined 3TAS Tier 1 Outcome Score Among Pilot Renewal Customers, and Proportion Retaining Their Driving Privilege

	Combined 3TAS score (*N* of customers)	Number unlicensed[1] (%)	Mean days to licensure[2] (*SD*)
Correctly-processed	Pass (8,444)	0 (0.0)	3.0 (12.5)
	SFail (1,759)	0 (0.0)	6.6 (20.1)
	XFail (621)	11 (1.8)	37.2 (46.9)
	Total unlicensed/overall mean	11 (0.1)	5.5 (19.3)
	Effective *N*	10,824	10,813
3-Tier eligible	Pass (8,887)	103 (1.2)	7.9 (42.4)
	SFail (2,028)	59 (2.9)	17.7 (55.7)
	XFail (1,241)	177 (14.3)	66.1 (95.0)
	Total unlicensed/overall mean	339 (2.8)	14.8 (54.2)
	Effective *N*	12,156	11,817

[1]The number unlicensed includes any customer whose driving privilege was expired, suspended, or revoked for the entire two-year follow-up period following their date of first contact.
[2]The mean days to licensure only includes those customers who eventually renewed their driving privilege.

This combined measure of limited functionality was constructed in the following manner: Pilot participants who passed all Tier 1 screening tests and passed the written renewal test on the first or second attempt were coded as “Passes.” The category “Somewhat functionally limited” (or SFail, for short), included the following groups:

- Customers who passed all Tier 1 tests, failed the written test twice (but passed it on the third attempt), and passed (or did not take) the PRT.
- Customers who somewhat failed on a single Tier 1 test, passed (or did not take) the PRT, and passed the written renewal test on the first, second, or third attempt.

Because of the way in which 3TAS was piloted, it was not possible for customers to take the PRT without having first failed some other screening test. Therefore, it was not possible for customers to be identified as somewhat functionally limited on the basis of the PRT alone.

The category “Extremely functionally limited” (or XFail, for short), included the following groups:

- Customers who somewhat failed two or more Tier 1 screening tests.
- Customers identified with two or more physical limitations.
- Customers who failed either the visual acuity standard or the Pelli-Robson contrast sensitivity chart (on lines 1 or 4) despite referral and correction.
- Customers who failed the written renewal test three or more times.
- Customers who failed or aborted the PRT.
- All combinations of the above five forms of screening test failure.

This categorization scheme is not precisely the same as that used by Hennessy and Janke (2009). However, it follows the method by which customers were routed to Tier 2 (the PRT) and Tier 3 (the on-road drive test) during the 3-Tier Pilot.

Because the number of written tests taken by erroneously-processed and lagging customers was usually missing or incomplete in the paperwork collected for the Pilot, this information was taken from the Audits database.

If an erroneously-processed or lagging customer was required to take the PRT and did not, their scores were not imputed. Such customers were coded as not having taken the PRT. This probably reduces the number of people who might otherwise have been categorized as XFails.

Because Driver Safety referral customers were not subjected to any Tier 1 tests, were never required to take the PRT, and their outcomes on the written renewal test and vision tests were not collected as part of the Pilot, they are generally excluded from most of the analyses presented in this section.

If a customer initially failed either the visual acuity standard, or the Pelli-Robson contrast sensitivity chart, but then was able to pass after referral and correction, this was coded as a "pass" on these individual screening tests. If they failed (or were not tested) due to the presence of a documented long-standing condition, they were coded as "fails" on these screening tests (and so coded as XFails for the purposes of the combined measure of limited functionality). However, if they passed these screening tests despite the presence of a documented long-standing condition (this happened in a few instances of monocular vision), they were coded as "passes" on these screening tests. Similarly, customers on limited-term licenses were scored according to their performance on these screening tests; if a customer passed the visual acuity standard or the Pelli-Robson contrast sensitivity chart, despite possession of a limited-term license, they were coded as passing these screening tests (this happened in a few instances where the limited-term license was related to a progressive mental or physical condition).

In some cases a customer failed the visual acuity standard and was issued a referral, but did not complete their application by the end of the Pilot. These lagging customers were coded as "fails" on the visual acuity standard (and so coded as XFails for the purposes of the combined measure of limited functionality). It is conceivable that some of these customers later passed the visual acuity standard with some form of correction (e.g., a new prescription for corrective lenses, or some type of surgery).

In a handful of cases (n=7), we lack all information on a customer's Tier 1 screening test outcomes (i.e., both their Tier 1 Score Sheet and 3-Tier Tracking Sheet are missing). In these cases, their combined score on the measure of limited functionality was set to missing. This has the effect of excluding them from the analyses discussed in this section.

Across all of the various kinds of driving-relevant limitations covered by 3TAS—physical function, cognition, and vision—the identification of limitations was, to some degree, associated with an increased likelihood that a driver would fail to renew their driving privilege. That said, nearly all customers—approximately 97%—renewed their driving privilege. Even among those customers who were identified as possessing severe limitations, the overwhelming majority (80-90%, depending on the nature and severity of the limitation) renewed their driving privilege, usually after demonstrating their ability to safely operate a motor vehicle though an on-road test.

It appears that at least some portion of customers who failed to renew their license did so because of issues unrelated to 3TAS. Among 3-Tier eligible customers with no identified limitations, about 1% (n=103) failed to renew their license in the follow-up period. Many of these customers had suspensions or revocations deriving from the excess accumulation of negligent operator "points" on their record, or because of a Driver Safety procedure; these issues may have led to a delay in renewal, or to a licensing action that prevented relicensure.

The identification of limitations under the 3-Tier Pilot was associated with an increased number of days necessary to process an application. This increased amount of processing time was most likely due to matters such as the scheduling of drive tests, obtaining a professional examination for visual acuity or contrast sensitivity, or studying the *California Driver Handbook* in between failures on the written renewal test.

Question #4: How Many Customers Used Certified Driving Rehabilitation Specialists?

During the Pilot, customers who were required to undertake a road test had an extensive interview with an LRE (or Manager I) prior to the test. As part of this interview, and prior to the on-road test of driving skill, the LRE (or Manager I) was trained to ask the following two-part question:

> "Since you found out that you needed to take a drive test, did you get some behind the wheel training? (Yes/No)
>
> Who did you drive with? (a) A Driving Instructor? (b) An Occupational/Rehabilitation Specialist? (c) A Friend or Relative, or (d) Other:____________________?"

Table 11 displays the answers to the first of these two questions, according to test type and attempt.

According to Table 11, most 3-Tier customers who were required to take a drive test reported that they had not used behind-the-wheel training (of any sort) prior to the test. While the proportion reporting the use of behind-the-wheel training generally rises after each failed attempt at an SDPE, it is not until the third attempt that the proportion of customers reporting that they used behind-the-wheel training exceeds 50%.

Table 11

Use of Behind-the-Wheel Training by Pilot Customers, by Test Type and Attempt

	Prepared for test with behind the wheel training?		
Drive test attempt/type (*N*)	Yes (%)	No (%)	No answer (%)
First SDPE (765)	119 (15.6)	527 (68.9)	119 (15.6)
Second SDPE (159)	50 (31.4)	87 (54.7)	22 (13.8)
Third SDPE (25)	18 (72.0)	7 (28.0)	0 (0.0)
First ADPE (16)	2 (12.5)	9 (56.3)	5 (31.3)
Second ADPE (2)	0 (0.0)	1 (50.0)	1 (50.0)

Note. Percentages may not total to 100% due to rounding. Includes both referral and renewal customers.

Table 12 reports the proportion using different kinds of behind-the-wheel training, for those who did so report preparing for their test.

Among those customers who reported engaging in behind-the-wheel training prior to taking a drive test, the overwhelming majority—regardless of test type or attempt—practiced with a friend or relative. The proportion who reported using a professional driving instructor rises with each SDPE failure, but was never more than a third. No 3-Tier Pilot customer reported using an occupational therapist or rehabilitation specialist to prepare for a drive test.

Table 12

Type of Behind-the-Wheel Training Used by Pilot Customers, by Test Type and Attempt

Drive test attempt/type (*N* reporting "yes" to training)	Type of behind-the-wheel training used: Driving instructor (%)	Occup./ rehab. specialist (%)	Friend or relative (%)	Other (%)	No answer (%)
First SDPE (119)	15 (12.6)	0 (0.0)	93 (78.2)	2 (1.7)	10 (8.4)
Second SDPE (50)	11 (22.0)	0 (0.0)	35 (70.0)	2 (4.0)	2 (4.0)
Third SDPE (18)	6 (33.3)	0 (0.0)	12 (66.7)	0 (0.0)	0 (0.0)
First ADPE (2)	0 (0.0)	0 (0.0)	2 (100.0)	0 (0.0)	0 (0.0)
Second ADPE (0)	N/A	N/A	N/A	N/A	N/A

Note. Customers could elect to answer in more than one category (e.g., with a friend/relative and with a driving instructor). Thus, the columns denoting the type of training used do not necessarily sum to the total number of customers reporting that they used training; similarly, the percentage reporting having taken a given type of training may not sum to 100%.

Question #5: What is the Relationship Between Screening Test Outcomes and Compensating Behaviors?

This section presents several confirmatory analyses testing the link between the various screening tests that compose 3TAS, and driving behavior. The driving-behavior variables were taken from the Driving Information Survey, which was constructed based on earlier survey instruments used in prior CA DMV studies on the relationships among various kind of vision problems and driving behavior (Hennessey, 1995) as well as other functional limitations and safe driving (Janke & Hersch, 1997). These surveys were somewhat longer, and included somewhat different questions, than the present survey instrument. The precise question wording is contained in Appendix A.

Table 13 presents simple descriptive statistics on the distribution of answers on the eight driving avoidance questions included in the Driving Information Survey. Table 14 presents the distribution of answers on the self-reported days of driving per week (customers who checked the box "most weeks do not drive" were coded as driving 0 days per week). For both tables, the data include all 3-Tier eligible customers. Missing data on any single question in this survey were left as missing, and not imputed.

Table 13

Distribution of Answers on the Driving Information Survey

	Never avoids (%)	Sometimes (%)	Often (%)	Always avoids (%)	Total *N*
Driving at night?	5419 (45.9)	4121 (34.9)	1516 (12.8)	749 (6.3)	11805
Driving in rain or fog?	4111 (34.8)	5428 (46.0)	1704 (14.4)	554 (4.7)	11797
Driving at sunrise/sunset?	7216 (61.1)	3260 (27.6)	983 (8.3)	345 (2.9)	11804
Driving alone?	8527 (72.2)	2431 (20.6)	641 (5.4)	205 (1.7)	11804
Making left turns?	6584 (56.0)	2804 (23.8)	985 (8.4)	1390 (11.8)	11763
Heavy traffic?	5214 (44.2)	4082 (34.6)	1874 (15.9)	623 (5.3)	11793
Freeways?	7642 (64.7)	3070 (26.0)	697 (5.9)	406 (3.4)	11815
Unfamiliar routes?	5335 (45.1)	4453 (37.7)	1416 (12.0)	617 (5.2)	11821

Note. Includes all 3-Tier eligible customers in the Pilot cohort who answered the Driving Information Survey.

Table 14

Distribution of Answers Regarding Number of Days of Driving per Week

N (%)							
0 days	1	2	3	4	5	6	7 days
351 (3.0)	297 (2.5)	568 (4.8)	1050 (8.9)	1101 (9.3)	2143 (18.1)	1395 (11.8)	4904 (41.5)

Note. Total valid $N = 11{,}809$. Includes all 3-Tier eligible customers in the Pilot cohort who answered the Driving Information Survey.

Table 15 presents a series of cross-tabulations, showing the relationship between overall 3TAS outcome score, and driving avoidance behaviors. For ease of interpretation (and to increase the cell size counts), the two categories of "always avoids" and "often avoids" were combined. It bears noting that these measures of self-reported driving exposure relate to behavior prior to enrollment in the Pilot, and do not capture any potential changes to driving exposure that may have been a result of Pilot participation.

The construction of the overall measure of limited functionality was discussed under Question #3 (pp. 47-49); the same categories are used here.

The relationship between overall level of limited functionality and driving avoidance behavior varies somewhat depending on the situation named in the question. In general, questions with lower χ^2-values (e.g., <80), show relatively little variation in the proportion avoiding a given situation "always" or "often." The proportion avoiding left turns often or always, for instance, hovers around 20-22%, regardless of the overall level of functional limitation as identified through the 3-Tier Assessment System. Similarly, the proportion avoiding heavy traffic always or often varies between 19-25%, the proportion avoiding driving in inclement weather always or often varies between 18-26%, and the proportion avoiding unfamiliar routes varies between 16-23%.

By contrast, the avoidance of other types of driving situations—night driving, sunrise/sunset, and freeways—shows a much stronger relationship to the overall level of functional limitation. In these cases, the proportion of customers who avoid these situations approximately doubles as the overall level of functional limitation goes from "pass" to "extremely functionally limited."

This relationship between driving avoidance and overall level of functionality obtains roughly the same order if we restrict the analysis to drivers aged 70 and older, as seen in Table 16. Although the χ^2-values generally decline, this is largely a function of the reduction in sample size. If anything, the shifting of answers towards "often/always avoids" as a function of overall functional score is somewhat sharper among older drivers than it is among all drivers considered together. However, it is worth noting that even among those older drivers who exhibit the highest levels of limited functionality, the proportion who report avoiding specific challenging situations is usually a small minority. About a third of drivers assessed as having extreme levels of limited functionality report avoiding driving at night often or always. This proportion declines for other questions—about 1 out of every 5, or 1 out of every 4, drivers report avoiding the named situations often or always. Fewer than 1 out of every 8 (11.8%) drivers aged 70 and older who were assessed as having extreme levels of limited functionality reporting avoiding driving alone often or always.

Table 15

Cross-Tabulation of Combined 3TAS Score and Driving Avoidance Behaviors
(All Pilot Customers)

		Pass (%)	SFail (%)	XFail (%)	χ^2 (two-tailed *p*-value)
Avoids night driving	Never	4210 (49.2)	798 (41.0)	370 (31.2)	249.7 (<.01)
	Sometimes	2962 (34.6)	686 (35.3)	435 (36.7)	
	Often/Always	1390 (16.2)	461 (23.7)	380 (32.1)	
	Total	8562	1945	1185	
Avoids driving in rain or fog	Never	3100 (36.2)	663 (34.2)	319 (27)	62.3 (<.01)
	Sometimes	3914 (45.7)	903 (46.5)	553 (46.8)	
	Often/Always	1547 (18.1)	375 (19.3)	310 (26.2)	
	Total	8561	1941	1182	
Avoids driving at sunrise/set	Never	5478 (64.0)	1123 (57.8)	545 (46.0)	177.6 (<.01)
	Sometimes	2232 (26.1)	593 (30.5)	411 (34.7)	
	Often/Always	852 (10.0)	227 (11.7)	230 (19.4)	
	Total	8562	1943	1186	
Avoids driving alone	Never	6315 (73.7)	1393 (71.8)	732 (61.7)	85.2 (<.01)
	Sometimes	1690 (19.7)	411 (21.2)	313 (26.4)	
	Often/Always	561 (6.5)	135 (7)	141 (11.9)	
	Total	8566	1939	1186	
Avoids left turns	Never	4784 (56.0)	1118 (57.7)	610 (51.8)	18.8 (<.01)
	Sometimes	2069 (24.2)	406 (21)	303 (25.7)	
	Often/Always	1684 (19.7)	413 (21.3)	264 (22.4)	
	Total	8537	1937	1177	
Avoids heavy traffic	Never	3881 (45.4)	864 (44.5)	419 (35.4)	47.3 (<.01)
	Sometimes	2878 (33.6)	701 (36.1)	467 (39.4)	
	Often/Always	1794 (21.0)	377 (19.4)	299 (25.2)	
	Total	8553	1942	1185	
Avoids freeways	Never	5757 (67.2)	1214 (62.5)	606 (51.0)	189.3 (<.01)
	Sometimes	2172 (25.3)	498 (25.6)	374 (31.5)	
	Often/Always	643 (7.5)	230 (11.8)	208 (17.5)	
	Total	8572	1942	1188	
Avoids unfamiliar routes	Never	4040 (47.1)	839 (43.2)	410 (34.5)	75.9 (<.01)
	Sometimes	3149 (36.7)	751 (38.6)	511 (43.0)	
	Often/Always	1388 (16.2)	354 (18.2)	267 (22.5)	
	Total	8577	1944	1188	

Note. Percentages may not sum to 100% due to rounding. See pp. 47-49 for a discussion of the definition of the categories "passed," "SFail," and "XFail."

Table 16

Cross-Tabulation of Combined 3TAS Score and Driving Avoidance Behaviors
(Pilot Customers Aged 70+)

		Pass (%)	SFail (%)	XFail (%)	χ^2 (two-tailed *p*-value)
Avoids night driving	Never	1327 (45.3)	527 (38.3)	278 (30.9)	97.0 (<.01)
	Sometimes	1009 (34.5)	477 (34.7)	317 (35.2)	
	Often/Always	592 (20.2)	371 (27.0)	306 (34.0)	
	Total	2928	1375	901	
Avoids driving in rain or fog	Never	1119 (38.3)	461 (33.6)	246 (27.4)	61.0 (<.01)
	Sometimes	1324 (45.3)	656 (47.9)	415 (46.3)	
	Often/Always	482 (16.5)	253 (18.5)	236 (26.3)	
	Total	2925	1370	897	
Avoids driving at sunrise/set	Never	1876 (64.0)	787 (57.4)	407 (45.1)	111.1 (<.01)
	Sometimes	728 (24.9)	411 (30.0)	317 (35.1)	
	Often/Always	325 (11.1)	174 (12.7)	178 (19.7)	
	Total	2929	1372	902	
Avoids driving alone	Never	2313 (79)	1007 (73.5)	578 (64.2)	91.5 (<.01)
	Sometimes	458 (15.6)	270 (19.7)	217 (24.1)	
	Often/Always	157 (5.4)	93 (6.8)	106 (11.8)	
	Total	2928	1370	901	
Avoids left turns	Never	1783 (61.1)	809 (59.2)	453 (50.7)	37.1 (<.01)
	Sometimes	647 (22.2)	283 (20.7)	241 (27.0)	
	Often/Always	486 (16.7)	275 (20.1)	200 (22.4)	
	Total	2916	1367	894	
Avoids heavy traffic	Never	1485 (50.7)	615 (44.8)	323 (35.9)	79.5 (<.01)
	Sometimes	1028 (35.1)	509 (37.1)	360 (40.0)	
	Often/Always	417 (14.2)	249 (18.1)	217 (24.1)	
	Total	2930	1373	900	
Avoids freeways	Never	1943 (66.2)	838 (61.1)	457 (50.6)	89.7 (<.01)
	Sometimes	716 (24.4)	357 (26.0)	278 (30.8)	
	Often/Always	274 (9.3)	176 (12.8)	168 (18.6)	
	Total	2933	1371	903	
Avoids unfamiliar routes	Never	1507 (51.3)	591 (43.0)	306 (33.9)	124.2 (<.01)
	Sometimes	1103 (37.6)	559 (40.7)	393 (43.5)	
	Often/Always	325 (11.1)	223 (16.2)	204 (22.6)	
	Total	2935	1373	903	

Note. Percentages may not sum to 100% due to rounding. See pp. 47-49 for a discussion of the definition of the categories "passed," "SFail," and "XFail."

Table 17 displays the relationship between overall level of limited functionality, and the mean number of days of driving. Although it is the case that the number of days of driving declines as a function of the overall measure of limited functionality, the group differences are relatively small—on the order of half a day of driving per level of impaired functionality. Even the most impaired drivers report driving, on average, at least every other day of the week.

The relationship between overall level of limited functionality and self-reported days of driving is roughly the same among older drivers (aged 70+) as it is among all drivers, as reported in Table 18. Even the most functionally limited older drivers report driving at least every other day.

Table 17

Mean Self-Reported Days of Driving per Week, by Combined 3TAS Score (All Pilot Customers)

	Pass	SFail	XFail
Mean days of driving per week (s.d.)	5.5 (1.9)	5.0 (1.9)	4.6 (2.0)
N of drivers	8567	1946	1183

Note. F-statistic for ANOVA test of between-groups differences: 142.4 ($p < 0.01$). See pp. 47-49 for a discussion of the definition of the categories "passed," "SFail," and "XFail."

Table 18

Mean Self-Reported Days of Driving per Week, by Combined 3TAS Score
(Pilot Customers Aged 70+)

	Pass	SFail	XFail
Mean days of driving per week (s.d.)	5.0 (1.9)	4.8 (1.9)	4.4 (2.0)
N of drivers	2933	1375	899

Note. F-statistic for ANOVA test of between-groups differences: 41.6 ($p < 0.01$). See pp. 47-49 for a discussion of the definition of the categories "passed," "SFail," and "XFail."

Question #6: What is the Inter-Correlation Among Outcomes on the Cognitive/Perceptual Screening Tests?

The screening tests of the 3-Tier Assessment System cover three broad areas of driving-relevant functionality: physical function, vision, and cognition/perception. The physical observation protocol covers multiple potential physical issues that might affect driving. The vision tests—the

Snellen standard and the Pelli-Robson contrast sensitivity chart screen—are for overlapping, but nevertheless distinct, problems associated with vision loss (Desapria et al., 2011). In the third domain of cognition/perception, 3TAS incorporated three screening tests: the memory recall test, the written renewal test, and the PRT.

This section presents a series of simple analyses that test the inter-correlation among the three tests related to cognition/perception. These analyses are complicated somewhat by the nested nature of 3TAS. Not every customer took the PRT, for instance—and the set of customers who did take the PRT were selected in part based on their performance on the memory recall test and the written renewal test. Thus, any potential correlation between the PRT and the other cognition/perception tests may in part be a result of the tiered nature of processing during the Pilot.

Also, it bears noting that there exists evidence of variation in the implementation of the memory recall test, as discussed in the Process Analysis Appendix (Camp, 2010a). This variation in implementation likely resulted in a lower failure rate on this test than would have occurred otherwise. To the extent that this may bias the results presented here, it would probably depress or obscure any observed statistical relationship between the memory recall test and other cognitive/perceptual screening tests. Thus, the results presented here may be construed as conservative estimates of any association between the memory recall screen and the other two tests related to cognition/perception.

Table 19 presents a cross-tabulation of results on the written renewal test, broken down by age cohort and outcomes on the memory recall test.

The written renewal test is, of course, primarily designed to be a knowledge test, focused on retention and recall of the laws and rules of the road. However, for the purposes of 3TAS it can also be conceptualized as a partial indicator of general cognitive health (Janke & Hersch, 1997). In other words, it is expected that drivers who suffer from various kinds of cognitive limitations—such as might occur with the early stages of dementia-type disorders—will fail both the memory recall test and the written renewal test. In the latter case, failing the test multiple times is expected to be associated with cognitive limitations.

In general, customers who failed the memory recall test had more difficulty passing the written renewal test. Due to small cell-size counts, the results are not reliable for customers younger than 70 y.o. However, for customers aged 70 and older, the relationship between outcomes on the two tests are statistically significant, and in the expected direction. It is true that the overwhelming majority of customers who failed the memory recall test passed the written renewal test on the first or second attempt. However, among customers aged 70 and over who failed the memory recall screen, over 20% had to take the written test at least three times. By comparison, among customers aged 70 and over who passed the memory recall scan, fewer than 10% had to take the written renewal test at least three times.

Table 19

Cross-Tabulation of Age, Memory Recall Test Outcome, and Number of Written Renewal Test Attempts

		N of written test attempts			
Age group	Memory recall test outcome	Passed on 1st or 2nd attempt	Passed on 3rd attempt	Required 4+ attempts	χ^2 (2-tailed *p*-value)
<40 y.o.	Pass (%)	3418 (93.1)	174 (4.7)	81 (2.2)	2.92 (0.23)
	Fail (%)	9 (81.8)	1 (9.1)	1 (9.1)	
40-69 y.o.	Pass (%)	2865 (93.2)	123 (4.0)	85 (2.8)	1.34 (0.51)
	Fail (%)	15 (93.8)	0 (0.0)	1 (6.3)	
70+ y.o.	Pass (%)	4799 (91.7)	279 (5.3)	155 (3.0)	33.07 (<0.00)
	Fail (%)	116 (78.4)	19 (12.8)	13 (8.8)	

Note. Includes all 3-Tier eligible Pilot customers.

The next two tables present the relationship between the PRT, and the memory recall and written renewal tests. In both tables, customers counted as failing on the PRT if they had a score of 24 ms or greater.[7] The analysis is restricted to those 3-Tier eligible customers who took the PRT. If a customer was required to take the PRT, but did not do so (as in the case of some erroneously-processed and lagging customers), they are excluded from the analysis due to missing data on this test.

