Peer Review of LBNL Statistical Analysis of the Effect of Vehicle Mass & Footprint Reduction on Safety (LBNL Phase 1 and 2 Reports)

United States Environmental Protection Agency

Peer Review of LBNL Statistical Analysis of the Effect of Vehicle Mass & Footprint Reduction on Safety (LBNL Phase 1 and 2 Reports)

Assessment and Standards Division
Office of Transportation and Air Quality
U.S. Environmental Protection Agency

Prepared for EPA by
Systems Research and Application Corporation
EPA Contract No. EP-C-11-007

Peer Review of *LBNL Statistical Analysis of the Effect of Vehicle Mass & Footprint Reduction on Safety (LBNL Phase 1 and 2 Reports)*

Table of Contents

Executive Summary

In September 2011, EPA contracted with SRA International (SRA) to conduct a peer review of *LBNL Statistical Analysis of the Effect of Vehicle Mass & Footprint Reduction on Safety) (LBNL Phase 1 and 2 Reports)*, prepared by Tom Wenzel, Lawrence Berkeley National Laboratory.

The peer reviewers selected by SRA were Donna Chen and Kara Kockelman (University of Texas at Austin), Charles Farmer (Insurance Institute for Highway Safety), David Greene (Oak Ridge National Laboratory), and Michael Van Auken (Dynamic Research, Inc.). EPA would like to extend its appreciation to all three reviewers for their efforts in evaluating this survey. The reviewers brought useful and distinctive views in response to the charge questions.

The first section of this document contains the final SRA report summarizing the peer review of the *LBNL Phase 1 and 2 Reports*, including the detailed comments of each peer reviewer and a compilation of reviewer comments according to the series of specific questions set forth in the peer review charge. The SRA report also contains the peer reviewers' resumes, completed conflict of interest and bias questionnaires for each reviewer, and the peer review charge letter. The second major section contains our responses to the peer reviewers' comments. In this section, we repeat the compiled comments provided by SRA and, after each section of comments, provide our response. We have retained the organization reflected in SRA's compilation of the comments to aid the reader in moving from the SRA report to our responses.

TO: Cheryl Caffrey, U.S. Environmental Protection Agency, Office of Transportation and Air Quality (OTAQ)

FROM: Brian Menard, SRA International

DATE: February 28, 2012

SUBJECT: Peer Review of *LBNL Statistical Analysis of the Effect of Vehicle Mass & Footprint Reduction on Safety) (LBNL Phase 1 and 2 Reports), prepar*ed by Tom Wenzel, Lawrence Berkeley National Laboratory.

1.   Background

In developing programs to reduce greenhouse gas (GHG) emissions and increase fuel economy of light-duty highway vehicles, the U.S. Environmental Protection Agency (EPA) and the National Highway Transportation Safety Administration (NHTSA) have to evaluate the safety of mass reduction technologies likely to be used to meet future standards. The U.S. Department of Energy (DOE) contracted with Lawrence Berkeley National Laboratory (LBNL) to perform a statistical analysis of the effect of vehicle mass and footprint reduction on safety. LBNL's analysis of the relationship between vehicle mass, footprint, and societal fatality and casualty risk consisted of two phases. Phase 1 was an assessment of the NHTSA report *Relationships between Fatality Risk, Mass, and Footprint in Model Year 2000-2007 Passenger Cars and LTVs.* This study used logistic regression analysis to estimate the relationship of changes in vehicle mass and footprint on U.S. fatality risk per vehicle mile traveled. Phase 2 was an independent logistic regression analysis to estimate the relationship between vehicle mass, footprint and total <u>casualty</u> (fatality plus serious injury) risk, per police-reported crash, using state-level data on all crashes.

This report documents the peer review of *Assessment of NHTSA's Report "Relationships Between Fatality Risk, Mass, and Footprint in Model Year 2000-2007 Passenger Cars and LTVs" (LBNL Phase 1 Report)* and *Analysis of the Relationship between Casualty Risk Per Crash and Vehicle Mass and Footprint for Model Year 2000-2007 Light-Duty Vehicles (LBNL Phase 2 Report)*. Section 2 of this memorandum describes the process for selecting reviewers, administering the review process, and closing the peer review. Section 3 summarizes reviewer comments according to the series of specific questions set forth in the peer review charge. The appendices to the memorandum contain the peer reviewers' resumes, completed conflict of interest and bias questionnaires for each reviewer, and the peer review charge letter.

2.   Description of Review Process

In September 2011, OTAQ contacted SRA International to facilitate the peer review of the *LBNL Phase 1 and 2 Reports*. The reports were prepared by Tom Wenzel, Lawrence Berkeley National Laboratory.

EPA provided SRA with a short list of subject matter experts from academia and industry to serve as a "starting point" from which to assemble a list of peer reviewer candidates. SRA selected three independent (as defined in Sections 1.2.6 and 1.2.7 of EPA's *Peer Review Handbook, Third Edition)* subject matter experts to conduct the requested reviews. SRA selected subject matter experts with

experience in statistics and statistical model analysis, vehicle mass safety, and vehicle crash and safety engineering. The coverage of these subject areas is shown below in Table A.

Table A:
Peer Reviewer Experience and Expertise

Name	Affiliation	Coverage				
		Statistics & methodology	Knowledge of past vehicle mass safety statistics study	Knowledge of vehicle crash/safety engineering	Knowledge of vehicle safety database (FAS, State Crash & Vehicle Attributes)	Statistical model analysis
David Green	Oak Ridge National Laboratory	Y	Y	/	Y	Y
Charles Farmer	IIHS	Y	Y	Y	Y	Y
Donna Chen& Kar Kockelman	University of Texas	Y	Y	Y	Y	Y
Michael Van Auken	Dynamic Research Inc.	Y	Y	Y	Y	Y

To ensure the independence and impartiality of the peer review, SRA was solely responsible for selecting the peer review panel. Appendix A of this report contains the resumes of the three peer reviewers. A crucial element in selecting peer reviewers was to determine whether reviewers had any actual or perceived conflicts of interest or bias that might prevent them from conducting a fair and impartial review of the CVCM and documentation. SRA required each reviewer to complete and sign a conflict of interest and bias questionnaire. Appendix B of this report contains an explanation of the process and standards for judging conflict and bias along with copies of each reviewer's signed questionnaire.

SRA provided the reviewers a copy of the most recent version of the *LBNL Phase 1 and 2 Reports* as well as the peer review charge containing specific questions EPA asked the reviewers to address. The charge included a matrix of questions issues upon which the reviewers were asked to comment. Reviewers were also encouraged to provide additional comments, particularly in their areas of expertise and work experience. Appendix C of this report contains the memo to reviewers from SRA with the peer review charge and response matrix.

EPA sought peer reviewers' expert opinions on the statistic methodologies used in the *LBNL Phase 1 and 2 Reports* and whether they are likely to yield realistic estimates of the relationship between vehicle mass, footprint, and total fatality or casualty risk. EPA requested that each reviewer comment on all aspects of the two LBNL studies, with particular emphasis on the methodologies employed, assumptions inherent to the analysis, sources of information employed, methods of calculation and any other key issues the reviewer may identify. Reviewers were encouraged to examine and evaluate the NHTSA study in helping them to understand the LBNL assessment analysis. Findings of this peer review may be

used toward validation and improvement of the statistical analysis conducted by LBNL, and to inform EPA staff on potential use of the regression results for predicting the safety effect of future standards in reducing mass and footprint.

A teleconference between EPA, LBNL, the reviewers, and SRA was held to allow reviewers the opportunity to raise any questions or concerns they might have about the *LBNL Phase 1 and 2 Reports*, and to raise any other related issues with EPA and SRA, including EPA's expectations for the reviewers' final review comments. SRA delivered the final review comments to EPA by the requested date. These reviews, contained in Appendix D of this report, included the reviewers' response to the specific charge questions and any additional comments they might have had.

3.   Compilation of Review Comments

The *LBNL Phase1 and 2 Reports* were reviewed by Donna Chen and Kara Kockelman (University of Texas at Austin), Charles Farmer (Insurance Institute for Highway Safety), David Greene (Oak Ridge National Laboratory), and Michael Van Auken (Dynamic Research, Inc.).   Appendix A contains detailed resumes for each of the reviewers. This section provides a compilation of their comments. The complete comments may be found in Appendix D.

Assessment of NHTSA's Report "Relationships Between Fatality Risk, Mass, and Footprint in Model Year 2000-2007 Passenger Cars and LTVs" (LBNL Phase 1 Report)	
1. ASSUMPTIONS	**COMMENTS**
Please comment on the validity of any assumptions embedded in the LBNL assessment analysis and the independent casualty analysis that could affect the projected relationship between vehicle mass/footprint reductions and fatality/casualty risk. Examples might include assumptions regarding whether recent historical relationships between vehicle weight, size, and safety will continue into the future; potential future improvements in vehicle technology and design may result in compensatory safety benefits; and the annual baseline fatality distribution.	**[Chen and Kockelman]** [1] The report does a nice job discussing recent trends in vehicles, such as the increase of ESC, side airbags, and light truck crash compatibility with passenger cars – which will improve safety outcomes for all vehicles, but perhaps most significantly the smaller and lighter vehicles. . It also mentions the phasing out of the lightest and smallest vehicles between model years 2000-2007 (but doesn't mention the makes and models somehow), which were particularly poor safety performers in the past. However, with the introduction of urban commuter vehicles, such as the SmartCar, Mini Cooper, and Fiat 500m, and the growing popularity of smaller, fuel-efficient compact vehicles following gas price increases, this trend does not seem so obvious. Such vehicles should be discussed.

[2] The simplistic logistic model employed in this analysis only accounts for two crash outcomes (fatal versus non-fatal) and so neglects the more detailed, and ordered nature of injury severity data, which is unfortunate. The model also assumes error-term homoscedasticity from one crash or individual to the next; in reality certain vehicle types (e.g., pickups) and crash contexts (e.g., high speed crashes) have more uncertainty associated with their severity outcomes. It would be good to point out such limitations for readers.

[Van Auken] The basic assumptions, methodology, and data are primarily the same as in the Kahane (2011) report. These include the following:

1) The probability of a crash fatality is proportional to the vehicle miles travelled (VMT), except as noted in Section 5.1
2) The logarithm of probability of fatality per VMT for a given curb weight, footprint, and control variable values varies as a linear combination of the curb weight, footprint, and control variables within the domain of the data.
3) The logistic regression methods determine a maximum likelihood estimate of model coefficients.
4) It is assumed that the above relationships remain constant in the recent past (i.e., 2000-2007 model year vehicles in the 2002-2008 calendar years), present, and near future (i.e., 2017-2025 model year vehicles).

The first assumption that crash fatalities are proportional to VMT rather than the number of vehicle |

| | registration years (VRY) is appropriate because the fatalities cannot occur if the vehicles are not driven on the road (i.e., VMT = 0). This assumption is qualified however because VMT is more difficult to measure than VRY and therefore may be less accurate. On the other hand the probability of a fatal crash or the number of fatalities in a crash may also depend on the vehicle occupancy. The analysis in Section 5.1 is a commendable attempt to explore the sensitivity to this assumption, however the Kahane (1997) and DRI (2003-2005) reports have shown that some driver, vehicle, and environmental factors may be underrepresented or overrepresented in unweighted induced-exposure data. VRY could have also been considered as a measure of exposure.

The second and third assumptions are appropriate provided that it is recognized that it is essentially impossible with currently available knowledge and information to model all of the factors that could affect the probability of fatality in a crash, and that the objective of the analysis is to identify overall trends versus vehicle weight and footprint. In general the probability of fatality depends on other many other factors which have not been modeled (e.g., driver behavior factors, vehicle design factors, roadway design factors, EMT factors) and these unmodeled factors are assumed to be uncorrelated with vehicle weight and footprint, and/or are represented by the other control variables. The latter assumption might or might not be valid.

The fourth assumption is perhaps the weakest because it assumes that future vehicles will have the same design characteristics as past vehicles, and that the characteristics of the vehicle population (e.g., collision partner weight, size, type) will also remain the same. A commendable attempt to partially address this effect is described in Section 6. These effects can be perhaps better addressed by the "Volpe model" described in Kahane (2011) of the Honda-DRI fleet systems model described in Refs (2324), which can be used to forecast the effects of mass reductions of individual makes and models on a year-by-year basis. |
| Please comment on any apparent unstated or implicit assumptions and related caveats or limitations. | **[Chen and Kockelman]** The role of driver behavior is briefly addressed in the report but not emphasized sufficiently. Fatality risk is a combination of driver, vehicle, and roadway characteristics. Driver behavioral differences are many and do not solely exist for pickup truck drivers versus car drivers. Socioeconomic data such driver household income, size, and education influence driver attitudes and driving environments. For example, Chen et al. (2010) found that crash risk increases for those living in socioeconomically disadvantaged areas (including households more likely to drive less expensive and older vehicles). Though such data is not typically available in state and national crash databases, the importance of these driver and environmental characteristics on crash rates (per mile driven) and fatality risk should be stressed in both reports. It is clearly very difficult to control for, but a major caveat to the NHTSA (& now LBNL) results. We expect that crash severity could be probably be lower for many of the small cars and pickups if they were driven by those who |

| | tend to drive more expensive vehicles, under the same settings (e.g., daytime, urban freeway). Similarly, in the second LBNL report (which uses VMT estimates), we expect that crash rates would probably be lower for these types of driver-vehicle-setting combinations.

[Farmer] The statistical models assume no interaction between the vehicle size/weight measures and any of the numerous covariates, but this may not be true. For example, size/weight reductions may differently affect vehicles with and without ESC if they affect vehicle handling. It is risky to make statements such as that on p. 11 of the Phase I report: *Therefore, the mass of a lighter car could be reduced by 800 lbs while adding ESC, without increasing fatality risk.*

[Van Auken] The induced-exposure data set provided by NHTSA is based on the "non-culpable" vehicle in two-vehicle crashes. It is assumed that the dataset is a reprehensive sample of the driver and environmental exposure factors for vehicle use. However, since these cases include moving vehicles, some vehicle-driver-environmental conditions may be under or over represented in this data depending on how they affect the ability of a non-culpable vehicle to avoid a crash. Results in Ref (17) indicate that the estimated effect of weight and size reduction are sensitive to whether the induced-exposure data are based on the Kahane (2003) non-culpable vehicle definition of the Kahane (1996) stopped vehicle definition.
Unfortunately it is not currently possible to test this sensitivity with the NHTSA-provided induced-exposure data. |
|---|---|

ADDITIONAL COMMENTS:

[Chen and Kockelman] Chen, H.Y., Ivers, R.Q., Mariniuk, A.L.C., Boufous, S., Senserrick, T., Woodward, M., Stevenson, M. and Norton R. Socioeconomic status and risk of car crash injury, independent of place of residence and driving exposure: Results from the DRIVE study. Journal of Epidemiology and Community Health 64(10), 2010, pp. 998-1003.

2. CONTROL AND DEPENDENT VARIABLES USED IN THE REGRESSION MODELS	COMMENTS
Please comment on the adequacy of control and dependent variables used in the assessment analysis and independent casualty analysis, and recommend any alternative control or dependent variables that are available for possible inclusion in the analysis. For example, what are the relative merits of the main dependent variables used, fatality risk per estimated VMT, and casualty risk per police-reported crash?	**[Chen and Kockelman]** [1] As alluded to above, a primary concern is that the NHTSA analysis (& thus the LBNL analyses) largely neglect the idea that vehicle type (make & model) is very much a proxy for driver type, and a vehicle's crash avoidance may have very little to do with vehicle type. It has a lot to do with the person behind the wheel, and gender & age simply aren't enough to control for such distinctions. Education, risk aversion, ability, wealth, etc., are important covariates. But existing data sets are quite limiting (though the MVOSS & FAR with 3-year driver violation history do offer some valuable insights, not discussed in these reports). In reality, small cars may be less crash prone than Kahane's & Wenzel's results suggest, because they are driven by lower-income, younger, less risk averse people driving in more crash prone settings (e.g., commercial strips rather than pricey residential suburbs). Such key caveats need thoughtful discussion. Four relevant papers on the topics of crash frequency and vehicle size-and-weight implications (by Knipling, Kweon & Kockelman, Wang and Kockelman, and Chen & Kockelman) have been sent to Tom Wenzel. These all include useful literature reviews for further connections to useful findings for citation in the reports, as time allows the contractor. [2] The grouping of the vehicles into heavier- and lighter-than-average weight categories essentially splits a "typical" weight vehicle of that type into two categories. The impacts of curb weight and footprint on fatality risk may be easier to interpret if the vehicles were grouped into 3 weight categories (light, average, and heavy) & by type (with the average category representing vehicles within one standard deviation of average weight). Furthermore, the grouping of CUVs and minivans into the same vehicle type category neglects the fact that these vehicles have faced rather different ground clearance requirements (impacting rollover potential), door types (sliding vs. standard), and, perhaps most importantly, can appeal to different types of drivers (as indicated in the market shift of car drivers to CUV drivers). **[Farmer]** One needs to restrict control variables to those that are available and reliable. A problem when combining state databases is that the states often are not consistent as to the variables coded and the definitions of those variables. This severely limits the list of possible control variables. **[Van Auken]** [1] The main metric used in both the Kahane (2011) and Wenzel (2011a) reports is the total number of fatalities (except as noted). Reducing the total number of fatalities, which includes both subject vehicle occupants and collision partner fatalities, is desirable from a societal viewpoint. Fatal crash occurrence is related to the total number of fatalities, which has been used by Kahane (2003, 2011) to address concerns

	about double counting. [2] VMT is a good measure of accident exposure provided that it can be accurately determined. [Note: This peer review does not address the Wenzel (2011b) companion report (Ref 3) which examines the risks per police-reported crash. See Ref 4 for comments on the companion report.]
What additional control variables, such as vehicle make or model, might be included in the regression models?	**[Chen and Kockelman]** [1] Vehicle height, a variable which may be more valuable than vehicle type for similarly structured vehicles such as sedans, wagons, CUVs, and minivans, would be a valuable control variable. In addition to a wider track, a lower center of gravity also increases vehicle stability, thereby reducing the risk of rollover. Relevant literature & findings exist, and should be cited. [2] Other variables which have been found in past studies to influence fatality risk such as seat belt use, roadway geometry and division type are not included in this study (which is largely a repeat of the NHTSA study, as specifically contracted by the EPA). [3] To account for driver characteristics that contribute to fatality risk, socioeconomic variables such as household income, education, household size, etc. would be valuable additions. Unfortunately, both state and national crash databases typically do not include such information (outside of MVOSS). Such issues should be flagged for readers. It seems the contractor has done his duty, and the key limitations lie with the original methodology he was to essentially duplicate.
Please comment on any caveats or limitations that these dependent variable or control variables entail with respect to use of the results as the basis for estimating the safety effect of mass reduction.	**[Chen and Kockelman]** Please see above comment (in Assumptions section) regarding driver behavior and environment. **[Farmer]** Model overspecification could be the reason for results that are non-intuitive, especially in the Phase II analyses of police-reported crashes. Control variables may be correlated with each other or with the size and weight variables. For example, Figure 2.9 of the Phase I report implies that torso side airbags increase fatality risk in CUVs.

ADDITIONAL COMMENTS:

[Chen and Kockelman] Table 2.1 has many indicator variables labeled as "C" for continuous variable (such as ABS, ESC, AWD, DRVMALE, etc). These C's should be removed.

[Van Auken] The underlying reasons for some of the estimated effects are unknown at this time, but presumably involve driver, vehicles, environment or

accident factors that have not been controlled for in the Kahane (2011) and Wenzel analyses. See, for example, Refs 17 and 25.

3. METHODOLOGY AND STATISTICS	COMMENTS
Please comment on the validity and applicability of the methodology LBNL used in assessing the NHTSA 2011 study and its analysis of the relationship between mass, footprint, and risks per police-reported crash.	**[Chen and Kockelman]** The report assesses the NHTSA 2011 study in a fair amount of detail and seeks to introduce some additional analyses to better examine the relationship between mass, footprint, and fatality risks. However, due to a lack of control for very specific vehicle differences (which vary by make & sub-model), the exclusion of driver characteristics and crash setting details (which cannot always be controlled for, but are often correlated with vehicle type), the effects of downweighting vehicles and/or shifting vehicle styles and sizes may be overestimated. Simply changing the vehicle on a risky driver in a high-risk setting is unlikely to influence outcomes significantly.

[Van Auken] The logistic regression methods seem to be appropriate. The confidence intervals are based on the logistic regression Wald Chi-Square statistic, which as Kahane (2003, 2011) has demonstrated does not include all sources of variation. However, these confidence intervals are useful because they do provide some indication of the uncertainty in the results. [Note: This peer review does not address the Wenzel (2011b) companion report, which examines the risks per police-reported crash. See Ref 4 for comments on the companion report.] |
| Please review other statistical methods LBNL has used in the analysis, in addition to the logistic regression methodology. Examples include the alternative approaches used by LBNL to assess NHTSA interval estimation results, and LBNL's linear regression analysis of actual, predicted, and residual risk by vehicle model. | **[Chen and Kockelman]** [1] In the alternative measures of exposure, the author examines the effect of vehicle manufacturer on fatality risk and treats the luxury models produced by Toyota, Honda, and Nissan as separate manufacturers. However, domestic luxury brands (such as Cadillac & Lincoln) are categorized with their nameplate manufacturers (GM and Ford), which appears inconsistent.

[2] The effect of calendar year variables on fatality risk may be overestimated here, since VMT is tracked by vehicle model and not by calendar year. The trend of greatest fatality risk reductions in light trucks, CUVs and minivans with increasing calendar year may simply be a reflection of rising gas prices in combination with the ailing economy contributing to lower VMT (in these relatively low-fuel-economy vehicles).

[3] It is unclear how the author determined the various percentage replacements of vehicle types in the aggressive vehicle market share shift scenario. (For example, why are 50% of SUVs replaced by CUVs and 60% of small pickups replaced by CUVs? The CUV is a more natural replacement for an SUV, and an SUV a more natural replacement for a pickup.)

[Van Auken] [1] The correlations in Section 3 appear to be assessed using the Coefficient of Multiple Determination (R^2) based on a linear fit to the data (e.g., the correlation between footprint versus curb weight in Figure 3.1 on p. 14). The linear regression model attributes the differences between the dependent variable |

(vertical axis_ and the linear fit to the independent variable (horizontal axis) to random effects. If there is no preference as to the choice of independent and dependent variables (e.g., footprint versus curb weight, or curb weight versus footprint), then the linear trend and R^2 result would be different if the two variables were interchanged, and having two different yet equally valid results would be undesirable.

[2] If the variation in the data can be attributed to both variables (e.g., footprint and curb weight), then it would be better to report the square of the sample correlation coefficient r^2, , where r is computed according to Eqn (1). The trend lines in these correlation figures should not be computed using a linear regression. Instead, the trend line should pass through the sample means (I.e. (x, y)), and have a slope equal to the ratio of the sample standard deviations in the data (i.e., s_y / s_x). Therefore, the reported correlation results do not depend on the ordering of the data variables.

Note this comments does not apply to linear trends indicated in Section 4, for which the Coefficient of Multiple Determination (R^2) seems appropriate.

[3] The Coefficient of Multiple Determination (R^2) is frequently used in the Wenzel (2011a) report as an indicator of the statistical importance of a linear trend (e.g., R^2 values in Tables 4.1 and 4.2 that are greater than 0.3 are shown in blue font). It would be better to report the standard error, confidence interval, and/or probability value as measures of the statistical significance of a linear trend.

| Please comment on caveats or limitations of using non-significant regression estimates to project the safety impact of mass reduction. | **[Chen and Kockelman]** First, the t-statistics are not provided in the report which makes it difficult for the reader to assess statistical significance of specific regression estimates (except where noted by the author). Second, inclusion of a statistically insignificant variable can influence the estimates of coefficients associated with related variables. Nevertheless, in general, it is best to keep insignificant estimates if one has a strong defense for their role, since removing such variables (& thus their parameters) will shift the burden of response to a correlated covariate's parameter, thus biasing the latter. We generally keep key covariates in a model up to a pvalue of 0.20 or 0.25 or so, especially in relatively small data sets (e.g., n < 1,000). Covariates for which we have no strong basis can be removed for pvalues > 0.10.

[Van Auken] Regression estimates are random numbers which have an unknown expected value and variance, and known sample value and standard error. If the sample value can be explained by a zero expected value and known standard error then the result is considered not statistically significantly different than zero and therefore the result is not considered to be statistically significant. However, If we can combine this estimate with other estimates then the unknown expected values and variances can also be combined using the same |

transformation, and the statistical significance of the combined result can be tested. Therefore, depending on the sample values and inter-correlation, the combined result may be statistically significant even if the individual estimates are not statistically significant.

For example, the results from each of the nine different crash types can be combined into an overall estimate and the standard error calculated assuming that the results for each crash type are independent of each other. Then the statistical significance of the combined effect can be determined.

However, and Kahane (2011) points out there are two sources of uncertainty in the regression results. The first is the FARS based sampling error which is uncorrelated across crash types because they are based on different fatal cases (Kahane 2011, p. 77). The second is the state based include-exposure sampling error which is correlated across crash types because they are based on the same included-exposure cases. Therefore a confidence interval estimated using the jackknife method described by Kahane (2011) and accounting for correlation of these two error sources would be more accurate than a simple estimate based on the Wald Chi-Square statistic and assumed independence.

How might the LBNL methodology be strengthened to better represent future vehicle designs and reduce multi-collinearity between mass and footprint in the regression analysis?	**[Chen and Kockelman]** Including more vehicle-specific characteristics (such as vehicle height and engine size) reduces the analysis' dependence on vehicle type, since vehicle shapes and structures will continue to evolve. There is also correlation with context (e.g., pickups are driven in more rural locations, with greater hazards [like less lighting, higher speed, & few medians]). Disaggregate data are almost always best, to avoid ecological fallacies & such. **[Van Auken]** [1] The effects of multi-collinearity can be mitigated by 1) obtaining more data, 2) pooling data from different crash type or vehicle types, or 3) reducing the number of regression variables. The first option would require more calendar years and/or model years, which would involve added newer data as it becomes available (or using older data). The second option might be to recombine the CUVs and minivans with truck based vans and adding a control variable to compensate for the differences in the vehicles types. The third option might involve removing statistically insignificant control variables or removing control variables that would not be expected to have an effect on the probability of fatality in the crash (e.g., the side airbag variable is not included in pedestrian crashes because it is not expected to affect pedestrian fatality risk). The number of driver age control variables might be reduced from eight to three (as in the Kahane (1997) and DRI (2002-2005) studies). Finally, a linear curb weight model instead of a two-piece linear model may help to better elucidate the general trend.

	[2] The Variance Inflation Factor (VIF) has been suggested as a measure of multi-collinearity in the Kahane (2010 and 2011) reports, however this diagnostic metric does not account for differences in database size (i.e., Options 1 and 2 above). The Wenzel (2011a) report does not discuss the Variance Inflation Factor or report any VIF results.

ADDITIONAL COMMENTS:

[Chen and Kockelman] On page 55, it is unclear what is meant by "however; if anything, reduction of this type of fatality will increase detrimental effect of mass reduction in cars."

4. DATA SETS	COMMENTS
Please comment on the validity and applicability of the datasets used to project changes in risk resulting from reduction in vehicle mass. LBNL's casualty analysis used police-reported crash data from 16 states, while the 2011 NHTSA study used national fatality data, combined with a subset of non-culpable vehicles involved in two-vehicle crashes from police-reported crash data from 13 states.	**[Chen and Kockelman]** [1] The acquisition of Polk data for VMT estimates by make & model is valuable, and a contribution to the literature. However, these estimates come from vehicles found in repair shops in non-attainment areas, and so will be biased towards problem-prone vehicles, wealthier households who service their vehicles more regularly, and/or urban (smoggier) areas. Such issues merit careful discussion in the paper, so that readers are well aware of caveats.

[2] Related to this, Tom Wenzel indicated (by phone) that he did take a look at CA's extensive odometer reads, which go into some semi-rural locations (not too rural), and he indicated that the VMT values by vehicle type (not controlling for HH attributes & such) are very similar (just 5% longer in rural areas) – except for vans (which are used much more extensively in rural areas). This is interesting to me, and is not that different from what we've seen in the past. For example, Kockelman & Zhao's JTS paper from 2000 (pre- print at http://www.ce.utexas.edu/prof/kockelman/public_html/ BTSJournalLDTs.pdf) suggests that, after controlling for various HH attributes & vehicle types, density is/was still very important (tables 1 & 2), but a shift from a density of 1k to 5k persons per sq mile (which is 1.5 vs. 7.8 persons per acre) means an increase of 750 mi/yr/vehicle (which is about 7.5% of annual VMT). Such differences, and their practical significance (or lack thereof) should be discussed in the reports.

[Van Auken] The inducted-exposure data set provided by NHTSA is based on the non-culpable vehicles in two-vehicle crashes. See the comments in Table 1 on the limitations of this data. In addition, there are also many differences in the coding variables and values used by the different states, which tend to make the recoding to a common data set imprecise. [Note: This peer review does not address the Wenzel (2011b) companion report, which examines the risks per police-reported crash. See Ref 4 for comments on the companion report.] |
| Please comment on any apparent, unstated, or implicit impact on estimated risks inherent in the two different approaches, and any related caveats or limitations. For example, what are the strengths and weaknesses of the two measures of vehicle exposure, miles of vehicle traveled scaled up from crash data from 13 states, and number of police-reported crashes? | **[Chen and Kockelman]** The use of non-culpable vehicles in two-vehicle crashes as a proxy for vehicles which are "just there" may be distorting the overall distribution of vehicle models. VMT may differ between vehicles that are more prone to run-off-road accidents, at-fault two-car crash vehicles, and non-culpable vehicles.

[Van Auken] [1] The number of fatal cases tends to be much less than the number of induced-exposure cases. Therefore the effective numbers of degrees-of-freedom in the statistical estimates tend to be limited by the available number of fatal cases. For example, it would not be possible to estimate the effects of two variables (e.g., just curb weight and footprint) if we had data for only one fatal case even if we had thousands of induced-exposure cases. Therefore it is desirable to use data for the entire US in order to get a large sample of fatal cases for the logistic regressions. This then requires the available induced-exposure data (i.e., from 13 |

	states) to be "scaled up" the US level using the method described in Kahane (2003 and 2011). The result is the best currently available estimate of vehicle exposure. [2] There may be some concerns about the accuracy of the vehicle miles-traveled data because the difficulty estimating the number of vehicle miles travelled at the make-model-year=state level of detail. [Note: This peer review does not address the Wenzel (2011b) companion report, which examines the risks per police-reported crash. See Ref 4 for comments on the companion report.]
ADDITIONAL COMMENTS:	

5. RECOMMENDATIONS	COMMENTS
Please comment on whether the LBNL assessment adequately addresses the NHTSA 2011 study and identifies the safety impact from mass reduction. Are the analytic methods and data used to assess the NHTSA study, and estimate the relationship between risk, mass, and footprint, appropriate? Is casualty risk per crash a legitimate measure of vehicle safety? What other methods or data could be used to better predict the effect of future vehicle designs on safety?	**[Chen and Kockelman]** While driver fatalities per crash seems a useful measure of vehicle design safety, and examination of fatal crash rates is very valuable (using Polk- based exposure estimates), there are many caveats to work of this type. As noted above: a primary concern remains a neglect of the notion that the type of car is very much a proxy for driver type, and a vehicle's crash avoidance may have very little to do with vehicle type. It has a lot to do with the person behind the wheel. Simply including gender and age variables cannot account for important covariates such as education, risk aversion, driving ability, wealth, etc. In reality, small cars may be less crash prone than Kahane's and Wenzel's results suggest, because they are driven by lower-income, younger, less risk averse people driving in more crash prone settings (e.g., commercial strips rather than pricey residential suburbs). Of course, as noted above, it is very difficult to control for all these variables, and the contractor was asked to rely on the original data. In reality, the best the report authors can do with such data sets is to explain how all the other, relevant attributes may factor in (e.g., quality of driver and typical driving settings), and how they can generate biased estimation (sometimes in either direction). Discussion of relevant literature that looks more deeply at crash outcomes (e.g., Wang or Chen's papers, mentioned above, allowing for heteroscedasticity and individual vehicle attributes, non-driver outcomes, etc.) will also be useful.

[Van Auken] [1] The basic methodology described by Kahane (2011) seems appropriate; however some results using this method and data are not well understood and need further diagnosis.

[2] The induced-exposure data set provided by NHTSA is based on the non-culpable vehicles in two-vehicle crashes. See the Table 1 comments on the limitations of this data. [Note: This peer review does not address the Wenzel (2011b) companion report, which examines the risks per police-reported crash.] |
| Please comment on the overall adequacy of LBNL's assessment of the 2011 NHTSA report and its independent study of casualty risk for predicting the effect of vehicle mass or footprint reduction on safety. Provide any recommended improvements that might reasonably be adopted by the author to improve the analysis. | **[Chen and Kockelman]** Overall, the study is a comprehensive assessment of the 2011 NHTSA report and introduces interesting additional analyses to examine the relationship of vehicle mass and footprint reduction on safety. However, as stated previously in the comments here, driver preference for specific car types (including size and mass) is related to driver socioeconomic characteristics and driving behavior. As vehicle, driver, and roadway environment characteristics all contribute to fatality risk, the effects of physical vehicle changes such as mass or footprint reduction on safety should not be overstated when the other two types of characteristics are not sufficiently accounted for.

[Farmer] Overall these are reasonably good studies. The Phase I report does a very good job of assessing the NHTSA report of fatality risk. |

	[Van Auken] The Wenzel (2011a) report provides a valuable supplement to the analysis and results in the Kahane (2011) report. [Note: This peer review does not address the Wenzel (2011b) companion report, which examines the risks per police-reported crash.]
ADDITIONAL COMMENTS: **[Van Auken]** See attached tables 6 and 7 below.	

Table 6. Additional General Comments and Recommendations
Mike Van Auken

Section	COMMENTS AND RECOMMENDATIONS
3	Use of R^2 is confusing. Suggest using lower case "r" when referring to the sample correlation coefficient (Box, Hunter, Hunter, 1978, P. 61); or upper case R when referring to the regression coefficient of multiple determination (Draper and Smith, 1981, p. 90).
All	In most cases the reported results are just estimates, but are not described as such. For example, "The effect of mass reduction on heavier cars and CUVs and minivans are not statistically significant" on p. iii should say "The *estimated* effect of mass reduction on heavier cars and CUVs and minivans are not statistically significant." This distinction is important when comparing results based on different models and assumptions because the different models and assumptions do not change the effect itself, but rather the estimate of the effect. For example, the statement "The first sensitivity, in dark purple, includes the weight variables in the regression model but excludes the footprint variable; this model tests the effect of mass reduction while allowing footprint to vary with vehicle mass. This sensitivity increases the risk from a 100-lb mass reduction in cars (from 1.43% to 2.64% for lighter cars, and from 0.48% to 1.94% for heavier cars) and CUVs/minivans (from a 0.47% decrease in risk to a 0.52% increase in risk); however, there is no change in fatality risk in light-duty trucks" on page 15 is misleading. It would be better to state that "The first sensitivity, in dark purple, includes the weight variables in the regression model but excludes the footprint variable; this model tests the *estimated* effect of mass reduction while allowing footprint to vary with vehicle mass. ~~This sensitivity~~*Removing the footprint variable from the regression model* increases the *estimated* risk from a 100-lb mass reduction in cars (from 1.43% to 2.64% for lighter cars, and from 0.48% to 1.94% for heavier cars) and CUVs/minivans (from a 0.47% decrease in risk to a 0.52% increase in risk); however, ~~there is no~~*the* change in *the estimated* fatality risk in light-duty trucks *is very small and not statistically significant*." This also applies to table and figure captions. For example, "Table ES.1. Effect of mass and footprint reduction on fatality risk, under alternative regression model specifications" should say "Table ES.1. *Estimated effects* of mass and footprint reduction on fatality risk, under alternative regression model specifications." "Figure 3.3 Effect of reduction in mass or footprint on US fatality risk per VMT, by vehicle type: mass only, footprint only, and both" should say "Figure 3.3 *Estimated effects* of reduction in mass or footprint on US fatality risk per VMT, by vehicle type: mass only, footprint only, and both." Overall the word "effect" appears over 200 times in this report with the "estimated" or other qualifier. In some cases this may be appropriate and in other cases it is not appropriate. It is recommended that the author review each instance and revise as appropriate.

4	Figures 4.1 through 4.17 do not control for the effect of vehicle size (e.g. footprint), which has been shown to be correlated with vehicle weight (e.g., Figure 3.1), and therefore these figures may be misleading. It is strongly suggested that the horizontal axis label be changed to "Curb weight (lbs) and corresponding changes in size," and/or a note such as the following be added to each figure: "Note these results do not control for the effect of vehicle size on fatality risk. Therefore the horizontal axis represents changes to both vehicle weight and vehicle size."
4	The statistical significance of the linear trends in Figures 4.1 through 4.17 are not reported. It would be helpful if the confidence intervals or statistical significance of the linear trends were reported, either in addition to or instead of R^2. The confidence intervals for the estimated slopes should be added to the results in Tables 4.1 and 4.2.