[7]During the Pilot, the computer algorithm used to program the PRT was designed to categorize as SFails those customers who scored between 24 and 40 ms. Fourteen customers scored in this range, out of a total of 2,677 customers required to take this test. Twelve of the customers who SFailed on the PRT were aged 70 or older. In all 14 cases of PRT SFails, they were routed to the PRT on the basis of their Tier 1 score, rather than on the basis of their performance on the written test. Twelve of these 14 PRT SFails were given drive tests (and so correctly processed); the remaining two were licensed without a required drive test. PRT SFails were combined with XFails for ease of interpretation; grouping them separately would not materially change any of the analyses presented in this section.

Table 20

Cross-Tabulation of Age, Memory Recall Test Outcome, and PRT Outcome

		PRT outcome		
Age group	Memory recall test outcome	Pass	Fail	χ^2 (Fisher's exact test, 2-sided)
<40 y.o.	Pass (%)	209 (97.2)	6 (2.8)	0.32 (1.00)
	Fail (%)	11 (100.0)	0 (0.0)	
40-69 y.o.	Pass (%)	358 (89.9)	40 (10.1)	0.17 (0.66)
	Fail (%)	13 (86.7)	2 (13.3)	
70+ y.o.	Pass (%)	1545 (81.1)	360 (18.9)	3.65 (0.06)
	Fail (%)	95 (74.2)	33 (25.8)	

Note. Includes all 3-Tier eligible Pilot customers who took the PRT. PRT "Fail" includes SFail (score of 24-40), XFail (score of 41-500), and abort (timed out).

Similar to the previous table, failure on the memory recall screen is associated with failure on the PRT. This is more obvious in the case of customers aged 70 and older, for whom the relationship between outcomes on these two tests approaches statistical significance.

The relationship between the written test and the PRT is quite strong, as seen in Table 21. Especially among those aged 70 and older, failing the written test multiple times is associated with a greater likelihood of failing the PRT. Among those who took the written test four or more times, over half also failed the PRT.

In sum, it appears that outcomes on the memory recall screen, the written renewal test, and the PRT correlate with one another. This correlation is most clearly seen among drivers aged 70 and older. This finding lends credence to the expectation that these tests all tap into a common domain of functionality. On the basis of prior research in this area, it seems plausible to assert that this common domain relates to potential limitations in cognition/perception.

The SDPE drive test also contains certain cognitive test elements. These include a destination (or way-finding) trip, and a section where a customer is given multiple direction at once, e.g., "merge into the center lane, then take a left turn at the next intersection when it is safe to do so." On the basis of prior research one would expect to find a correlation between outcomes on the

Table 21

Cross-Tabulation of Age, Number of Written Renewal Test Attempts, and PRT Outcome

		PRT outcome		
Age group	*N* of written test attempts	Pass	Fail	χ^2 (2-tailed *p*-value)
<40 y.o.	Passed on 1st or 2nd attempt (%)	80 (98.8)	1 (1.2)	4.49 (0.11)
	Passed on 3rd attempt (%)	103 (98.1)	2 (1.9)	
	Required 4+ attempts (%)	37 (92.5)	3 (7.5)	
40-69 y.o.	Passed on 1st or 2nd attempt (%)	261 (91.6)	24 (8.4)	3.41 (0.18)
	Passed on 3rd attempt (%)	68 (87.2)	10 (12.8)	
	Required 4+ attempts (%)	42 (84.0)	8 (16.0)	
70+ y.o.	Passed on 1st or 2nd attempt (%)	1408 (84.1)	267 (15.9)	109.22 (<0.00)
	Passed on 3rd attempt (%)	174 (74.4)	60 (25.6)	
	Required 4+ attempts (%)	60 (47.2)	67 (52.8)	

Note. Includes all 3-Tier eligible Pilot customers who took the PRT. PRT "Fail" includes SFail (score of 24-40), XFail (score of 41-500), and abort (timed out).

cognitive screening tests and errors on an on-road test of driving skill (Anstey & Wood, 2011; Hennessy & Janke, 2009; Janke & Hersch, 1997; Selander, Lee, Johansson, & Falkmer, 2011). However, the outcomes on the Tier 1 (memory recall) and Tier 2 (written renewal test, PRT) cannot be correlated with the SDPE cognitive test elements using the current data set. Only a select set of customers—all with multiple and/or severe functional limitations—were given a drive test. Therefore, the observed variance on the Tier 1 and Tier 2 tests in this case would be rather limited. Moreover, in the 3-Tier Pilot the administration of the SDPE test was often terminated at the point at which a customer committed a Critical Driving Error (or CDE). A CDE is defined as the commission of certain kinds of particularly dangerous driving errors, including disobeying traffic signs/signals, lane violations, striking objects or curbs, etc. The commission of a CDE always results in failure on the overall drive test, regardless of performance on other test elements. In cases where a CDE resulted in immediate termination of the test,[8] this often meant that the cognitive portions of the SDPE were either truncated (e.g., the multiple directions task), or not administered at all (as in the case of the destination task, normally done as the last element of the road test). For this reason, the subset of customers who were subjected to the full battery of cognitive elements of the SDPE excludes almost all customers who failed the test because of a CDE. And while it may be true that commission of a CDE could indicate potential cognitive/perceptual limitations, this cannot readily be analyzed using the present dataset.

Question #7: What is the Relationship Between Screening Test Outcomes and Prior Crash Record?

The constituent elements of the 3-Tier Assessment System were chosen, on the basis of prior research (Hennessy & Janke, 2009) for, among other reasons, their validity for identifying drivers at risk of crashing. This section presents a series of confirmatory analyses which test the validity of the component screening tests, separately and in combination, at identifying drivers with an increased individual risk of crashing. Also included is an analysis which tests the validity of the component screening tests, in combination, at identifying drivers at risk of failing an on-road test of driving skill.

For each model related to crashes, a logistic regression was estimated, calculating the differences in the odds of crashing one or more times in the 3 years prior to enrollment in the Pilot. This general approach is similar to that used by Hennessey and Janke (2009); however, due to differences in sample sizes, somewhat different statistical methods are available for use in the present analysis. For each model the primary independent variable of interest consists of one of the following: Tier 1 screening test scores (modeled separately), the number of written knowledge test attempts, the outcome on the PRT, and the combined measure of limited functionality which incorporates all of the Tier 1 and Tier 2 screening tests. This combined measure was discussed in the previous subsection (see pp. 47-49).

[8]Termination of the test upon commission of a CDE is usually done to preserve the safety of the driver, the examiner, and other road users.

As a general matter, this approach assumes that the observation of limited functionality measured at one point in time (enrollment in the Pilot) is related to crashes that may have occurred up to 3 years previously. This assumption may be problematic, for several reasons. It conflates conditions that may have been recently acquired with those that are of long-standing duration. Also, using this approach does not allow for examining changes in crash risk over time that may occur, for instance, with progressive diseases, where drivers may steadily lose their ability to safely compensate for some limitation. Nor does it allow for examining changes in crash risk over time that may occur with stable conditions, where drivers may acquire over time the ability to safely compensate for some limitation. Due to the nature of the available data, as well as the study design, the problems associated with this assumption cannot be explored.

A series of sub-models are presented for each screening test and for the combined measure of limited functionality. The first sub-model shows the unadjusted differences in odds ratios, with only the predictor variable entered, for all 3-Tier eligible customers. The second sub-model then adds the following control variables to adjust for known correlates of crash risk: prior record (suspensions/revocation in the 3 years prior to enrollment in the Pilot), sex, and exposure. Exposure was measured using the self-reported number of days of driving per week. The third sub-model then enters the results of the Driving Information Survey. For this last sub-model, backwards elimination was used to trim the model to only those questions from the Driving Information Survey that produced statistically significant results. The inclusion of information from the Driving Information Survey does not appear to alter the size or direction of the other coefficients in the models presented.

Along with these sub-models, a table is provided with a simple descriptive cross-tabulation showing the raw crash rates, stratified by age and limitation. The tables in this section include all 3-Tier eligible customers.

The Validity of the Individual Screening Tests at Predicting Prior Crashes

Table 22 displays a cross-tabulation of the relationship between the number of observed physical limitations, age, and crashes in the 3 years prior to enrollment in the Pilot. A customer's score on this screening test (zero, one, or two) captures the number of driving relevant physical limitations observed by staff during the license renewal process. This score was capped at two, both for ease of interpretation and due to training protocols in place during the Pilot (i.e., staff were trained to score only up to two physical limitations).

The findings suggest that the relationship between observed physical limitations and crash risk may be age-dependent. Customers in the middle-age group (40-69 years of age) with one observed physical limitation appear to be somewhat more likely to crash than other customers in the same age cohort. However, these findings should be treated with a great deal of caution due to the small cell-size counts for some of the sub-groups. There are a mere baker's dozen (n=13) customers younger than 70 y.o. who were identified with two or more physical limitations. In addition, it must be stressed that comparing crash rates across age groups is problematic for the following reasons: while all customers over the age of 70 must renew in person in a DMV field office, for customers younger than 70 y.o. only those customers with limited-term licenses, recent at-fault crashes, or negligent operator points are required to renew in person. This means that there are proportionally fewer younger customers with clean records in the 3-Tier Pilot sample. While it is true that younger drivers (especially those in their teens and twenties) generally present higher crash risk than more experienced drivers, nevertheless the higher crash propensity of younger customers in Table 22 is potentially an artifact of CA DMV policy and state law in place at the time of the 3-Tier Pilot.

The next set of analyses consists of a series of logistic regressions, modeling the relationship between the number of observed physical limitations and the odds ratio of crashing in the prior 3 years.

Table 22

Cross-Tabulation of Prior Crashes, Age, and Observed Physical Limitations

Age group	Observed physical limitations	N of crash-free customers (%)	N of crashed 1+ times in prior 3 years (%)	χ^2 (2-tailed p-value)
<40 y.o.	0 limitations	2867 (78.2)	797 (21.8)	2.66 (0.27)
	1 limitation	11 (64.7)	6 (35.3)	
	2+ limitations	3 (100.0)	0 (0.0)	
40-69 y.o.	0 limitations	2469 (82.1)	537 (17.9)	6.04 (0.05)
	1 limitation	55 (73.3)	20 (26.7)	
	2+ limitations	10 (100.0)	0 (0.0)	
70+ y.o.	0 limitations	4359 (88.7)	556 (11.3)	0.64 (0.73)
	1 limitation	358 (88.2)	48 (11.8)	
	2+ limitations	55 (91.7)	5 (8.3)	

Note. Includes all 3-Tier eligible renewal customers in the Pilot.

Table 23

Logistic Regression Predicting Odds of Crashing in the 3 Years Prior to Pilot Enrollment, Based on Observed Driving-Relevant Physical Limitations

	Sub-model A		Sub-model B		Sub-model C	
	β (*SE*)	Exp(β) (CI)	β (*SE*)	Exp(β) (CI)	β (*SE*)	Exp(β) (CI)
S/R indicator			0.37** (0.06)	1.45 (1.29 - 1.62)	0.33** (0.06)	1.39 (1.24 - 1.57)
Gender (male)			0.06 (0.05)	1.06 (0.96 - 1.17)	0.06 (0.05)	1.06 (0.96 - 1.18)
Days of driving per week			0.13** (0.02)	1.14 (1.10 - 1.17)	0.14** (0.02)	1.15 (1.11 - 1.18)
Sometimes avoids driving at night (ref group: never avoids driving at night)					-0.06 (0.07)	0.94 (0.83 - 1.07)
Often/always avoids driving at night (ref group: never avoids driving at night)					-0.22* (0.09)	0.80 (0.67 - 0.96)
Sometimes avoids driving in the rain (ref group: never avoids driving in the rain)					0.10 (0.07)	1.12 (0.97 - 1.26)
Often/always avoids driving in the rain (ref group: never avoids driving in the rain)					0.35** (0.10)	1.42 (1.18 - 1.70)
Sometimes avoids making left turns (ref group: never avoids making left turns)					0.11† (0.07)	1.12 (0.99 - 1.27)
Often/always avoids making left turns (ref group: never avoids making left turns)					0.21** (0.07)	1.23 (1.07 - 1.41)
1 physical limitation (ref. group: no physical limitations)	-0.11 (0.13)	0.90 (0.70 – 1.15)	0.09 (0.13)	1.10 (0.84 - 1.42)	0.10 (0.13)	1.11 (0.85 - 1.44)
2+ physical limitations (ref. group: no physical limitations)	-0.98* (0.46)	0.38 (0.15 – 0.94)	-0.73 (0.47)	0.48 (0.19 - 1.21)	-1.24* (0.59)	0.29 (0.09 - 0.92)

Note. *N* for Sub-model A = 12,156; *N* for Sub-model B = 11,696; *N* for Sub-model C = 11,582. All sub-models include all 3-Tier eligible customers. All confidence intervals (CI) were calculated at a 95% confidence level.
*: $p \leq 0.05$. **: $p \leq 0.01$. †: $p \leq 0.10$.

Customers identified with two or more driving-relevant physical limitations were less likely to have crashes on their record during the 3 years prior to renewal, as compared to customers with no such limitations. Customers identified with one driving-relevant physical limitation were statistically indistinguishable in their prior crash record from customers with no identified physical limitations.

The next two tables present a similar set of analyses, for the memory recall test.

Table 24

Cross-Tabulation of Prior Crashes, Age, and Outcome on the Memory Recall Test

Age group	Memory recall test outcome	*N* of crash-free customers (%)	*N* of crashed 1+ times in prior 3 years (%)	χ^2 (Fisher's exact test, 2-sided)
<40 y.o.	Pass	2871 (78.2)	802 (21.8)	1.05 (0.47)
	Fail	10 (90.9)	1 (9.1)	
40-69 y.o.	Pass	2519 (81.9)	556 (18.1)	1.51 (0.33)
	Fail	15 (93.8)	1 (6.3)	
70+ y.o.	Pass	4644 (88.7)	589 (11.3)	0.73 (0.36)
	Fail	128 (86.5)	20 (13.5)	

Note. Table includes all 3-Tier eligible renewal customers in the Pilot.

Unlike the previous analysis (for physical limitations), there does not appear to be a statistically-significant relationship among age, crash risk, and outcomes on the memory recall test. It is possible that the lack of significance is related to certain problems in implementation identified in the Process Analysis Appendix (Camp, 2010a). Qualitative evidence collected as part of the staff surveys indicated that there was variation across staff in the fail rate on this screening test. The net effect of this variation in implementation may have resulted in a lower fail rate than would have occurred under different circumstances. A suppressed fail rate on this screening test would, in this table, count some customers in the "pass" category that should have been listed in the "fail" category, had the test been administered consistently. However, it is impossible to determine with certainty whether, and in what direction, this would alter the association between crash record and outcomes on the memory recall test.

Table 25

Logistic Regression Predicting Odds of Crashing in the 3 Years Prior to Pilot Enrollment, Based on Outcomes on the Memory Recall Test

	Sub-model A		Sub-model B		Sub-model C	
	β (*SE*)	Exp(β) (CI)	β (*SE*)	Exp(β) (CI)	β (*SE*)	Exp(β) (CI)
S/R indicator			0.37** (0.06)	1.44 (1.29 – 1.62)	0.33** (0.06)	1.39 (1.24 - 1.57)
Gender (male)			0.06 (0.05)	1.06 (0.96 – 1.17)	0.06 (0.05)	1.06 (0.96 - 1.18)
Days of driving per week			0.13** (0.02)	1.14 (1.11 - 1.17)	0.14** (0.02)	1.14 (1.11 - 1.18)
Sometimes avoids driving at night (ref group: never avoids driving at night)					-0.06 (0.07)	0.94 (0.83 - 1.07)
Often/always avoids driving at night (ref group: never avoids driving at night)					-0.22* (0.09)	0.80 (0.67 - 0.96)
Sometimes avoids driving in the rain (ref group: never avoids driving in the rain)					0.10 (0.07)	1.11 (0.97 - 1.27)
Often/always avoids driving in the rain (ref group: never avoids driving in the rain)					0.35** (0.10)	1.42 (1.18 - 1.71)
Sometimes avoids making left turns (ref group: never avoids making left turns)					0.11† (0.07)	1.12 (0.98 - 1.27)
Often/always avoids making left turns (ref group: never avoids making left turns)					0.20** (0.07)	1.23 (1.07 - 1.41)
Memory recall fail (ref group: pass)	-0.30 (0.23)	0.74 (0.47 – 1.16)	-0.12 (0.24)	0.89 (0.56 - 1.42)	-0.19 (0.25)	0.83 (0.51 - 1.34)

Note. *N* for Sub-model A = 12,156; *N* for Sub-model B = 11,696; *N* for Sub-model C = 11,582. All confidence intervals (CI) were calculated at a 95% confidence level.
*: $p \leq 0.05$. **: $p \leq 0.01$. †: $p \leq 0.10$.

Table 25 presents a logistic regression predicting the relative odds of having crashed in the prior 3 years, as a result of outcomes on the memory recall test, with and without statistical controls. Regardless of the sub-model, customers who fail the memory recall test have directionally lower odds of having crashed in the 3 years prior to enrollment. However, this finding does not achieve conventional standards of statistical significance.

The next two tables present a similar set of analyses, for the visual acuity screening test.

Table 26

Cross-Tabulation of Prior Crashes, Age, and Outcome on the Visual Acuity Test

Age group	Visual acuity test	*N* of crash-free customers (%)	*N* of crashed 1+ times in prior 3 years (%)	χ^2 (Fisher's exact test, 2-sided)
<40 y.o.	Pass	2859 (78.3)	793 (21.7)	1.69 (0.20)
	Fail	22 (68.8)	10 (31.3)	
40-69 y.o.	Pass	2486 (82.1)	543 (17.9)	0.89 (0.32)
	Fail	48 (77.4)	14 (22.6)	
70+ y.o.	Pass	4493 (88.7)	570 (11.3)	0.30 (0.58)
	Fail	279 (87.7)	39 (12.3)	

Note. Table includes all 3-Tier eligible renewal customers in the Pilot.

Across age groups, customers who fail the visual acuity standard appear to be more likely to have been involved in a crash in the previous 3 years. However, these differences are quite small for drivers aged 70 and older, and for no age group does the association between failure on the visual acuity standard and prior 3-year crash record meet standard tests of statistical significance. This may in part be due to small cell-size counts.

Table 27 presents a logistic regression predicting the relative odds of having crashed in the prior 3 years, as a result of outcomes on the visual acuity standard, with and without statistical controls. Regardless of the sub-model, outcomes on the visual acuity standard are not associated with 3-year prior crash risk.

Table 27

Logistic Regression Predicting Odds of Crashing in the 3 Years Prior to Pilot Enrollment, Based on Outcomes on the Visual Acuity Test

	Sub-model A		Sub-model B		Sub-model C	
	β (*SE*)	Exp(β) (CI)	β (*SE*)	Exp(β) (CI)	β (*SE*)	Exp(β) (CI)
S/R indicator			0.37** (0.06)	1.45 (1.29 – 1.63)	0.33** (0.06)	1.40 (1.24 - 1.57)
Gender (male)			0.06 (0.05)	1.06 (0.96 – 1.17)	0.06 (0.05)	1.06 (0.96 - 1.18)
Days of driving per week			0.13** (0.02)	1.14 (1.11 - 1.17)	0.14** (0.02)	1.15 (1.11 - 1.18)
Sometimes avoids driving at night (ref group: never avoids driving at night)					-0.06 (0.07)	0.94 (0.82 - 1.07)
Often/always avoids driving at night (ref group: never avoids driving at night)					-0.22* (0.09)	0.80 (0.67 - 0.96)
Sometimes avoids driving in the rain (ref group: never avoids driving in the rain)					0.10 (0.07)	1.11 (0.97 - 1.27)
Often/always avoids driving in the rain (ref group: never avoids driving in the rain)					0.35** (0.10)	1.42 (1.18 - 1.71)
Sometimes avoids making left turns (ref group: never avoids making left turns)					0.11† (0.07)	1.12 (0.98 - 1.27)
Often/always avoids making left turns (ref group: never avoids making left turns)					0.21** (0.07)	1.23 (1.07 - 1.41)
Fail on visual acuity test (ref group: pass)	-0.07 (0.14)	0.93 (0.71 – 1.22)	0.08 (0.14)	1.08 (0.82 - 1.44)	0.03 (0.15)	1.03 (0.77 - 1.38)

Note. *N* for Sub-model A = 12,156; *N* for Sub-model B = 11,696; *N* for Sub-model C = 11,582. All sub-models include all 3-Tier eligible customers. All confidence intervals (CI) were calculated at a 95% confidence level.
*: $p \leq 0.05$. **: $p \leq 0.01$. †: $p \leq 0.10$.

The next two tables present a similar set of analyses, for the Pelli-Robson contrast sensitivity chart. For these analyses, customers who failed the visual acuity test and were never tested on the Pelli-Robson chart were included as "extreme fails." This happened with a number of lagging customers, especially in those cases where the referral to a vision specialist delayed relicensure past the final date of the 3-Tier Pilot. Please note: in a few cases, it was possible for customers to fail the visual acuity test, and then pass on the Pelli-Robson chart. This appears to have occurred for certain categories of customers with long-standing vision conditions (e.g., monocular vision) for which the department had record, and for which the customer had at some point in the past demonstrated their ability to safely operate a motor vehicle in an on-road test.

Table 28

Cross-Tabulation of Prior Crashes, Age, and Outcome on the Pelli-Robson Chart

Age group	Pelli-Robson chart	*N* of crash-free customers (%)	*N* of crashed 1+ times in prior 3 years (%)	χ^2 (2-tailed *p*-value)
<40 y.o.	Pass	2815 (78.3)	781 (21.7)	0.54 (0.76)
	Somewhat fail	45 (75.0)	15 (25.0)	
	Extreme fail	21 (75.0)	7 (25.0)	
40-69 y.o.	Pass	2299 (81.9)	508 (18.1)	0.87 (0.65)
	Somewhat fail	186 (83.8)	36 (16.2)	
	Extreme fail	49 (79.0)	13 (21.0)	
70+ y.o.	Pass	3194 (88.6)	409 (11.4)	1.99 (0.37)
	Somewhat fail	1281 (89.3)	154 (10.7)	
	Extreme fail	297 (86.6)	46 (13.4)	

Note. Table includes all 3-Tier eligible renewal customers in the Pilot.

Across age groups, customers who fail the Pelli-Robson contrast sensitivity chart appear to be more likely to have been involved in a crash in the previous 3 years. This is especially true for customers who "extreme fail" this test, meaning they missed one or more letters on lines 1 or 4 of the chart. However, these differences do not meet standard tests of statistical significance. There is thus no convincing evidence of a relationship between screening outcomes on this test and prior crashes—but this very well may be a result of small cell-size counts in some of the groupings of outcome by age.

Table 29 presents a logistic regression predicting the relative odds of having crashed in the prior 3 years, as a result of outcomes on the Pelli-Robson chart, with and without statistical controls. As with the previous table, customers who failed the visual acuity standard and were not tested on the Pelli-Robson chart are coded here as “extreme fails.”

Customers who somewhat failed the Pelli-Robson chart are less likely to have a crash on their record, as compared to customers who passed this test. Customers who extreme failed the Pelli-Robson chart are not statistically distinguishable from those who passed.

The Process Analysis (Camp, 2010a) reported evidence that there existed some amount of variation across staff in the implementation of this screening test. The net effect of this variation probably resulted in a lower failure rate than would have occurred otherwise. In other words, some customers who are here coded as “pass” may, under other circumstances, have been coded as either “somewhat fails” or “extreme fails.” Also, some customers who are here coded as “somewhat fails” may, under other circumstances, have been coded as “extreme fails.” However, it is unclear whether, and to what degree, this would alter the (negative) association found in this analysis regarding prior crash record.

Tables 30 and 31 present a similar set of analyses, for outcomes on the written renewal test. To simplify the modeling, the number of attempts has been truncated to three ordinal categories: (a) made one or two attempts to pass, (b) passed on the third attempt, or (c) required four or more attempts before passage. This coding parallels the practical use of the written knowledge test as a screening tool into different assessment tiers. Those who passed on the first or second attempt, and who had no other identified limitations, were generally issued their license with no other required tests. Those who failed the written test twice were required to take the PRT. Those who failed the test three or more times were required to take an on-road drive test.[9]

[9]Two Pilot customers did not take the written test at all, choosing instead to surrender their driving privilege. For this analysis, these customers’ data values on the written test were set to missing. This has the effect of excluding them from the analysis.