Section	Page	COMMENTS AND RECOMMENDATIONS

Table 7. Additional Specific Comments and Recommendations
Mike Van Auken

Section	Page	COMMENTS AND RECOMMENDATIONS
Executive Summary, 7	iii, 65	2nd paragraph refers to "our analysis," however the results are the same as the NHTSA analysis. The author should clarify who or what "our analysis" refers to and how it relates to the NHTSA analysis. Perhaps the statement "LBNL was able to reproduce the NHTSA analysis, which finds that…" would be more appropriate.
Executive Summary	iv	Last bullet – suggest changing the statement that "Logistic regression does not allow a statistic" to "Logistic regression ***methods do not have*** a statistic."
Executive Summary, 4	iv, v, 22, 66	Suggest changing "variance in risk" to ***variation*** in risk" throughout.
Executive Summary, 7	viii, 69	The numerical results for the NHTSA preferred model in Tables ES.1 and 7.1 are slightly different than the results reported in the NHTSA report. For example 1.43%/0.48%/0.52%/-0.40%/-0.47% should be ***1.44***%/***0.47***%/0.52%/***-0.39***%/***-0.46***%
4	31-32	Figures 4.12 through 4.14 have the results for small and heavy-duty pickups combined, which is inconsistent with the results in Table 4.1
4	35	The R^2 values in Table 4.1 are different than the values in Figures 4.6, 4.8, 4.9, 4.11.
5.1	37	The subsection title should be "Alternative measures of exposure ***and outcome***" because fatal crashes and fatalities are measures of the crash outcome, not exposure.
5.2	39	It would be helpful to list the 18 manufacturer dummy variables in a table.
5.2	39	It is unclear why Lexus, Acura, and Infinity are treated as separate manufacturers, but Cadillac and Lincoln are not.
5.2	39	It is unclear why AM General is considered a Chrysler brand. The AM General Hummer was sold by GM beginning with the 2001 model year.
5.3	44	It would be helpful if the figures include error bars or shading to indicate the confidence intervals.
6.4	63	Table 6.3 – Suggest adding a note that the 72,316 total includes fatalities that are counted more than once in crashes involving more than one vehicle type.

<table>
<tr><td colspan="2" align="center">An Analysis of the Relationship between Casualty Risk Per Crash and Vehicle Mass and Footprint
for Model Year 2000-2007 Light-Duty Vehicles
(LBNL Phase 2 Report)</td></tr>
<tr><td align="center">1. ASSUMPTIONS</td><td align="center">COMMENTS</td></tr>
<tr><td>Please comment on the validity of any assumptions embedded in the LBNL assessment analysis and the independent casualty analysis that could affect the projected relationship between vehicle mass/footprint reductions and fatality/casualty risk. Examples might include assumptions regarding whether recent historical relationships between vehicle weight, size, and safety will continue into the future; potential future improvements in vehicle technology and design may result in compensatory safety benefits; and the annual baseline fatality distribution.</td><td>[Chen and Kockelman] [1] The Phase 2 report serves as a complimentary document to the Phase 1 report by isolating the effect of vehicle mass and footprint on crashworthiness. Whereas the Phase 1 report analyzes fatality risk per estimated VMT, the Phase 2 report analyzes casualty risk per crash. The parallel structure of the two reports makes it easy for the reader to compare the results of the two analyses.

[2] The binary logistic model employed in this analysis can only account for two injury outcome categories; here it is used to distinguish crashes resulting in serious injury or death from all other crash outcomes. Thus, the model does not account for the ordinal nature of injury severity and neglects the difference between a serious injury and a death.

[3] The report states that "a serious incapacitating injury can be just as traumatic to the victim and her family, and costly from an economic perspective, as a fatality." While serious injuries are very costly to society (and may have similar economic cost implications as deadly crashes), willingness to-pay estimates (which include pain and suffering) price the cost of a fatality at almost 20 times the cost of an incapacitating injury (NSC 2010). Thus, it is difficult to assess the economic cost of the estimates of increases in casualty risk per crash without distinguishing whether that outcome is a serious injury or a death. This limitation of the model should be addressed in the report.

[4] The logistic model also assumes error-term homoscedasticity and cannot account for increases and decreases in the variation of injury outcomes due to vehicle and driver type, for example. Such limitations of the model should be discussed.

[Farmer] The report concludes "that much of the detrimental effect of mass or footprint reduction on risk can be attributed to the tendency for mass or footprint reduction to increase crash frequency, rather than to reduce vehicle crashworthiness (risk once a crash has occurred)." However, the interpretation of casualties per crash as inversely proportional to crashworthiness ignores the possibility that injury severity also depends upon the circumstances of the crash. Casualties per crash must be divided into casualties per severe crash and severe crashes per crash, where a severe crash would be one involving more energy, e.g., high-speed or</td></tr>
</table>

rollover. It could be that weight reduction <u>increases</u> casualties per severe crash (i.e., reduces crashworthiness), but reduces the likelihood that a crash is severe.

[Van Auken] The basic assumptions, methodology, and data are primarily the same as in the Kahane (2011) report, but have been extended to include serious injuries as well as fatalities, and also address crash involvement (i.e., fatalities and serious injuries per accident, and also accidents per VMT).. These include the following:

1) The probability of a crash fatality or serious injury is proportional to the number of accidents (provided the crash conditions remain the same); and the probability of an accident is proportional to the vehicle miles travelled (VMT).
2) The logarithm of probabilities of fatality or serious injury per accident, and accidents per VMT for a given curb weight, footprint, and control variable values varies as a linear combination of the curb weight, footprint, and control variables within the domain of the data.
3) The logistic regression methods determine a maximum likelihood estimate of model coefficients.
4) It is assumed that the above relationships remain constant in the recent past (i.e., 2000-2007 model year vehicles in the 2002-2008 calendar years), present, and near future (i.e., 2017-2025 model year vehicles).

The first assumption that crash fatalities and serious injuries are proportional to the number of accidents provided the crash conditions remain the same seems self-evident (e.g., if two fatal crashes had exactly the same conditions, then the expected number of fatalities for the two crashes would be twice the value for just one of the crashes). The assumption that the number of accidents are proportional to VMT rather than the number of vehicle registration years (VRY) is also appropriate because accidents cannot occur if the vehicles are not driven on the road (i.e., VMT = 0). This assumption is qualified however because VMT is more difficult to measure than VRY and therefore may be less accurate. On the other hand the probability of a fatal crash or the number of fatalities in a crash may also depend on the vehicle occupancy.

The second and third assumptions are appropriate provided that it is recognized that it is essentially impossible with currently available knowledge and information to model all of the factors that could affect the probability of fatality in a crash, and that the objective of the analysis is to identify overall trends versus vehicle weight and footprint. In general the probability of fatality depends on other many other factors which have not been modeled (e.g., driver behavior factors, vehicle design factors, roadway design factors, EMT factors), and these unmodeled factors are assumed to be uncorrelated with vehicle weight and footprint, and/or are represented by the other control variables. The latter assumption might or might not be valid.

	The fourth assumption is perhaps the weakest because it assumes that future vehicles will have the same design characteristics as past vehicles, and that the characteristics of the vehicle population (e.g., collision partner weight, size, type) will also remain the same.
Please comment on any apparent unstated or implicit assumptions and related caveats or limitations.	**[Chen and Kockelman]** The role of driver behavior is briefly addressed in the report but not emphasized sufficiently. Casualty risk is a combination of driver, vehicle, and roadway characteristics. Whereas vehicle characteristics significantly influence crashworthiness, driver behavioral differences play a significant, if not primary, role in determining crash frequency. Socioeconomic data such driver household income, size, and education influence driver attitudes and driving environments. For example, Chen et al. (2010) found that crash risk increases for those living in socioeconomically disadvantaged areas (including households more likely to drive less expensive and older vehicles). Though such data is not typically available in state and national crash databases, the importance of these driver and environmental characteristics on crash rates (per mile driven) and casualty risk should be stressed in both reports. It is clearly very difficult to control for, but is a major caveat to the NHTSA (& now LBNL) results. We expect that crash severity could be probably be lower for many of the small cars and pickups if they were driven by those who tend to drive more expensive vehicles, under the same settings (e.g., daytime, urban freeway). Thus, statements like "a 100-lb reduction in the mass of lighter cars leads to a 1.84% increase in crash frequency" should be accompanied by an explanation of the possibility of the mass variable accounting/proxying for effects of lower income households owning smaller vehicles. **[Van Auken]** [1] The induced-exposure data set provided by NHTSA is based on the "non-culpable" vehicle in two-vehicle crashes. It is assumed that the dataset is a reprehensive sample of the driver and environmental exposure factors for vehicle use. However, since these cases include moving vehicles, some vehicle-driver-environmental conditions may be under or over represented in this data depending on how they affect the ability of a non-culpable vehicle to avoid a crash. Results in Ref (17) indicated that the estimated effect of weight and size reduction are sensitive to whether the induced-exposure data are based on the Kahane (2003) non-culpable vehicle definition of the Kahane (1997) stopped vehicle definition. Unfortunately it is not currently possible to test this sensitivity with the NHTSA-provided induced-exposure data. [2] It is also assumed that the accident data from the 13 or 16 states are representative of all US states. Figure 2.1 in Wenzel (2011b) provides a useful comparison of the distribution of fatalities in the US and 13 states by the nine different crash types.
ADDITIONAL COMMENTS:	

[Chen and Kockelman] National Safety Council (2010) Estimating the Costs of Unintentional Injuries. Available online at http://www.nsc.org/news_resources/injury_and_death_statistics/Pages/ EstimatingtheCostsofUnintentionalInjuries

Chen, H.Y., Ivers, R.Q., Mariniuk, A.L.C., Boufous, S., Senserrick, T., Woodward, M., Stevenson, M. and Norton R. Socioeconomic status and risk of car crash injury, independent of place of residence and driving exposure: Results from the DRIVE study. *Journal of Epidemiology and Community Health* 64(10), 2010, pp. 998-1003.

[Farmer] NHTSA's fatality analysis covered calendar years 2001-08, but the casualty analysis excludes 2008. Such exclusion is understandable given that 2008 data were at the time unavailable for a majority of the states (I think they are available now). However, 2008 was an unusual year and may have affected the size and weight effect estimates. The footnote on p. 6 of the Phase II report states that an analysis including the available 2008 data will be summarized in Appendix A. I don't see Appendix A. Is an analysis planned including 2008 data?

2. CONTROL AND DEPENDENT VARIABLES USED IN THE REGRESSION MODELS	COMMENTS
Please comment on the adequacy of control and dependent variables used in the assessment analysis and independent casualty analysis, and recommend any alternative control or dependent variables that are available for possible inclusion in the analysis. For example, what are the relative merits of the main dependent variables used, fatality risk per estimated VMT, and casualty risk per police-reported crash?	**[Chen and Kockelman]** [1] As alluded to above, a primary concern is that the NHTSA analysis (& thus the LBNL analyses) largely neglect the idea that vehicle type (make & model) is very much a proxy for driver type, and a vehicle's crash frequency may have very little to do with physical vehicle characteristics. It has a lot to do with the person behind the wheel, and gender and age simply aren't enough to control for such distinctions. Education, risk aversion, ability, wealth, etc., are important covariates. But existing data sets are quite limiting (though the MVOSS & FAR with 3-year driver violation history do offer some valuable insights, not discussed in these reports). In reality, small cars may be less crash prone than Kahane's & Wenzel's results suggest, because they are driven by lower-income, younger, less risk averse people driving in more crash prone settings (e.g., commercial strips rather than pricey residential suburbs). Such key caveats need thoughtful discussion. Four relevant papers on the topics of crash frequency and vehicle size-and-weight implications (by Knipling, Kweon & Kockelman, Wang and Kockelman, and Chen & Kockelman) have been sent to Tom Wenzel. These all include useful literature reviews for further connections to useful findings for citation in the reports, as time allows the contractor.
	[2] Independent variables such as vehicle mass and footprint may be accounting for effects of driver socioeconomic factors as discussed in the Assumptions section. Furthermore, vehicle option variables such as AWD and side curtain airbags may be reflecting the effects of driver environment (e.g., those living in areas with icy winters opting for AWD) and attitude (e.g., more risk-averse drivers opting for side curtain airbags) rather than the vehicle technology themselves. While extremely heavy and extremely large vehicles may have significantly different handling and braking characteristics which influence crash frequency and casualty risk, it is unlikely that given the same driver in the same environment, a small change in vehicle mass or footprint would influence the driver's crash proneness.
	[Farmer] One needs to restrict control variables to those that are available and reliable. A problem when combining state databases is that the states often are not consistent as to the variables coded and the definitions of those variables. This severely limits the list of possible control variables.
	[Van Auken] [1] Reducing the total number of fatalities and serious injuries is desirable from a societal viewpoint. This includes both subject vehicle occupant and collision partner (e.g., other vehicle occupant, pedestrian) fatalities and serious injuries.

	[2] VMT is a good measure of accident exposure provided that it can be accurately determined. [3] The number of fatalities and serious injuries per accident is a measure of vehicle crashworthiness (i.e., effect of a crash on the subject vehicle occupants) and crash compatibility (i.e., effect of a crash on the other vehicle occupants or vulnerable road users). Subject vehicle occupant fatalities and serious injuries per accident are a measure of the subject vehicle crashworthiness. Collision partner fatalities and serious injuries per accident are a measure of vehicle crash compatibility. [4] The number of accidents per VMT is a measure of the crash avoidance capabilities of a given vehicle.
What additional control variables, such as vehicle make or model, might be included in the regression models?	**[Chen and Kockelman]** [1] Vehicle height, a variable which may be more valuable than vehicle type for similarly structured vehicles such as sedans, wagons, CUVs, and minivans, would be a valuable control variable. In addition to a wider track, a lower center of gravity also increases vehicle stability, thereby reducing the risk of rollover. Relevant literature & findings exist, and should be cited. [2] Other variables which have been found in past studies to influence fatality risk such as seat belt use, roadway geometry and division type are not included in this study. [3] To account for driver characteristics that contribute to casualty risk, socioeconomic variables such as household income, education, household size, etc. would be valuable additions. Unfortunately, both state and national crash databases typically do not include such information (outside of MVOSS). Such issues should be flagged for readers. **[Farmer]** I think that already there are too many control variables in the regression models. Instead I would consider defining different classifications of crash types. Table 2.2 of the Phase II report shows that the distribution of crash types for casualty crashes is very different from that for fatal crashes.
Please comment on any caveats or limitations that these dependent variable or control variables entail with respect to use of the results as the basis for estimating the safety effect of mass reduction.	**[Chen and Kockelman]** Please see above comment (in Assumptions section) regarding driver behavior and environment. **[Farmer]** Model overspecification could be the reason for results that are non-intuitive, especially in the Phase II analyses of police-reported crashes. Control variables may be correlated with each other or with the size and weight variables. For example, Figure 2.9 of the Phase I report implies that torso side airbags increase fatality risk in CUVs.

ADDITIONAL COMMENTS:

[Chen and Kockelman] Table 2.1 has many indicator variables labeled as "C" for continuous variable (such as ABS, ESC, AWD, DRVMALE, etc). These C's should be removed.

[Farmer] The sensitivity results of Chapter 5 (Phase II) point out the extreme differences in results when changing the control variables. For example, including vehicle make changes the effect of a 100-lb reduction in heavier cars from -0.91% to +0.55% (see Fig 5.3).

[Van Auken] The underlying reasons for some of the estimated effects are unknown at this time, but presumably involve driver, vehicle, environment or accident factors than have not been controlled for in the Kahane (2011) and Wenzel (2011b) analyses. See, for example, Refs 17 and 23.

3. METHODOLOGY AND STATISTICS	COMMENTS
Please comment on the validity and applicability of the methodology LBNL used in assessing the NHTSA 2011 study and its analysis of the relationship between mass, footprint, and risks per police-reported crash.	**[Chen and Kockelman]** The Phase 2 report enhances the findings in the Phase 1 report by isolating the effect of vehicle mass and footprint on crashworthiness. Like the Phase 1 report, this analysis goes into a fair amount of detail and seeks to introduce additional analyses to better examine the relationship between mass, footprint, and casualty risks. However, due to a lack of control for very specific vehicle differences (which vary by make & sub-model), the exclusion of driver characteristics and crash setting details (which cannot always be controlled for, but are often correlated with vehicle type), the effects of downweighting vehicles and/or shifting vehicle styles and sizes may be overestimated. Simply changing the vehicle mass or footprint on a risky driver in a high-risk setting is unlikely to influence crash outcomes significantly.

[Farmer] Figure 2.11 of the Phase II report implies that NHTSA's fitting of a separate regression model for each of the 9 crash types was unnecessary, at least for the analysis of casualty risk per crash. I don't recall seeing a similar analysis for fatality risk per VMT. Is it possible to get essentially the same results as the NHTSA study using a single regression model?

[Van Auken] [1] The logistic regression methods seem to be appropriate. The confidence intervals are based on the logistic regression Wald Chi-Square statistic, which as Kahane (2003, 2011) has demonstrated does not include all sources of variation. However, these confidence intervals are useful because they do provide some indication of the uncertainty in the results.

[2] The two-stage results for the 13-state fatalities per crash and 13-state crashes per VMT, and the one-stage result for 13-state fatalities per VMT reported in Tables ES.1 and 6.1 and Figure 2.7 were computed using independent logistic regressions. The differences between the two-stage results and the one-stage results for fatalities per crash could have been eliminated by using the "simultaneous three-way" logistic regression method described in DRI (2003). This method imposes the constraint that the combined two-stage estimated and the one-stage estimated are equal. |
| Please review other statistical methods LBNL has used in the analysis, in addition to the logistic regression methodology. Examples include the alternative approaches used by LBNL to assess NHTSA interval estimation results, and LBNL's linear regression analysis of actual, predicted, and residual risk by | **[Chen and Kockelman]** In the alternative measures of exposure, the author examines the effect of vehicle manufacturer on fatality risk and treats the luxury models produced by Toyota, Honda, and Nissan as separate manufacturers. However, domestic luxury brands (such as Cadillac & Lincoln) are categorized with their nameplate manufacturers (GM and Ford), which appears inconsistent.

[Van Auken] [1] The correlations in Section 3 appear to be assessed using the Coefficient of Multiple Determination (R^2) based on a linear fit to the data (e.g., the correlation between footprint versus curb weight |

vehicle model.	in Figure 3.1 on p. 32). The linear regression model attributes the differences between the dependent variable (vertical axis) and the linear fit to the independent variable (horizontal axis) to random effects. If there is no preference as to the choice of independent and dependent variables (e.g., footprint versus curb weight, or curb weight versus footprint), then the linear trend and R^2 result would be different if the two variables were interchanged, and having two different yet equally valid results would be undesirable.
	[2] If the variation in the data can be attributed to both variables (e.g., footprint and curb weight), then it would be better to report the square of the sample correlation coefficient r^2, where r is computed according to Eqn (1). The trend lines in these correlation figures should not be computed using a linear regression. Instead, the trend line should pass through the sample means (i.e., (x, y)), and have a slope equal to the ratio of the sample standard deviations in the data (i.e., s_y / s_x). Therefore, the reported correlation results do not depend on the ordering of the data variables.
	Note this comments does not apply to linear trends indicated in Section 4, for which the Coefficient of Multiple Determination (R^2) seems appropriate.
	[3] The Coefficient of Multiple Determination (R^2) is frequently used in the Wenzel (2011b) report as an indicator of the statistical importance of a linear trend (e.g., R^2 values in Tables 4.1 and 4.2 were compared to 0.3). It would be better to report the standard error, confidence interval, and/or probability value as measures of the statistical significance of a linear trend.
Please comment on caveats or limitations of using non-significant regression estimates to project the safety impact of mass reduction.	**[Chen and Kockelman]** First, the t-statistics are not provided in the report which makes it difficult for the reader to assess statistical significance of specific regression estimates (except where noted by the author). Second, inclusion of a statistically insignificant variable can influence the estimates of coefficients associated with related variables. Nevertheless, in general, it is best to keep insignificant estimates if one has a strong defense for their role, since removing such variables (& thus their parameters) will shift the burden of response to a correlated covariate's parameter, thus biasing the latter. We generally keep key covariates in a model up to a pvalue of 0.20 or 0.25 or so, especially in relatively small data sets (e.g., n < 1,000). Covariates for which we have no strong basis can be removed for pvalues > 0.10.
	[Farmer] Making projections from non-significant regression estimates is proper so long as the resulting confidence intervals are constructed conservatively (to account for the accumulated imprecision). In that sense, I prefer NHTSA's jackknife approach to the standard errors produced by SAS (see p. 13 of Phase II).

| | **[Van Auken]** [1] Regression estimates are random numbers which have an unknown expected value and variance, and known sample value and standard error. If the sample value can be explained by a zero expected value and known standard error then the result is considered not statistically significantly different than zero and therefore the result is not considered to be statistically significant. However, If we can combine this estimate with other estimates then the unknown expected values and variances can also be combined using the same transformation, and the statistical significance of the combined result can be tested. Therefore, depending on the sample values and inter-correlation, the combined result may be statistically significant even if the individual estimates are not statistically significant.

For example, the results from each of the nine different crash types can be combined into an overall estimate and the standard error calculated assuming that the results for each crash type are independent of each other. Then the statistical significance of the combined effect can be determined.

[2] However, and Kahane (2011) points out there are two sources of uncertainty in the regression results. The first is the FARS based sampling error which is uncorrelated across crash types because they are based on different fatal cases (Kahane 2011, p. 77). The second is the state based induced-exposure sampling error which is correlated across crash types because they are based on the same induced-exposure cases. Therefore a confidence interval estimated using the "jackknife" method described by Kahane (2011) and accounting for correlation of these two error sources would be more accurate than a simple estimate based on the Wald Chi-Square statistic and assumed independence. |
| How might the LBNL methodology be strengthened to better represent future vehicle designs and reduce multi-collinearity between mass and footprint in the regression analysis? | **[Chen and Kockelman]** Including more vehicle-specific characteristics (such as vehicle height and engine size) reduces the analysis' dependence on vehicle type, since vehicle shapes and structures will continue to evolve. There is also correlation with context (e.g., pickups are driven in more rural locations, with greater hazards [like less lighting, higher speed, & few medians]). Disaggregate data are almost always best, to avoid ecological fallacies & such.

[Van Auken] [1] The effects of multi-collinearity can be mitigated by 1) obtaining more data, 2) pooling data from different crash type or vehicle types, or 3) reducing the number of regression variables. The first option would require more states (for serious injuries and police-reported accidents), calendar years and/or model years, which would involve added newer data as it becomes available (or using older data). The second option might be to recombine the CUVs and minivans with truck based vans and adding a control variable to compensate for the differences in the vehicles types. The third option might involve removing statistically insignificant control variables or removing control variables that would not be expected to have an effect on |

	the probability of crash or crash outcome (e.g., the side airbag variable is not included in pedestrian crashes because it is not expected to affect pedestrian fatality risk). The number of driver age control variables might be reduced from eight to three (as in the Kahane (1997) and DRI (2002-2005) studies). Finally, a linear curb weight model instead of a two-piece linear model may help to better elucidate the general trend. [2] The Variance Inflation Factor (VIF) has been suggested as a measure of multi-collinearity in the Kahane (2010 and 2011) reports, however this diagnostic metric does not account for differences in database size (i.e., Options 1 and 2 above). The Wenzel (2011b) report does not discuss the Variance Inflation Factor or report any VIF results.
ADDITIONAL COMMENTS:	

4. DATA SETS	COMMENTS
Please comment on the validity and applicability of the datasets used to project changes in risk resulting from reduction in vehicle mass. LBNL's casualty analysis used police-reported crash data from 16 states, while the 2011 NHTSA study used national fatality data, combined with a subset of non-culpable vehicles involved in two-vehicle crashes from police-reported crash data from 13 states.	**[Chen and Kockelman]** [1] The Phase 2 report uses an unusually extensive data set of police-reported crash data from 13 states which the author compares in detail to national data sets to illustrate similarities and differences. The author is very thorough in addressing the difference in definitions of "serious" and "incapacitating" injuries across different states and the effects of such inconsistency on the regression results. [2] Since casualty risk in the report accounts for serious injuries but not minor injuries, the author should note that police-reported injury levels may also be poor indicators of the actual or Modified Abbreviated Injury Scale (MAIS) level, following medical evaluation. Farmer (2003) found that 41% of injuries reported by U.S. police as incapacitating received MAIS ratings of "minor injury" by health care professionals using NASS Crashworthiness Data System (CDS). Thus, the results of the estimated casualty risk increases and decreases rely heavily on the assumption that police errors in reporting actual MAIS ratings are consistent across states. **[Farmer]** A major limitation of the Phase II analysis is a bias that may be due to the patterns of missing data. In particular, the vehicle identification number (VIN) is missing or mistyped for many crash records. High-severity crashes (especially fatal) are more likely to have detailed police investigation, so VINs (and other variables) in these crashes may be more complete. State crash files are therefore much less reliable than FARS. **[Van Auken]** [1] The induced-exposure data set provided by NHTSA is based on the non-culpable vehicles in two-vehicle crashes. See the comments in Table 1 on the limitations of this data. [2] The use of property damage accident data and cases with serious injury from the 13 states seems appropriate (with the noted qualification that the different states may have different accident reporting thresholds and injury reporting criteria). The concerns about the use of data for the 3 additional states (Georgia, Illinois, and New Mexico) have also been noted. [3] In addition, there are also many differences in the coding variables and values used by the different states, which tend to make the recoding to a common data set (either induced-exposure, police-reported accident, or severe injury) imprecise.

Please comment on any apparent, unstated, or implicit impact on estimated risks inherent in the two different approaches, and any related caveats or limitations. For example, what are the strengths and weaknesses of the two measures of vehicle exposure, miles of vehicle traveled scaled up from crash data from 13 states, and number of police-reported crashes?	**[Chen and Kockelman]** The Phase 1 analysis used non-culpable vehicles in two-vehicle crashes as a proxy for induced exposure crashes. In contrast, Phase 2 analysis uses data from vehicles involved in one-car crashes and the responsible vehicle in two-car crashes. The exclusion of the not-at-fault vehicle in two-car crashes may be distorting the distribution of crash frequency and casualty risk across different vehicle makes and models if crash-prone drivers are more likely to drive certain types of vehicles.
	[Farmer] The VMT weights provided by NHTSA were scaled to represent the entire US. Comments on pp. 9 and 18 of the Phase II report seem to acknowledge this deficiency, promising to adjust these to the 13 states in the future. Was any adjustment made, such as multiplying the weights by the proportion of annual US VMT accounted for by each of these states? The accuracy of the VMT weights is critical is we are to believe the somewhat surprising results concerning crashes per VMT.
	[Van Auken] [1] The number of fatal or serious injury cases tends to be much less than the number of induced-exposure cases (and the number of police-reported accidents). Therefore the effective numbers of degrees-of-freedom in the statistical estimates tend to be limited by the available number of fatal or serious injury cases. For example, it would not be possible to estimate the effects of two variables (e.g., just curb weight and footprint) if we had data for only one fatal of serious injury case even if we had thousands of induced-exposure cases. Therefore it is desirable to use data for the entire US in order to get a large sample of fatal cases for the logistic regressions. This then requires the available induced-exposure data (i.e., from 13 states) to be "scaled up" the US level using the method described in Kahane (2003 and 2011). The result is the best currently available estimate of vehicle exposure.
	[2] There may be some concerns about the accuracy of the vehicle miles-travelled data because the difficulty estimating the number of vehicle miles travelled at the make-model-year-state level of detail.

ADDITIONAL COMMENTS:

[Chen and Kockelman] Farmer, C.M. Reliability of police-reported information for determining crash and injury severity. *Traffic Injury Prevention* 4(1), 2003, pp.38-44.

[Farmer] Statements above Figure 2.7 in the Phase II report imply that the effects of weight reduction on crashes per VMT and fatalities per crash should add up to the effect on fatalities per VMT. This is not the case. For example, a 1.43% increase in crashes per VMT and a 0.76% decrease in fatalities per crash would imply a 2.16% decrease in fatalities per VMT (i.e., 1 - 0.9924/1.0143). The fact that the model on fatalities per VMT yields an estimated 1.08% increase should be a cause for concern. Either the VMT weights are inaccurate or the control variables have different effects on crash frequency and crashworthiness.

5. RECOMMENDATIONS	COMMENTS
Please comment on whether the LBNL assessment adequately addresses the NHTSA 2011 study and identifies the safety impact from mass reduction. Are the analytic methods and data used to assess the NHTSA study, and estimate the relationship between risk, mass, and footprint, appropriate? Is casualty risk per crash a legitimate measure of vehicle safety? What other methods or data could be used to better predict the effect of future vehicle designs on safety?	**[Chen and Kockelman]** As noted above, a primary concern remains a neglect of the notion that the type of car is very much a proxy for driver type, and a vehicle's crash avoidance may have very little to do with vehicle type. It has a lot to do with the person behind the wheel. Simply including gender and age variables cannot account for important covariates such as education, risk aversion, driving ability, wealth, etc. In reality, small cars may be less crash prone than Kahane's and Wenzel's results suggest, because they are driven by lower-income, younger, less risk averse people driving in more crash prone settings (e.g., commercial strips rather than pricey residential suburbs). Alas, it is very difficult to control for all these variables, since they are not readily available in data sets. In reality, the best the report authors can do with such data sets is to explain how all the other, relevant attributes may factor in (e.g., quality of driver and typical driving settings), and how they can generate biased estimation (sometimes in either direction). Discussion of relevant literature that looks more deeply at crash outcomes (e.g., Wang or Chen's papers, mentioned above, allowing for heteroscedasticity and individual vehicle attributes, non-driver outcomes, etc.) will also be useful.

[Farmer] Casualty risk per crash does not fully measure the effects of vehicle size and weight reductions on society. Casualty risk per VMT best coincides with the NHTSA analysis of fatalities per VMT. The breakdown of casualty risk per VMT into the crash frequency and crashworthiness components is of interest. However, the surprising results reported here make everything suspect. For example, the Phase II report concludes that "the detrimental effect of male drivers has to do with their higher tendency of getting into a serious crash rather than their sensitivity to injury once a serious crash has occurred" (p. 24). A few pages later it concludes that "male drivers have essentially no effect on crash frequency, but cause a statistically significant increase in fatality risk once a crash occurs" (p. 28).

[Van Auken] [1] The basic methodology described by Kahane (2011) seems appropriate; and the extension by Wenzel (2011b) are also appropriate. However some results using these methods and data are not well understood and need further diagnosis.

[2] The induced-exposure data set provided by NHTSA is based on the non-culpable vehicles in two-vehicle crashes. See the Table 1 comments on the limitations of this data.

[3] The state accident data files tend to have different database variable and coding definitions and criteria, which could confound the results. |
| Please comment on the overall adequacy of | **[Chen and Kockelman]** Overall, the study is an enriching complementary document to the Phase 1 assessment |

LBNL's assessment of the 2011 NHTSA report and its independent study of casualty risk for predicting the effect of vehicle mass or footprint reduction on safety. Provide any recommended improvements that might reasonably be adopted by the author to improve the analysis.	of the 2011 NHTSA report. The parallel structure of the two reports allows the reader to easily compare and contrast the various additional analyses which examine the relationship of vehicle mass and footprint reduction on safety. However, as stated previously in the comments here, driver preference for specific car types (including size and mass) is related to driver socioeconomic characteristics and driving behavior. As vehicle, driver, and roadway environment characteristics all contribute to fatality risk, the effects of physical vehicle changes such as mass or footprint reduction on safety should not be overstated when the other two types of characteristics are not sufficiently accounted for.
	[Farmer] Overall these are reasonably good studies. The Phase I report does a very good job of assessing the NHTSA report of fatality risk. However, the Phase II report should be more cautious in its conclusions concerning casualty risk. The casualty analysis is based solely on police-reported data from 13 states, which:
	1. May not be representative of the US as a whole.
	2. Are inconsistent in the information given and the way in which it is coded.
	3. Suffer from information that is missing, inaccurate, or unclear.
	[Van Auken] The Wenzel (2011b) report provides a valuable supplement to the analysis and results in the Kahane (2011) report.

ADDITIONAL COMMENTS:

[Farmer] Column G of Table 6.1 in the Phase II report provides the most appropriate comparison to results from the NHTSA report (Column A). For both fatalities and casualties per VMT, a 100-lb weight reduction is most harmful in lighter cars, less harmful in heavier cars and lighter light trucks, and slightly beneficial in heavier light trucks, minivans, and crossovers.

[Van Auken] See attached tables 6 and 7 below.

Table 6. Additional General Comments and Recommendations Mike Van Auken	
Section	**COMMENTS AND RECOMMENDATIONS**
4	Use of R^2 is confusing. Suggest using lower case "r" when referring to the sample correlation coefficient (Box, Hunter, Hunter, 1978, P. 61); or upper case R when referring to the regression coefficient of multiple determination (Draper and Smith, 1981, p. 90).
All	In most cases the reported results are just estimates, but are not described as such. The word "effect" appears several hundred times in this report with the "estimated" or other qualifier. In some cases this may be appropriate and in other cases it is not appropriate. It is recommended that the author review each instance and revise as appropriate.
All	"Crashworthiness" in most instances should be changed to "crashworthiness and crash compatibility" because the fatalities and/or serious injuries may either be in the subject vehicle (crashworthiness effect) or collision partner (crash compatibility effect).
4	The statistical significance of the linear trends in Figures 4.1 through 4.9 are not reported. It would be helpful if the confidence intervals or statistical significance of the linear trends were reported, either in addition to or instead of R^2. The confidence intervals for the estimated slopes should be added to the results in Tables 4.1 and 4.2.

Table 7. Additional Specific Comments and Recommendations Mike Van Auken		
Section	**Page**	**COMMENTS AND RECOMMENDATIONS**
Executive Summary, 4	iv, v, 22, 66	Suggest changing "variance in risk" to "*variation* in risk" throughout.
Executive Summary, 6	vii, 63	The statement "In conclusion, casualty risk per crash is not necessarily a better metric than fatality risk per VMT for evaluating the effect of mass or footprint reduction on risk; rather, it provides a different perspective in assessing the benefits or drawbacks of mass and footprint reduction on safety in vehicles. However, it does allow the separation of risk per VMT to be separated into its two components, crash frequency and risk per crash" suggests that the casualty risk per crash metric was needed in order to assess the crash frequency and risk per crash, which is incorrect. The DRI (2003-2012) methods have also estimated the effects of weight and size on crash frequency (A/E) and risk per crash (F/E) in terms of fatalities.
5.3	49	It would be helpful to list the 18 manufacturer dummy variables in a table.
5.3	48-49	It is unclear why Lexus, Acura, and Infinity are treated as separate manufacturers, but Cadillac and Lincoln are not.
5.3	49	It is unclear why AM General is considered a Chrysler brand. The AM General Hummer was sold by GM beginning with the 2003 model year.

Review of "An Analysis of the Relationship between Casualty Risk per Crash and Vehicle Mass and Footprint for Model Year 2000-2007 Light-Duty Vehicles"

David L. Greene

January 10, 2012

Summary

The Phase I and Phase II analyses by Tom Wenzel of LBNL have been executed diligently and consistently in accord with the methods and data used in the original NHTSA analysis. The studies contain many valuable, new insights. The phase I study highlights the weakening relationship between vehicle mass and highway fatalities. This is not only seen in decreasing coefficient estimates but in the very large number of results that are not statistically significant. When regressions were done separately by footprint deciles, vehicle mass was statistically significantly positively related to fatalities only for light-duty trucks in rollovers, there were almost as many cases in which mass was negatively related to fatalities (9 vs. 13 out of 27) and there were more instances of statistically significant negative relationships than positive relationships. Given that so many tests are being jointly conducted, it is quite possible that when joint probabilities are considered, there is no significant relationship between mass and fatalities (more on joint probabilities later). Showing the weakness and inconsistency of these results is an important contribution.

Another meaningful contribution of the phase I study, and one that deserves more emphasis, is a logical inference from the following findings: 1) much of the variance in risk remains unexplained even by the most complete models, 2) control variables explain 1 to 2 orders of magnitude more of the variance than the variables of interest (mass and size), 3) when key control variables are removed or changed it strongly influences the coefficients of mass and size. These results have very important implications for the robustness of the results and the likelihood that some or all of the apparently statistically significant relationships are due to spurious correlations with omitted or imperfectly controlled factors. Noting that exposure measures are control variables with constrained coefficients, the following observation from the phase I study is especially perceptive.

> "Calculating risk as total fatalities per induced exposure crash, rather than per vehicle mile traveled, reverses the sign of mass reductions on risk in cars and the lighter light trucks, with mass reduction leading to a <u>reduction</u> in risk in all vehicle types."

Finally, the phase I report notes that if only the control variables are included in the regression and not size or mass, the resulting residuals from the regression are uncorrelated with size or mass. Given these findings (as well as those of phase II) the conclusions that,

> "The 2011 NHTSA study, and this report, conclude that the effect of mass reduction on US fatality risk is small."

should be revised with the following emendation, "… and probably non-existent."

Both studies, like the NHTSA analysis, have shortcomings in terms of interpreting the results and the language used to describe the results, and acknowledging the limitations of the data and methodology. The limitations are extensive. The interpretation of the results of the LBNL studies commits two important, related errors. The first is to attribute inferred coefficients of mass and size as representing only the effects of vehicle mass and size when, as the phase I and II study results indicate, there is a virtual certainty that aliasing effects are present due to a combination of omitted variables, errors in variables and correlations among variables. Given that estimated driver and environmental factors tend to have 1-2 orders of magnitude larger impacts on safety outcomes than vehicle factors, the almost certain presence of aliasing effects must be explicitly acknowledged as severely limiting the ability to draw inferences about the effects of vehicle attributes. Second, the language used in interpreting results fails to acknowledge that the analysis does not address the effects of down-weighting or down-sizing specific vehicles or vehicle designs, but instead relies on correlations between vehicle weight and size in existing vehicle designs. In existing vehicles, weight and size are correlated with each other and many other vehicle attributes (and driver and environmental attributes, as well). Thus, the study is not actually measuring the effects of down-weighting via the material substitution and design changes likely to occur as a consequence of fuel economy and emissions standards. An early example of the kind of misleading language referred to here can be found on page iv.

> "For example, a 100-lb reduction in the mass of lighter cars leads to a 1.84% increase in crash frequency (columns B), while mass reduction leads to a 0.76% decrease in the number of fatalities per crash (column C);"

This statement is misleading in that it implies causality rather than correlation, and it is additionally misleading in that it implies that the inference applies to removing weight from specific vehicles. Neither is correct. A better statement would be the following.

> "For example, vehicles in the lighter class that are 100 lbs. lighter are correlated with a 1.84% increase in crash frequency.…"

There are so many examples of this misleading language that it is not feasible to list them all. All should be corrected, however. Failure to correct them could lead to serious misinterpretation of the studies' findings.