Table 29

Logistic Regression Predicting Odds of Crashing in the 3 Years Prior to Pilot Enrollment, Based on Outcomes on the Pelli-Robson Chart

	Sub-model A		Sub-model B		Sub-model C	
	β (*SE*)	Exp(β) (CI)	β (*SE*)	Exp(β) (CI)	β (*SE*)	Exp(β) (CI)
S/R indicator			0.33** (0.06)	1.40 (1.24 - 1.57)	0.30** (0.06)	1.35 (1.19 - 1.52)
Gender (male)			0.06 (0.05)	1.06 (0.96 - 1.17)	0.06 (0.05)	1.06 (0.96 - 1.18)
Days of driving per week			0.13** (0.02)	1.13 (1.10 - 1.17)	0.13** (0.02)	1.14 (1.11 - 1.18)
Sometimes avoids driving at night (ref group: never avoids driving at night)					-0.05 (0.07)	0.95 (0.83 - 1.08)
Often/always avoids driving at night (ref group: never avoids driving at night)					-0.20* (0.09)	0.82 (0.69 - 0.98)
Sometimes avoids driving in the rain (ref group: never avoids driving in the rain)					0.10 (0.07)	1.10 (0.97 - 1.26)
Often/always avoids driving in the rain (ref group: never avoids driving in the rain)					0.34** (0.10)	1.40 (1.17 - 1.69)
Sometimes avoids making left turns (ref group: never avoids making left turns)					0.11 (0.07)	1.11 (0.98 - 1.27)
Often/always avoids making left turns (ref group: never avoids making left turns)					0.20** (0.07)	1.22 (1.07 - 1.41)
Somewhat fail the Pelli-Robson chart (ref group: pass)	-0.41** (0.08)	0.66 (0.57 – 0.78)	-0.28** (0.08)	0.76 (0.64 - 0.89)	-0.28** (0.08)	0.76 (0.65 - 0.89)
Extreme fail the Pelli-Robson chart (ref group: pass)	-0.13 (0.14)	0.88 (0.67 – 1.15)	0.02 (0.14)	1.02 (0.77 - 1.34)	-0.02 (0.15)	0.98 (0.73 - 1.30)

Note. *N* for Sub-model A = 12,156; *N* for Sub-model B = 11,696; *N* for Sub-model C = 11,582. All confidence intervals (CI) were calculated at a 95% confidence level.
*: $p \leq 0.05$. **: $p \leq 0.01$. †: $p \leq 0.10$.

Table 30

Cross-Tabulation of Prior Crashes, Age, and Written Renewal Test Attempts

Age group	*N* of written test attempts	*N* of crash-free customers (%)	*N* of crashed 1+ times in prior 3 years (%)	χ^2 (2-tailed *p*-value)
<40 y.o.	Passed on 1st or 2nd attempt	2682 (78.2)	746 (21.8)	1.09 (0.58)
	Passed on 3rd attempt	133 (76.0)	42 (24.0)	
	Required 4+ attempts	67 (81.7)	15 (18.3)	
40-69 y.o.	Passed on 1st or 2nd attempt	2370 (81.9)	512 (17.8)	2.70 (0.26)
	Passed on 3rd attempt	99 (83.8)	24 (19.5)	
	Required 4+ attempts	65 (79.0)	21 (24.4)	
70+ y.o.	Passed on 1st or 2nd attempt	4356 (88.6)	560 (11.4)	0.21 (0.90)
	Passed on 3rd attempt	268 (89.3)	32 (10.7)	
	Required 4+ attempts	150 (86.6)	18 (10.7)	

Note. Table includes all 3-Tier eligible renewal customers in the Pilot.

Although Driver Safety referral customers often took the written renewal test as part of their hearing procedures, the outcomes of any such tests were not reliably transmitted to the 3-Tier Pilot database. Therefore, these customers are excluded from this analysis. Otherwise, these tables include all 3-Tier eligible customers.

As a general matter, there is no consistent relationship among age, failure on the written renewal test, and prior crashes. Among middle-aged customers (those 40-69 years of age), customers who failed the test more often may be somewhat more likely to crash; however, these differences do not rise to conventional levels of statistical significance. Among older customers (those aged 70 and above), crash propensities are relatively consistent across the failure categories.

Table 31

Logistic Regression Predicting Odds of Crashing in the 3 Years Prior to Pilot Enrollment, Based on Written Renewal Test Attempts

	Sub-model A		Sub-model B		Sub-model C	
	β (*SE*)	Exp(β) (CI)	β (*SE*)	Exp(β) (CI)	β (*SE*)	Exp(β) (CI)
S/R indicator			0.37** (0.06)	1.45 (1.29 - 1.62)	0.33** (0.06)	1.40 (1.24 - 1.57)
Gender (male)			0.06 (0.05)	1.06 (0.95 - 1.17)	0.06 (0.05)	1.06 (0.96 - 1.18)
Days of driving per week			0.13** (0.02)	1.14 (1.11 - 1.17)	0.14** (0.02)	1.15 (1.11 - 1.18)
Sometimes avoids driving at night (ref group: never avoids driving at night)					-0.06 (0.07)	0.94 (0.83 - 1.07)
Often/always avoids driving at night (ref group: never avoids driving at night)					-0.22* (0.09)	0.80 (0.67 - 0.96)
Sometimes avoids driving in the rain (ref group: never avoids driving in the rain)					0.10 (0.07)	1.11 (0.97 - 1.27)
Often/always avoids driving in the rain (ref group: never avoids driving in the rain)					0.35** (0.10)	1.42 (1.18 - 1.71)
Sometimes avoids making left turns (ref group: never avoids making left turns)					0.11† (0.07)	1.12 (0.98 - 1.27)
Often/always avoids making left turns (ref group: never avoids making left turns)					0.20** (0.07)	1.23 (1.07 - 1.41)
Passed on 3rd attempt (ref group: passed on 1st or 2nd attempt)	0.14 (0.11)	1.01 (0.81 – 1.27)	0.06 (0.12)	1.06 (0.84 - 1.33)	0.03 (0.12)	1.03 (0.82 - 1.30)
Required 4 or more attempts (ref group: passed on 1st or 2nd attempt)	-0.10 (0.15)	0.99 (0.74 – 1.33)	0.00 (0.16)	1.00 (0.74 - 1.37)	-0.04 (0.16)	0.96 (0.70 - 1.32)

Note. *N* for Sub-model A = 12,160; *N* for Sub-model B = 11,699; *N* for Sub-model C = 11,585. All confidence intervals (CI) were calculated at a 95% confidence level.
*: $p \leq 0.05$. **: $p \leq 0.01$. †: $p \leq 0.10$.

Approximately 11% of this age group has a crash on their 3-year prior record, regardless of performance on this test.

Table 31 presents a logistic regression predicting the relative odds of having crashed in the prior 3 years, as a result of outcomes on the written renewal test, with and without statistical controls. There does not appear to be a relationship between outcomes on this test and prior crash history.

The next two tables present a similar set of analyses, for outcomes on the PRT. For these analyses, customers who somewhat failed (had a score of 24-40), extreme failed (had a score of 41-500), or aborted (i.e., timed out on the test) are coded in a single category. Customers who passed the test are the referent category. The analysis is restricted to those who took this screening test.

Table 32

Cross-Tabulation of Prior Crashes, Age, and Outcome on the PRT

Age group	PRT outcome	*N* of crash-free customers (%)	*N* of crashed 1+ times in prior 3 years (%)	χ^2 (Fisher's exact test, 2-sided)
<40 y.o.	Pass	172 (78.2)	48 (21.8)	0.09 (1.00)
	Fail	5 (83.3)	1 (16.7)	
40-69 y.o.	Pass	302 (81.4)	69 (18.6)	0.47 (0.67)
	Fail	36 (85.7)	6 (14.3)	
70+ y.o.	Pass	1473 (89.7)	170 (10.3)	5.37 (0.03)
	Fail	338 (85.6)	57 (14.4)	

Note. "Fail" includes SFail (score of 24-40), XFail (score of 41-500), and abort (timed out). Table includes all 3-Tier eligible renewal customers in the Pilot who took the PRT.

Due to small cell-size counts, the findings presented in Table 32 should be treated with caution. That said, it appears that among those aged 70 and above, failure on the PRT is associated with a higher rate of crashes in the 3 years prior to enrollment in the Pilot, a finding that is statistically significant.

Table 33

Logistic Regression Predicting Odds of Crashing in the 3 Years Prior to Pilot Enrollment, Based on PRT Outcome

	Sub-model A		Sub-model B		Sub-model C	
	β (*SE*)	Exp(β) (CI)	β (*SE*)	Exp(β) (CI)	β (*SE*)	Exp(β) (CI)
S/R indicator			0.47* (0.21)	1.59 (1.05 – 2.42)	0.38† (0.22)	1.47 (0.96 - 2.25)
Gender (male)			-0.04 (0.12)	0.96 (0.76 - 1.22)	-0.04 (0.12)	0.97 (0.76 - 1.23)
Days of driving per week			0.11** (0.03)	1.12 (1.05 - 1.19)	0.13** (0.03)	1.14 (1.07 - 1.22)
Sometimes avoids driving at night (ref group: never avoids driving at night)					-0.08 (0.16)	0.92 (0.67 - 1.26)
Often/always avoids driving at night (ref group: never avoids driving at night)					-0.30* (0.20)	0.74 (0.50 - 1.09)
Sometimes avoids driving in the rain (ref group: never avoids driving in the rain)					0.27 (0.17)	1.31 (0.95 - 1.81)
Often/always avoids driving in the rain (ref group: never avoids driving in the rain)					0.60** (0.22)	1.82 (1.19 - 2.77)
Sometimes avoids making left turns (ref group: never avoids making left turns)					0.02 (0.16)	1.02 (0.75 - 1.39)
Often/always avoids making left turns (ref group: never avoids making left turns)					0.22 (0.15)	1.25 (0.92 - 1.69)
Fail PRT (ref group: pass)	0.14 (0.15)	1.15 (0.86 – 1.54)	0.18 (0.16)	1.20 (0.88 - 1.63)	0.15 (0.16)	1.16 (0.85 - 1.58)

Note. *N* for Sub-model A = 2,677; *N* for sub-model B = 2,583; *N* for Sub-model C = 2,557. All sub-models include all 3-Tier eligible customers. All confidence intervals (CI) were calculated at a 95% confidence level.
*: $p \leq 0.05$. **: $p \leq 0.01$. †: $p \leq 0.10$.

Table 33 presents a logistic regression predicting the relative odds of having crashed in the prior 3 years, as a result of outcomes on the PRT, with and without statistical controls. As with the previous table, the analysis is restricted to those who actually took this screening test. Although customers who failed the test are at a directionally greater risk of crashing as compared to those who passed this test, this difference does not rise to standard levels of statistical significance. This may be due in part to the fact that the test was not administered to all customers: only those who had failed one or more Tier 1 screening tests, or who failed the written renewal test at least twice, took the PRT. This reduces the relevant N quite substantially, and consequently the amount of statistical power.

The Validity of the Screening Tests, in Combination, at Predicting Prior Crashes

The final three tables in this section present a set of analyses using the overall measure of limited functionality that combines all of the Tier 1 and Tier 2 screening tests. Customers who passed all tests are the referent category. As with the previous analyses, the tables include all 3-Tier eligible customers. If a customer did not take a given screening test (for instance, the PRT), their scores were not imputed. However, missing information regarding the number of written renewal test attempts was available from the Audits database; these values were used where such information was missing from Pilot paperwork.

Due to small cell-size counts, the findings presented in Table 34 should be treated with caution. That said, it appears that among customers aged 40 and older, there is a slight rise in the proportion crashing in the prior 3 years, as a function of combined 3TAS score. Those coded as "extreme fails" are slightly more likely to have crashes on their record, as compared to customers who passed all screening tests. However, this pattern does not hold true for younger customers (<40 y.o.), and the differences do not rise to standard levels of statistical significance for any age cohort.

Table 34

Cross-Tabulation of Prior Crashes, Age, and Combined 3TAS Score

Age group	3-TAS score	*N* of crash-free customers (%)	*N* of crashed 1+ times in prior 3 years (%)	χ^2 (2-tailed *p*-value)
<40 y.o.	Pass	2601 (78.3)	720 (21.7)	0.66 (0.72)
	SFail	189 (76.2)	59 (23.8)	
	XFail	91 (79.1)	24 (20.9)	
40-69 y.o.	Pass	2095 (82.3)	451 (17.7)	1.81 (0.40)
	SFail	290 (81.7)	65 (18.3)	
	XFail	149 (78.4)	41 (21.6)	
70+ y.o.	Pass	2676 (88.6)	344 (11.4)	3.98 (0.14)
	SFail	1280 (89.8)	145 (10.2)	
	XFail	816 (87.2)	120 (12.8)	

Note. See pp. 47-49 for more details on the construction of the combined measure of limited functionality. Table contains all 3-Tier eligible renewal customers in the Pilot.

These patterns differ quite markedly from those presented in the 3-Tier Technical Report (Hennessy & Janke, 2009). These differences may be due to differences in the sampled populations (e.g., Northern versus Southern California), sample size, or to effects of history (the studies were done approximately 6 years apart). They may also be due to the fact that the two studies involved somewhat different batteries of tests. In the prior study, all participating customers were subjected to all tests; in the current Pilot, only a small minority of customers were required to take the PRT. This likely produced some difference in the number of people categorized as either "somewhat" or "extremely" functionally limited. However, it does not appear that the differences are due to the nature of the comparisons. Restricting the sample to full-term renewal customers—i.e., excluding limited-term license holders—does not substantially change the relationship between combined 3TAS score and crash propensity, as can be seen in Table 35.

Table 35

Cross-Tabulation of Prior Crashes, Age, and Combined 3TAS Score (Full-Term Licensees Only)

Age group	3-TAS score	*N* of crash-free customers (%)	*N* of crashed 1+ times in prior 3 years (%)	χ^2 (2-tailed *p*-value)
<40 y.o.	Pass	2600 (78.3)	720 (21.7)	0.74 (0.69)
	SFail	189 (76.2)	59 (23.8)	
	XFail	90 (79.6)	23 (20.4)	
40-69 y.o.	Pass	2094 (82.3)	449 (17.7)	1.36 (0.51)
	SFail	289 (81.9)	64 (18.1)	
	XFail	134 (78.8)	36 (21.2)	
70+ y.o.	Pass	2661 (88.6)	341 (11.4)	3.16 (0.21)
	SFail	1251 (89.8)	142 (10.2)	
	XFail	624 (87.3)	91 (12.7)	

Note. See pp. 47-49 for more details on the construction of the combined measure of limited functionality.

Table 36 presents a logistic regression predicting the relative odds of having crashed in the prior 3 years, as a result of overall 3TAS score, with and without statistical controls. Customers who somewhat failed one screening test appear to have substantially and significantly lower odds of crashing, as compared to those who passed all tests. Customers who extreme failed one test, or who failed multiple screening tests, have directionally lower risk of crashing as compared to those who passed all tests. However, the effect size is small, and the difference in odds does not rise to standard levels of statistical significance.

An alternative model to that presented in Table 36 was also run using just those customers aged 70+. The results are somewhat different: somewhat failing one or more assessment tests is associated with a directionally decreased risk of crashes in the prior 3 years, while extreme fails are associated with a directionally higher risk of crashes. However, for none of the sub-models (A, B, or C) does the estimation of the beta coefficients for overall 3TAS outcome score rise to conventional standards of statistical significance. This model is available upon request from the author.

Table 36

Logistic Regression Predicting Odds of Crashing in the 3 Years Prior to Pilot Enrollment, Based on Combined 3TAS Score

	Sub-model A		Sub-model B		Sub-model C	
	β (*SE*)	Exp(β) (CI)	β (*SE*)	Exp(β) (CI)	β (*SE*)	Exp(β) (CI)
S/R indicator			0.35** (0.06)	1.42 (1.26 – 1.59)	0.31** (0.06)	1.36 (1.21 – 1.54)
Gender (male)			0.06 (0.05)	1.06 (0.96 – 1.17)	0.06 (0.05)	1.06 (0.96 – 1.18)
Days of driving per week			0.13** (0.02)	1.13 (1.10 – 1.17)	0.13** (0.02)	1.14 (1.11 – 1.18)
Sometimes avoids driving at night (ref group: never avoids driving at night)					-0.05 (0.07)	0.95 (0.83 – 1.08)
Often/always avoids driving at night (ref group: never avoids driving at night)					-0.21* (0.09)	0.81 (0.68 – 0.97)
Sometimes avoids driving in the rain (ref group: never avoids driving in the rain)					0.10 (0.07)	1.11 (0.97 – 1.26)
Often/always avoids driving in the rain (ref group: never avoids driving in the rain)					0.34** (0.10)	1.41 (1.17 – 1.70)
Sometimes avoids making left turns (ref group: never avoids making left turns)					0.11 (0.07)	1.11 (0.98 – 1.27)
Often/always avoids making left turns (ref group: never avoids making left turns)					0.21** (0.07)	1.23 (1.07 – 1.41)
SFail (ref group: pass)	-0.30** (0.07)	0.74 (0.65 – 0.86)	-0.18* (0.07)	0.84 (0.72 – 0.97)	-0.18* (0.08)	0.84 (0.74 – 0.99)
XFail (ref group: pass)	-0.16† (0.08)	0.85 (0.72 – 1.01)	-0.02 (0.09)	0.98 (0.82 – 1.17)	-0.04 (0.09)	0.96 (0.80 – 1.15)

Note. *N* for Sub-model A = 12,156; *N* for sub-model B = 11,696; *N* for Sub-model C = 11,582. All confidence intervals (CI) were calculated at a 95% confidence level.
*: $p \leq 0.05$. **: $p \leq 0.01$. †: $p \leq 0.10$.

In sum, the analyses presented in this section do not confirm a consistent predictive relationship between 3-year prior crash record and outcomes on the 3TAS screening tests, whether taken singly or in combination. If anything, customers who fail the various screening tests appear to have somewhat better prior crash records, when controlling for gender, suspension/revocation indicators, and self-reported days of driving per week. This appears to be particular true for the physical observation protocol and the Pelli-Robson chart. Failures on other screening tests, when examined in isolation, do not indicate a consistent relationship with prior crashes, though failure on the PRT may be associated with greater crash risk among drivers aged 70 and older. Taken in combination, customers classified as "somewhat functionally limited" appear to have a lower risk of crashing in the 3 years prior to renewal compared to customers who passed all screening tests. Customers classified as "extremely functionally limited" may also have a lower risk of crashing compared to customers who passed all screening tests; however, this association does not rise to conventional standards of statistical significance.

Finally, it appears that some driving avoidance behaviors, as indicated by the Driving Information Survey, are associated with reduced prior crash risk—especially avoiding driving at night. However, other driving avoidance behaviors—especially avoiding driving in the rain, and avoiding left turns—appear to be associated with increased prior crash risk. Controlling for self-reported driving avoidance behaviors does not appear to alter the modeled relationship between screening test outcomes (either singly or in combination) and prior crash record.

It is quite possible that the various screening test outcomes of the 3TAS battery, singly or in combination, bear a predictive relationship to *subsequent* crashes, rather than prior crash involvement. Staplin, Lococo, Gish, and Decina (2003b), for instance, generally found statistically significant relationships between subsequent crash records and some of the tests used in the 3TAS battery (e.g., the PRT, see especially Figure 33 of their Technical Report). However, due to the non-randomized/quasi-experimental nature of the 3-Tier Pilot, it is not possible in the current analysis to distinguish between individual screening test outcomes and the simple effect of testing.

Question #8: What is the Relationship Between Screening Test Outcomes and Drive Test Outcomes?

This section presents a confirmatory analysis regarding the relationship between functional status, as measured by the components of the 3-Tier Assessment System, and performance on an on-road test of driving skill. Comparing drive-test fail rates across Pilot customers may be interpreted as an estimation of the utility of 3TAS for identifying customers with impaired or reduced driving skill, in comparison with other (non 3-Tier) methods by which CA DMV identifies such customers.

For the purposes of the following analysis, three groups of customers are compared. If a Pilot customer was required to take a drive test because of a Driver Safety referral, they are coded as "Driver Safety." This includes a pair of customers who were referred to Driver Safety during the course of a renewal transaction, and returned during the Pilot for their required road test. If a customer was required to take a road test due to their participation in the 3-Tier Pilot, they are coded as "3-Tier generated." This includes customers whose required road test may have been triggered in part due to a vision condition, as (for instance) indicated by failure on the visual acuity standard. If a Pilot customer was required to take a drive test in part because of their possession of a pre-existing limited-term license, they are coded as "limited-term." However, this third category does not include customers who were issued limited-term licenses specifically as a result of participation in the Pilot; these latter are instead categorized as "3-Tier generated."

The analysis presented here is somewhat different from that conducted in prior studies that led to the development of the 3-Tier Assessment System (Hennessy & Janke, 2009; Janke & Hersch, 1997). The results discussed here are not strictly comparable with the findings in those reports regarding the relationship between screening outcomes and drive test results, for several reasons. First, only a small minority of 3-Tier Pilot participants were required to take a drive test. All customers who were required to take a drive-test in the Pilot had some kind of major functional limitation—either identified in the course of the Pilot, or by a medical professional as in the case of limited-term license holders. There is, therefore, relatively little variation in the severity of overall functional status among those 3-Tier Pilot customers required to take drive tests. Secondly, because of the nested nature of the screening tests used in the Pilot, it is not really possible to segregate customers who failed the visual acuity standard, for comparison with other types of 3-Tier Pilot participants. Third, in Hennessy and Janke (2009) the screening tests were given to all participants—including Driver Safety referrals. In the present 3-Tier Pilot, Driver Safety referral customers were not subjected to any 3-Tier Assessment System screening tests. Therefore, it is not possible in the present analysis to group Driver Safety referrals according to level of functional impairment as determined by the 3-Tier Assessment System. Fourth, the present analysis includes information on all road tests (i.e., both SDPE and ADPE tests). Due to the rarity of the ADPE—especially as a first test—excluding such customers from the analysis would not meaningfully change the results presented here. Finally, it is worth noting that in some cases the possession of a limited-term license is associated with a progressive mental or physical condition, rather than a vision condition. Therefore, those customers categorized here as limited-term customers should not be confused with the category of "visual acuity referrals" as discussed in the Technical Report (Hennessy & Janke, 2009).

Table 37

Proportion Failing the First Drive Test, by Age and Reason for Drive-Test Requirement

Age group	Reason for drive test	Passed 1st drive test (%)	Failed 1st drive test (%)	χ^2 (2-tailed *p*-value)
<40 y.o.	3-Tier generated	18 (81.8%)	4 (18.2%)	0.22 (0.90)
	Driver Safety	14 (82.4%)	3 (17.6%)	
	Limited-term	1 (100.0)	0 (0.0%)	
40-69 y.o.	3-Tier generated	41 (73.2%)	15 (26.8%)	1.05 (0.59)
	Driver Safety	39 (79.6%)	10 (20.4%)	
	Limited-term	15 (83.3%)	3 (16.7%)	
70+ y.o.	3-Tier generated	342 (77.9%)	97 (22.1%)	25.44 (0.00)
	Driver Safety	25 (46.3%)	29 (53.7)	
	Limited-term	128 (75.3%)	42 (24.7%)	

Table 37 includes all 3-Tier eligible customers who were required to take drive tests, including erroneously-processed and lagging customers. Because many lagging and erroneously-processed customers later took drive tests (after the conclusion of the Pilot), and because the outcomes on these tests can be confirmed through the use of the Audits database, the number of "3-Tier generated" drive tests differs substantially from similar figures presented in the Process Analysis

(Camp 2010b). The column "failed drive test" refers only to the first drive test taken; customer outcomes on second (or subsequent) drive tests are not included here.

Among drivers younger than 70 y.o., there are no statistically significant differences in drive-test fail rates across the three types of customers. However, this finding should be interpreted with some caution, due to small cell-size counts, especially for limited-term drivers, of whom there are fewer than 20 cases among drivers younger than 70 y.o. Among drivers aged 70 and older, Driver Safety referrals have a substantially and significantly higher fail rate on the drive test, as compared to 3-Tier generated or limited-term renewal customers. If one excludes Driver Safety referrals, there is no significant difference in fail rates between 3-Tier generated drive-test customers and limited-term renewal customers aged 70 and older (χ^2 value =0.47, with an associated two-sided Fisher's Exact Test significance of 0.52). This suggests that, among drivers aged 70 and older, customers required to take a road test because of 3TAS screening have approximately the same likelihood of failing a road test as customers who have been assigned a limited-term license on the basis of a formal medical evaluation and recommendation from a health professional.

Based on methodological problems reported in the Process Analysis (Camp, 2010b), it must be acknowledged that the results discussed in this particular section may involve serious bias. Every effort was made to reconstruct drive-test results for Pilot customers from the Audits database; this resulted in the imputation of several dozen data points for those lagging and erroneously-processed customers who eventually took an SDPE or ADPE. However, this only captured those customers so tested. It remains true that several hundred customers were required to take a drive test under 3-Tier Pilot procedures, and yet never did take such a test, according to the available evidence. Given the available data, it is not possible to speculate on how these untested customers would have performed on an SDPE or ADPE. The results presented in Table 37 may therefore have been quite different if these customers had actually been given on-road drive-tests. This form of bias likely affects the results of rows titled "3-Tier generated" drive tests more so than the other two categories.

Question #9: What Was the Effect of the 3-Tier Pilot on Reducing Subsequent Crashes?

Overall Program Effects on Subsequent Crashes: *Pilot Cohort vs. Baseline II Cohort and Pilot Cohort vs. Nearby Cohort*

One of the main goals of the 3-Tier Assessment System was to reduce the subsequent risk of crashing for individual drivers. This reduction in risk was intended to occur primarily through the mechanism of making drivers aware of any driving-relevant limitations they may possess, and educating them about how to compensate for these limitations so as to improve the safety of their driving. In addition, 3TAS also aimed to reduce crashes by restricting or revoking the driving privilege of those drivers who are no longer able to drive safely.