Following in the footsteps of the seminal study by DRI, the NHTSA and LBNL studies contribute to the literature in three important ways: 1) the LBNL and NHTSA studies recognize that the societal safety perspective is the correct perspective to when assessing the impacts of fuel economy and emissions regulations, 2) they recognize that vehicle dimensions and vehicle mass may have separate and potentially different impacts on both the likelihood of a crash and the outcomes of the crash and, 3) the LBNL phase two analysis makes an additional contribution by attempting to disentangle factors affecting the likelihood of a crash and factors affecting the outcomes of a crash.

Speaking of the DRI study, I am puzzled about why there are no references cited in the phase I study and only a handful all by Kahane and Wenzel, in the phase II study. This is perhaps due to the scope of work defined for the two studies but there are highly relevant studies in the literature that could have been cited, those by DRI foremost among them.   Making use of the insights from these studies would have been helpful in interpreting the results of both phase I and phase II.

Lack of a Theory or Model of the Phenomenon

Both the NHTSA and LBNL studies lack a rigorous theory of the process by which down-weighting at constant size or down-sizing at constant mass affect societal safety either through crash avoidance or crashworthiness. This is not a trivial shortcoming because it affects the ability to formulate hypotheses and interpret results. Prior to the dissenting report on safety of the NRC 2002 CAFÉ report, the physics of elastic collisions between objects was typically cited as the underlying physical model.   That report showed how taking the societal perspective renders that model inappropriate. What remains appears to be far more complex, involving the quantity of kinetic energy, the ability of vehicle designs to absorb that energy so as to minimize maximum deceleration rates, stability, maneuverability, safety technologies, and more.

The consequences of the lack of a rigorous theory are that it is not known, a priori, what the signs of coefficients are expected to be, let alone what their quantitative relationships should be. Hypotheses must be formulated based on intuition and the interpretation of results is likewise ad hoc. One implication of this is that results that suggest that lower fatalities are associated with lower vehicle mass have equal standing, a priori, with results that indicate that higher fatalities are associated with lower vehicle mass, and similarly for vehicle size. There are no surprising or unsurprising results, in theory.

This also makes it difficult to develop a plan for statistically testing the model or theory and its implications. It would have been helpful to the reader to have been presented early on in the report with such a plan of analysis.

On the Virtual Certainty of Aliasing

The LBNL report typically attributes causal effects to correlations between mass or size and safety. In fact, most or all of the observed correlations are almost certainly affected by aliasing effects. There is ample evidence for this inference in the results presented in the LBNL phase II report.

The coefficients of mass and size change in important ways when different model formulations are estimated. Removing and adding control variables changes the magnitudes and sometimes the signs of the mass and size variables. This means that, at a minimum, the mass and size variables alias the effects of the omitted control variables. The question is whether the aliasing is eliminated entirely by the inclusion of the control variables available or whether some aliasing remains either because not all relevant and correlated control variables have been included or because the included control variables are imperfect measures of the factors they are intended to represent.

The latter seems highly likely for the following reasons. First, the overall explanatory power of the full models (including control variables), as measured by their R^2 is low. Most of the variance in casualties and fatalities remains unexplained. Second, at least some of the important included control variables are only crude approximations of the factors they are intended to represent. For example, dummy variables represent differences in state reporting practices, age and gender represent risky driving behavior differences among owners of different sizes and masses of vehicles, the presence or absence of a kind of safety equipment represents both its performance and use in a particular vehicle, and calendar year dummy variables represent unknown factors associated with the respective calendar years. Such practices are common and their use is appropriate. Third, the control variables generally account for 1-2 orders of magnitude more variance in the casualty and fatality variables than do vehicle weight and size. To recap, the amount of unexplained variability in the dependent variables is larger relative to the variance statistically explained by the most complete models. Control variables are correlated with size and mass, and they account for 1-2 orders of magnitude more variability in the dependent variables than the variables of interest, mass and size. Therefore, even small correlations of size and mass with omitted variables or with errors (imprecision) of the control variables could easily result in biased estimates for the effects of size and mass on the dependent variables.

Tom Wenzel is to be commended for providing the results that definitively demonstrate the three key points made above. The above is not a criticism of the analysis nor of the results, per se. It is a criticism of their interpretation. In light of the above, the results should be interpreted in light of the virtual certainty that many of the estimated coefficients are likely to be biased in ways that make their interpretation highly uncertain. The implication is that phrases such as "down-weighting or down-sizing caused" to "mass (or size) and unobserved correlated factors are associated with…"

On Joint Probabilities

The NHTSA and LBNL studies do not correctly interpret their results as joint statistical tests. When testing a hypothesis on, for example, 5 vehicle classes simultaneously, a result for one equation that might be statistically significant on its own may not be statistically significant as one of five related tests.

Statistical analyses comprised of multiple regressions too often overlook the fact that tests of statistical significance designed for individual regressions may not apply in the case of multiple regressions. That is the case here. NHTSA conducted 5 analyses to infer relationships between mass differences among vehicles holding footprint constant for 5 classes of cars. The results showed one relationship out of five was statistically significant. As table 1 illustrates, using a simple example, if one conducts 5 trials, each with a 0.05 probability of given result, there is a 22.6% probability of finding at least one such result in the five trials. Thus, the joint significance level of the overall result (1 statistically significant regression out of 5) is 0.226, rather than 0.05.

Table 1. Simplified Illustration of the Joint Probability of Inferences in Multiple Regressions

# Significant Regressions	Combinations	0.95	0.05	Joint Probability
0	1	0.773780938	1	77.37809%
1	5	0.81450625	0.05	20.36266%
2	10	0.857375	0.0025	2.14344%
3	10	0.9025	0.000125	0.11281%
4	5	0.95	0.00000625	0.00297%
5	1	1	3.125E-07	0.00003%
				100.0%
				22.6%

So there is between a 1:4 and a 1:5 chance of getting one statistically significant result by pure chance. In fact, the actual significance level of the results is more complicated to calculate, and probably a bit smaller than 0.226. Thus, it is very appropriate for Dr. Kahane to add the qualifier "if any" to his conclusions about the relationship between the societal highway fatalities and mass reduction, holding footprint constant. Had appropriate tests of joint statistical significance been used to evaluate the results in the NHTSA and LBNL studies, the significance levels very likely would not meet accepted criteria for statistical significance. This could change the conclusions of the studies from the inference that mass is correlated with fatalities or casualties in some case but not others to the lack of statistically significant evidence that mass is correlated with fatalities or injuries on the highway. This is an important difference.

Page-by-Page Comments

I will make page by page comments on the phase II study only, since that contains the overwhelming share of original contributions and the key findings of the phase I study are recapitulated there.

p. iii   Paragraph 3. This would be a very good place to acknowledge the importance of driver behavior and environment on crash avoidance especially.

p. iv   Para. 3. This would be a good place to discuss probability inference in joint tests.

Para 4. The statement about a 100-lb reduction in the mass, etc., is a good example of misleading language.

p. v   Para. 1. Again, it is misleading to say that mass reduction increases crash frequency, for reasons stated above.

Para. 2. It is more accurate to describe the association of lower vehicle mass with casualty risk than the "effect of mass reduction on..." casualty risk.
Para. 3. Would benefit greatly from joint probability inferences.

Para. 4. As noted above, this shows how much more important the control variables are than the variables of interest.

Para 5. Again, these are correlations not necessarily effects.

p. vi   Para. 1. Again, mass reduction is misleading terminology and you do not know if it increases casualty risk or not, you know only a correlation. Why is this so important? It is the virtual certainty of spurious correlations, or aliasing, as noted above.

Para. 4. (1st bullet) This is clear evidence of aliasing. Take variables out of the regression and the coefficients of interest change in important ways. Are there no important factors still missing? Are the variables included perfect measures of the factors of interest? Of course not. Thus, there must be remaining aliasing. How bad is it? We don't know.

The third bullet shows the same effect with a different set of variables.

p. vii   Para. 4. No, your analysis does not indicate "…that much of the detrimental effect of mass or footprint reduction on risk can be attributed to the tendency for mass or footprint reduction to increase crash frequency." Again, you have correlation, not causation and you have good reason to believe that what you are seeing is affected by spurious correlations.

Para. 5. The "effect" is small, 1-2 orders of magnitude smaller than correlations with other control variables, and IS strongly affected by which variables are in the equation, as stated on the previous page, and there is a great deal of unexplained variance. Please reconsider the meaning of these results in light of the comments above.

Finally, as the last paragraph of the ES implies, it would be far better not to speak in terms of "reducing" mass or size. That is not what is happening in your data set.

p. 1   Para. 4. Risk per VMT includes the effects of how well vehicles are driven as well as how well they can be driven. I think there is no chance that you have fully accounted for how well vehicles are driven.

p. 2   Para. 2. Exposure measures are explanatory variables whose elasticity is constrained to 1. That is, it is assumed that an increase in vehicle use of 1 vehicle mile produces a 1 unit increase in the chance of a fatality (or casualty as the case may be). This is actually a maintained hypothesis. If this hypothesis is incorrect, it can bias the other coefficients in the equation. Thus, the change from fatalities/VMT to fatalities/registration-year, to fatalities per crash not only changes the meaning of the analysis, it also may bias coefficients in the event that the true relationship between fatalities and VMT is not an elasticity of 1.

p. 3   Line 1. Please acknowledge that your "accounting for differences in driver characteristics, crash locations, and other vehicle attributes" is incomplete and that this could affect your inferences about size and mass.

p. 3   Para. 2. NHTSA's use of "non-culpable" vehicles involved in two vehicle crashes as an exposure measure raises its own issues. How non-culpable was the non-culpable vehicle. Often this is a matter of degree, rather than black or white. Driver behavior may also be involved. It seems to

me this is just another potential measure of exposure that may or may not be better than any other measure and may introduce new sources of bias in the analysis.

p. 3   Para. 4. Induced exposure needs to be defined. What is it intended to mean? This needs to be explicit.

            Also, I am startled that there are no equations in these reports. Equations can provide an unequivocal explanation of the assumed relationships that cannot be adequately accomplished by words, in many cases. Why no equations?

p. 7   Para 3. CUVs and minivans are involved in fewer crashes with stationary objects than cars. Why? Is it the drivers, the vehicles, or the passengers? How well can you control for such differences? Not well. What does this mean for your analysis?

p. 9   Para 1. Here an equation showing how the weighting was done would be very helpful.

p. 9   Bullet 1. Excluding these vehicle types implies that the control variables in the model are not adequate to account for whatever makes these vehicle types different from the vehicles included in the analysis. First, this is an admission that the model is not adequate to explain the fatalities associated with these vehicles. Second, it is an admission that if they were included the coefficients on the variables of interest would likely be biased by spurious correlations. Clearly, it would not even be sufficient to include the vehicles along with a control variable (e.g., X = 1 if vehicle is a police car, 0 otherwise). This is yet another indication that the model suffers generally from omitted variables, errors in variables and correlation among right-hand side variables.

p. 10   Line 3. Sentence does not make sense. Please correct.

p. 12   Sect. 2.3. Please provide an equation.

p. 13   Were the confidence intervals calculated using e^x-1? Please state explicitly or, better, show an equation.

            Para. 2. Here another instance where you say "mass reduction increases societal fatality risk" but you really are not entitled to say that. It is misleading. Also, the NHTSA CI's are larger, as they should be in a joint test.

            Para. 3. These results require an underlying theory for interpretation. The lack of one makes it seem like there is just no consistency in the results.

            Para. 4. The fact that the results for fatalities per crash differ substantially from fatality per VMT may be very important. Taken at face value, it would imply that any negative effect of reduced mass is due to its effect on crash avoidance (crash probability) rather than crashworthiness. This is where the lack of a theory is most troubling. Why would that be? Are lighter vehicles less easily controlled, etc.? Or, as seems much more likely, is there a spurious correlation between mass and other omitted or imperfectly measured factors (including driver behavior) that lead to an increased probability of a crash? Consider, for example, driver age. Driver age is related to crash involvement. Driver age is a control variable. But are all young drivers the

same? Is it possible that young drivers more prone to risky behavior tend to drive lighter vehicles? If so, this could partly explain the result observed. Of course, this is just speculative, but the point is that correlations with imperfectly measured and omitted factors are highly likely to be present in the data and, if there, could easily affect the statistical inferences.

p. 14   Here we see that changing the exposure measure influence the effect of mass and size on fatalities and casualties, which is more evidence that spurious correlations are likely to be biasing estimated coefficients for mass and size.

The bar graphs with confidence intervals are well done and convey a great deal of information effectively. The patterns of magnitude and statistical significance are difficult to interpret, partly because there is no explicit theory of what should happen and partly, perhaps, because the relationships are actually not real.

p. 16   Para. 1. Reduction in the mass of lighter cars increases crash frequency but reduces fatalities per crash. This is contradictory to the previously maintained theory that mass protects due to the physics of velocity changes in elastic collisions. Indeed, there is no theoretical explanation for these results, only speculation.

p. 18   Para. 1. Here is a good example of such casual speculation.

p. 19   Para. 1. Developing VMT weights for the 13 states is a good idea, given the effect of exposure measures on inferences. Still the results would not be definitive.

p. 20   Para. 2. Mass reduction leads to a large reduction in risk only in crashes with objects for heavier cars?   There is only one type of crash in which the simple physics of collisions leads to an unambiguous benefit for increased mass, and that is collisions with moveable or breakable objects. This finding contradicts even that.

Para. 3. More speculation, this time about rollovers.

p. 21   Para. 3. "Curiously,…" Curiouser and curiouser.

Figure 2.15 printed without labels. Could be my computer but the other graphs were fine.

p. 24   Para. 2. This is probably a very important finding that needs further investigation and explanation. As figure 2.16 illustrates well, the correlations with mass and size are orders of magnitude smaller than the correlations with driver and environmental factors. This is why even small correlations with omitted or imperfectly measured control factors could be, are even likely to be, predominantly responsible for the estimated coefficients of mass and size.

p. 25   Para. 2. The results for minivans discussed here could be due to what is going on inside the vehicle as much or more than the vehicle itself. How could these results be explained in terms of the vehicle itself. Figure 2.17 shows this again. The effects of calendar year dummies, which can only be considered rough approximations to unknown and various time-related factors, have much large effects than size or mass. Again in figure 2.20.

p. 28   Para. 1. "Surprisingly,…" How can side airbags, which deploy only in a crash, reduce crash frequency but not fatality risk in a crash? Only if the real effect is a reflection of who buys a CUV/minivan with side airbags and how and where they drive. There are more surprising inferences in paragraph 2 about male and female drivers. Surprising relative to what theory?

p. 32   The problem here is not numerical multicollinearity (numerical difficulties inverting the cross-product matrix) but the more complex problem of correlations among right-hand side variables, omitted variables, errors in variables, and correlations of included variables with omitted and imperfectly measured variables. This leads to biased estimates.

p. 36   Table 3.1 cries out for inference based on joint probabilities. What is the probability of observing "success" in at least 3 out of 27 trials when the true probability of success is only 0.05. See discussion above. The probability is certainly much higher than 0.05 and probably closer to 0.5. The implication is that, taken together, these results do not show any statistically significant relationship between mass or size and risk per crash. If there were a rigorous underlying theory, the interpretation might be different (patterns of significance could matter) but there is none. Again, good graphs on succeeding pages.

p. 41   The statistical significance of such a relationship should be the same whether bins are used or not. Is it?

p. 42   Para. 3. R-squared is not the correct measure of statistical significance. Is the coefficient of weight significantly different from zero?

p. 43   Para. 2. This is perhaps the key finding of the phase I and II analyses. Control variables explain 1-2 orders of magnitude more variance than size or mass. Still, most of the variance remains unexplained and is uncorrelated with mass or size. It is very likely there is nothing going on here.

p. 47   Para. 2. More evidence of correlation of mass and size with control variables and how changing definitions or excluding control variables results in important changes in the coefficient of mass and size. Such results are considered unstable.

p. 49   Para. 1. The rationale for the grouping of manufacturers is not obvious. Can you explain it?

         Para. 2. Yet more evidence for the instability of the model and likelihood that variables still missing from the model, plus errors in measuring the included control variables are likely biasing the inferences. The results described in this paragraph do not make sense to me. How can they be explained other than random results?

p. 50   Para. 1. More casual interpretation of results. OK, maybe, maybe not. Same for paragraph 2. The economy faltered in 2008 but the big negative effect was in 2007. The downward trend started in 2004. This correlates neither with vehicle mass changes over time nor economic growth as measured by real GDP. Idle speculation.

p. 51   Para. 1. Good discussion of the gratuitous speculation by NHTSA about the meaning of the observed correlations. This is more a Rorschach test than statistical analysis.

p. 51   Para. 2. "We have no explanation for why…" More of this kind of honest appraisal is needed in studies like these.

p. 52   More results for which there is no explanation.

p. 53   Why the interaction between calendar year dummy variables and safety equipment? The presence of the safety equipment on a particular vehicle is established. What has calendar year (not model year) to do with it? Again, one suspects spurious correlations.

p. 55   One needs to think carefully about the reasons why vehicles would be excluded. It does not appear that NHTSA did that. First line of first paragraph "used" rather than "sused".

p. 57   Well reasoned. It is interesting that NHTSA resisted including footprint or size in previous analyses on the grounds of correlation with mass. These results show that assertion was groundless.

p. 59   Para. 3. Rather than say risk per VMT accounts for two effects, it is better to say it includes or comprises two effects. But this statement also ignores the important influences of drivers and environment and their potential correlations with other factors. Yes, it includes how well a vehicle can be driven, but more importantly it includes how well a vehicle IS driven. That is in there too and is very likely to be correlated with make, model, and other vehicle attributes.

p. 60   Para. 3. Here again, the conclusions are misstated. It is not a genuine "reduction" in mass, but an association with the mass of vehicles. And how does it "lead" to and increase in crash frequency? What is the theory or model that predicts this? Driver and environment are very likely mixed up in these results to an unknown but likely substantial degree. So what can we really conclude? Not this.

As I read the conclusions and inferences I find myself asking, why?, why?, repeatedly without any sound explanations. Page 61, paragraph 3 contains more "surprising" results. Surprising because they are contrary to theory? Surprising because they are contrary to intuition? Surprising because they are random? To what can we attribute so many "surprising" results, and how many must there be before one concludes that the analysis is not revealing what we had hoped it would.

Para. 4. Again, this cries out for joint statistical inference. Three statistically significant results out of 27 is probably nothing statistically significant at all.

p. 62   Para. 1. What this shows, again, is that the coefficients of mass and size are strongly influenced by which control variables are included in the model and how they are defined. These results and their implications need explanation. The bottom line is that the effects of mass and size are likely to be (after the necessary joint significance calculations are done) not statistically significant, not consistent, and not robust.

p. 63   How do mass and footprint reduction (again, it's not really reduction in the sense of designing lighter vehicles to increase fuel economy or reduce GHG emissions, the issue at hand) increase crash frequency. What is the theory? I don't find a theory in either the NHTSA report or the phase I and phase II studies. Absent a theory, these results seem sufficiently unstable and

inconsistent to be highly questionable as evidence of any relationship between mass or size and crashworthiness or crash avoidance. I think joint estimation of significance levels would provide additional support for this view.

4. References

Tom Wenzel. *Assessment of NHTSA's Report "Relationships Between Fatality Risk, Mass, and Footprint in Model Year 2000-2007 Passenger Cars and LTVs"*. 2011.

Tom Wenzel. *Analysis of the Relationship between Casualty Risk Per Crash and Vehicle Mass and Footprint for Model Year 2000-2007 Light-Duty Vehicles*. 2011.

Appendix A: Resumes of Peer Reviewers

T. Donna Chen, P.E.

1206 Arcadia Ave, Austin, TX 78757   •   (501) 908-9620   •   chen.donna@gmail.com

EDUCATION

University of Texas, Austin, TX

Doctorate of Philosophy, Civil Engineering – Transportation (3.9 GPR)   *August 2010 – Present*
- Thrust Fellow
- Advanced Institute for Transportation Infrastructure Engineering and Management Fellow

University of Texas-Arlington, Arlington, TX

Master of Engineering, Civil Engineering – Transportation (3.9 GPR)   *December 2008*

Texas A&M University, College Station, TX

Bachelor of Science, Civil Engineering (3.8 GPR)   *May 2005*
- National Merit Scholar
- Recipient of the Texas A&M President's Endowed and Merit Plus scholarships
- Recipient of the Dallas/Fort Worth Women's Transportation Seminar scholarship
- Recipient of Society of Women Engineers A&M Student Chapter scholarship
- Alumni of Texas A&M Engineering Scholars Program

WORK EXPERIENCE

Center for Transportation Research, Austin, TX

Graduate Research Assistant   *August 2010 – Present*
- Developing a reference on transportation economics for TxDOT

Richland College, Dallas, TX

Instructor, College of Business, Engineering, and Technology   *January 2010 – August 2010*
- Teaching statics and dynamics as part of core engineering curriculum.

HNTB Corp., Plano, TX

Transportation Planning Engineer II   *December 2008 - November 2009*

Transportation Planning Engineer I   *June 2005 - December 2008*

Performed geometric design, cost estimation, and operational analysis for projects such as:

New Route Studies
- Lavon Lake Bridge Study, Collin County, Texas – collected constraints data; developed and evaluated alignment alternatives for ten mile bridge route; assisted in developing GIS exhibits; answered citizen questions at public meeting.
- Border Highway-East, El Paso, Texas – developed four alignment and two interchange alternatives; evaluated alternatives and selected preferred alternative based on constraints; assisted in development of technical documents.
- Collin County Outer Loop – developed alternative corridors and alignments; evaluated alternatives and selected locally preferred alignment for 40 mile study from US 75 to Rockwall County Line.

Traffic Analyses
- I-95/395 HOT Lanes, Arlington, Virginia – performed LOS analysis for revised managed lane access for 24 mile corridor.
- Beltway 8/SH 249 Connector, Houston, Texas – modeled two proposed ramping alternatives in CORSIM.
- I-30, Greenville, Texas – performed LOS analysis using HCS and prepared Interstate Access Justification report.
- I-35, Lorena, Texas – modeled traffic along corridor using CORSIM and prepared Interstate Access Justification report.

Design Schematics
- Dallas North Tollway Ramping Modifications, Plano, Texas – designed various ramping improvements along DNT between PGBT and SH 121 to improve congestion and accommodate for an additional mainlane in each direction.
- Dallas North Tollway/SH 121 Interchange, Plano, Texas – designed four ramping options for the retrofit of a fully directional interchange between existing mainlanes; conducted LOS analysis for toll and non-toll options along SH 121.
- I-10, El Paso, Texas – developed two alternatives for the addition of two managed lanes to the existing four-lane facility.
- US 67, Cedar Hill and Midlothian, Texas – produced schematic plans for the addition of center reversible HOV lane.

Participated in leadership development activities such as:
- Leadership HNTB – completed professional development/project management training program for younger staff.
- H4 Community Outreach Program – worked with Leadership HNTB classmates to implement a community outreach program which yielded over 900 hours of combined service and over $13,000 in charitable donations in one year.
- HNTB College Recruitment Team – represented HNTB as college recruiter at Texas A&M Engineering Career Fair and Society of Women Engineers National Conference Career Fair.

Texas Transportation Institute, College Station, TX
Undergraduate Research Assistant   *May 2003 - May 2004*
- Analyzed travel behavior data for the Houston QuickRide (HOT lane) project using Excel and SPSS
- Received the Dwight Look College of Engineering Undergraduate Summer Research Grant and published paper in Southwest Regional University Transportation Center Compendium

PROFESSIONAL MEMBERSHIPS/HONOR SOCIETIES

Intelligent Transportation Society of America
- UT Student Chapter President   *November 2010 – Present*
  - Coordinated new student and intern peer support activities
  - Organized prospective student recruitment activities
  - Served as UT liaison for the 2011 Southwest Region University Transportation Center Student Symposium
  - Represented ITS as Explore UT day volunteer

American Society of Civil Engineers
- Dallas Branch History and Heritage Committee Chair   *2008 – 2010*

Society of Women Engineers
- Dallas Section Career Night Chair   *2006 – 2008*

Tau Beta Pi Member   *2004 – Present*
Chi Epsilon Member   *2004 – Present*

CHARLES M. FARMER
1005 North Glebe Road, Suite 800
Arlington, Virginia 22201
cfarmer@iihs.org
(703)247-1590

EDUCATION

December, 1986	Iowa State University, Ames, Iowa Doctor of Philosophy Major: Statistics
August, 1981	Old Dominion University, Norfolk, Virginia Master of Science Major: Applied Mathematics Concentration: Statistics
May, 1979	St. John Fisher College, Rochester, New York Bachelor of Arts Major: Mathematics Minor: Political Science

PROFESSIONAL EXPERIENCE

2004-Present	*Director, Statistical Services*
1997-2004	*Senior Statistician*
1994-1997	*Statistician* Insurance Institute for Highway Safety Arlington, Virginia
1987-1994	*Assistant Professor* Department of Mathematics James Madison University Harrisonburg, Virginia
1986-1987	*Visiting Assistant Professor* Department of Statistics University of Kentucky Lexington, Kentucky

PROFESSIONAL AFFILIATIONS

American Statistical Association

RECENT PUBLICATIONS

June, 2011   Farmer, C.M. Methods for estimating driver death rates by vehicle make and series. Arlington, VA: Insurance Institute for Highway Safety.

October, 2010   Farmer, C.M., Braitman, K.A., and Lund, A.K. Cell phone use while driving and attributable crash risk. *Traffic Injury Prevention* 11:466-470.

August, 2010   McCartt, A.T., Farmer, C.M., and Jenness, J.W. Perceptions and experiences of participants in a study of in-vehicle monitoring of teenage drivers. *Traffic Injury Prevention* 11:361-370.

May, 2010   Farmer, C.M. Effects of electronic stability control on fatal crash risk. Arlington, VA: Insurance Institute for Highway Safety.

April, 2010   McCartt, A.T., Hellinga, L.A., Strouse, L.M., and Farmer, C.M. Long-term effects of hand-held cell phone laws on driver hand-held cell phone use. *Traffic Injury Prevention* 11:133-141.

February, 2010   Farmer, C.M. and Wells, J.K. Effect of enhanced seat belt reminders on driver fatality risk. *Journal of Safety Research* 41:53-57.

February, 2010   Farmer, C.M., Kirley, B.B., and McCartt, A.T. In-vehicle monitoring and the driving behavior of teenagers. *Journal of Safety Research* 41:39-45.

DAVID L. GREENE

Home: *212 Way Station Trail • Farragut, Tennessee 37922 • (865) 966-0891*
Work: *Oak Ridge National Laboratory • National Transportation Research Center • 2360 Cherahala Boulevard •*
Knoxville, Tennessee 37932 • (865) 946-1310

PERSONAL

Born: November 18, 1949, New York, New York
Married, two children

EDUCATION

THE JOHNS HOPKINS UNIVERSITY
Ph.D., Geography and Environmental Engineering, 1973–78

UNIVERSITY OF OREGON
M.A., 1972–73

COLUMBIA UNIVERSITY
B.A., 1967–71

EMPLOYMENT

UNIVERSITY OF TENNESSEE, KNOXVILLE   2010–PRESENT

1/2010–Present   *Senior Fellow, Howard H. Baker, Jr. Center for Public Policy*
Research Professor, Department of Economics

INSTITUTE FOR TRANSPORTATION STUDIES, UNIVERSITY OF CALIFORNIA, DAVIS   2008–2009

9/2008–6/2009   *Visiting Research Faculty*

OAK RIDGE NATIONAL LABORATORY (ORNL)   1977–PRESENT

1999–Present   *Corporate Fellow, Oak Ridge National Laboratory*
1989–1999   *Senior Research Staff Member II and Manager of Energy Policy Research*
Programs, Center for Transportation Analysis
1988–1989   *Senior Research Analyst, Office of Policy Integration, U.S. Department of*
Energy (On assignment from ORNL)
1987–1988   *Head, Transportation Research Section*
1984–1987   *Senior Research Staff Member I*
1982–1984   *Research Staff Member*
1980–1982   *Leader, Transportation Energy Group*
1977–1980   *Research Associate*

AWARDS AND HONORS

2011 DOE Vehicle Technologies Program R&D Award, U.S. Department of Energy (with Z. Lin)
2011 Edward L. Ullman Award, Association of American Geographers
2009 Alliance to Save Energy, Energy Efficiency Hall of Fame
2008 Science Communicator Award, UT-Battelle
Recognition by the Intergovernmental Panel on Climate Change for Contributions to the Award of the 2008
Nobel Peace Prize to the IPCC

2007 Department of Energy Hydrogen Program R&D Award (with P.N. Leiby)
Barry D. McNutt Award for Excellence in Automotive Policy Analysis, Society of Automotive Engineers, 2007
Member Emeritus, Transportation Research Board Committee on Alternative Fuels, 2006
Barry D. McNutt Award for best paper of 2004, Energy Committee, Transportation Research Board
Lifetime National Associate of the National Academies, 2002
UT-Battelle Award for Excellence in Science and Technology, 2001
Oak Ridge National Laboratory Significant Event Award, 2001
Corporate Fellow of Oak Ridge National Laboratory, 1999
Outstanding Paper of 1999, *The Energy Journal*, International Association for Energy Economics
Lockheed-Martin Significant Event Award, 1999
Member Emeritus, Transportation Research Board Committee on Transportation Energy, 1998
Lockheed-Martin Significant Event Award, 1996
Distinguished Service Certificate, Transportation Research Board, 1993
ORNL Special Achievement Award, 1991
Distinguished Service Certificate, Transportation Research Board, 1989
Energy Specialty Group Paper Award, Association of American Geographers, 1986
ORNL Special Recognition Award, Oak Ridge National Laboratory, 1986
Technical Achievement Award, Martin Marietta Energy Systems, 1985
Pyke Johnson Award, Transportation Research Board, 1984

PROFESSIONAL ACTIVITIES

- Board of Directors, American Council for an Energy Efficiency Economy
- Board of Advisors, Institute for Transportation Studies, University of California, Davis
- Editorial Advisory Board, *Transportation Research Part D*, 1996–2006
- Editorial Board Member, *Energy Policy*, 2001–present
- Editorial Board Member, *Journal of Transportation and Statistics*, 2001–2006
- Editorial Board Member, *Transportation Quarterly*, 1999–2005
- Editor-in-Chief, *Journal of Transportation and Statistics*, 1997–2000
- Editorial Board Member, Macmillan Encyclopedia of Energy, 1998–2001
- Editorial Advisory Board, *Transportation Research A*, 1986–1997
- National Research Council
    - Transportation Research Board Standing Committees:
        - Committee on Transportation and Sustainability, Member, 2006–present
        - Committee on Energy, A1F01, Chairman 1983–1986, 1986–1990; Member, 1993–1998; Member Emeritus, 1999–present
        - Subcommittee on Forecasting Transportation Energy Demand, A1F01(2), Chairman, 1982–1983
        - Section F, Energy and Environmental Concerns, Chairman, 1990–1992
        - Committee on Alternative Fuels, A1F05, Member, 1993–2006, Member Emeritus, 2006–present
        - Task Force on Freight Transportation Data, A1B51, Secretary, 1989–1996
        - Committee on Transportation Information Systems and Data Requirements, Member, 1983–1986, 1986–1989
    - Ad Hoc Committees:
        - Committee on Transitions to Alternative Vehicles and Fuels, 2011-2012
        - Special Task Force on Energy and Climate Change, 2008–2009
        - Committee on the Assessment of Fuel Economy Technologies for Light-Duty Vehicles, 2007–2010
        - Planning Group for Workshop on Issues Related to Peaking of Global Oil Production, 2005
        - Committee on State Practices in Setting Mobile Source Emissions Standards, 2004–2006
        - Chair, Committee for the Symposium on Introducing Sustainability into Surface Transportation Planning, 2003–2004
        - Panel on Combating Global Warming through Sustainable Surface Transportation Policy, TCRP Project Panel H-21A, 2002–2005

<div style="padding-left: 4em;">

Committee on Effectiveness and Impacts of Corporate Average Fuel Economy (CAFE) Standards, 2001

Committee for the Study of the Impacts of Highway Capacity Improvements on Air Quality and Energy Consumption, 1993–1994

Committee on Fuel Economy of Automobiles and Light Trucks, Energy Engineering Board, Commission on Engineering and Technical Systems, 1991–1992

Committee for the Study of High-Speed Surface Transportation in the United States, 1990

Planning Group on Strategic Issues in Domestic Freight Transportation, 1990

Steering Committee for Conference on Transportation, Urban Form, and the Environment, 1990

National Cooperative Highway Research Program, Panel on "Evaluating Alternative Methods of Highway Finance," 1991–1992

</div>

- Intergovernmental Panel on Climate Change
    - Lead Author, Working Group III, Fourth Assessment Report, 2007
    - Lead Author, Working Group III, Third Assessment, 2001
    - Lead Author, Working Group III, Aviation and the Global Atmosphere, 1999
    - Principal Lead Author, Working Group II, Second Assessment Report, 1995
- Association of American Geographers
    - Board of Directors, Transportation Specialty Group, 1989–1991
    - Secretary-Treasurer, Transportation Geography Specialty Group, 1980–1982
    - Editor, *Transportation Geography Newsletter*, 1980–1982
- Society of Automotive Engineers, member, 1985–present
- International Association for Energy Economics, member
- Consulting
    - International Council for Clean Transportation, 2011
    - International Transport Forum, 2007
    - Addx Corporation, 2007
    - United Nations Framework Convention on Climate Change, 2007
    - Securing America's Future Energy, 2007
    - Center for Clean Air Policy, 2007
    - Pollution Probe Canada, 2006-2007
    - The Energy Foundation China Project, 2005—present
    - The Pew Center on Global Climate Change, 2004—present
    - Eno Transportation Foundation, 1991–1996
    - Transportation Research Board, 1996–1997

BOOKS

and D.W. Jones and Mark Delucchi, eds., *The Full Costs and Benefits of Transportation*, Springer-Verlag, Heidelberg, 1997.

Transportation and Energy, Eno Foundation for Transportation, Lansdowne, Virginia, 1996.

and D. J. Santini, eds., *Transportation and Global Climate Change*, American Council for an Energy Efficient Economy, Washington, DC, 1993.

ARTICLES IN PROFESSIONAL JOURNALS

"What's Greener than a VMT Tax? The Case for an Indexed Energy User Fee to Finance U.S. Surface Transportation," *Transportation Research D-Environment*, vol. 16, pp. 451-458.

"Uncertainty, Loss Aversion and Markets for Energy Efficiency," *Energy Economics*, vol. 33, pp. 608-616, 2011.

Z. Lin and D.L. Greene, *Predicting Individual On-road Fuel Economy Using Simple Consumer and Vehicle Attributes*, SAE Technical Paper Series No. 11SDP-0014, Society of Automotive Engineers, Warrendale, PA, April 12, 2011.

and P.R. Boudreaux, D.J. Dean, W. Fulkerson, A.L. Gaddis, R.L. Graham, R.L. Graves, J.L. Hopson, P. Hughes, M.V. Lapsa, T.E. Mason, R.F. Standaert, T.J. Wilbanks and A. Zucker, "The Importance of Advancing Technology to America's Energy Goals," *Energy Policy*, vol. 38, no. 8, pp. 3886-3890, March 2010.

Rubin, J., P.N. Leiby and D.L. Greene, "Tradable Fuel Economy Credits: Competition and Oligopoly," *Journal of Environmental Economics and Management*, vol. 58, no. 3, pp. 315-328, 2009.

"Measuring Energy Security: Can the United States Achieve Oil Independence?" *Energy Policy*, 2009, Vol. 38, No. 4, pp. 1614-1621.

"Feebates, Footprints and Highway Safety," *Transportation Research Part D*, vol. 14, pp. 375-384, 2009.

"Vehicles and E85 Stations Needed to Achieve to Achieve Ethanol Goals," *Transportation Research Record No. 2058*, pp. 172-178.

and P.N. Leiby, P.D. Patterson, S.E. Plotkin and M. Sing, "Oil Independence: Achievable National Goal or Empty Slogan?" *Transportation Research Record,* No. 2017, pp. 47-53, Washington, DC, 2007.

and J.L. Hopson, R. Goeltz and J. Li, "Analysis of In-Use Fuel Economy Shortfall Based on Voluntarily Reported Mile-per-Gallon Estimates," *Transportation Research Record*, No. 1983, pp. 99-105, 2007.

Leiby, P.N., D.L. Greene, D. Bowman and E. Tworek, "Systems Analysis of Hydrogen Transition with HyTrans," *Transportation Research Record*, No. 1983, pp. 129-139, 2007.

and J.L. Hopson and J. Li, "Have We Run Out of Oil Yet? Oil Peaking Analysis from an Optimist's Perspective," *Energy Policy*, vol. 34, pp. 515–531, 2006.

S. Ahmad and D.L. Greene, "The Effect of Fuel Economy on Automobile Safety: A Reexamination," *Transportation Research Record No. 1941*, pp. 1-7, Washington, DC, January 2005.

and J.L.Hopson and J. Li, "Running Out of and Into Oil: Analyzing Global Depletion and Transition Through 2050, *Transportation Research Record 1880*, pp. 1-9, Transportation Research Board, Washington, DC, 2005.

and P.D. Patterson, M. Singh and J. Li, "Feebates, Rebates and Gas-Guzzler Taxes: A Study of Incentives for Increased Fuel Economy," *Energy Policy*, vol. 33, no. 6, pp. 721-827, 2005.

Sheffield, J., et al., "Energy Options for the Future," *Journal of Fusion Energy*, vol. 23, no. 2, pp. 63-109, 2004.

and J. Hopson, "An Analysis of Alternative Forms of Automotive Fuel Economy Standards for the United States," *Transportation Research Record No. 1842*, pp. 20-28, Transportation Research Board, Washington, DC, 2003.

H.L. Hwang, S.M. Chin and D.L. Greene, "In, Out, Within and Through: Geography of Truck Freight in the Lower 48," *Transportation Research Record,* no. 1768, pp. 18–25, Transportation Research Board, Washington, DC, 2001.

and S.E. Plotkin, "Energy Futures for the U.S. Transportation Sector," *Energy Policy*, vol. 29, no. 14, pp. 1255–1270, 2001.

and N. Tishchishyna, "The Costs of Oil Dependence: A 2000 Update," *Transportation Quarterly*, vol. 55, no. 3, pp. 11–32, 2001.