In this section, a series of analyses are conducted, comparing the subsequent 2-year crash rates of 3-Tier Pilot participants with customers in the Nearby and Baseline II cohorts. The primary statistical technique used is Cox proportional hazards modeling. The outcome of interest consists of a simple dummy variable (crashed versus crash-free) in the 2 years subsequent to the date of first contact. Crashes are included regardless of fault status. It is hypothesized that customers in the 3-Tier Pilot will have a lower hazard rate of crashing, other effects held constant, than customers in the Nearby and Baseline II cohorts.

In order to model the observed time (i.e., the number of days to event), the following procedure was used. The reference date for all customers—regardless of cohort—consisted of the start date of their application (if a renewal customer) or the date of the first drive test (if a Driver Safety referral customer).[10] The maximum number of days to event was then set to 731 (365 x 2, +1), to account for the leap year in 2008.[11] If a customer crashed, the value of this variable was reduced to the number of days between the reference date and the date of their first subsequent

[10]Setting the reference date for Driver Safety referral customers in this manner implicitly assumes that the main effect of treatment for referral customers consists of the on-road drive test, rather than whatever interviews and processing may have occurred prior. These latter forms of processing may have been quite involved: for instance, the initial incident that triggered the referral, the submission of medical records if those were required, taking the written renewal test if that was required, the scheduling of the drive test, etc. Because of the way data were collected for the Pilot, however, it was not possible to construct reference dates for these “pre-drive” Driver Safety procedures, especially given that such procedures may have varied a great deal depending on the nature, source, and type of referral (e.g., medical versus law-enforcement referrals, priority re-examination versus other types of re-examinations, etc.). Hence, because the drive test was a common requirement for those Driver Safety referrals enrolled in the Pilot, and because it was possible to collect a firm date for the taking of such a test, this was used as the reference date for processing.

[11]Customers in the Baseline II cohort did not experience a leap year in their observation period. Nevertheless, their days to event was set to 731 to ensure comparability.

crash. If a customer died or switched license class (i.e., to a commercial or motorcycle license), the value of this variable was also reduced, to the number of days between the reference date and the date of death (as reported to CA DMV), or the switching of license class, as appropriate to the individual customer. In a handful of cases, customers appear to have crashed on the reference date; these customers had their days to event set to “1”, which involved assuming that they crashed after their visit to the field office, rather than before.

Each model controls for age, sex, prior record (suspension/revocation indicator, number of convictions, and number of crashes in the 3 years prior to the reference date), Driver Safety referral status, potential delays in licensure, and the annualized crash rate of the zip code of residence for the driver. The first model—Table 38—compares drivers in the Baseline II cohort to drivers in the Pilot cohort. The second model—Table 39—compares drivers in the Nearby cohort to drivers in the Pilot cohort.

Table 38 suggests that, controlling for age, sex, prior record, Driver Safety referral status, potential delays in licensure, and an ecological measure of exposure, participants in the 3-Tier Pilot had approximately 15% lower odds (main effect) of crashing in the 2 years subsequent to participation, as compared to drivers in the Baseline II cohort. All control variables operate in the expected direction. However, interaction terms—which would capture any potential difference in the effect of 3-Tier Pilot participation across age strata—show no effect. Thus, any potential correlation between Pilot participation and subsequent crash savings appears to be evenly distributed across the age spectrum.

That said, this finding must be interpreted with a great deal of caution. As can be seen in the next figure, there was a secular decline in the number of crashes in California (and in the Sacramento metropolitan area) between the exposure period of the Baseline II cohort, and the exposure period of the Pilot and Nearby cohorts.

Table 38

Cox Proportional Hazards Model of Days to First Subsequent Crash (Within 2 Years of Reference Date), Pilot vs. Baseline II

	Model A: main effects		Model B: w/ interaction terms	
	β (*SE*)	Exp(β) (CI)	β (*SE*)	Exp(β) (CI)
Age 40-69 (reference = age<40)	-0.20** (0.05)	0.82 (0.74 – 0.91)	-0.20** (0.06)	0.82 (0.72 – 0.92)
Age 70+ (reference = age<40)	-0.34** (0.05)	0.71 (0.64 – 0.79)	-0.39** (0.07)	0.68 (0.59 – 0.77)
Sex (male)	0.11** (0.04)	1.12 (1.03 – 1.21)	0.11** (0.04)	1.12 (1.03 – 1.21)
Suspension/Revocation in prior 3 years?	0.18** (0.05)	1.20 (1.09 – 1.32)	0.19** (0.05)	1.21 (1.09 – 1.33)
# of convictions in prior 3 years	0.09** (0.01)	1.09 (1.06 – 1.12)	0.09** (0.01)	1.09 (1.06 – 1.12)
# of crashes in prior 3 years	0.28** (0.03)	1.33 (1.26 – 1.40)	0.28** (0.03)	1.33 (1.26 – 1.40)
Driver Safety referral?	0.42† (0.22)	1.52 (0.98 – 2.34)	0.42† (0.22)	1.52 (0.99 – 2.35)
Number of valid days of licensure in follow-up	0.00** (0.00)	1.00 (1.00 – 1.00)	0.00** (0.00)	1.00 (1.00 – 1.00)
Crashes per driver per year in zip code of residence	0.14** (0.05)	1.15 (1.05 – 1.25)	0.14** (0.05)	1.15 (1.05 – 1.25)
3-Tier Pilot participant?	-0.16** (0.04)	0.85 (0.79 – 0.92)	-0.20** (0.06)	0.82 (0.73 – 0.92)
3-Tier Pilot X age 40-69			0.02 (0.10)	1.02 (0.84 – 1.24)
3Tier Pilot X age 70+			0.11 (0.09)	1.12 (0.93 – 1.35)

Note. *N*=27,176 (Baseline II = 14,901, Pilot = 12,275). Models include all 3-Tier eligible customers. All confidence intervals (CI) were calculated at a 95% confidence level.
*: $p \le 0.05$. **: $p \le 0.01$. †: $p \le 0.10$.

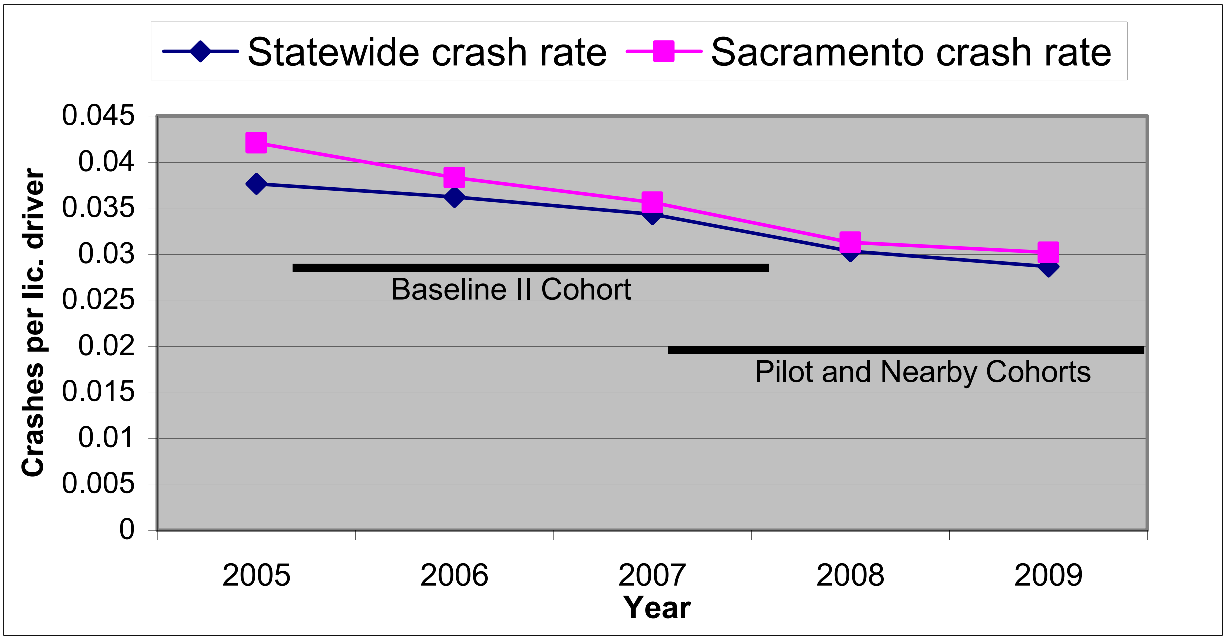

Figure 1: Crashes per licensed driver in California and greater Sacramento metropolitan area, 2005-2009. Source: SWITRS. Crashes include all severity levels.

The Baseline II cohort was enrolled from September 2005 through January 2006; thus, their follow-up period ranges from September 2005 through January 2008. The Pilot cohort was enrolled from June 2007 through October 2007; thus, their follow-up period is from June 2007 through October of 2009. In general, crashes statewide—the bottom line in this graph—declined by 8.7% during the Baseline II follow-up period, but by almost twice that (16.6%) during the Pilot follow-up period. For the Sacramento metro area—the top line in this graph—crashes declined during both the Pilot and Baseline II follow-up periods by about 15%.

For the purposes of these calculations, the Sacramento metro area includes the following counties: Sacramento, El Dorado, Placer, Yolo, and Solano.

The next model compares drivers in the Nearby cohort to those of the Pilot cohort. Again, the analysis includes all 3-Tier eligible customers.

Table 39 suggests that, controlling for age, sex, prior record, potential delays in licensure, and an ecological measure of exposure, participants in the 3-Tier Pilot were not statistically different in their odds of crashing during the 2 years subsequent to participation, as compared to drivers in the Nearby cohort. All control variables operate in the expected direction, and are of approximately the same size as the coefficients found in the previous model. As

Table 39

Cox Proportional Hazards Model of Days to First Subsequent Crash (Within 2 Years of Reference Date), Pilot vs. Nearby

	Model A: main effects		Model B: w/ interaction terms	
	β (*SE*)	Exp(β) (CI)	β (*SE*)	Exp(β) (CI)
Age 40-69 (reference = age<40)	-0.21** (0.06)	0.82 (0.72 – 0.92)	-0.23** (0.09)	0.79 (0.67 – 0.94)
Age 70+ (reference = age<40)	-0.29** (0.06)	0.75 (0.66 – 0.84)	-0.30** (0.09)	0.74 (0.62 – 0.88)
Sex (male)	0.10* (0.05)	1.10 (1.00 – 1.21)	0.10* (0.05)	1.10 (1.00 – 1.21)
Suspension/Revocation in prior 3 years?	0.15** (0.06)	1.17 (1.03 – 1.32)	0.15** (0.06)	1.17 (1.03 – 1.32)
# of convictions in prior 3 years	0.10** (0.02)	1.11 (1.07 – 1.15)	0.10** (0.02)	1.11 (1.07 – 1.15)
# of crashes in prior 3 years	0.24** (0.04)	1.27 (1.18 – 1.37)	0.24** (0.04)	1.27 (1.18 – 1.37)
Number of valid days of licensure in follow-up	0.00** (0.00)	1.00 (1.00 – 1.00)	0.00** (0.00)	1.00 (1.00 – 1.00)
Crashes per driver per year in zip code of residence	0.16** (0.05)	1.17 (1.06 – 1.30)	0.16** (0.05)	1.17 (1.06 – 1.30)
Driver Safety referral?	0.64** (0.22)	1.89 (1.23 – 2.92)	0.64** (0.22)	1.89 (1.22 – 2.92)
3-Tier Pilot participant?	0.02 (0.05)	1.02 (0.93 – 1.13)	0.01 (0.08)	1.01 (0.87 – 1.17)
3-Tier Pilot X age 40-69			0.05 (0.12)	1.05 (0.83 – 1.32)
3Tier Pilot X age 70+			0.01 (0.11)	1.01 (0.82 – 1.26)

Note. *N*=22,824 (Nearby = 10,549, Pilot = 12,275). Models include all 3-Tier eligible customers. All confidence intervals (CI) were calculated at a 95% confidence level.
**: $p \leq 0.01$. *: $p \leq 0.05$. †: $p \leq 0.10$.

with the Baseline II group, interaction terms—which would capture any potential difference in the effect of 3-Tier Pilot participation across age strata—show no effect.

Both of the analyses presented in this section present substantial methodological problems. The follow-up period for the Baseline II cohort only partially overlaps with that of the Pilot cohort. As shown in Figure 1, there was a strong secular decline in crashes between the two follow-up periods. This produces the appearance of net crash savings associated with the 3-Tier Pilot that may, in fact, be largely an artifact of extraneous and uncontrollable bias in the modeling. Due to the near-complete separation in follow-up periods (i.e., the narrow window of overlap), it is not really possible to incorporate a variable that would allow statistical separation of the effect of history from the potential effect of the 3-Tier Pilot.

On the other hand, the comparison between the Nearby and Pilot cohorts is also problematic, not least because the driving environments of these groups of drivers are likely quite different. As noted in Table 1 (p. 19), the mean number of crashes per licensed driver per year in the zip code of residence is somewhat different across the Nearby and Pilot cohorts. And while this variable is included as a control, this may not control for all differences in driving environment, or for that matter other, unmeasured, variables associated with the zip code of residence.

In a separate series of models focused on the Nearby and Pilot cohorts, the author constructed a hierarchical design, with individuals nested within offices, to test for potential differences in crash risk associated with variation across field offices. These models demonstrated that there was no significant residual variation in crash risk associated with field offices as a grouping variable. Using a nested design did not substantially alter the size, direction, or significance levels of any of the other coefficients of interest (e.g., the effect of program participation, age, etc.), in comparison to the Cox modeling reported in Table 39. For the sake of brevity these models are omitted from the present manuscript; however, they are available upon request from the author.

Despite these potential problems in the comparisons, nevertheless it can be stated that there exists little evidence for an overall program effect of the 3-Tier Pilot on reducing subsequent crashes among participating customers. Similarly, there exists no evidence of an overall effect of the Pilot on the crash risk of older drivers, as indicated by the lack of significance in the interaction terms included in Tables 38 and 39.

Overall Program Effects on Subsequent At-Fault Crashes: *Pilot Cohort vs. Baseline II Cohort and Pilot Cohort vs. Nearby Cohort*

An alternative method of analysis focuses on the fault status of subsequent crashes. It is possible that, while the 3-Tier Assessment System Pilot had no overall effect on crash propensity, it may have made drivers safer by reducing their likelihood of *causing* a crash.

In the following analyses, the proportional hazards of being at fault for an injury or fatal crash are calculated for Pilot cohort drivers in comparison to drivers in (i) the Nearby and (ii) the Baseline II cohorts. The analysis is restricted to injury/fatal crashes for the simple reason that CA DMV records list fault status for property-damage-only (PDO) crashes only if they are reported to the agency by law enforcement. PDO crashes reported by drivers (required in cases where the damage exceeds $500), or by insurance companies, are not assigned a fault status on the permanent driver record. Unfortunately, restricting the analysis to injury/fatal crashes radically reduces statistical power.

Customers found to be “most at fault” or to have “contributed to the cause of the accident” were coded as being at fault. Customers found to be “not at fault,” or “not responsible following DMV hearing” were coded as not being at fault. If the fault status was “fault not determined” or if the crash involved an emergency vehicle—which occurred in 23 cases for injury/fatality crashes—these drivers were coded as “not at fault.” Drivers who were not in a fatal/injury crash were excluded from the analysis.

Table 40 compares Pilot cohort customers to customers in the Baseline II cohort in their likelihood of being found to be at fault for an injury/fatal crash in the 2 years subsequent to participation in the study. The model does not include interaction terms for age due to lack of statistical power.

Table 40

Cox Proportional Hazards Model of Days to First Subsequent At-Fault Injury/Fatal Crash (Within 2 Years of Reference Date), Pilot vs. Baseline II

	β (*SE*)	Exp(β) (CI)
Age 40-69 (reference = age<40)	-0.06 (0.14)	0.94 (0.71 – 1.25)
Age 70+ (reference = age<40)	0.15 (0.17)	1.17 (0.84 – 1.61)
Sex (male)	-0.03 (0.12)	0.97 (0.77 – 1.22)
Suspension/Revocation in prior 3 years?	0.08 (0.13)	1.08 (0.83 – 1.41)
# of convictions in prior 3 years	0.08* (0.04)	1.08 (1.01 – 1.17)
# of crashes in prior 3 years	0.17* (0.08)	1.18 (1.01 – 1.39)
Driver Safety referral?	0.34 (0.53)	1.41 (0.50 – 3.99)
Number of valid days of licensure in follow-up	-0.00† (0.00)	1.00 (1.00 – 1.00)
Crashes per driver per year in zip code of residence	-0.08 (0.13)	0.92 (0.71 – 1.20)
3-Tier Pilot participant?	-0.24* (0.12)	0.79 (0.63 – 0.99)

Note. N = 683 (Baseline II = 415, Pilot = 268). Model includes all 3-Tier eligible customers whose first subsequent crash involved an injury or fatality. All confidence intervals (CI) were calculated at a 95% confidence level.
*: $p \leq 0.05$. **: $p \leq 0.01$. †: $p \leq 0.10$.

The model displayed in Table 40 shows that, compared to drivers in the Baseline II cohort, 3-Tier Pilot participants had approximately 20% lower odds of being at fault for an injury/fatal crash. Although the obtained beta coefficient meets the standard threshold for statistical significance ($p<0.05$), the reader should remember the history effects discussed on p. 88. Drivers in the Baseline II cohort appear to have driven under much different conditions than Pilot cohort participants, and findings in Table 40 may be largely an artifact of history. Not only did overall crash rates go down both statewide and in the Sacramento metropolitan area (see Figure I), but fatal/injury crash rates also declined. During the Baseline II follow-up period, fatal/injury crashes (combining injury severity levels), declined statewide by about 10%; during the Pilot cohort follow-up period, fatal/injury crashes declined statewide by about 14%.

Table 41

Cox Proportional Hazards Model of Days to First Subsequent At-Fault Injury/Fatal Crash (Within 2 Years of Reference Date), Pilot vs. Nearby

	β (*SE*)	Exp(β) (CI)
Age 40-69 (reference = age<40)	-0.38* (0.18)	0.69 (0.48 – 0.97)
Age 70+ (reference = age<40)	-0.04 (0.20)	0.97 (0.65 – 1.44)
Sex (male)	0.16 (0.15)	1.18 (0.88 – 1.57)
Suspension/Revocation in prior 3 years?	0.17 (0.17)	1.19 (0.85 – 1.66)
# of convictions in prior 3 years	0.10* (0.05)	1.11 (1.01 – 1.21)
# of crashes in prior 3 years	0.06 (0.10)	1.06 (0.87 – 1.28)
Driver Safety referral?	0.32 (0.60)	1.37 (0.43 – 4.44)
Number of valid days of licensure in follow-up	0.00 (0.00)	1.00 (1.00 – 1.00)
Crashes per driver per year in zip code of residence	-0.22 (0.16)	0.81 (0.59 – 1.10)
3-Tier Pilot participant?	-0.23 (0.15)	0.80 (0.59 – 1.07)

Note. N = 442 (Nearby = 174, Pilot = 268). Model includes all 3-Tier eligible customers whose first subsequent crash involved an injury or fatality. All confidence intervals (CI) were calculated at a 95% confidence level.
*: $p \leq 0.05$. **: $p \leq 0.01$. †: $p \leq 0.10$.

Table 41 displays a similar analysis, comparing drivers in the Nearby cohort to 3-Tier Pilot participants. The model shows that, compared to drivers in the Nearby cohort, 3-Tier Pilot participants had directionally lower odds of being at fault for an injury/fatal crash. The effect is similar in size and direction as that found in Table 40; however, the estimate does not meet conventional thresholds of statistical significance. The *p*-value for cohort status ("3-Tier Pilot participant?") in Table 41 is 0.14, and the 95% confidence interval includes the value of 1; taken together, this means that the null hypothesis cannot be rejected, and that the direction of the effect could, in fact, be the opposite of what is indicated here.

Together, this suggests that at best there exists weak, but inconsistent, evidence for a positive effect of the 3-Tier Pilot in reducing the likelihood of being found at fault for subsequent injury/fatal crashes. Due to the lack of statistical power, as well as severe methodological problems discussed earlier in this manuscript, this finding must be interpreted with a great deal of caution.

Question #10: What is the Overall Program Effect of the 3-Tier Pilot on Reducing Convictions?

Although the 3-Tier Pilot may have had no effect on overall crashes, there is weak evidence of a potential effect in reducing at-fault crashes. Given this latter finding, the Pilot may also have had an effect on other kinds of behavior normally associated with negligent operators. The next two analyses test the idea that the Pilot may have had a beneficial safety effect by reducing subsequent convictions.

Tables 42 and 43 contain Cox proportional hazards models predicting days to first subsequent conviction, controlling for age, sex, prior record (suspension/revocation indicator, number of crashes, and number of convictions, all for the 3 years prior to enrollment), Driver Safety referral status, number of valid licensure days in the follow-up period, and the annualized number of convictions per licensed driver in the zip code of residence. This latter variable is slightly different from the exposure variable used to predict crashes, but is nevertheless intended to capture some amount of the differences associated with the driving environment.

Compared to drivers in the Baseline II cohort, participation in the 3-Tier Pilot program does not appear to be associated with a statistically significant difference in the relative odds of receiving a conviction for a traffic violation in the 2 years subsequent to enrollment.

Table 42

Cox Proportional Hazards Model of Days to First Subsequent Conviction (Within 2 Years of Reference Date), Pilot vs. Baseline II

	β (*SE*)	Exp(β) (CI)
Age 40-69 (reference = age<40)	-0.45** (0.03)	0.64 (0.60 – 0.68)
Age 70+ (reference = age<40)	-1.67** (0.05)	0.19 (0.17 – 0.21)
Sex (male)	0.35** (0.03)	1.41 (1.34 – 1.49)
Suspension/Revocation in prior 3 years?	0.32** (0.03)	1.38 (1.31 – 1.46)
# of convictions in prior 3 years	0.22** (0.01)	1.24 (1.22 – 1.26)
# of crashes in prior 3 years	0.09** (0.02)	1.09 (1.05 – 1.14)
Number of valid days of licensure in follow-up	0.00 (0.00)	1.00 (1.00 – 1.00)
Convictions per driver per year in zip code of residence	0.03** (0.01)	1.04 (1.02 – 1.05)
Driver Safety referral?	-0.31 (0.20)	0.73 (0.50 – 1.08)
3-Tier Pilot participant?	-0.01 (0.03)	0.99 (0.94 – 1.04)

Note. *N*=27,176 (Baseline II = 14,901, Pilot = 12,275). Model includes all 3-Tier eligible customers. All confidence intervals (CI) were calculated at a 95% confidence level.
*: $p \leq 0.05$. **: $p \leq 0.01$. †: $p \leq 0.10$.

Compared to drivers in the Nearby cohort, participation in the 3-Tier Pilot appears to be associated with a small, but statistically significant, greater odds of receiving a conviction for a traffic violation in the 2 years subsequent to participation in the program. Given the differences in driving environments among customers in these two cohorts—differences which may include the severity of law enforcement regimes across different jurisdictions—these differences may simply be an artifact of uncontrolled bias in the estimates. This finding should therefore be treated with a great deal of caution.

Table 43

Cox Proportional Hazards Model of Days to First Subsequent Conviction (Within 2 Years of Reference Date), Pilot vs. Nearby

	β (*SE*)	Exp(β) (CI)
Age 40-69 (reference = age<40)	-0.45** (0.03)	0.64 (0.60 – 0.68)
Age 70+ (reference = age<40)	-1.62** (0.05)	0.20 (0.18 – 0.22)
Sex (male)	0.34** (0.03)	1.40 (1.32 – 1.49)
Suspension/Revocation in prior 3 years?	0.36** (0.03)	1.43 (1.34 – 1.53)
# of convictions in prior 3 years	0.23** (0.01)	1.26 (1.23 – 1.28)
# of crashes in prior 3 years	0.11** (0.02)	1.12 (1.07 – 1.17)
Number of valid days of licensure in follow-up	0.00 (0.00)	1.00 (1.00 – 1.00)
Convictions per driver per year in zip code of residence	0.04** (0.01)	1.04 (1.03 – 1.05)
Driver Safety referral?	-0.37† (0.21)	0.69 (0.46 – 1.05)
3-Tier Pilot participant?	0.06* (0.03)	1.07 (1.01 – 1.13)

Note. *N*=22,824 (Nearby = 10,549, Pilot = 12,275). Model includes all 3-Tier eligible customers. All confidence intervals (CI) were calculated at a 95% confidence level.
*: $p \leq 0.05$. **: $p \leq 0.01$. †: $p \leq 0.10$.

In sum, there is weak, but inconsistent, evidence that participation in the 3-Tier Pilot is associated with a small increase in the odds of receiving a conviction in the 2 years subsequent to participation.