H.L. Hwang, D.L. Greene, S.M. Chin, J. Hopson and A.A. Gibson, "Real-time Indicators of VKT and Congestion: One Year of Experience," *Transportation Research Record*, no. 1719, pp. 209–214, Transportation Research Board, Washington, DC, 2000.

and J.M. DeCicco, "Engineering-Economic Analyses of Automotive Fuel Economy Potential in the United States," *Annual Review of Energy and the Environment*, vol. 25, pp. 477–536, 2000.

L.A. Greening, D.L. Greene and C. Difiglio, "Energy Efficiency and Consumption—The Rebound Effect—A Survey," *Energy Policy*, vol. 28, pp. 389–401, 2000.

R.N. Schock, W. Fulkerson, M.L. Brown, R.L. San Martin, D.L. Greene and J. Edmonds, "How Much Is Energy R&D Worth as Insurance?" *Annual Review of Energy and the Environment*, vol. 24, pp. 487–512, Annual Review, Palo Alto, California, 1999.

S.M. Chin, D.L. Greene, J. Hopson, H.L. Hwang and B. Thompson, "Towards Real-Time Indices of U.S. Vehicle Travel and Traffic Congestion," *Transportation Research Record*, no. 1660, pp. 132–139, National Academy Press, Washington, DC, 1999.

and J. Kahn and R. Gibson, "Fuel Economy Rebound Effect for U.S. Household Vehicles*," The Energy Journal*, vol. 20, no. 3, pp. 1–31, 1999.

"Survey Evidence on the Importance of Fuel Availability to Choice of Alternative Fuels and Vehicles," *Energy Studies Review*, vol. 8, no. 3, pp. 215–231, 1998.

"Why CAFE Worked," *Energy Policy*, vol. 26, no. 8, pp. 595–614, 1998.

and Donald W. Jones and Paul N. Leiby, "The Outlook for U.S. Oil Dependence," *Energy Policy*, vol. 26, no. 1, pp. 55–69, 1998.

and Michael Wegener, "Sustainable Transport," in *Journal of Transport Geography*, vol. 5, no. 3, pp. 177–190, 1997.

Steven E. Plotkin and David Greene, "Prospects for Improving the Fuel Economy of Light-Duty Vehicles," *Energy Policy*, vol. 25, no. 14-15, pp. 1179–1188, 1997.

"Economic Scarcity: Monopoly, Not Geology, Threatens Global Supply," *Harvard International Review*, vol. XIX, no. 3, Summer, 1997.

"Environmental Impacts," *Journal of Transport Geography*, vol. 5, no. 1, pp. 28–29, 1997.

"Energy for Transportation," *Journal of Transport Geography*, vol. 5, no. 1, pp. 30–32, 1997.

and Y. Fan, "Transportation Energy Intensity Trends, 1972–1992," *Transportation Research Record*, no. 1475, pp. 10–19, Energy and Environment, Transportation Research Board, Washington, DC, 1995.

M.A. Deluchi, D.L. Greene and Quanlu Wang, "Motor Vehicle Fuel Economy: The Forgotten Hydrocarbon Control Strategy?" *Transportation Research A*, vol. 28A, no. 3, pp. 223–244, 1994.

"Transportation and Energy," *Transportation Quarterly*, vol. 48, no. 1, pp. 91–101, Winter, 1994.

and K.G. Duleep, "Costs and Benefits of Automotive Fuel Economy Improvement," *Transportation Research*, vol. 27A, no. 3, pp. 217–236, May, 1993.

"Transportation and Energy: The Global Environmental Challenge," *Transportation Research*, vol. 27A, no. 3, pp. 163–166, May, 1993.

"Energy Efficiency Improvement Potential of Commercial Aircraft," *Annual Review of Energy and Environment*, vol. 17, pp. 537–573, 1992.

"Vehicle Use and Fuel Economy: How Big is the Rebound Effect?" *The Energy Journal*, vol. 13, no. 1, pp. 117–143, April 1992.

"A Note on OPEC Market Power and Oil Prices," *Energy Economics*, vol. 13, no. 2, pp. 123–129, April 1991.

"The Cost of Short-Run Pricing Strategies to Increase Corporate Average Fuel Economy," *Economic Inquiry*, vol. XXIX, no. 1, pp. 101–114, January 1991.

"Fuel Choice for Multifuel Vehicles," *Contemporary Policy Issues*, vol. VIII, no. 4, pp. 118–137, October 1990.

"CAFE or PRICE? An Analysis of the Effects of Federal Fuel Economy Regulations and Gasoline Price on New Car MPG, 1978–89," *The Energy Journal*, vol. 11, no. 3, pp. 37-57, September 1990.

"Technology and Fuel Efficiency," *Forum for Applied Research and Public Policy*, vol. 5, no. 1, pp. 23–29, University of Tennessee, Spring 1990.

Carmen Difiglio, K.G. Duleep and D.L. Greene, "Cost Effectiveness of Future Fuel Economy Improvements," *The Energy Journal*, vol. 11, no. 1, 1990.

"Short-Term Options for Controlling CO_2 Emissions of Light-Duty Vehicles," *SAE Technical Paper Series 901111*, Society of Automotive Engineers, 1990.

"Motor Fuel Choice: An Econometric Analysis," *Transportation Research A*, vol. 23A, no. 3, pp. 243–253, 1989.

"Fuel Choice for Dual-Fuel Vehicles: An Analysis of the Canadian Natural Gas Vehicles Survey," *SAE Technical Paper Series 892067*, Society of Automotive Engineers, Warrendale, Pennsylvania, 1989.

J.J. Erickson, D.L. Greene and A.J. Sabadell, "An Analysis of Transportation Energy Conservation Projects in Developing Countries," *Transportation*, vol. 15, no. 3, pp. 163–189, 1988.

and J.T. Liu, "Automotive Fuel Economy Improvements and Consumers' Surplus," *Transportation Research A*, vol. 22A, no. 3, pp. 203–218, 1988.

"Advances in Automobile Technology and the Market for Fuel Efficiency, 1978–1985," *Transportation Research Record 1155*, pp. 18–27, Transportation Research Board, National Research Council, Washington, DC, 1987.

and Anthony Araya Jacome, Robert Kowalski and Patricia S. Hu, "Road Transport Energy Conservation in Costa Rica," *Energy*, vol. 12, no. 12, pp. 1299–1308, 1987.

and P.S. Hu, "A Functional Form Analysis of the Short-Run Demand for Travel and Gasoline by One-Vehicle Households," *Transportation Research Record*, no. 1092, pp. 10–15, Transportation Research Board, Washington, DC, 1986.

and N. Meddeb and J.T. Liu, "Vehicle Stock Modeling of Highway Energy Use: Tunisian and U.S. Applications," *Energy Policy*, pp. 437–446, October 1986.

"Efficiency-Related Changes in Automobile and Light Truck Markets," *SAE Technical Paper Series*, no. 861423, Society of Automotive Engineers, Warrendale, Pennsylvania, September 1986.

"The Market Share of Diesel Cars in the U.S., 1979–83," *Energy Economics*, vol. 8, no. 1, pp. 13–21, January 1986.

and P.S. Hu and L. Till, "An Analysis of Trends in Automotive Fuel Economy from 1978 to 1984," *Transportation Research Record*, no. 1049, pp. 51–56, Washington, DC, 1985.

"Estimating Daily Vehicle Usage Distributions and the Implications for Limited-Range Vehicles," *Transportation Research B*, vol. 19B, no. 4, pp. 347–358, 1985.

and P.S. Hu, "Vehicle Usage in Multi-Vehicle Households and the Price of Gasoline," *Transportation Research Record*, no. 988, pp. 19–24, Washington, DC, 1984.

and P.S. Hu and G.F. Roberts, "An Analysis of Geographical and Temporal Variation in Vehicles Classification Count Statistics," *Transportation Research Record*, no. 987, pp. 21–28, Washington, DC, 1984.

and G.F. Roberts, "A Comment on Fuel Consumption for Road Transport in the U.S.A," *Energy Economics*, vol. 6, no. 2, pp. 145–147, April 1984.

"A Derived Demand Model of Regional Highway Diesel Fuel Use," *Transportation Research B*, vol. 18B, no.1, pp. 43–61, 1984.

P.D. Patterson, F.W. Westbrook, D.L. Greene and G.F. Roberts, "Reasons for Changes in MPG Estimates, Model Year 1978 to the Present," SAE Technical Paper Series, no. 840500, Society of Automotive Engineers, Warrendale, Pennsylvania, February/March, selected for inclusion in *1984 SAE Transactions*, 1984.

J. Soderstrom, E. Hirst, D. Greene and J. Trimble, "Have Department of Energy Conservation Programs Saved Energy?" *Evaluation Review*, vol. 8, no. 1, pp. 93–112, February 1984.

E. Hirst, R. Marlay, D. Greene and R. Barnes, "Recent Changes in U.S. Energy Consumption: What Happened and Why?" *Annual Review of Energy*, vol. 8, pp. 193–243, Annual Reviews, Inc., Palo Alto, California, 1983.

"Streamlining the Collection and Processing of Traffic Count Statistics, A Comment," *Transportation Research Record*, no. 928, pp. 18–19, 1983.

"A Note on Implicit Consumer Discounting of Automobile Fuel Economy: Reviewing the Available Evidence," *Transportation Research*, vol. 17B, no. 6, pp. 491–499, 1983.

and C.K. Chen, "A Time Series Analysis of State Gasoline Demand, 1975–1980," *The Professional Geographer*, vol. 35, no. 1, pp 40–51, February 1983.

G.F. Roberts and D.L. Greene, "A Method for Assessing the Market Potential of New Energy-Saving Technologies," *Transactions on Systems, Man, and Cybernetics*, Institute of Electrical and Electronics Engineers, vol. SMC-13, no. 1, pp. 3–7, January/February 1983.

Eric Hirst et al., "Effects of Improved Energy Efficiency on U.S. Energy Use: 1973–1980," *Energy*, vol. 7, no. 11, pp. 897–907, 1982.

and G. Kulp, "An Analysis of the 1979–1980 Decline in gasoline Consumption in the United States," *Energy*, vol. 7, no. 4, pp. 367–375, 1982.

and G. Walton, "Data and Methodological Problems in Establishing State Gasoline Conservation Targets," *Transportation Research Record*, no. 815, pp. 24–30, 1981.

and E. Chen, "Scrappage and Survival Rates of Passenger Cars and Trucks in the U.S., 1966–77," *Transportation Research*, vol. 15A, no. 5, pp. 383–389, 1981.

"City Size Distribution and Income Distribution in Space: A Comment," *Regional Development Dialogue*, vol. II, no. 1, pp. 124–126, Spring 1981.

"A State Level Stock System Model of Gasoline Demand," *Transportation Research Record*, no. 801, pp. 44–50, 1981.

"Estimated Speed/Fuel Consumption Relationships for a Large Sample of Cars," *Energy*, vol. 6, pp. 441–446, 1981.

"The Spatial Dimension of Gasoline Demand," *Geographical Survey*, vol. 9, no. 2, pp. 19–28, April 1980.

"Regional Demand for Gasoline: Comment," *Journal of Regional Science*, vol. 20, no. 1, pp. 103–109, 1980.

"Urban Subcenters: Recent Trends in Urban Spatial Structure," *Growth and Change*, vol. 11, no. 1, pp. 103–109, January 1980.

R. Dubin, D.L. Greene and C. Begovich, "Multivariate Classification of Automobiles Using an Automobile's Characteristics' Data Base," *Transportation Research Record*, no. 726, pp. 29–27, 1979.

"State Differences in the Demand for Gasoline: An Econometric Analysis," *Energy Systems and Policy*, vol. 3, no. 2, pp. 191–212, 1978.

and Joern Barnbrock, "A Note on Problems in Estimating Urban Density Models," *Journal of Urban Economics*, vol. 5, April 1978.

and Rolf R. Schmitt, "An Alternative Derivation of the Intervening Opportunities Model," *Geographical Analysis*, vol. 10, no. 1, January 1978.

Joern Barnbrock and D.L. Greene, "A Comment on Population Density and Trend Surface Analysis," *Land Economics*, vol. 53, no. 2, May 1977.

CONTRIBUTIONS TO NATIONAL RESEARCH COUNCIL REPORTS

"Assessment of Fuel Economy Technologies for Light-Duty Vehicles," Report of the Committee on the assessment of Technologies for Improving Light-duty Vehicle Fuel Economy, National Academies Press, Washington, D.C., 2010.

J. Zucchetto, *Trends in Oil Supply and Demand, Potential for Peaking of Conventional Oil Production, and Possible Mitigation Options*, a summary report of the Modeling the Oil Transition workshop, Member, Planning Group and Keynote Speaker, Washington, DC, April 2006.

"State and Federal Standards for Mobile Source Emissions," Report of the Committee on State Practices in Setting Mobile Source Emissions Standards, National Research Council, National Academies Press, Washington, DC, March 2006.

"Integrating Sustainability into the Transportation Planning Process," *Conference Proceedings 37*, Transportation Research Board of the National Academies, Washington, DC, 2005.

"Effectiveness and Impact of Corporate Average Fuel Economy (CAFE) Standards," Report of the Committee, National Research Council, National Academy Press, Washington, 2002.

"Ecological, Environmental and Energy-Related Issues, in *The Future Highway Transportation System and Society*, Transportation Research Board, National Research Council, National Academy Press, Washington, DC, 1997.

Expanding Metropolitan Highways: Implications for Air Quality and Energy Use, Special Report 245, Transportation Research Board, National Research Council, Washington, DC, July 1995.

Automotive Fuel Economy: How Far Can We Go? Report of the Committee on Automobile and Light Truck Fuel Economy, National Research Council, National Academy Press, Washington, DC, 1992.

In Pursuit of Speed: New Options for Intercity Passenger Transport, Special Report 233, Transportation Research Board, National Research Council, Washington, DC, 1991.

and D. Sperling and B. McNutt, "Transportation Energy to the Year 2020," pp. 207–231, in *A Look Ahead: Year 2020*, Special Report 220, Transportation Research Board, National Research Council, Washington, DC, 1988.

CONGRESSIONAL TESTIMONY

"Technology-Neutral Incentives for Energy Efficient Low Greenhouse Gas Emitting Vehicles," Testimony to the Finance Committee, United States Senate, April 23, 2009.

"Near-Term Options to Increase Fuel Economy and Reduce Petroleum Demand," *Testimony to the Committee on Energy and Natural* Resources, United States Senate, July 23, 2008.

"Facing the Challenges of Oil Dependence and Climate Change: What Will It Take?" *Testimony to the Subcommittee on Energy and Water Development*, U.S. House of Representatives Committee on Appropriations, February 14, 2008.

"Policies to Increase Passenger Car and Light Truck Fuel Economy," *Testimony to the U.S. Senate Committee on Energy and Natural Resources*, January 30, 2007.

"Is Cap-and-Trade a Sufficient Carbon Policy for Transportation?" Testimony to the U.S. Senate Committee on Environment and Public Works, November 13, 2007, Washington, DC.

"Energy Challenges for Transportation in the 21st Century," Testimony to the National Surface Transportation Policy and Revenue Study Commission, March 19, 2007, Washington, DC.

"Corporate Average Fuel Economy (CAFE) Standards," Testimony to the U.S. Senate Commerce Committee, March 6, 2007, Washington, DC.

"Policies to Increase Passenger Car and Light Truck Fuel Economy," Testimony to the U.S. Senate Committee on Energy and Natural Resources, January 30, 2007, Washington, DC .

 "Observations on the H-Prize Act of 2006 (H.R. 5143)," Testimony to the U.S. House of Representatives Committee on Science, April 27, 2006, Washington, DC.

 "Improving the Nation's Energy Security: Can Cars and Trucks be Made More Fuel Efficient?" Testimony to the U.S. House of Representatives Committee on Science, February 9, 2005, Serial No. 109-3, U.S.G.P.O, Washington, DC.

"The Outlook for Surface Transportation Growth," Testimony to the Subcommittee on Surface Transportation of the Committee on Transportation Infrastructure of the United States House of Representatives, March 28, 1996.

CONTRIBUTIONS TO BOOKS

"Measuring Energy Sustainability: Making Progress," in *Linkages of Sustainability*, T.E. Graedel and E. van der Voet, eds., Strüngmann Forum Reports, MIT Press, Cambridge, MA, 2010.

Löschel, A., J. Johnston, M.A. Delucchi, T.N. Demayo, D.L. Gauthier, D.L. Greene, J. Ogden, S. Rayner and E. Worrell, "Energy: Stocks, Flows, and Prospects," in *Linkages of Sustainability*, T.E. Graedel and E. van der Voet, eds., Strüngmann Forum Reports, MIT Press, Cambridge, MA, 2010.

D.L. Greene, "What Will Replace Liquid Hydrocarbon Fuels?" in M.J. Benton, ed., *The Seventy Great Mysteries of the Natural World*, Thames and Hudson, Inc., New York, 2008.

and J. German and M. A. Delucchi, "Fuel Economy: The Case for Market Failure," in *Reducing Climate Impacts of the Transportation Sector*, D. Sperling and J.S. Cannon, eds., Springer Science+Business Media, springer.com, 2008.

Ribeiro, S.K., S. Kobayashi, D.L. Greene et al., "Transport and its Infrastructure," Chapter 5 in *Climate Change 2007: Mitigation of Climate Change*, Working Group III contribution to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change, Cambridge Press, Cambridge, United Kingdom.

A. Dumas, D.L. Greene and A. Bourbeau, "North American Feebate Analysis Model," in *Driving Climate Change*, D. Sperling and J.S. Cannon, eds., Academic Press, San Francisco, CA, 2007.

"Transportation and Energy," in *The Geography of Urban Transportation*, S. Hanson and G. Giuliano, eds., The Guilford Press, New York, 2004.

"Transportation and Energy Overview," in *Encyclopedia of Energy*, Cutler J. Cleveland, ed., Elsevier, Oxford, United Kingdom, 2004.

M. Wegener and D.L. Greene, "Sustainable Transport" in *Social Change and Sustainable Transport*, pp. 35-42, W.R. Black and P. Nijkamp, eds., Indiana University Press, Bloomington, IN, 2002.

"Sustainable Transportation," in *International Encyclopedia of the Social and Behavioral Sciences*, N.J. Smelser and P.B. Baltes, eds., Elsevier, Oxford, pp. 15,335–15,339, 2002.

W.R. Moomaw, J.R. Moreira, K. Blok, D.L. Greene et al., "Technological and Economic Potential of Greenhouse Gas Emissions Reduction," Chapter 3 in *Climate Change 2001: Mitigation*, contribution of Working Group III to the Third Assessment Report of the Intergovernmental Panel on Climate Change, Cambridge University Press, Cambridge, UK, 2001.

"Transportation's Oil Dependence and Energy Security in the 21st Century," in *Environmental Change, Adaptation, and Security*, S.C. Lonergan, ed., NATO Advanced Science Institute Series, Klumer Academic Publishers, Boston, 1999.