Question #11: What Effect Did the Educational Material Distributed As Part of Tier 3 Have on Individual Risk of Crashing?

As noted in the 3-Tier Process Analysis (Camp, 2010b), it cannot be determined with any degree of confidence which individuals received the educational materials distributed as part of the 3-Tier Pilot, and which did not. The original design of the Pilot called for a randomizing method of distribution, whereby those customers who possessed a driver license with an odd-numbered final digit, and who either somewhat failed the contrast sensitivity screen or who somewhat failed the PRT (or both) would receive a package of educational materials that discussed specifically how to adjust one's driving habits to compensate for potential limitations in these areas. In practice, it appears that these educational materials were found to be quite useful in preparing customers for their upcoming drive test, reducing test-related anxiety, in answering customer questions about the drive-test requirement, and generally in providing good customer service. This meant that these materials were likely distributed to a much broader set of customers—including many who extreme failed on the PRT, or who were required to take an on-road drive test for reasons unrelated to contrast sensitivity or perceptual limitations (e.g., multiple physical limitations, or multiple failures on the written test). Unfortunately, the notation of which customers received the materials was inconsistent, and qualitative evidence collected as part of the staff interviews—see the 3-Tier Process Analysis Appendix (Camp, 2010a) for more details—indicates the existence of multiple types of errors in these data. In other words, some customers who are <u>not</u> recorded as having received the educational intervention probably did receive it, while other customers who <u>are</u> recorded as having received the educational intervention may not in fact have received it. It therefore cannot be determined whether the educational materials used as part of the 3-Tier Pilot had any effect on subsequent crash risk.

Question #12: What Effect Did the 3-Tier Pilot Have on Licensing/Mobility Options in the Post-Participation Period?

As demonstrated in the analyses under Question #3 (pp. 37-50), the overwhelming majority of 3-Tier Pilot customers retained their driving privilege upon completion of the pilot. Even among the most severely functionally-limited group—those who extreme failed one screening test, or who failed multiple screening tests—over 85% of 3-Tier eligible customers successfully renewed their driving privilege. Nonetheless, it appears that customers who were found to possess driving-relevant limitations generally took longer to renew their license.

In this subsection, the comparison of licensing and mobility outcomes is extended to customers in the Nearby and Baseline II cohorts. Briefly stated, was participation in the Pilot associated with delays in renewal of the driving privilege? Were customers in the Pilot cohort more (or less) likely to have restrictions placed upon the driving privilege? Were customers in the 3-Tier Pilot more (or less) likely to retain their driving privilege?

Overall Program Effects on Time to Renewal: Pilot vs. Baseline II and Pilot vs. Nearby

Table 44 first displays a means comparison of days to licensure for the three cohorts, broken down by age. The reference date for renewal customers is taken as the start date of their application; for Driver Safety customers the reference is taken as the date of the first drive test. The licensure date for renewal customers is taken as the issue date of the renewed license. For Driver Safety customers, the licensure date is taken as the lifting of whatever action was placed on the license (if suspended), or the date that no action was taken (in the case of investigations). If customers failed to renew during the two-year follow-up period, their days to licensure was set to missing, and so are effectively excluded from this analysis.

Two patterns are worth noting. First, it is clear that there exist substantial differences across the three cohorts—including differences between the Baseline II and Nearby cohorts. The Baseline II cohort is consistently closer to the Pilot in mean days to licensure, indicating some unmeasured differences between customers using the Nearby offices and customers using the field offices associated with the Baseline II and Pilot cohorts. Secondly, the mean days to licensure for Pilot customers is consistently higher, by about 3.5 days (relative to the Baseline II cohort), regardless of the age group. Any effect of delay in licensure associated with the Pilot appears to be evenly spread across age groups.

Table 45 contains the results of a series of Ordinary Least Squares (OLS) regressions predicting days to licensure, based upon cohort status. Control variables include age, sex, prior 3-year record (suspension/revocation indicator, number of crashes, and number of convictions), whether a customer possessed a limited-term license, whether a customer was required to take a drive test, how many written renewal tests the customer took, and the annualized crash rate of the zip code of residence. Driver Safety customers are excluded from this analysis due to missing data on outcomes on the written renewal test.

Table 44

Days to Licensure by Age and Cohort

Age group	Cohort (*N*)	Mean (*SD*)
<40 y.o.	Baseline II (5,404)	12.31 (48.41)
	Pilot (3,626)	15.82 (61.15)
	Nearby (3,047)	9.67 (47.40)
40-69 y.o.	Baseline II (3,781)	14.59 (57.90)
	Pilot (3,045)	17.93 (66.78)
	Nearby (2,879)	9.32 (44.38)
70+ y.o.	Baseline II (5,444)	8.98 (37.85)
	Pilot (5,253)	12.81 (39.74)
	Nearby (4,506)	7.21 (32.54)

Note. Only includes those customers who renewed their driving privilege at some point during the 2-year follow-up period.

The differences in the models occur in two areas. The first two columns (Models A1 and A2) compare the Pilot cohort with the Baseline II cohort. The second two columns (Models B1 and B2) compare the Pilot cohort with the Nearby cohort. The difference between the “1” and “2” models lies in the inclusion of variables related to field office processing: whether or not a customer took a drive test, and the number of written tests they took before successful renewal of their license.

Table 45

Ordinary Least Squares Regression Predicting Days to Licensure

	Unstd B (*SE*)			
	Model A1: Pilot vs. Baseline II	Model A2: w/o processing controls	Model B1: Pilot vs. Nearby	Model B2: w/o processing controls
Age 40-69 (reference = age<40)	2.20** (0.79)	2.61** (0.83)	1.33 (0.85)	1.87* (0.89)
Age 70+ (reference = age<40)	-4.01** (0.80)	-2.93** (0.85)	-2.83** (0.86)	-1.47† (0.90)
Sex (male)	-3.09** (0.59)	-4.50** (0.63)	-2.34** (0.63)	-3.48** (0.66)
Suspension/Revocation in prior 3 years?	2.38** (0.82)	3.11** (0.87)	4.04** (0.92)	4.68** (0.97)
N of convictions in prior 3 years	0.44 (0.27)	0.92** (0.28)	0.75* (0.31)	1.18** (0.33)
N of crashes in prior 3 years	0.09 (0.53)	0.54 (0.56)	0.79 (0.62)	1.05 (0.65)
Limited-Term license?	27.72** (2.85)	34.31** (2.30)	30.97** (2.72)	38.14** (2.21)
N of written renewal tests	24.37** (0.42)		21.93** (0.47)	
Took drive test?	6.78** (2.10)		7.30** (2.01)	
Crashes per driver per year in zip code of residence	1.21† (0.67)	2.79** (0.71)	1.40* (0.70)	2.40** (0.73)
3-Tier Pilot participant?	0.63 (0.59)	3.29** (0.62)	1.00 (0.68)	4.92** (0.71)
Constant	-28.54** (2.87)	-7.69* (3.03)	-26.37** (2.80)	-7.36* (2.90)

Note. Models A1 and A2: $N = 26{,}411$; Models B1 and B2: $N = 22{,}237$. Both models include all 3-Tier eligible renewal customers.
*: $p \leq 0.05$. **: $p \leq 0.01$. †: $p \leq 0.10$.

There are some differences in the estimates provided by these four models. These indicate the existence of unmeasured differences in the comparison groups (Nearby and Baseline II) that affect the size and significance of some of the control variables. It is not entirely clear why this should occur, though it does underscore the necessity of treating the results with some caution.

Models A1 and B1 indicate that controlling for age, sex, prior record, length of license term, and processing variables (written tests and drive tests), there are no significant differences between the Pilot cohort and the two comparison cohorts. However, it bears noting that substantially more drive tests were administered during the Pilot, as compared to the Baseline II and Nearby cohorts. Removing the variables associated with standard renewal processing—as in Models A2 and B2—increases the size (and makes significant) the Beta coefficient associated with 3-Tier Pilot participation. In other words, participation in the 3-Tier Pilot did lengthen the time to renewal—by somewhere between 3 and 5 days—and this increase in processing time appears to be associated with both the number of drive tests and the number of written tests that were administered. The latter point (regarding written tests) is somewhat surprising, yet nevertheless is confirmed by a basic means comparison provided in Table 46.

Table 46

Mean Number of Written Renewal Test Attempts, by Cohort

Cohort (N)	Mean of written test attempts (*SD*)	F-value of ANOVA (sig.)
Baseline II (14,861)	1.26 (0.66)	Pilot vs. Baseline II 143.37 (<0.01)
Pilot (12,162)	1.37 (0.79)	
Nearby (10,539)	1.17 (0.54)	Pilot vs. Nearby 498.84 (<0.01)

Note. Table includes only 3-Tier eligible renewal customers.

It is not entirely clear why, precisely, participation in the 3-Tier Pilot is associated with an increase in the number of written renewal tests taken by customers. It is possible that this is an artifact of the data-collection methods used to calculate the number of written tests (manual recording by field office staff in the case of the Pilot cohort, electronic recording in the case of the two comparison cohorts). However, if the difference in recorded number of written tests is a methodological artifact, this would not necessarily explain why such an artifact is associated with a delay in licensure. It is possible that this is a product of uncontrolled demographic differences among the cohorts, for instance the fact that the Pilot cohort was restricted to customers taking the written renewal test in English, while the two comparison cohorts may have included some unknown number of customers taking the written renewal test in other languages. It is also possible that field office staff during the Pilot were somewhat more rigorous in deciding what constituted failure on the written renewal test (e.g., in the case of verbally restating a missed question).

Overall Program Effects on Restrictions: Pilot vs. Baseline II and Pilot vs. Nearby

Table 47 displays a cross-tabulation of the proportion of customers in each cohort assigned various kinds of restrictions. The numbers in parentheses in the left-most column denotes the code used by CA DMV to assign such a restriction on the permanent driver record. The restrictions discussed in this section include those most clearly associated with the kinds of driving-relevant limitations for which 3TAS was designed to screen. Not included are the host of restrictions CA DMV uses for the regulation of negligent operators, DUI offenders, and holders of commercial licenses.

Unfortunately, there is no assignment date associated with license restrictions in the CA DMV driver record master file. Neither is there a date associated with the lifting of restrictions. For these reasons, it cannot be determined with any certainty whether restrictions were assigned because of participation in the 3-Tier Pilot. Instead, we must rely upon comparisons among the three cohorts, comparisons which are necessarily imperfect. The data for these analyses are drawn from the Audits files associated with the month of license issue, in the case of renewal customers. For Driver Safety referral customers, these data were drawn from the driver record master file at the time of the extraction of data regarding crashes.

Across the three cohorts, it appears that participation in the pilot is associated with the assignment of restrictions regarding corrective lenses, freeway driving, the use of additional right-side mirrors, and area driving. Pilot participation is also possibly associated with the assignment of restrictions regarding driving at night; however, this association is significant only in comparison to the Baseline II cohort, and not in comparison to the Nearby cohort. For all other kinds of restrictions discussed in Table 47, there is no statistically significant association between the possession of restrictions on the license and participation in the 3-Tier Pilot.

Table 47

Cross-Tabulation of Assigned Restrictions, by Cohort

Restriction (DMV code)		Baseline II (%)	Pilot (%)	Nearby (%)
Must wear corrective lenses (01)	Unrestricted	7,783 (52.2)	5,863 (47.7)	5,330 (50.5)
	Restricted	7,124 (47.8)	6,416 (52.3)	5,221 (49.5)
	χ^2 (Fisher's Test, 2-sided)	Pilot vs. Baseline II 53.63 (<0.01)		Pilot vs. Nearby 17.40 (<0.01)
No freeway driving (02)	Unrestricted	14,859 (99.7)	12,187 (99.3)	10,516 (99.7)
	Restricted	48 (0.3)	92 (0.7)	35 (0.3)
	χ^2 (Fisher's Test, 2-sided)	Pilot vs. Baseline II 24.0 (<0.01)		Pilot vs. Nearby 17.88 (<0.01)
Restricted to driving with additional right side mirror (06)	Unrestricted	14,868 (99.7)	12,224 (99.6)	10,536 (99.9)
	Restricted	39 (0.3)	55 (0.4)	15 (0.1)
	χ^2 (Fisher's Test, 2-sided)	Pilot vs. Baseline II 6.78 (<0.05)		Pilot vs. Nearby 17.36 (<0.01)
Restricted to driving from sunrise to sunset (07)	Unrestricted	14,860 (99.7)	12,215 (99.5)	10,506 (99.6)
	Restricted	47 (0.3)	64 (0.5)	45 (0.4)
	χ^2 (Fisher's Test, 2-sided)	Pilot vs. Baseline II 7.02 (<.01)		Pilot vs. Nearby 1.07 (0.34)
Area restriction (13)	Unrestricted	14,905 (100.0)	12,266 (99.9)	10,550 (100.0)
	Restricted	2 (<0.0)	13 (0.1)	1 (<0.0)
	χ^2 (Fisher's Test, 2-sided)	Pilot vs. Baseline II 10.44 (<0.01)		Pilot vs. Nearby 8.60 (<0.01)
Special Instruction Permit (50)	Unrestricted	14,889 (99.9)	12,272 (99.9)	10,554 (99.9)
	Restricted	18 (0.1)	7 (0.1)	7 (0.1)
	χ^2 (Fisher's Test, 2-sided)	Pilot vs. Baseline II 2.98 (0.11)		Pilot vs. Nearby 0.08 (0.80)
Other P&M-related restrictions[1]	Unrestricted	14,886 (99.9)	12,262 (99.9)	10,539 (99.9)
	Restricted	21 (0.1)	17 (0.1)	12 (0.1)
	χ^2 (Fisher's Test, 2-sided)	Pilot vs. Baseline II 0.00 (1.00)		Pilot vs. Nearby 0.27 (0.71)

Note. Table includes all 3-Tier eligible customers (renewals and referrals).

[1]"Other" P&M-related restrictions include the following: adequate signaling devices (09), automatic transmission (10), adequate support to ensure proper driving position (11), hand-controlled brakes (16), knob attachment on steering wheel (17), full hand controls (22), and bioptic lens (44).

Table 48 provides a simple logistic regression predicting the relative odds of having any restriction related to a physical or vision condition—but excluding a corrective lens requirement—on the basis of cohort status, age, and gender.[12]

Table 48

Logistic Regression Model Predicting Odds of Possessing a Restricted License (Excluding Corrective Lenses)

	Model A: Pilot vs. Baseline II		Model B: Pilot vs. Nearby	
	β (*SE*)	Exp(β) (CI)	β (*SE*)	Exp(β) (CI)
Age 40-69 (reference = age<40)	0.96** (0.25)	2.62 (1.60 - 4.27)	1.28** (0.35)	3.61 (1.83 - 7.12)
Age 70+ (reference = age<40)	2.10** (0.22)	8.14 (5.34 - 12.40)	2.61** (0.31)	13.61 (7.42 - 24.99)
Sex (male)	-0.22† (0.12)	0.80 (0.64 – 1.00)	-0.16 (0.13)	0.85 (0.67 - 1.09)
3-Tier Pilot participant?	0.41** (0.12)	1.51 (1.21 - 1.89)	0.66** (0.13)	1.94 (1.50 - 2.53)

Note. Model A: N = 27,186 (Baseline II = 14,907, Pilot = 12,279); Model B: N = 22,830 (Nearby = 10,551, Pilot = 12,279). Both models include all 3-Tier eligible customers. All confidence intervals (CI) were calculated at a 95% confidence level.
*: $p \leq 0.05$. **: $p \leq 0.01$. †: $p \leq 0.10$.

As with other analyses across cohorts, there appears to remain some unmeasured differences that cannot be controlled for with the current data. Nevertheless, it appears that participation in the Pilot is associated with an increase in the relative odds of having restrictions placed upon the

[12]Corrective lens restrictions were excluded from the analysis for two reasons. First, their assignment is of less substantive interest for the kinds of conditions expected to be identified through the 3-Tier Assessment System. Also, because they are so much more commonly assigned in comparison to other restrictions, the modeling would essentially reduce to a prediction of the likelihood of a corrective lens restriction. Substantively, it is perhaps somewhat surprising that the likelihood of receiving a vision restriction increased under the 3-Tier Pilot, by approximately 3-5% points in comparison to the other two cohorts. Two potential explanations suggest themselves. In the first case, it may be that the additional staff training associated with preparation for the Pilot induced on the part of staff a greater amount of attention to all forms of competency testing, not just those that were strictly associated with the new 3-Tier screening tools. On the other hand, it may simply be the case that there are unmeasured population differences among the cohorts that express themselves in these data in the form of wearing corrective lenses. Determining between these two explanations is somewhat beyond the scope of the analysis. In logistic regressions (not shown, but available from the author) that include controls for age and gender, Pilot cohort status is significantly associated with increased odds (by approximately 15%) of possessing a vision restriction in comparison to the Nearby cohort. A comparison to the Baseline II cohort yields no significant effect, however.

license—especially restrictions other than for corrective lenses—by something like 50-90%, controlling for age and gender. The magnitude of these differences appears relatively small in absolute terms (see Table 47). However, in some cases (e.g., restricted from freeway driving, requiring additional right-side mirrors) the proportion of drivers assigned these restrictions basically doubled in comparison to the two comparison cohorts.

Overall Program Effects on Licensure: Pilot vs. Baseline II and Pilot vs. Nearby

Table 49 first displays a cross-tabulation of the number of customers in each cohort who failed to renew their license during the follow-up period, by age and cohort status. To remove some of the effects of different kinds of processing, the analysis is restricted to renewal customers (regardless of license term), and excludes Driver Safety referrals.

It was noted in the Process Analysis (Camp, 2010b) that surprisingly few suspension and revocation actions were taken specifically as a result of 3-Tier Pilot processing. Due to the rarity of such actions, the analyses in this section include as unlicensed all those who failed to renew their driving privilege at any time during the 2-year follow-up period. This may include individuals who moved out of state, or whose license privilege was suspended or revoked. It also, crucially, may include drivers who have informally surrendered their driving privilege. To the extent that it is possible to identify the full range of individuals whose licensing status was impacted by participation in the 3-Tier Pilot, this is the best measure available given the obtainable data.

Given the substantial differences between the Baseline II and Nearby cohorts, the risk of delicensure appears to be a product of more than simply participation in the 3-Tier Pilot. Regardless of these differences, it appears that, among renewal customers, participation in the 3-Tier Pilot is consistently associated with an increased likelihood of being unlicensed in the 2 years following participation. Furthermore, this effect is spread relatively evenly across the age spectrum. If anything, it appears that 3-Tier customers in the middle-age range (aged 40-69) have a somewhat elevated rate of delicensure compared to customers in the other two cohorts.

Table 49

Cross-Tabulation of Age, Cohort, and Licensure Status During the 2-Year Follow-Up Period (Renewal Customers Only)

Age group	Cohort	Licensed (%)	Unlicensed (%)	χ^2 (2-sided Fisher's Exact Test)
<40 y.o.	Baseline II	5399 (98.6)	74 (1.4)	Pilot vs. Baseline II 6.90 (<0.01)
	Pilot	3609 (97.9)	76 (2.1)	
	Nearby	3041 (99.2)	23 (0.8)	Pilot vs. Nearby 19.92 (<0.01)
40-69 y.o.	Baseline II	3773 (98.1)	69 (1.8)	Pilot vs. Baseline II 10.50 (<0.01)
	Pilot	3001 (96.9)	92 (3.0)	
	Nearby	2875 (99.2)	23 (0.8)	Pilot vs. Nearby 37.80 (<0.01)
70+ y.o.	Baseline II	5426 (97.6)	120 (2.2)	Pilot vs. Baseline II 10.80 (<0.01)
	Pilot	5213 (96.7)	171 (3.2)	
	Nearby	4504 (98.4)	73 (1.6)	Pilot vs. Nearby 25.88 (<0.01)

Table 50 displays the results of two logistic regression models, predicting the odds of delicensure for the entire 2 years following the date of first contact with the field office. Customers who failed to renew their license at all during this time are coded as "1," while customers who eventually renewed their license during the follow-up period are coded as "0."[13] The primary independent variable of interest is cohort status (Pilot vs. Baseline II and Pilot vs. Nearby). Control variables include sex, age, a dummy variable capturing whether an individual had a suspension or revocation during the prior 3 years, and the number of crashes per year in the zip

[13]It is possible that some customers here coded as "1," (unlicensed) eventually renewed their license sometime subsequent to the 2-year follow-up period.

code of residence. To eliminate potential effects associated with different kinds of processing, the analysis is restricted to renewal customers.

Table 50

Logistic Regression Model of the Relative Odds of Delicensure for the 2 Years Following Enrollment (Renewal Customers Only)

	Model A: Pilot vs. Baseline II		Model B: Pilot vs. Nearby	
	β (*SE*)	Exp(β) (CI)	β (*SE*)	Exp(β) (CI)
Age 40-69 (ref. = <40)	0.40** (0.12)	1.48 (1.18 – 1.86)	0.37** (0.14)	1.45 (1.10 – 1.92)
Age 70+ (ref. = <40)	0.64** (0.12)	1.89 (1.50 – 2.38)	0.72** (0.14)	2.04 (1.55 – 2.70)
Sex (male)	-0.28** (0.08)	0.76 (0.64 – 0.89)	-0.43** (0.10)	0.65 (0.54 – 0.79)
S/R in prior 3 years?	0.48** (0.12)	1.62 (1.29 – 2.03)	0.53** (0.14)	1.69 (1.28 – 2.23)
Zip code crash exposure	0.17† (0.10)	1.18 (0.97 – 1.43)	0.17 (0.11)	1.19 (0.95 – 1.48)
3-Tier Pilot participant?	0.42** (0.08)	1.53 (1.30 – 1.80)	0.83** (0.12)	2.30 (1.83 – 2.89)

Note. Model A: N = 27,013 (Baseline II = 14,855, Pilot = 12,158); Model B: N = 22,695 (Nearby = 10,537, Pilot = 12,158). Both models include all 3-Tier eligible renewal customers. All confidence intervals (CI) were calculated at a 95% confidence level.
*: $p \leq 0.05$. **: $p \leq 0.01$. †: $p \leq 0.10$.

Depending on the comparison made, participation in the 3-Tier Pilot is associated with an increase in the relative odds of delicensure among renewal customers by something like 50-130%. By using a population attributable fraction method (Rockhill, Newman, & Weinberg, 1998) we can calculate the number of "excess" persons who failed to renew their license in the follow-up period:

$$\text{Attributable fraction} = \text{pd}(\text{rr}-1/\text{rr})$$

Where "pd" equals the proportion of the total number of unlicensed drivers who were 3-Tier Pilot customers, and "rr" equals the adjusted relative risk, or in other words the relevant exponentiated Beta coefficient found in Models A and B of Table 50. The pd value for Model A is 339/602 (or 0.56) while the pd value for Model B is 339/458 (or 0.74). Plugging these values into the attributable fraction equation above yields an estimate of between 117 renewal drivers (compared to the Baseline II cohort) and 192 renewal drivers (compared to the Nearby cohort) who may have failed to complete their renewal application as a result of participation in the 3-Tier Pilot.

Although it is usually assumed that delicensure is synonymous with cessation of driving, this may not actually be true. There exists evidence of driving among some unlicensed drivers in all three cohorts, as indicated in Table 51 by the accumulation of convictions in the 2 years following enrollment. As with the previous models, Driver Safety referral customers are excluded from this analysis. The comparisons have been stratified by age. Due to small cell-size counts, statistical significance tests were not undertaken. These comparisons are therefore provided for descriptive purposes only.

Table 51

Number of Convictions in 2-Year Follow-Up Period, by Cohort and Licensure Status
(Renewal Customers Only)

Cohort	Licensure status (*N*)	Age (*N*)	*N* of drivers with 1+ convictions? (%)
Baseline II	Licensed (14,598)	<70 y.o. (9,172)	3,314 (36.1)
		70+ y.o. (5,426)	359 (6.6)
	Unlicensed (263)	<70 y.o. (143)	46 (32.2)
		70+ y.o. (120)	7 (5.8)
Pilot	Licensed (11,823)	<70 y.o. (6,610)	2,340 (35.4)
		70+ y.o. (5,213)	309 (5.9)
	Unlicensed (339)	<70 y.o. (168)	55 (32.7)
		70+ y.o. (171)	9 (5.3)
Nearby	Licensed (10,420)	<70 y.o. (5,916)	1,790 (30.3)
		70+ y.o. (4,504)	284 (6.3)
	Unlicensed (119)	<70 y.o. (46)	14 (30.4)
		70+ y.o. (73)	3 (4.1)

Among both younger (<70 y.o.) and older (70+ y.o.) customers, delicensure—failure to complete the license renewal process—is generally associated with a reduction in the accumulation of citations in the subsequent 2 years. However, the accumulation of citations among delicensed drivers does not cease entirely, suggesting that some delicensed customers continue to drive, despite the fact that they lack a valid license. Among younger and middle-age customers (i.e., those younger than 70 y.o.), approximately one third of those who are delicensed show evidence of driving, as indicated by the accumulation of traffic convictions during the follow-up period. Among older customers, a much smaller proportion (4-6%) of the unlicensed accumulate convictions. This may indicate that very, very few unlicensed older persons in the three cohorts continue to drive after failing to renew. Stated more simply, it is possible that delicensure, for most older drivers, indicates cessation of driving. On the other hand, it may also indicate that those older unlicensed customers who drive without proper licensure do so in an otherwise very law-abiding manner (and so do not accumulate traffic convictions). Without information about actual exposure (e.g., miles or even days driven per week), it is impossible to distinguish between these possibilities with the present data.[14]

Question #13: What Are the Cost Savings Associated With Projected Reductions in Crashes?

It was anticipated that the 3-Tier Pilot would produce some demonstrable safety effect in terms of a reduction in crash rates among participants, relative to customers who did not participate in the Pilot. Due to severe methodological constraints discussed at the beginning of this monograph, it is somewhat difficult to test this idea rigorously. Nevertheless, using the best available data, with two comparison groups (the Baseline II and Nearby cohorts), there is no evidence for an *overall* program effect of 3-Tier in reducing the crash propensity of participants.

That said, participation in the 3-Tier Pilot was associated with an increased likelihood of receiving a restricted license. However, the available data do not lend themselves to estimating

[14]Whatever else may be occurring with these unlicensed older drivers, it does not seem likely that they are unaware of their license status. Although the cell-size counts are small, the mean time to conviction for unlicensed older drivers is substantial: for unlicensed drivers aged 70+ in the Pilot cohort, for instance, it is 218 days. This is a good deal longer than the typical term (60 days) of a temporary license that is often given to customers who do not finish their license renewal in the first office visit. The traffic violations for which these drivers were convicted varied somewhat by age. Among younger drivers the violations were most often for DUI-related offenses, or for driving on a suspended or revoked license. Among older drivers, the cell-size counts are too small to be reliable, but the violations across the three cohorts include: driving without a license, driving on a suspended/revoked license, failure to carry proof of insurance, reckless driving, speeding, failure to comply with the basic seat belt law, disobeying signs/signals, and failure to comply with laws regarding proper vehicle equipment (e.g., taillights, headlights).

whether possession of a restricted license is associated with reduced crash risk.[15] Nor is it possible to determine whether the assignment of restrictions extended the valid license status of drivers who otherwise would have received a suspension or revocation. Given the rarity of customers who were assigned a Special Instruction Permit (the most likely scenario for customers who are assigned a restriction in lieu of revocation/suspension), this outcome (restrictions instead of revocation/suspension) likely occurred quite rarely, if at all, among Pilot cohort customers.

It does appear that some small number of 3-Tier Pilot cohort participants (approximately 117-192 customers) surrendered their driving privilege in excess of what would have been expected, based on delicensure rates among the comparison cohorts. Delicensure may be associated with reduced exposure, as indirectly suggested by Table 51 (p. 108), where violation rates among the delicensed are generally smaller than among same-age validly licensed drivers in the same cohort. This reduced exposure may take the form of driving cessation, especially among older DMV customers, though that is not really possible to determine from the available data. To the extent that delicensure—or its logical extension, driving cessation—is associated with reduced exposure, it is also likely to be associated with reduced crash risk. However, because this effect is concentrated among such a small number of drivers, these crash savings are invisible within the overall program effect estimates.

For all of these reasons—the lack of an overall program effect, the impossibility of estimating crash-risk reduction among restricted drivers, the difficulty of estimating crash-risk reduction among the delicensed—it is not really possible to estimate cost savings that may be associated with the 3-Tier Pilot. As a logical corollary, it therefore remains possible that there exist unmeasured benefits associated with the 3-Tier Assessment System, as piloted by CA DMV in 2006-2007.

[15]This is unfortunately the case simply because there is no date associated with the assignment (or lifting) of restrictions in CA DMV's permanent driver record. Hence, especially among Nearby and Baseline II customers, it is impossible to tell when a restriction was assigned (and hence, how the safety of an individual's driving did or did not change, as a result).

DISCUSSION AND CONCLUSIONS

The 3-Tier Pilot involved the first large-scale effort to test, in a real-time public agency setting in California, the 3-Tier Assessment System (3TAS). Given the original twin program goals of (a) preserving and extending the safe driving years of California drivers of all ages, and (b) reducing crashes and violations, what are the results of the 3-Tier Pilot?

In regards to the first goal, it appears that the overall effect of the 3-Tier Pilot likely reduced the driving years of a small, but nonetheless non-trivial, number of drivers. This reduction occurred mostly through the mechanism of informal surrender of the driving privilege (failure to renew) rather than through formal licensing actions taken by CA DMV (e.g., suspensions or revocations). The quasi-experimental nature of the pilot makes for quite strenuous methodological constraints on the estimation of any program effects. That said, as far as can be determined, participation in the 3-Tier Pilot was associated with delicensure among some 117-192 drivers (see pp. 106-109), over what might have been expected under alternate circumstances. This is not to say that all of these unlicensed customers ceased driving entirely. Some proportion of the customers who failed to complete their license renewal process continued to drive, as indicated by the accumulation of traffic violations in the subsequent observation period (see Table 51, p. 108). Nevertheless, inasmuch as the failure to renew the driving privilege constitutes an indicator of surrender of the driving privilege, the 3-Tier Pilot appears to be associated with an increase in the number of delicensed customers.

In regards to the second goal, it must again be emphasized that the quasi-experimental nature of the pilot makes for quite strenuous methodological constraints on the estimation of any program effects. In particular, there remain substantial differences between Pilot customers and the two comparison groups that cannot entirely be statistically eliminated. These differences appear in the form of (a) a marked decline in overall crash risk entirely unassociated with the 3-Tier Pilot (in the case of the Baseline II cohort, see Figure 1, p. 88), and (b) the likely existence of both unmeasured demographic differences and uncontrolled differences in driving environments (in the case of the Nearby cohort, see Table 1, p. 19). Given these differences, any conclusions regarding program effects on crash risk must be made with extreme caution.

That said, two somewhat contradictory points may be made. First, it appears that there exists no evidence for an overall program effect of the Pilot in reducing the likelihood of crashes and convictions among participants. Compared to the Baseline II cohort, the overall reduction in crash risk is of basically the same magnitude as the non-Pilot reduction in crash risk seen statewide and in the Sacramento metropolitan area (see Table 38 and Figure 1, pp. 87-88). Compared to the Nearby cohort, there was no statistically significant reduction in crash risk (see Table 39, p. 89). There was weak, but inconsistent, evidence that 3TAS may be associated with a reduction in the likelihood of being found at fault for a subsequent injury/fatal crash. The proportional hazards of being found at-fault for an injury/fatal crash were lower for the Pilot cohort compared to the other two cohorts, but these differences reached conventional standards of statistical significance in only one case, and then only barely (see Tables 40 and 41, pp. 92 and 93). There was weak, but inconsistent, evidence that 3TAS may be associated with an increase in the likelihood of receiving a traffic conviction. The proportional hazards of receiving a conviction were greater for Pilot customers as compared to customers in the Nearby cohort, but this difference reached conventional standards of statistical significance only barely. The comparison to the Baseline II cohort yielded a result that was not statistically significant. In a certain sense, this lack of an overall program effect is not particularly surprising. For the overwhelming majority of Pilot participants, the 3-Tier Pilot consisted of processing that was not all that different (from the customer's perspective, that is) from what was experienced by non 3-Tier CA DMV customers.[16] This was intended, and in accord with the original design (Hennessy & Janke, 2009): for those customers who were screened as having no identifiable limitations, there was no educational intervention, and no additional processing at Tiers 2 or 3, precisely because there was no apparent need for such intervention. With no educational intervention, and only a relatively minor degree of obvious differences in first-line screening tests, it is unlikely that the 3-Tier Pilot bore any lasting impact on subsequent driving habits for the modal 3-Tier customer.

On the other hand, a second point regarding crash savings may also be made. Among a small minority of 3-Tier customers, the Pilot may in fact have had a lasting impact on subsequent driving habits. Given that delicensure is presumably associated with a reduction in driving—though not, in all cases, with cessation of driving—the unalloyed effect of the Pilot in inducing delicensure may itself have had a (presumably small, certainly unmeasured) effect on reducing the number of crashes. How this effect actually played out among this small group of drivers cannot really be determined with certainty using these data. It may be the case that the Pilot encouraged some extremely functionally limited drivers to cease driving altogether, when

[16]It remains true, of course, that there were differences in processing from the perspective of CA DMV; such differences occurred in the need for additional staff to handle the 5-7 minutes of additional processing time per license renewal transaction (Camp, 2010b). However, as indicated by customer surveys (Camp, 2010a) the difference in processing time was of relatively minor concern to customers.

otherwise they would have crashed. At the same time, it must also be acknowledged that the Pilot may have induced some less-limited drivers to cease driving prematurely. Not all those who failed to renew their license had been designated as extremely functionally limited. In any case, any potentially beneficial effect of delicensure on subsequent crash risk cannot really be estimated at the present time. And while it is entirely possible that the educational intervention may have improved the safety of Pilot customers' driving, for a variety of reasons this cannot really be analyzed using the data at hand.

Because delicensure largely occurred informally—rather than through a formal licensing action such as suspension or revocation actions—those customers who failed to renew their driving privilege in all likelihood did so without any formal consultation and advice regarding the availability of alternative transportation options. It is unknown whether, and to what extent, this group of informally delicensed drivers suffered from the same kinds of locus-of-control problems identified by Hersch in her study of drivers with revoked licenses (Janke & Hersch, 1997; see also Windsor et al., 2007). It is also unknown whether delicensure was, for these customers, associated with the negative social and health consequences normally associated with driving cessation (Bauer, Rottunda, & Adler, 2003; Bryanton, Weeks, & Lees, 2010; Burkhardt, Berger, & McGavock, 1996; Edwards, Lunsman, Perkins, Rebok, & Roth, 2009; Marottoli et al., 2000; Shope, 2003; Siren, Hakamies-Blomqvist, & Lindeman, 2004; Taylor & Tripodes, 2001). On the other hand, it is entirely possible that the drivers who failed to renew under 3-Tier constituted precisely the kind of voluntary, self-determined, driving cessation identified as a goal by some prior authors (Jett, Tappen, & Rosselli, 2005; Johnson, 2008; Malfetti & Winter, 1990; Persson, 1993). It is certainly the case that the informal delicensure associated with the 3-Tier Pilot appears to be similar to the modal form of driving cessation, which in many cases is apparently voluntary and otherwise unassociated with formal licensing actions or even major medical diagnoses (Anstey, Windsor, Luszcz, & Andrews, 2006; Campbell, Bush, & Hale, 1993; Dellinger, Sehgal, Sleet, & Barrett-Connor, 2001).

The 3-Tier Assessment System, as piloted by CA DMV in 2006-2007, constituted only one form that a tiered system of driver competency assessment might take. Other jurisdictions that are developing (or currently using) tiered assessment systems include Florida, whose licensing authority in some cases refers drivers, through their medical review process, to occupational therapists/certified driver rehabilitation specialists (OT/CDRS) prior to a road test (Florida At-Risk Driver Council, 2004). Maryland also incorporates medical review by that state's Medical Advisory Board and, in some cases, referral to outside health professionals (including OTs; see Staplin et al., 2003a). This is somewhat different from what occurs currently in California, where there is no formal procedure for recommending the use of OT/CDRS specialists. There are procedures for requiring formal review of at-risk drivers by a licensed health professional through CA DMV's Driver Safety procedures, when indicated by a referral of a driver to the department for a physical or mental (P&M) condition. However, these reviews are used by the department for assessment, not for rehabilitation. There also exist proposals for revising the procedures used to assess potentially unsafe drivers in jurisdictions in Australia and New Zealand (Fildes et al., 2008). These latter proposals would incorporate tiered systems of assessment, and among other changes include (in some cases) referral for assessment by medical/health care specialists prior to an on-road drive test. The key distinction in all of these cases is the potential incorporation of medical-professional and occupational therapist advice and assessment prior to an on-road test of driving skill by the licensing authority (a version of this idea was part of the initial conceptualization of a tiered system of assessment, as reported in Janke, 1994). Although drivers in Tier 3 of the 3-Tier Pilot were encouraged verbally to seek out such advice and assessment as might be offered by trained driving instructors and occupational therapists, all indications are that such professional advice was rarely (if ever) used (see Table 12, p. 51). It may be that formal referral to professional advice and assessment is needed where extremely functionally limited or somewhat functionally limited customers may yet improve their ability to operate a motor vehicle, and thus avoid premature surrender of the driving privilege. It may also be that some sort of referral to professional advice and assessment is needed even in cases where drivers informally surrender their driving privilege through failure to renew their license.

RECOMMENDATIONS

Given these conclusions, and the substantive findings detailed in the rest of this report, the following recommendations are offered. These recommendations cover two general topical areas: (1) whether to implement the 3-Tier Assessment System statewide in California, in whole or in part, and (2) directions for future research related to the assessment of driving competency in respect to age-related declines in functional ability.

Recommendation #1: Whether to Implement the 3-Tier Assessment System, As a Whole, Statewide

The analyses presented here found no evidence for an overall program effect of 3TAS for reducing subsequent crashes, and weak evidence for an overall program effect in reducing subsequent at-fault injury/fatal crashes. These findings are subject to rather strenuous methodological constraints, and so must be treated with caution. There are three related, but somewhat distinct, issues to be raised as to how these methodological constraints affect interpretation of the results regarding the overall safety effects of the Pilot.

First, it is possible that the 3-Tier Pilot produced beneficial crash savings that remain obscured or unmeasured because of various methodological problems and potential sources of bias discussed elsewhere in this report and in the Process Analysis (Camp, 2010a and 2010b). In other words, if it were possible to pilot 3TAS under different circumstances, with more rigorous experimental controls to eliminate potential confounding sources of bias, it might be possible to determine with more confidence whether it in fact reduces future crash risk among participants.

Second, it should be acknowledged that to a certain extent some of the methodological problems and potential sources of bias noted in this report and in the Process Analysis constituted deviations from the protocols developed for implementation. In other words, it is possible that the analyses presented here found no evidence for an overall program effect because the Pilot was not implemented perfectly. Had it been possible to implement the Pilot differently, with more rigorous quality controls, it is possible that participating customers would have demonstrated reduced subsequent crash risk, as originally hypothesized. That said, it must be emphasized that every reasonable effort was made to enact the 3-Tier Pilot according to those protocols developed by the 3-Tier Task Force, to train staff appropriately in the new Pilot procedures, and to undertake substantial efforts at quality control and ongoing monitoring of implementation during the Pilot, both by Field Office Division and the Research & Development Branch of CA DMV (see especially pp. 7-9, 13-16, and 32-33 of Camp, 2010b for details on the 3-Tier Task Force, the training and organizational coordination enacted prior to implementation, and the quality control measures undertaken while the Pilot was in operation). These efforts were possible in part because of the physical location of the Pilot offices (within reasonable driving distance of DMV state headquarters in Sacramento), and in part because of a certain amount of bureaucratic control made possible by the location of the Pilot field offices within one region of Field Office Division. Thus, while it may be the case that the Pilot was not implemented perfectly, it remains an open question as to whether it would be possible to implement the 3TAS statewide with as high a degree of integrity as was achieved during the Pilot.

Third, it should also be acknowledged that the 3-Tier Pilot was implemented somewhat differently from the model proposed originally by Hennessy and Janke (2009); these differences in implementation were due in part to certain organizational constraints attendant upon altering CA DMV field office procedures, in quite substantial ways, for a temporary period of time. Therefore, the results presented here are not precisely a test of the impact of 3TAS as proposed by Hennessy & Janke (2009); rather, the results presented here more strictly relate to 3TAS as piloted by CA DMV in 2006-2007.

The analyses presented here did show evidence of a reduction in licensure rates associated with 3TAS. These findings are also subject to strenuous methodological constraints, although the evidence regarding delays in licensure, the assignment of license restrictions, and retention of the driving privilege are generally stronger, and more consistent, than the evidence regarding crash rates. However, it remains unclear whether those who failed to renew their driving privilege did so in a manner that preserved their safety (i.e., they otherwise might have crashed), or if they did so in a manner that was potentially premature (i.e., they otherwise would have continued to drive safely).

Given these findings regarding crash rates and retention of the driving privilege, it is not recommended to implement the 3-Tier Assessment System statewide at this point in time.

Recommendation #2: Whether to Implement Separately Any of the Constituent Components of the 3-Tier Assessment System

The 3-Tier Assessment System was piloted by CA DMV in 2006-2007 as an integrated whole, where each screening test contributed information to an overall score, and the overall score indicated whether a given customer required additional assessment of their ability to safely operate a motor vehicle (i.e., a drive test), and/or education regarding how to compensate for some identified functional limitation. Some of these tests were already part of CA DMV's standard repertoire of validated procedures associated with driver license screening: this includes the written knowledge test, the visual acuity standard, and the on-road drive test (Chapman & Masten, 2002; Hagge, 1994; Hagge, 1995; Masten, 1998a; Masten, 1999; Reiner & Hagge, 2006; Romanowicz & Hagge, 1995). Certain other 3TAS screening tests were not then part of CA DMV's repertoire of driver license screening standards, though in all cases the "new" tests had been previously piloted, both here at CA DMV and elsewhere. This included the physical observation protocol, the memory recall test, the Pelli-Robson chart, and the PRT.

It is not possible to determine at this time, on the basis of data collected for the 3-Tier Pilot, whether any of the latter screening tools—the physical observation protocol, memory recall test, Pelli-Robson chart, and PRT—would operate effectively as "stand-alone" mechanisms for identifying drivers in need of additional education and assessment. It is entirely possible that one or more of these new screening tests would constitute a valuable addition to CA DMV's current array of validated driver competency screening tools (the visual acuity test, the written knowledge test, and the on-road drive test). However, such a determination would require a quite different research design than was used for the present pilot project. Such a determination would also, it must be noted, run contrary to the original "ecological" perspective on safe driving that formed the theoretical basis for the development of 3TAS, as explicated in Hennessy and Janke (2009).

Given these methodological and theoretical constraints, it is not recommended to implement at this time, separately and on a stand-alone basis, any of the new screening tests used in the 3-Tier Pilot (the physical observation protocol, memory recall test, Pelli-Robson chart, and PRT).

Recommendation #3: Future Research: Cognitive Screening Tests

Screening for cognitive and perceptual limitations are likely to be a critical component of any system for identifying persons suffering from age-related limitations in abilities that may affect driving. Two screening tests—the memory recall test and the PRT—were included in the Pilot, with the intention that they should directly capture information indicating potential limitations in cognition and perception.

As noted in the Process Analysis (Camp, 2010a and 2010b), staff reported substantial variation in the implementation of the memory recall test. Furthermore, some staff openly questioned the validity of the memory recall test as a screening mechanism for identifying limitations in cognitive abilities that may affect driving. A similar criticism of this test was also raised by some members of a convened panel of experts at the 2011 Transportation Research Board meetings, where it was suggested that the memory recall test may not be the most effective means of flagging potential cognitive limitations that may affect driving.[17] Unfortunately, these questions regarding the empirical validity of the memory recall test cannot be addressed directly using data collected during the 3-Tier Pilot, especially given the reported variation in implementation. However, it is the case that 3-Tier customers failed this test quite rarely, and—keeping in mind the methodological problems associated with the reported variation in implementation—outcomes on this test were not statistically associated with prior crash risk.

It was also noted in the Process Analysis that there existed widespread misunderstanding on the part of staff as to the purpose, and traffic-safety relevance, of the PRT. However, because this was a computer-based test (rather than administered in a face-to-face fashion by a trained staff person), there does not appear to have been the kind of variation in implementation as was reported for the memory recall test. Among Pilot cohort customers aged 70+, there exists evidence that outcomes on this test are associated with prior crash risk (see Table 32, p. 75). This is consistent with the results of prior studies, where the PRT has been shown to be reliably predictive of crash risk, especially among drivers suffering from cognitive and perceptual impairments.

Given these findings, it is recommended that additional research be conducted to identify, if possible, a Tier 1 screening test for the kinds of cognitive and perceptual deficits associated with dementia-type disorders, mild cognitive impairment, and early-stage Alzheimer's disease.

[17]Specifically, the panelists at TRB (and some audience members) suggest (a) that the memory recall test may identify deficits in long-term memory, rather than declines to executive function or other sorts of cognitive problems more directly related to safe driving, (b) that the fail rate on this test was substantially lower than would be expected, given the incidence of dementia in the general population of older adults, and (c) that the test does not appear to have the kinds of psychometric properties (e.g., specificity and sensitivity) that are desirable in this type of front-line screening tool. The author is indebted to Jack Joyce, of TransAnalytics, Elin Schold Davis, of the American Occupational Therapy Association, and Dr. Nina Silverstein, of the University of Massachusetts – Boston, for these valuable insights.

Although the PRT appears to possess utility as a screening tool for (among other things) drivers at risk of crashing due to cognitive and perceptual deficits, the 3-Tier Pilot incorporated this test at Tier 2, rather than at Tier 1. Administering the PRT to all renewal customers would involve substantial disruptions to current office procedures; therefore it would not necessarily be appropriate as a Tier 1 screening tool.

Such a Tier 1 screening test for functional limitations in cognition or perception must, if at all possible, possess face-validity properties that connect the testing mechanism (and its outcomes) to safe driving. In other words, staff must be able, quickly and straightforwardly, to explain to customers the connection between the test and safe driving. Furthermore, such a test must, if possible, be capable of being administered by front-line staff working in an agency setting—keeping in mind that such staff persons will not be making diagnoses of any kind.[18] There already exist a number of tests that predict both cognitive deficits and crash risk (e.g., Trails B, the Clock-Drawing Test); however, many of these tests take several minutes to apply, and the results usually require interpretation by someone with some kind of clinical training (rather than front-line staff at a public agency). Ideally, this test would also be readily incorporated into whatever office procedures are at that time in place for processing driver license renewal customers.

Recommendation #4: Future Research: Educational Intervention

The 3-Tier Assessment System incorporated certain educational components that were intended to serve several goals. Among these goals were preparation for an on-road drive test, and encouraging drivers to adopt changes to their driving habits in the face of identified limitations to certain functional areas (viz., perceptual speed and contrast sensitivity). It was not possible, on the basis of data collected during the Pilot, to determine whether these educational materials had any effect on pass/fail rates on the drive test. Nor, for somewhat different reasons, was it possible to determine whether these materials had any effect on subsequent crash risk. It is entirely possible that these educational materials improved the safe driving of 3-Tier customers; however, any positive effects cannot be estimated using the data at hand.[19]

[18]Alternatively, the test might be delivered using some kind of computer-based format; however, this would likely depend upon other changes to current CA DMV field office processing—e.g., automation of the written knowledge test—that are outside the scope of the present work.

[19]Anecdotal reports from staff (See Camp 2010a, esp. pp. 72-76) suggest that the educational materials were useful especially as customer service tools: for answering questions and managing test-related anxiety. Such positive

It is desirable that any system of screening for driving-relevant functional limitations, as well as any system of assessment of drivers' abilities to safely compensate for identified functional limitations, incorporate educational materials. It is therefore recommended that additional research be conducted to identify appropriate materials that encourage safe driving, and to identify the appropriate mechanism(s) for delivering such materials to CA DMV customers. Such research might fruitfully build upon research already published by CA DMV (Kelsey & Janke, 2005).

Recommendation #5: Future Research: Use of Occupational Therapists/Certified Driving Rehabilitation Specialists

There was substantial missing data regarding whether customers who were required to take drive tests undertook any formal drive-test preparation, and if so then with what type of instructor: a family member or friend, professional instructor, or occupational therapist/certified driving rehabilitation specialist (OT/CDRS). There remains, therefore, much that is unknown about the effect of different kinds of drive-test preparation on (a) pass/fail rates on an on-road test, and (b) subsequent safe driving.

It is therefore recommended that additional research be conducted regarding the relationship between different kinds of drive-test preparation—of which referral to an OT/CDRS might be one form—and drive-test outcomes, subsequent driving habits, and subsequent safety outcomes.

Recommendation #6: Future Research: Informal Delicensure

There exists suggestive evidence that when delicensure occurred in the 3-Tier Pilot, it was informal, rather than formal. Instead of a suspension or revocation order (of which there were very few, as reported in the Process Analysis, Camp 2010b), delicensure among Pilot cohort customers appears to have more commonly taken the form of letting the renewal application lapse. Furthermore, there appears to have been a statistically significant increase in the odds of delicensure associated with participation in the Pilot, compared to the two comparison cohorts.

It is not precisely clear what happened with these informally delicensed customers once they let their applications lapse. Certainly some of them appear still to be driving in California, as

effects might be kept in mind in the design of educational materials, as well as in the analysis of any potential effects associated with the distribution of such materials.

indicated by the accumulation of later traffic convictions. Others may have ceased driving altogether, and still others may have moved to another jurisdiction.[20] However, it is unknown in what proportion these forms of informal delicensure (unlicensed driving, driving cessation, and moving outside the state) occur. It is unknown to what extent different kinds of informal delicensure are associated with known demographic variables of interest, such as age and gender. It is unknown if informal delicensure is associated with organizational variables of interest (e.g., the office at which a renewal is conducted). It is unknown if informal delicensure is associated with changes to crash risk, for those who still drive. It is also unknown if informal delicensure is associated with other kinds of social and health outcomes, especially in the case of those who cease driving.

For all of these reasons, it is recommended that additional research be conducted regarding informal delicensure, driving habits and exposure, individual demographic variables (e.g., age, gender, health status), organizational variables (office processing), and safety outcomes such as crash risk.

[20]While different jurisdictions generally notify California when a customer moves and applies for another license, official notification would not cover cases where someone moves, but does not apply for a driver license in the new jurisdiction.

REFERENCES

Adler, G., & Kuskowski, M. (2003). Driving habits and cessation in older men with dementia. *Alzheimer Disease and Associated Disorders, 17*, 67-71.

Adler, G., & Rottunda, S. (2006). Older adults' perspectives on driving cessation. *Journal of Aging Studies, 20*, 227-235.

Adler, G., Rottunda, S., & Kuskowski, M. (1999). The impact of dementia on driving: Perceptions and changing habits. *The Clinical Gerontologist, 20*, 23-34.

Adler, G., & Silverstein, M. (2008). At-risk drivers with Alzheimer's Disease: Recognition, response, and referral. *Traffic Injury Prevention, 9*, 299-303.

Alvarez, F. J., & Fierro, I. (2008). Older drivers, medical condition, medical impairment and crash risk. *Accident Analysis and Prevention, 40*, 55-60.

Anstey, K. J., Windsor, T. D., Luszcz, M. A., & Andrews, G. R. (2006). Predicting driving cessation over 5 years in older adults: Psychological well-being and cognitive competence are stronger predictors than physical health. *Journal of the American Geriatrics Society, 54,* 121-126.

Anstey, K. J., & Wood, J. (2011). Chronological age and age-related cognitive deficits are associated with an increase in multiple types of driving errors in late life. *Neuropsychology (May)*.

Anstey, K. J., Wood, J., Lord, S., & Walker, J. G. (2005). Cognitive, sensory and physical factors enabling driving safety in older adults. *Clinical Psychology Review, 25*, 45-65.

Attebo, K., Mitchell, P., & Smith, W. (1996). Visual acuity and the causes of visual loss in Australia: The Blue Mountains eye study. *Ophthalmology, 103*, 357-364.

Baker, T., Falb, T., Voas R., & Lacey, J. (2003). Older women drivers: Fatal crashes in good conditions. *Journal of Safety Research, 34*, 399-405.

Baldock, M. R. J., Mathias, J. L., McLean, A. J., & Berndt, A. (2006). Self-regulation of driving and its relationship to driving ability among older adults. *Accident Analysis and Prevention, 38*, 1038-1045.

Ball, K., Owsley, C., Sloane, M. E., Roenker, D. L., & Bruni, J. R. (1993). Visual attention problems as a predictor of vehicle crashes in older drivers. *Investigative Ophthalmology and Visual Science, 34,* 3110-3124.

Ball, K., Owsley, C., Stalvey, B., Roenker, D., Sloane, M., & Graves, M. (1998). Driving avoidance and functional impairment in older drivers. *Accident Analysis and Prevention, 30*, 313-322.

Ball, K. K., Roenker, D., Wadley, V. G., Edwards, J. D., Roth, D. L., McGwin, G., Jr., Raleigh, R., Joyce, J. J., Cissell, G. M., & Dube, T. (2006). Can high-risk older drivers be identified through performance-based measures in a Department of Motor Vehicles setting? *Journal of the American Geriatrics Society, 54*, 77-84.

Banister, B., & Bowling, A. (2004). Quality of life for the elderly: The transport dimension. *Transport Policy, 11*, 105-115.

Barancik, J. I., Chatterjee, B. F., Greene-Cradden, Y. C., Michenzi, E. M., Kramer, C. F., Thode, H. C., Jr., & Fife, D. (1986). Motor vehicle trauma in northeastern Ohio, I: Incidence and outcome by age, sex, and road-use category. *American Journal of Epidemiology, 123*, 846-861.

Barr, R. A. (1991). Recent changes in driving among older adults. *Human Factors, 33*, 597-600.

Bauer, M. J., Rottunda, S., & Adler, G. (2003). Older women and driving cessation. *Qualitative Social Work, 2*, 309-325.

Bédard, M., Guyatt, G. H., Stones, M. J., & Hirdes, J. P. (2002). The independent contributions of driver, crash, and vehicle characteristics to driver fatalities. *Accident Analysis and Prevention, 34*, 717-727.

Bédard, M., Isherwood, I., Moore, E., Gibbons, C., & Lindstrom, W. (2004). Evaluation of a re-training program for older drivers. *Canadian Journal of Public Health/Revue Canadienne de Santé Publique, 95*, 295-298.

Bédard, M., Weaver, B., Dārzinš P., & Porter, M. (2008). Predicting driving performance in older adults: We are not there yet! *Traffic Injury Prevention, 9*, 336-341.

Benekohal, R. F., Michaels, R. M., Shim, E., & Resende, P. T. V. (1994). Effects of aging on older drivers' travel characteristics. *Transportation Research Record, 1438*, 91-98.

Blanchard, R. A., & Myers, A. M. (2010). Examination of driving comfort and self-regulatory practices in older adults using in-vehicle devices to assess natural driving patterns. *Accident Analysis and Prevention, 42,* 1213-1219.

Blanchard, R. A., Myers, A. M., & Porter, M. (2010). Correspondence between self-reported and objective measures of driving exposure and patterns in older drivers. *Accident Analysis and Prevention, 42*, 523-529.

Braitman, K. A., Chaudhary, N. K., & McCartt, A. T. (2010). Restricted licensing among older drivers in Iowa. *Journal of Safety Research, 41,* 481-186.

Braver, E., & Trempel, R. (2004). Are older drivers actually at higher risk of involvement in collisions resulting in deaths or non-fatal injuries among their passengers and other road users? *Injury Prevention, 10*, 27-32.

Brook, M. M., Qustad, K. A., Patterson, D. R., & Valois, T. A. (1992). Driving evaluation after traumatic brain injury. *The American Journal of Physical Medicine and Rehabilitation, 71*, 177-182.

Bryanton, O., Weeks, L. E., & Lees, J. M. (2010). Supporting older women in the transition to driving cessation. *Activities, Adaptation & Aging, 34,* 181-195.

Burkhardt, J., Berger A. M., & McGavock, A. T. (1996). The mobility consequences of the reduction or cessation of driving by older women. *Women's Travel Issues: Proceedings from the Second Conference* (pp. 440-454). Federal Highway Administration Publication FHWA-PL-97-024.

California Department of Finance. (2007). *Population projections for California and its counties 2000-2050, by age, gender and race/ethnicity*. Sacramento, CA: Department of Finance.

Camp, B. J. (2010a). *California's three-tier pilot process appendix* (Report #229). Sacramento, CA: California Department of Motor Vehicles.

Camp, B. J. (2010b). *California's three-tier driving centered assessment system: Process analysis* (Report #323). Sacramento, CA: California Department of Motor Vehicles.

Campbell, M. C., Bush, T. L., & Hale, W. E. (1993). Medical conditions associated with driving cessation in community-dwelling, ambulatory elders. *Journal of Gerontology: Social Sciences, 48*, S230-S234.

Carr, D., Schmader, K., Bergman, C., Simon, T. C., Jackson, T. W., Haviland, S., & O'Brien, J. (1991). A multi-disciplinary approach in the evaluation of demented drivers referred to geriatric assessment centres. *Journal of the American Geriatric Society, 39*, 1132-1136.

Chapman, E. A., & Masten, S. V. (2002). *Development and evaluation of revised Class C driver license written knowledge tests* (Report No. 196). Sacramento, CA: California Department of Motor Vehicles.

Charlton, J., Koppel, S., O'Hare, M., Andrea, D., Smith, G., Khodr, B., Langford, J., Odell, M., & Fildes, B. (2004). *Influence of chronic illness on crash involvement of motor vehicle drivers* (Report No. 213). Clayton, Victoria, Australia: Monash University Accident Research Centre.

Charlton, J. L., Oxley, J., Fildes, B., Oxley, P., & Newstead, S. (2003). Self-regulatory behaviors of older drivers. *Annual Proceedings of the Association for the Advancement of Automotive Medicine, 47*, 181-194.

Charlton, J. L., Oxley, J., Fildes, B., Oxley, P., Newstead, S., Koppel, S., & O'Hare, M. (2006). Characteristics of older drivers who adopt self-regulatory driving behaviours. *Transportation Research Part F, 9*, 517-521.

Cheung, I., & McCartt, A. T. (2011). Declines in fatal crashes of older drivers: Changes in crash risk and survivability. *Accident Analysis and Prevention, 43,* 666-674.

Classen, S., McCarthy, D. P., Shechtman, O., Awadzi, K. D., Lanford D. N., Okun, M. S., Rodriguez, R. L., Romrell, J., Bridges, S., Kluger, B., & Fernandez, H. H. (2009). Useful Field of View as a reliable screening measure of driving performance in people with Parkinson's Disease: Results of a Pilot Study. *Traffic Injury Prevention, 10,* 593-598.

Clay, O. J., Wadley, V. G., Edwards, J. D., Roth, D. L., Roenker, D. L., & Ball, K. K. (2005). Cumulative meta-analysis of the relationship between Useful Field of View and driving performance in older adults: Current and future implications. *Optometry and Vision Science 82*, 1-8.

Dellinger, A. M., Kresnow, M., White, D., & Sehgal, M. (2004). Risk to self versus risk to others: How do older drivers compare to others on the road? *American Journal of Preventive Medicine, 26*, 217-221.

Dellinger, A. M., Sehgal, M., Sleet, D. A., & Barrett-Connor, E. (2001). Driving cessation: What older former drivers tell us. *Journal of the American Geriatrics Society,49*, 431-435.

Di Stefano, M., & McDonald, W. (2003). Assessment of older drivers: relationships among on-road errors, medical conditions and test outcome. *Journal of Safety Research, 34*, 415-429.

Diller, E. M., Cook, L. J., Leonard, D., Reading, J. C., Deon, J. M., & Vernon, D. D. (2001). *Further analysis of drivers licensed with medical conditions in Utah* (Technical Report No. DOT HS 809 211). Springfield, VA: US Department of Transportation.

Dobbs, B. M. (2005). *Medical conditions and driving: A review of the scientific literature (1960-2000), technical report*. Washington, DC: National Highway Traffic Safety Administration and the Association for the Advancement of Automotive Medicine Project.

Dobbs, B. M., & Carr, D. (2005). Screening and assessment of medically at-risk drivers. *Public Policy and Aging Report, 15*, 6-12.

Dobbs, B. M., Carr, D. B., & Morris, J. C. (2002). Evaluation and management of the driver with dementia. *Neurologist, 8*, 61-70.

Donorfio, L. K. M., D'Ambrosio, L. A., Coughlin, J. F., & Mohyde, M. (2009). To drive or not to drive that *isn't* the question – the meaning of self-regulation among older drivers. *Journal of Safety Research, 40*, 221-226.

Duchek, J., Hunt, L., Ball, K., Buckles, V., & Morris, J. (1998). Attention and driving performance in Alzheimer's Disease. *Journals of Gerontology Series B: Psychological Sciences and Social Sciences, 53,* S130-S141.

Dulisse, B. (1997). Driver age and traffic citations resulting from motor vehicle collisions. *Accident Analysis and Prevention, 29*, 779-783.

Eberhard, J. (2008). Older drivers' "high per-mile crash involvement": The implications for licensing authorities. *Traffic Injury Prevention, 9*, 284-290.

Eberhard, J., & Mitchell, C. G. B. (2009). Recent changes in driver licensing rates, fatality rates, and mobility options for older men and women in the United States and Great Britain. *Topics in Geriatric Rehabilitation, 25*, 88-98.

Eby, D., Molnar, L., Shope, J., Vivoda, J., & Fordyce, T. (2003). Improving older driver knowledge and self-awareness through self-assessment: The Driving Decisions Workbook. *Journal of Safety Research, 34*, 371-381.

Edwards, J. D., Bart, E., O'Connor, M. L., & Cissell, G. (2009), Ten years down the road: Predictors of driving cessation. *The Gerontologist, 50*¸393-399.

Edwards, J. D., Lunsman, M., Perkins, M., Rebok, G. W., & Roth, D. O. (2009). Driving cessation and health trajectories in older adults. *Journal of Gerontology A: Biological Sciences and Medical Sciences, 64,* 1290-1295.

Eisenhandler, S. A. (1993). The asphalt identikit: Old age and the driver's license. *International Journal of Aging and Human Development, 30*, 1-14.

Evans, L. (1988a). Older driver involvement in fatal and severe traffic crashes. *Journal of Gerontology, 43*, 186-193.

Evans, L. (1988b). Risk of fatality from physical trauma versus sex and age. *Journal of Trauma, 28*, 368-378.

Evans, L. (1991). Age and fatality risk from similar severity impacts. *Journal of Traffic Medicine, 29*, 10-19.

Evans, L. (2000). Risks older drivers face themselves and threats they pose to other road users. *International Journal of Epidemiology, 29*, 315-322.

Fildes, B., Pronk, N., Langford, J., Hull, M., Frith, B., & Anderson, R. (2000). *Model license re-assessment procedure for older and disabled drivers* (Publication No. AP-R176/00). Sydney, NSW, Australia: Austroads, Inc.

Fildes, B., Charlton, J., Pronk, N., Langford, J., Oxley, J., & Koppel, S. (2008). An Australasian model license reassessment procedure for identifying potentially unsafe drivers. *Traffic Injury Prevention, 9*, 350-359.

Fisk, G.D., Novack, T., Mennemeier, M., & Roenker, D. (2002). Useful Field of View after traumatic brain injury. *Journal of Head Trauma Rehabilitation, 17,* 16-25.

Foley, D. J., Masaki, K. H., Ross, G. W., & White, L. R. (2000). Driving cessation in older men with incident dementia. *Journal of the American Geriatrics Society, 48*, 928-930.

Fonda, S., Wallace, R., & Herzog, A. (2001). Changes in driving patterns and worsening depressive symptoms among older adults. *Journals of Gerontology Series B: Psychological Sciences & Social Sciences, 56*, S343-S351.

Fox, G. K., Bowden, S. C., & Smith, D. S. (1998). On-road assessment of driving competence after brain impairment: Review of current practice and recommendations for a standardized examination. *Archives of Physical Medicine and Rehabilitation, 79*, 1288-1296.

Florida At-Risk Driver Council. (2004). *The Effects of Aging on Driving Ability*. Tallahassee, FL: Department of Highway Safety & Motor Vehicles.

Freeman, E. E., Gange, S. J., Muñoz, B., & West, S. K. (2006). Driving status and risk of entry into long-term care in older adults. *American Journal of Public Health, 96*, 1254-1259.

Freund, B., Colgrove, L. A., Burke, B. L., & McLeod, R. (2005). Self-rated driving performance among elderly drivers referred for driving evaluation. *Accident Analysis and Prevention, 37*, 613-618.

Friedland, R. P., Koss, E., Kumar, A., Gaine, S., Metzler, D., Haxby, J. V., & Moore, A. (1988). Motor vehicle crashes in dementia of the Alzheimer's type. *Annals of Neurology, 24*, 782-786.

Gallo, J., Rebok, G., & Lesiker, S. (1999). The driving habits of adults aged 60 years and older. *Journal of the American Geriatrics Society, 47*, 335-341.

Galski, T., Bruni, R. L., & Ehle, H. T. (1992). Driving after cerebral damage: A model with implications for evaluation. *American Journal of Occupational Therapy, 46*, 324-332.

Gebers, M. A. (forthcoming). *Predicting traffic crash involvement using individual driver habits, driving record, and territorial risk indices*. Sacramento, CA: California Department of Motor Vehicles.

Gebers, M. A., & Peck, R. C. (1994). *An inventory of California driver accident risk factors* (Report No. 144). Sacramento, CA: Department of Motor Vehicles.

George, C. F., & Smiley, A. (1999). Sleep apnea and automobile crashes. *Sleep, 22*, 790-795.

Gilley, D. W., Wilson, R. S., Bennett, D. A., Stebbins, G. T., Bernard, B. A., Whalen, M. E., & Fox, J. H. (1991). Cessation of driving and unsafe motor vehicle operation by dementia patients. *Archives of Internal Medicine, 151*, 941-946.

Grabowski, D., Campbell, C., & Morrisey, M. (2004). Elderly licensure laws and motor vehicle fatalities. *Journal of the American Medical Association, 291*, 2840-2846.

Hagge, R. A. (1994). *The California driver performance evaluation project: An evaluation of a new driver licensing road test* (Report No. 150). Sacramento, CA: California Department of Motor Vehicles.

Hagge, R. A. (1995). *Evaluation of California's special drive test program* (Report No. 160). Sacramento, CA: California Department of Motor Vehicles.

Hakamies-Blomqvist, L. (1993). Fatal accidents of older drivers. *Accident Analysis and Prevention, 25*, 19-27.

Hakamies-Blomqvist, L., Johansson, K., & Lundberg, C. (1996). Medical screening of older drivers as a traffic safety measure – A comparative Finnish-Swedish evaluation study. *Journal of the American Geriatrics Society, 44*, 650-653.

Hakamies-Blomqvist, L., Raitanen, T., & O'Neil, D. (2002). Driver ageing does not cause higher accidents rates per km. *Transportation Research Part F, 5*, 271-274.

Hakamies-Blomqvist, L., & Wahlstrom, B. (1998). Why do older drivers give up driving? *Accident Analysis and Prevention, 30*, 305-312.

Hanrahan, R. B., Layde, P. M., Zhu, S., Guse, C. E., & Hargarten, S. W. (2009). The association of driver age with traffic injury severity in Wisconsin. *Traffic Injury Prevention, 10*, 361-367.

Hanson, T. R., & Hildebrand, E. D. (2011a). Are rural older drivers subject to low-mileage bias? *Accident Analysis and Prevention, 43*, 1872-1877.

Hanson, T. R., & Hildebrand, E. D. (2011b). Are age-based licensing restrictions a meaningful way to enhance rural older driver safety? The need for exposure considerations in policy development. *Traffic Injury Prevention, 12,* 24-30.

Harper, J., & Schatz, S. (1998). *The premature reduction or cessation of driving*. Chapel Hill, NC: Highway Safety Research Center, University of North Carolina.

Harris, A. (2000). *Transport and mobility in rural Victoria* (Report No. PP 00/02). Melbourne, Australia: Royal Automobile Club of Victoria.

Harrison, A., & Ragland, D. (2003). Consequences of driving reduction or cessation for older adults. *Transportation Research Record, 1843*, 96-104.

Hennessy, D. F. (1995). *Vision testing of renewal applicants: Crashes predicted when compensation for impairment is inadequate* (Report No. 152). Sacramento, CA: California Department of Motor Vehicles.

Hennessy, D. F., & Janke, M. K. (2005). *Clearing a road to driving fitness by better assessing driving wellness: California's three-tier driving-centered assessment system – Summary report* (Report No. 215). Sacramento, CA: California Department of Motor Vehicles.

Hennessy, D. F., & Janke, M. K. (2009). *Clearing a road to being driving fit by better assessing driving wellness: Development of California's prospective three-tier driving centered assessment system – Technical Report* (Report No. 216). Sacramento, CA: California Department of Motor Vehicles.

Huebner, K., Porter, M. M., & Marshall, S. C. (2006). Validation of an electronic device for measuring driving exposure. *Traffic Injury Prevention, 7,* 76-80.

Holland, C. A., & Rabbitt, P. M. A. (1990). People's awareness of their age-related sensory and cognitive deficits and the implications for road safety. *Applied Cognitive Psychology, 6*, 217-231.

Hu, P., Jones, D., Reuscher, T., Schmoyer, R., & Truett, L. (2000). *Projecting fatalities in crashes involving older drivers, 2000-2025* (No. ORNL-6963). Oak Ridge, TN: Oak Ridge National Laboratory.

Hu, P., Trumble, D., Foley, D., Eberhard, J., & Wallace, R. (1998). Crash risks of older drivers: A panel data analysis. *Accident Analysis and Prevention, 30*, 569-581.

Hunt, L. A., Murphy, C., Carr, D. V., Duchek, J., Buckles, V., & Morris, J. (1997). Reliability of the Washington University road test: A performance-based assessment for drivers with dementia of the Alzheimer type. *Archives of Neurology, 54*, 707-712.

Janke, M. K. (1991). Accidents, mileage, and the exaggeration of risk. *Accident Analysis and Prevention, 23*, 183-188.

Janke, M. K. (1994). *Age-related disabilities that may impair driving and their assessment*: *Literature review* (Report No. 156). Sacramento, CA: California Department of Motor Vehicles.

Janke, M. K. (2001). Assessing older drivers: Two studies. *Journal of Safety Research, 32*, 43-74.

Janke, M., & Eberhard, J. (1998). Assessing medically impaired older drivers in a licensing agency setting. *Accident Analysis and Prevention, 30*, 347-361.

Janke, M. K., & Hersch, S. W. (1997). *Assessing the older driver: Pilot studies* (Report No. 172). Sacramento, CA: California Department of Motor Vehicles.

Janke, M. K., Masten, S. V., McKenzie, D. M., Gebers, M. A., & Kelsey, S. L. (2003). *Teen and Senior Drivers* (Report No. 194). Sacramento, CA: California Department of Motor Vehicles.

Jett, K., Tappen, R. M., & Rosselli, M. (2005). Imposed versus involved: Different strategies to effect driving cessation in cognitively impaired older adults. *Geriatric Nursing, 26,* 111-116.

Johnson, J. E. (1998). Older rural adults and the decision to stop driving: The influence of family and friends. *Journal of Community Health Nursing, 15*, 205-216.

Johnson, J. E. (2002). Why rural elders drive against advice. *Journal of Community Health Nursing, 19*, 237-244.

Johnson, J. E. (2008). Informal social support networks and the maintenance of voluntary driving cessation by older rural women. *Journal of Community Health Nursing, 25,* 65-72.

Justiss, M. D., Mann, W. C., Stab, W. B., & Velozo, C. (2006). Development of a behind-the-wheel driving performance assessment for older adults. *Topics in Geriatric Rehabilitation, 22*, 121-128.

Kantor, B., Mauger, L., Richardson, V. E., & Tschantz-Unroe, K. (2004). An analysis of an older driver evaluation program. *Journal of the American Geriatrics Society, 52*, 1326-1330.

Kazniak, A. W., Keyl, P. M., & Albert, M. S. (1991). Dementia and the older driver. *Human Factors, 33*, 527-537.

Kelsey, S. L., & Janke, M. (2005). *Pilot educational outreach to high-risk elderly drivers* (Report No. 213). Sacramento, CA: California Department of Motor Vehicles.

Kington, R., Reuben, D., Rogowski, J., & Lillard, L. (1994). Sociodemographic and health factors in driving patterns after 50 years of age. *American Journal of Public Health, 84*, 1327-1329.

Klavora, P., & Heslegrave, R. (2002). Senior drivers: An overview of problems and intervention strategies. *Journal of Aging and Physical Activity 10*, 332-335.

Korner-Bitensky, N., Kua, A., von Zweck, C., & Van Benthem, K. (2009). Older driver retraining: An updated systematic review of evidence of effectiveness. *Journal of Safety Research, 40*, 105-111.

Kostyniuk, L., & Molnar, L. J. (2005). Driving self-restriction among older adults: Health, age, and sex effects. *The Gerontologist*, *45*, 143.

Kostyniuk, L., Shope, J., & Molnar, L. (2000). *Reduction and cessation of driving among older drivers in Michigan: Final report* (Report No. UMTRI-2000-06). Ann Arbor, MI: University of Michigan Transportation Research Institute.

Kua, A., Korner-Bitensky, N., Desrosiers, J., Man-Son-Hing, M., & Marshall, S. (2007). Older driver retraining: A systematic review of evidence of effectiveness. *Journal of Safety Research, 38*, 81-90.

Lange, J. E., & McKnight, A. J. (1996). Age-based road test policy evaluation. *Transportation Research Record, 1550*, 81-87.

Langford, J. (2008). Usefulness of off-road screening tests to licensing authorities when assessing older driver fitness to drive. *Traffic Injury Prevention, 9*, 328-335.

Langford, J., & Koppel, S. (2011). License restrictions as an under-used strategy in managing older driver safety. *Accident Analysis and Prevention, 43,* 487-493.

Langford, J., Bohensky, M., Koppel, S., & Newstead, S. (2008a). Do age-based mandatory assessments reduce older drivers' risk to other road users? *Accident Analysis and Prevention, 40,* 1913-1918.

Langford, J., Bohensky, M., Koppel, S., & Newstead, S. (2008b). Do older drivers pose a risk to other road users? *Traffic Injury Prevention, 9*, 181-189.

Langford, J., Fitzharris, Koppel, S., & Newstead S. (2004). Effectiveness of mandatory license testing for older drivers in reducing crash risk among urban older Australian drivers. *Traffic Injury Prevention, 5*, 326-335.

Langford, J., Fitzharris, M., Newstead, S., & Koppel, S. (2004). Some consequences of different older driver licensing procedures in Australia. *Accident Analysis and Prevention, 36*, 993-1001.

Langford, J., Methorst, R., & Hakamies-Blomqvist, L. (2006). Older drivers do not have a high crash risk – A replication of low mileage bias. *Accident Analysis and Prevention, 38*, 574-578.

Levy, D., Vernick, J., & Howard K. (1995). Relationship between driver's license renewal policies and fatal crashes involving drivers 70 years or older. *Journal of the American Medical Association, 274*, 1026-1030.

Li, G., Braver, E. R., & Chen, L. H. (2003). Fragility versus excessive crash involvement as determinants of high death rates per vehicle-mile of travel among older drivers. *Accident Analysis and Prevention, 35*, 227-235.

Lococo, K. H., & Staplin L. (2005). *Strategies for medical advisory boards and licensing review* (DOT HS 809 874). Washington, DC: Office of Research and Traffic Records, National Highway Traffic Safety Administration.

Lyman, S., Ferguson, S. A., Braver, E. R., & Williams A. F. (2002). Older driver involvement in police reported crashes and fatal crashes: Trends and projections. *Injury Prevention, 8*, 116-120.

Lyman, J., McGwin, G., Jr., & Sims, R. (2001). Factors related to driving difficulty and habits in older drivers. *Accident Analysis and Prevention, 33*, 413-421.

Malfetti, J. L., & Winter, D. J. (1990). *A graded license for special elderly drivers: Premise and guidelines*. Washington, DC: AAA Foundation for Traffic Safety.

Man-Son-Hing, M., Marshall, S. C., Molnar, F. J., & Wilson, K. G. (2007). Systematic review of driving risk and the efficacy of compensatory strategies in persons with dementia. *Journal of the American Geriatrics Society, 55*, 878-884.

Marottoli, R. A., Allore, H., Araujo, K. L. B., Iannone, L. P., Acampora, D., Gottschaulk, M., Charpentier, P., Kasl, S., & Peduzzi, P. (2007). A randomized trial of a physical conditioning program to enhance the driving performance of older persons. *Society of General Internal Medicine, 22*, 590-597.

Marottoli, R. A., Cooney, L. M., Wagner, D. R., Doucette, J., & Tinetti, M. E. (1994). Predictors of automobile crashes and moving violations among elderly drivers. *Annals of Internal Medicine, 121*, 842-846.

Marottoli, R. A., Mendes de Leon, C. F., Glass, T. A., Williams, C. S., Cooney L. M., Jr., & Berkman, L. F. (2000). Consequences of driving cessation: Decreased out-of-home activity levels. *Journals of Gerontology Series B: Psychological Sciences & Social Sciences, 55*, S334-S340.

Marottoli, R. A., Mendes de Leon, C. F., Glass, T. A., Williams, C. S., Cooney L. M., Jr., Berkman, L. F., & Tinetti, M. E. (1997). Driving cessation and increased depressive symptoms: Prospective evidence from the New Haven EPESE. *Journal of the American Geriatrics Society, 45*, 202-206.

Marottoli, R. A., Ostfeld, A. M., Merril, S. S., Perlman, G. D., Foley, D. J., & Cooney, L.M., Jr. (1993). Driving cessation and changes in mileage driven among elderly individuals. *Journals of Gerontology Series B: Psychological Sciences and Social Sciences, 48*, S255-S260.

Marottoli, R., & Drickamer, M. (1993). Psychomotor ability and the elderly driver. *Clinics in Geriatric Medicine, 9*, 403-411.

Marshall, S. C. (2008). The role of reduced fitness to drive due to medical impairments in explaining crashes involving older drivers. *Traffic Injury Prevention, 9*, 291-298.

Massie, D. L., & Campbell, K. L. (1993). *Analysis of accident rates by age, gender and time of day based on the 1990 Nationwide Personal Transportation Survey* (Publication No. UMTRI-93-7). Ann Arbor: University of Michigan Transportation Research Institute.

Masten, S. V. (1998a). *Evaluation of the Class C driver license written knowledge test* (Report No. 173). Sacramento, CA: California Department of Motor Vehicles.

Masten, S. V. (1998b). *Preliminary evaluation of the referral driving performance evaluation program* (Report No. 176). Sacramento, CA: California Department of Motor Vehicles.

Masten, S. V. (1998c). *Evaluation of the referral driving performance evaluation program – Follow-Up Report* (Report No. 177). Sacramento, CA: California Department of Motor Vehicles.

Masten, S. V. (1999). *Evaluation of the Class C license written knowledge tests* (Report No. 182). Sacramento, CA: California Department of Motor Vehicles.

Mazer, B. L., Sofer, S., Korner-Bitensky, N., Gelinas, I., Hanley, J., & Wood-Dauphinee, S. (2003). Effectiveness of a visual attention retraining program on the driving performance of clients with stroke. *Archives of Physical Medicine and Rehabilitation, 84,* 541-550.

McCoy, G. F., Johnston, R. A., & Duthie, R. B. (1989). Injury to the elderly in road traffic accidents. *Journal of Trauma, 29*, 494-497.

Meuleners, L. B., Harding, A., Lee, A. H., & Legge, M. (2006). Fragility and crash over-representation among older drivers in Western Australia. *Accident Analysis and Prevention, 38*, 1006-1010.

Molnar, F. J., Patel, A., Marshall, S., Man-Son-Hing, M., & Wilson, K.G. (2006). Clinical utility of office-based cognitive predictors of fitness to drive in persons with dementia: A systematic review. *Journal of the American Geriatrics Society, 54*, 1809-1824.

Molnar, L. J., & Eby, D. W. (2008). The relationship between self-regulation and driving-related abilities in older drivers: an exploratory study. *Traffic Injury Prevention, 9*, 314-319.

Molnar, L. J., Eby, D. W., Kartje, P. S., & St. Louis, R. M. (2010). Increasing self-awareness among older drivers: The role of self-screening. *Journal of Safety Research, 41,* 367-373.

Nasvadi, G. C., & Wister, A. (2009). Do restricted driver's licenses lower crash risk among older drivers? A survival analysis of insurance data from British Columbia. *The Gerontologist, 49,* 474-478.

National Highway Traffic Safety Administration. (2005). *Traffic safety facts: 2005 data. Older Population (*DOT HS 810 622). Washington, DC: National Highway Traffic Safety Administration.

Naumann, R. B., Dellinger, A. M., & Kresnow, M. (2011). Driving self-restriction in high-risk conditions: How do older drivers compare to others? *Journal of Safety Research, 42,* 67-71.

Newgard, C. D. (2008). Defining the "older" crash victim: The relationship between age and serious injury in motor vehicle crashes. *Accident Analysis and Prevention, 40*, 1498-1505.

Ni, R., Kang, J. J., & Andersen, G. J. (2010). Age-related declines in car following performance under simulated fog conditions. *Accident Analysis and Prevention, 42,* 818-826.

O'Neill, D. (1997). Predicting and coping with the consequences of stopping driving. *Alzheimer Disease and Associated Disorders, 11*, 70-72.

Organization for Economic Cooperation and Development (OECD). (2001). *Aging and transport: Mobility needs and safety issues*. Paris: OECD.

Oulad Daoud, S., & Tashima, H. (2011). *2011 annual report of the California DUI Management Information System* (Report No. 233). Sacramento, CA: California Department of Motor Vehicles.

Owsley, C., Ball, K. K., Sloane, M. E., Roenker, D. L., & Bruni, J. R. (1991). Visual/cognitive correlates of vehicle accidents in older drivers. *Psychology and Aging, 6*, 403-415.

Owsley, C., Stalvey, B. T., & Phillips, J. M. (2003). The efficacy of an educational intervention in promoting self-regulation among high-risk older drivers. *Accident Analysis and Prevention, 35,* 393-400.

Owsley, C., Stalvey, B., Wells, J., & Sloane, M. (1999). Older drivers and cataract: driving habits and crash risk. *Journals of Gerontology Series A: Biological Sciences and Medical Sciences, 54*, M203-M211.

Oxley, J. A., Charlton, J. L., Koppel, S., Scully, J., & Fildes, B. N. (2005). Crash risk of older female drivers – contributing factors. *Annual Proceedings of the Association for the Advancement of Automotive Medicine, 49*, 345-360.

Oxley, J., & Whelan, M. (2008). It cannot be all about safety: The benefits of prolonged mobility. *Traffic Injury Prevention, 9*, 367-378.

Peck, R. C., & Gebers, M. A. (1991). *The traffic safety impact of traffic violator school citation dismissals* (Report No. 133). Sacramento, CA: California Department of Motor Vehicles.

Persson, D. (1993). The elderly driver: Deciding when to stop. *Gerontologist, 33*, 88-91.

Preusser, D. F., Williams, A. F., Ferguson, S. A., Ulmer, R. G., & Weinstein, H. B. (1998). Fatal crash risk for older drivers at intersections. *Accident Analysis and Prevention, 30*, 151-159.

Rabbit, P. (1993). Does it all go together when it goes? The nineteenth Bartlett Memorial Lecture. *Quarterly Journal of Experimental Psychology – A, 30*, 385-484.

Rabbitt, P., Carmichael A., Shilling, V., & Sutcliffe, P. (2002). *Age, health and driving: Longitudinally observed changes in reported general health, mileage, self-rated competence and in attitudes of older drivers*. Manchester, UK: AA Foundation for Road Safety Research.

Ragland, D., Satariano, W., & MacLeod, K. (2005). Driving cessation and increased depressive symptoms. *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences 60*, 399-403.

Raitanen, T., Tormakangas, T., Mollenkopf, H., & Marcellini, F. (2003). Why do older drivers reduce driving? Findings from three European countries. *Transportation Research Part F, 6*, 81-95.

Ranney, T. A., & Hunt, L. A. (1997). Researchers and occupational therapists can help each other to better understand what makes a good driver: Two perspectives. *Work, 8*, 293-7.

Reiner, T. W., & Hagge, R. A. (2006). *Evaluation of the class C driver license written knowledge tests* (Report No. 221). Sacramento, CA: California Department of Motor Vehicles.

Retchin, S., Cox, J., Fox, M., & Irwin, L. (1988). Performance-based measurements among elderly drivers and nondrivers. *Journal of the American Geriatric Society, 36*, 813-819.

Reuben, D., Silliman, R., & Traines, M. (1988). The aging driver: Medicine, policy and ethics. *Journal of the American Geriatrics Society, 36*, 813-819.

Rizzo, M., Reinach, S., McGehee, D., & Dawson, J. (1997). Simulated car crashes and crash predictors in drivers with Alzheimer Disease. *Archives of Neurology, 54,* 545-551.

Rock, S. M. (1998). Impact from changes in Illinois drivers license renewal requirements for older drivers. *Accident Analysis and Prevention, 30*, 69-74.

Rockhill, B., Newman, B., & Weinberg, C. (1998). Use and misuse of population attributable fractions. *American Journal of Public Health, 88,* 15-19.

Romanowicz, P. R., & Hagge, R. A. (1995). *An evaluation of the validity of California's driving performance evaluation road test* (Report No. 154). Sacramento, CA: California Department of Motor Vehicles.

Ruechel, S., & Mann, W. C. (2005). Self-regulation of driving by older persons. *Physical and Occupational Therapy in Geriatrics, 23*, 91-101.

Ryan, G. A., Legge, M., & Rosman, D. (1998). Age-related changes in drivers' crash risk and crash type. *Accident Analysis and Prevention, 30* 379-387.

Schultheis, M. T., Garay, E., & DeLuca J. (2001). The influence of cognitive impairment on driving performance in Multiple Sclerosis. *Neurology, 56,* 1089-1094.

Selander, H., Lee, H. C., Johnsson, K., & Falkmer, T. (2011). Older drivers: On-road and off-road test results. *Accident Analysis and Prevention, 43,* 1348-1354.

Shope, J.T. (2003). What does giving up driving mean to older adults, and why is it so hard to do? *Generations: Journal of the American Society on Aging, 27,* 57-59.

Simms, B. (1985). The assessment of the disabled for driving: A preliminary report. *International Rehabilitation Medicine, 7*, 187-192.

Siren, A., Hakamies-Blomqvist, L., & Lindeman, M. (2004). Driving cessation and health in older women. *Journal of Applied Gerontology, 23*, 68-69.

Soderstrom, C. A., & Joyce, J. J. (2008). Medical review of fitness to drive in older drivers: The Maryland experience. *Traffic Injury Prevention, 9*, 342-349.

Sommer, S. M., Falkmer, T., Bekiaris, E., & Panou, M. (2004). Toward a client-centered approach to fitness-to-drive assessment of elderly drivers. *Scandinavian Journal of Occupational Therapy, 11*, 62-69.

Stalvey, B. T., & Owsley, C. (2000). Self-perceptions and current practices of high-risk older drivers: implications for driver safety interventions. *Journal of Health Psychology, 5*, 441-456.

Staplin, L., Gish, K., & Wagner, E. (2003). MaryPODS revisited: Updated crash analysis and implications for screening program implementation. *Journal of Safety Research, 34*, 389-397.

Staplin, L., Lococo, K., Gish, K., & Decina, L. (2003a). *Model driver screening and evaluation program, final technical report, volume I: Project summary and model program recommendations* (Report No. DOT HS 809 582). Washington, DC: National Highway Traffic Safety Administration.

Staplin, L., Lococo, K., Gish, K., & Decina, L. (2003b). *Model driver screening and evaluation program, final technical report, volume II: Maryland Pilot Older Driver Study* (Report No. DOT HS 809 583). Washington, DC: National Highway Traffic Safety Administration.

Stav, W. B., Justiss, M. D., McCarthy, D. P., Mann, W. C., & Lanford, D. N. (2008). Predictability of clinical assessments for driving performance. *Journal of Safety Research, 39,* 1-7.

Stelmach, G., & Nahom, A. (1992). Cognitive-motor abilities of the elderly driver. *Human Factors , 34*, 53-65.

Stutts, J. C. (1998). Do older drivers with visual and cognitive impairments drive less? *Journal of the American Geriatrics Society, 46*, 854-861.

Stutts, J. C., Stewart, J., & Martell C. (1998). Cognitive test performance and crash risk in an older driver population. *Accident Analysis and Prevention, 30*, 337-346.

Stutts, J. C., Wilkins, J. W., Reinfurt, D. W., Rodgman, E. A., & Van Heusen-Causey, S. (2001). *The premature reduction and cessation of driving by older men and women*. Chapel Hill, NC: University of North Carolina Highway Safety Research Center.

Sundström., A. (2011). The validity of self-reported driver competence: Relations among measures of perceived driver competence and actual driving skill. *Transportation Research Part F, 14,* 155-163.

Taylor, B. D., & Tripodes, S. (2001). The effects of driving cessation on the elderly with dementia and their caregivers. *Accident Analysis and Prevention, 33*, 519-528.

Torpey, S. (1986). *License re-testing of older drivers*. Melbourne, Australia: Road Traffic Authority.

United States Department of Transportation. (1993). *Addressing the safety issues related to younger and older drivers*. Washington, DC: U.S. Department of Transportation.

United States Department of Transportation. (2003). *Motor vehicle traffic crash fatality and injury estimates for 2002*. Washington, DC: U.S. Department of transportation.

Vaa, T. (2004). *Impairment, diseases, age, and their relative risks of accident involvement: Results from meta-analysis* (TØI Report 690). Oslo, Norway: Institute of Transport Economics.

Vance, D. E., Roenker, D., Cissell, G., Edwards, J., Wadley, V., & Ball, K. K. (2006). Predictors of driving exposure and avoidance in a field study of older drivers from the state of Maryland. *Accident Analysis and Prevention, 38*, 823-831.

Vernon, D. D., Diller, E. M., Cook, L. J., Reading, J. C., Suruda, A. J., & Dean, J. M. (2002). Evaluating the crash and citation rates of Utah drivers licensed with medical conditions, 1992-1996. *Accident Analysis and Prevention, 34*, 237-246.

Viano, D. V., Culver, C. C., Evans, L., Frick, M., & Scott, R. (1990). Involvement of older drivers in multivehicle side-impact crashes. *Accident Analysis and Prevention, 22*, 177-188.

Waller, J. (1992). Research and other issues concerning effects of medical conditions on elderly drivers. *Human Factors, 34*, 3-24.

West, C. G., Gildengorin, G., Haegerstrom-Portnoy, G., Lott, L. A., Schneck, M. E., & Brabyn, J. A. (2003). Vision and driving self-restriction in older adults. *Journal of the American Geriatrics Society, 51*, 1348-1355.

Wheatley, C. J., & DiStefano, M. (2008). Individualized assessment of driving fitness for older individuals with health, disability, and age-related concerns. *Traffic Injury Prevention, 9*, 320-327.

Williams, A. F., & Carsten, O. (1989). Driver age and crash involvement. *American Journal of Public Health, 79*, 326-7.

Windsor, T., Anstey, K. J., Butterworth, P., Luszcz, M., & Andrews, G. (2007). The role of perceived control in explaining depressive symptoms associated with driving cessation in a longitudinal study. *Gerontologist, 47*, 215-223.

Withaar, F. K., Brouwer, W. H., & Van Zomeren, A. H. (2000). Critical review: Fitness to drive in older drivers with cognitive impairment. *Journal of the International Neuropsychological Society, 6*, 480-490.

Wood, J., & Mallon, K. (2001). Comparison of driving performance of young and old drivers (with and without visual impairment) measured during in-traffic conditions. *Optometry and Vision Science, 78*, 343-349.

Yasuda, M., Wilson, J., & von Mering, O. (1997). Driving cessation: The perspective of senior drivers. *Educational Gerontology, 23*, 525-538.

GLOSSARY OF TERMS AND ACRONYMS

3TAS:	3-Tier Assessment System
ADPE:	Area Driving Performance Evaluation
CA DMV:	California Department of Motor Vehicles
LRE	Licensing Registration Examiner
LOD:	Licensing Operations Division
PDO:	Property Damage Only
PRT:	Perceptual Response Test
R&D:	California DMV Research and Development Branch
SDPE:	Supplemental Driving Performance Evaluation
SWITRS:	Statewide Integrated Traffic Records System

APPENDICES

Appendix A

The Driving Information Survey

DL Number:_____________ Last Name: _______________________________

Please check only one box for each question. Thank you!

1. How many days per week do you normally drive a motor vehicle?
□ 1 □ 2 □ 3 □ 4 □ 5 □ 6 □ 7
□ Check here, if in most weeks you do not drive.

2. Do you avoid driving at night?
□ 1 Never □ 2 Sometimes □ 3 Often □ 4 Always

3. Do you avoid driving when it's raining or foggy?
□ 1 Never □ 2 Sometimes □ 3 Often □ 4 Always

4. Do you avoid driving during sunrise or sunset hours?
□ 1 Never □ 2 Sometimes □ 3 Often □ 4 Always

5. Do you avoid driving alone?
□ 1 Never □ 2 Sometimes □ 3 Often □ 4 Always

6. Do you avoid making left-hand turns across oncoming traffic?
□ 1 Never □ 2 Sometimes □ 3 Often □ 4 Always

7. Do you avoid driving in heavy traffic?
□ 1 Never □ 2 Sometimes □ 3 Often □ 4 Always

8. Do you avoid driving on the freeway?
□ 1 Never □ 2 Sometimes □ 3 Often □ 4 Always

9. Do you avoid driving on unfamiliar routes?
□ 1 Never □ 2 Sometimes □ 3 Often □ 4 Always

Appendix B

Correlation Matrices

Table B1

Correlation Matrix, Pilot Cohort (All 3-Tier Eligible Customers)

	Age	Sex	S/R? (prior 3 yrs)	Conv. (prior 3 yrs)	Crashes (prior 3 yrs)	Valid license days	Renewal?	Limited-term?	Zip code crash index	Zip code conv. index	Unlicensed?	Crashed? (subsequent 2 yrs
Age	1	—	—	—	—	—	—	—	—	—	—	—
Sex	.145**	1	—	—	—	—	—	—	—	—	—	—
S/R? (prior 3 yrs)	-.482**	-.150**	1	—	—	—	—	—	—	—	—	—
Convictions (prior 3 yrs)	-.461**	-.153**	.389**	1	—	—	—	—	—	—	—	—
Crashes (prior 3 yrs)	-.140**	-.020*	.044**	.212**	1	—	—	—	—	—	—	—
Valid license days	-.027**	-.046**	-.033**	-.007	-.008	1	—	—	—	—	—	—
Renewal?	-.033**	.005	-.059**	-.006	-.060**	.074**	1	—	—	—	—	—
Limited-term?	.158**	.015	-.072**	-.060**	-.010	-.032**	.005	1	—	—	—	—
Zip code crash index	.018*	-.004	.030**	-.002	.051**	-.024**	.009	-.011	1	—	—	—
Zip code conv. index	-.162**	-.043**	.136**	.101**	.055**	-.052**	.021*	-.027**	.492**	1	—	—
Unlicensed?	.042**	.033**	.015	-.010	-.001	-.912**	-.063**	-.005	.013	.032**	1	—
Crashed? (subsequent 2 yrs)	-.074**	-.026**	.063**	.082**	.083**	.030**	-.027**	.002	.021*	.025**	-.032**	1

Note. Total $N = 12,279$.
* = $p<0.05$. ** = $p<0.01$.

Table B2

Correlation Matrix, Baseline II Cohort (All 3-Tier Eligible Customers)

	Age	Sex	S/R? (prior 3 yrs)	Conv. (prior 3 yrs)	Crashes (prior 3 yrs)	Valid license days	Renewal?	Limited-term?	Zip code crash index	Zip code conv. index	Unlicensed?	Crashed? (subsequent 2 yrs)
Age	1	—	—	—	—	—	—	—	—	—	—	—
Sex	.090**	1	—	—	—	—	—	—	—	—	—	—
S/R? (prior 3 yrs)	-.391**	-.124**	1	—	—	—	—	—	—	—	—	—
Convictions (prior 3 yrs)	-.449**	-.155**	.343**	1	—	—	—	—	—	—	—	—
Crashes (prior 3 yrs)	-.184**	-.014	.035**	.210**	1	—	—	—	—	—	—	—
Valid license days	-.030**	-.016*	-.014	-.005	.007	1	—	—	—	—	—	—
Renewal?	-.041**	.000	-.008	.017*	-.010	.128**	1		—	—	—	—
Limited-term?	.121**	.001	-.050**	-.049**	-.017*	-.021**	-.102**	1	—	—	—	—
Zip code crash index	.023**	-.004	.002	-.002	.031**	-.026**	-.006	.006	1	—	—	—
Zip code conv. index	-.158**	-.025**	.101**	.096**	.045**	-.060**	.001	-.007	.478**	1	—	—
Unlicensed?	.051**	.005	.006	-.008	-.015*	-.895**	-.101**	-.009	.019*	.044**	1	—
Crashed? (subsequent 2 yrs)	-.098**	-.037**	.066**	.103**	.086**	.006	.003	-.006	.021**	.035**	-.006	1

Note. Total N = 15,879.
* = $p<0.05$. ** = $p<0.01$.

Table B3

Correlation Matrix, Nearby Cohort (All 3-Tier Eligible Customers)

	Age	Sex	S/R? (prior 3 yrs)	Conv. (prior 3 yrs)	Crashes (prior 3 yrs)	Valid license days	Renewal?	Limited-term?	Zip code crash index	Zip code conv. index	Unlicensed?	Crashed? (subsequent 2 yrs)
Age	1	—	—	—	—	—	—	—	—	—	—	—
Sex	.100**	1	—	—	—	—	—	—	—	—	—	—
S/R? (prior 3 yrs)	-.435**	-.186**	1	—	—	—	—	—	—	—	—	—
Convictions (prior 3 yrs)	-.431**	-.169**	.405**	1	—	—	—	—	—	—	—	—
Crashes (prior 3 yrs)	-.150**	-.035**	.072**	.241**	1	—	—	—	—	—	—	—
Valid license days	-.022*	-.013	-.013	-.008	.007	1	—	—	—	—	—	—
Renewal?	.015	.005	-.061**	-.024*	-.011	.002	1	—	—	—	—	—
Limited-term?	.131**	-.011	-.053**	-.050**	-.011	-.091**	.005	1	—	—	—	—
Zip code crash index	-.099**	-.016	.059**	.065**	.073**	-.010	.005	-.005	1	—	—	—
Zip code conv. index	-.078**	-.020*	.075**	.093**	.054**	-.013	-.011	-.003	.536**	1	—	—
Unlicensed?	.042**	.017	-.008	-.012	-.017	-.883**	.004	.030**	.004	.012	1	—
Crashed? (subsequent 2 yrs)	-.073**	-.031**	.051**	.091**	.044**	.003	-.011	.005	.028**	.015	-.004	1

Note. Total N = 10,564.
* = p<0.05. ** = p<0.01.