S.L. Baughcum, J.J. Begin, F. Franco, D.L. Greene et al., "Aircraft Emissions: Current Inventories and Future Scenarios," Chapter 9 in *Aviation and Global Atmosphere*, Intergovernmental Panel on Climate Change, Cambridge University Press, Cambridge, UK, 1999.

and Robert Gibson, "Transportation, Energy and the Environment," Chapter 4 in *Transportation Statistics Annual Report 1998*, BTS98-S-01, Bureau of Transportation Statistics, U.S. Department of Transportation, Washington, DC, September 1998.

"Commercial Air Transport Energy Use and Emissions: Is Technology Enough?" pp. 207–228 in DeCicco and Delucchi, eds., *Transportation, Energy, and the Environment: How Far Can Technology Take Us?* American Council for an Energy Efficient Economy, Washington, DC, 1997.

B. McNutt, L. Fulton and D.L. Greene, "Is Technology Enough? A Synthesis of Views Expressed at the Conference," pp. 251–256 in DeCicco and Delucchi, eds., *Transportation, Energy, and Environment: How Far Can Technology Take Us?* American Council for an Energy Efficient Economy, Washington, DC, 1997.

and Robert Gibson, "Transportation, Energy, and the Environment," Chapter 4 in *Transportation Statistics Annual Report 1997*, BTS97-5-01, Bureau of Transportation Statistics, U.S. Department of Transportation, Washington, DC, October 1997.

and Donald W. Jones, "The Full Costs and Benefits of Transportation: Conceptual and Theoretical Issues," Chapter 1 of *The Full Costs and Benefits of Transportation*, Springer-Verlag, Heidelberg, 1997.

"The Cost of Transportation's Oil Dependence," in O. Hohmeyer, R.L. Ottinger and K. Rennings, eds., *Social Costs and Sustainability*, Springer-Verlag, Berlin, 1996.

"Twenty Years of Energy Policy: What Should We Have Learned?" in David L. Feldman, ed., *The Energy Crisis*, The Johns Hopkins University Press, Baltimore, 1996.

"Environmental Impacts of Transportation," Chapter 6 in *Transportation Statistics Annual Report 1996*, Bureau of Transportation Statistics, U.S. Department of Transportation, Washington, DC, October 1996.

R. Gibson and D.L. Greene, "Environmental Trends and the U.S. Transportation System," Chapter 7 in *Transportation Statistics Annual Report 1996*, Bureau of Transportation Statistics, U.S. Department of Transportation, Washington, DC, October 1996.

"Energy and Transportation," Chapter 4 in *Transportation Statistics Annual Report 1996*, Bureau of Transportation Statistics, U.S. Department of Transportation, Washington, DC, October 1996.

"Transportation," Chapter 5 in *Policies and Measures for Reducing Energy Related Greenhouse Emissions*, DOE/PO-0047, U.S. Department of Energy, Office of Policy and International Affairs, Washington, DC, July 1996.

Guest editor, "Energy and Global Climate Change," *Transportation Research*, Special Issue, vol. 27A, no. 3, May 1993.

"Regional Demand Implication for Gasoline Supply Shortages," in T.R. Lakshmanan and P. Nijkamp, eds., *Systems and Models for Energy and Environmental Analysis*, pp. 206–233, Gower, UK, 1983.

G. Samuels, A.B. Rose, D.L. Greene and J.N. Hooker, "Energy Conservation in Transportation," a chapter in Vol. III of *Advances in Energy Systems and Technology*, Peter Auer, ed., Academic Press, New York, 1982.

"A Regional Stock System Model of Highway Gasoline Demand," in *Changing Energy Use Futures, Volume 1*, Proceedings of the Second International Conference on Energy Use and Management, Los Angeles, California, October 22–26, 1979.

and K.G. Duleep, and G. Upreti, "Status and Outlook for the U.S. Non-Automotive Fuel Cell Industry: Impacts of Government Policies and Assessment of Future Opportunities", ORNL/TM-2011/101, Oak Ridge National Laboratory, Oak Ridge, Tennessee.

Bunch, D.S. and D.L. Greene, "Potential Design, Implementation, and Benefits of a Feebate Program for New Passenger Vehicles in California," State of California Air Resources Board and the California Environmental Protection Agency, Sacramento, California, 2011, available at http://76.12.4.249/artman2/uploads/1/Feebate_Program_for_New_Passenger_Vehicles_in_California.pdf.

"Oil Peak or Panic?" Book review, *Science*, vol. 328, no. 5980, p. 828, May 14, 2010

"What's Greener than a VMT Tax? Discussion Paper, Howard H. Baker, Jr. Center for Public Policy, The University of Tennessee, Knoxville, TN, May 2010.

and S.E. Plotkin, *Reducing Greenhouse Gas Emissions from U.S. Transportation*, Pew Center on Global Climate Change, Arlington, VA, 2011.

How Consumers Value Fuel Economy: A Literature Review, EPA-420-R-10-008, U.S. Environmental Protection Agency, March 2010.

"Why the Market for New Passenger Cars Generally Undervalues Fuel Economy," Discussion Paper No. 2010-6, International Transport Forum, OECD, Paris, January 2010.

D. McCollum, G. Gould and D. Greene, *Greenhouse Gas Emissions from Aviation and Marine Transportation: Mitigation Potential and Policies*, Pew Center on Global Climate Change, Arlington, Virginia, available at http://www.pewclimate.org/technology/report/aviation-an-marine, December 2009.

"Energy Assurance: Essential Energy Technologies for Climate Protection and Energy Security," ORNL/TM-2009/314, Oak Ridge National Laboratory, Oak Ridge, TN, December 2009.

S. Plotkin et al., *Multipath Transportation Futures Study: Vehicle Characterization and Scenario Analyses*, ANL/ESD/09-5, Argonne National Laboratory, Argonne, Illinois, July 2009.

and K.G. Duleep, "Bootstrapping a Sustainable North American PEM Fuel Cell Industry: Could a Federal Acquisition Program Make a Difference?" ORNL/TM-2008/183, October, 2008.

"Future Prices and Availability of Transport Fuels," *International Transport Forum*, Research Round Table on Oil Dependence: Is Transport Running Out of Affordable Fuel? November 14-15, 2007, OECD, Paris.

American Physical Society Study Group, *Energy Future: Think Efficiency*, American Physical Society, College Park, Maryland, 2008.

D.L. Greene et al., *Analysis of the Transition to Hydrogen Fuel Cell Vehicles and the Potential Hydrogen Infrastructure Requirements*, ORNL/TM-2008/30, Oak Ridge National Laboratory, Oak Ridge, Tennessee, March, 2008.

"Future Prices and Availability of Transport Fuels," Discussion Paper No. 2007-15, International Transport Forum, Joint Transportation Research Center, OECD, Paris,
 http://www.internationltransport.forum.org/jtre/DiscussionPapers/DiscussionPaper15.pdf.

"Transportation," in *The First State of the Carbon Cycle Report (SOCCR): The North American Budget and Implications for the Global Carbon Cycle. A Report by the U.S. Climate Change Science Program and the Subcommittee on Global Change Research* [King, A.W., L. Dilling, G.R. Zimmerman, D.M. Ciarman,

R.A. Houghton, G. Marland, A.Z. Rose, and T.J. Wilbanks (eds)]. National Oceanic and Atmospheric Administration, National Climate Data Center, Asheville, North Carolina, USA, pp. 73-84.

and P.N. Leiby and D. Bowman, *Integrated Analysis of Market Transformation Scenarios with HyTrans*, ORNL/TM-2007/094, Oak Ridge National Laboratory, Oak Ridge, Tennessee, June 2007.

D. Gordon, D.L. Greene, M.H. Ross and T.P. Wenzel, "Sipping Fuel and Saving Lives: Increasing Fuel Economy without Sacrificing Safety," International Council on Clean Transportation, Washington, DC, June 2007.

and P.N. Leiby, "Oil Independence: Realistic Goal or Empty Slogan," published as "Expert Commentary" on The Lugar Energy Initiative website, http://lugar.senate.gov/energy/links/commentary/08_greene_summary.html, 2007.

Editor, *Modeling the Oil Transition: A Summary of the Proceedings of the DOE/EPA Workshop on the Economic and Environmental Implications of Global Energy Transitions*, ORNL/TM-2007/014, Oak Ridge National Laboratory, Oak Ridge, Tennessee, February 2007.

and P.N. Leiby, *The Oil Security Metrics Model*, ORNL/TM-2006/505, Oak Ridge National Laboratory, Oak Ridge, Tennessee, May 2006.

and S. Ahmad, *Costs of U.S. Oil Dependence: 2005 Update*, ORNL/TM-2005/45, Oak Ridge National Laboratory, Oak Ridge, Tennessee, March 2005.

"Climate Change Policy for Transportation While Waiting for H2," in *The 10-50 Solution: Technologies and Policies for a Low-Carbon Future*, Pew Center on Global Climate Change, National Commission on Energy Policy, Washington, DC, March 2004.

and K.G. Duleep and W. McManus, *Future Potential of Hybrid and Diesel Powertrains in the U.S. Light-Duty Vehicle Market*, ORNL/TM-2004/181, Oak Ridge National Laboratory, Oak Ridge, Tennessee, August 2004.

and J.L. Hopson, *Running Out of and Into Oil: Analyzing Global Depletion and Transition Through 2050*, ORNL/TM-2003/259, Oak Ridge National Laboratory, Oak Ridge, Tennessee, October 2003.

and R.C. Gibson and K.G. Duleep, *Energy Star Concepts for Highway Vehicles*, ORNL/TM-2003/37, Oak Ridge National Laboratory, Oak Ridge, Tennessee, June 2003.

R. Nye, D.L. Greene and J.W. Saulsbury, *Providing Consumers with Web-based Information on the Environmental Effects of Automobiles*, ORNL/TM-2003/166, Oak Ridge National Laboratory, Oak Ridge, Tennessee, June 2003.

and A. Schafer, "Reducing Greenhouse Gas Emissions from U.S. Transportation," Pew Center on Global Climate Change, Arlington, Virginia, May 2003.

and S. Plotkin and K.G. Duleep, *Examining the Potential for Voluntary Fuel Economy Standards in the United States and Canada*, ANL/ESD/02-5, Center for Transportation Research, Argonne National Laboratory, Argonne, Illinois, October 2002.

S.M. Chin, O. Franzese, D.L. Greene, H.L. Hwang and R.C. Gibson, *Temporary Losses of Highway Capacity and Impacts on Performance*, ORNL/TM-2002/3, Oak Ridge National Laboratory, Oak Ridge, Tennessee, May 2002.

TAFV Alternative Fuels and Vehicles Choice Model Documentation, ORNL/TM-2001/134, Oak Ridge National Laboratory, Oak Ridge, Tennessee, July 2001.

R.N. Schock, W. Fulkerson, M.L. Brown, R.L. San Martin, D.L. Greene and J.A. Edmonds, *The Insurance Value of Energy R&D*, UCEL-ID-141815, CGSR-2001, 003, Center for Global Security Research, Lawrence Livermore National Laboratory, Livermore, California, May 2001.

"Energy," in Chapter 5 of *The Changing Face of Transportation*, U.S. Department of Transportation, Bureau of Transportation Statistics, BTS00-007, Washington, DC, 2000.

with S. Plotkin, "Transportation Sector," Chapter 6 in *Scenarios for a Clean Energy Future*, Interlaboratory Working Group on Energy-Efficient and Clean Energy Technologies, Oak Ridge National Laboratory, ORNL/CON-476, Lawrence Berkeley National Laboratory, LBNL-44029, November 2000.

R.N. Schock, W. Fulkerson, M.L. Brown, R.L. San Martin, D.L. Greene and J. Edmonds, "How Much Is Energy R&D Worth?" *Energy*, vol. 25, no. 2, pp. 4–8, 2000.

Costs of Oil Dependence: A 2000 Update, ORNL/TM-2000/152, Oak Ridge National Laboratory, Oak Ridge, Tennessee, May 2000.

and J.M. DeCicco, *Engineering-Economic Analyses of Automotive Fuel Economy Potential in the United States*, ORNL/TM-2000/26, Oak Ridge National Laboratory, Oak Ridge, Tennessee, February 2000.

and J.M. DeCicco, "Energy and Transportation Beyond 2000," Millennium Paper Series, Transportation Research Board, National Research Council, Washington, DC, January 2000.

and many others, *National Transportation Statistics 1999*, BTS99-04, U.S. Department of Transportation, Bureau of Transportation Statistics, Washington, DC, 1999.

and many others, *Transportation Statistics Annual Report 1999*, BTS99-03, U.S. Department of Transportation, Bureau of Transportation Statistics, Washington, DC, 1999.

An Assessment of Energy and Environmental Issues Related to the Use of Gas-to-Liquid Fuels in Transportation, ORNL/TM-1999/258, Oak Ridge National Laboratory, Oak Ridge, Tennessee, November 1999.

and J. Kahn and R. Gibson, *An Econometric Analysis of the Elasticity of Vehicle Travel with Respect to Fuel Cost per Mile Using RTEC Survey Data*, ORNL-6950, Oak Ridge National Laboratory, Oak Ridge, Tennessee, March 1999.

"Why CAFE Worked," *Transportation Research Circular No. 492*, Transportation Research Board, National Research Council, Washington, DC, 1999.

R.N. Schock, W. Fulkerson, M.L. Brown, R.L. San Martin, D.L. Greene and J. Edmonds, *How Much is Energy R&D Worth as Insurance?* UCRL-JC-131205, Lawrence Livermore National Laboratory, Livermore, California, October 1998.

"Transportation and Energy Policy," in *Twenty-Five Years of Energy and Environmental Policy*, pp. 125–140, University of Illinois at Chicago, Energy Resources Center, Chicago, Illinois, 1998.

"Ecological, Environmental and Energy-Related Issues," pp. 127–149 in *The Future Highway Transportation System and Society*, Transportation Research Board, National Research Council, 1997.

and S. Plotkin, "Transportation Sector," Chapter 5 *in Scenarios of U.S. Carbon Reductions*, Interlaboratory Working Group on Energy-Efficient and Low-Carbon Technologies, ORNL/CON-444, report to the Office of Energy Efficiency and Renewable Energy, U.S. Department of Energy, Washington, DC, September 1997.

and many others, *Energy Technology R&D: What Could Make a Difference*, ORNL-6921/V2-P1, Oak Ridge National Laboratory, Oak Ridge, Tennessee, September 1997.

"Energy and Environmental Consequences of Transportation: Indicators of Sustainability," *51st Session of the International Statistical Institute Proceedings*, Istanbul, Turkey, August 18-26, 1997.

"Oil Dependence: The Value of R&D," *32nd Intersociety Energy Conversion Engineering Conference Proceedings*, volume 3, pp. 2148–2153, Honolulu, Hawaii, July 28-30, 1997.

and Steve Plotkin and K.G. Duleep, "The Potential for Energy-Efficient Technologies to Reduce Carbon Emissions in the United States: Transport Sector," *32nd Intersociety Energy Conversion Engineering Conference Proceedings*, volume 3, pp. 2114–2119, Honolulu, Hawaii, July 28-30, 1997.

"Grin and Obey," *Energy*, vol. XXII, p. 6, April 1997.

"Consequences of U.S. Oil Dependence," *Energy*, vol. XXI, no. 5, pp. 3–8, November 1996.

D. Streets and many others, "Inventory of Technologies, Methods, and Practices for Reducing Emissions of Greenhouse Gases," technical appendix to *Climate Change 1995: Impacts, Adaptations, and Mitigation of Climate Change: Scientific-Technical Analyses*, Argonne National Laboratory, Argonne, Illinois, May 1996.

one of Principal Lead Authors, "Industry, Energy and Transportation: Impacts and Adaptation," in *Climate Change 1995: Impacts, Adaptations and Mitigation of Climate Change: Scientific-Technical Analyses*, R.T. Watson, M.C. Zinyowera, R.H. Moss and D.J. Dokken, eds., Intergovernmental Panel on Climate Change, Cambridge University Press, 1996.

one of Contributing Authors, "Mitigation Options for the Transportation Sector," in *Climate Change 1995: Impacts, Adaptations and Mitigation of Climate Change: Scientific-Technical Analyses*, R.T. Watson, M.C. Zinyowera, R.H. Moss and D.J. Dokken, eds., Intergovernmental Panel on Climate Change, Cambridge University Press, 1996.

"Commercial Air Transport Energy Use and Emissions: Is Technology Enough?" in *Sustainable Transportation: Is Technology Enough?* Proceedings of the 1995 Asilomar Conference on Transportation and Energy, American Council for an Energy-Efficient Economy, Washington, DC, 1995.

"Transportation, Energy and Environment: Global Policy Issues for the Early 21st Century," in *International Workshop on Motor Vehicles and Global Environmental Problems*, Japan Automotive Research Institute, Tokyo, Japan, October–November 1995.

P.N. Leiby, D.L. Greene and Harry Vidas, *Market Potential and Impacts of Alternative Fuel use in Light Duty Vehicles: A 2000/2010 Analysis*, DOE/PO-0042, Office of Policy Analysis, U.S. Department of Energy, Washington, DC, January 1996.

and X. Han, "The Unintended Consequences of Transportation," Chapter 3 in *Transportation Statistics Annual Report 1995*, pp. 43–86, Bureau of Transportation Statistics, U.S. Department of Transportation, Washington, DC, 1995.

with D.W. Jones and P.N. Leiby, *The Outlook of U.S. Oil Dependence*, ORNL-6783, Oak Ridge National Laboratory, Oak Ridge, Tennessee, May 1995.

with S.M. Chin and R. Gibson, *Aggregate Vehicle Travel Forecasting Model*, ORNL-6872, Oak Ridge National Laboratory, Oak Ridge, Tennessee, May 1995.

D. Belzer, J. Roop, R. Sands and D.L. Greene, *Energy Conservation Trends, Energy Efficiency in the U.S. Economy, Report Three*, DOE/PO-0034, U.S. Department of Energy, Office of Energy Efficiency and Alternative Fuels Policy, Washington, DC, April 1995.

Transportation Energy Efficiency Trends, 1972–1992, ORNL-6828, Oak Ridge National Laboratory, Oak Ridge, Tennessee, December 1994.

Alternative Fuels and Vehicles Choice Model, ORNL/TM-12738, Oak Ridge National Laboratory, Oak Ridge, Tennessee, October 1994.

and E. Hillsman, L. Fulton, A.K. Wolfe and J. Niles, *Energy Emissions and Social Consequences of Telecommuting*, DOE/PO-0026, Technical Report One, Energy Efficiency in the U.S. Economy Series, U.S. Department of Energy, Office of Policy, Planning and Program Evaluation, Washington, DC, June 1994.

R.R. Schmitt, D.L. Greene et al., *Transportation Statistics Annual Report*, Bureau of Transportation Statistics, U.S. Department of Transportation, Washington, DC, 1994.

and Paul N. Leiby, *The Social Costs to the U.S. of Monopolization of the World Oil Market, 1972–1991*, ORNL-6744, Oak Ridge National Laboratory, Oak Ridge, Tennessee, March 1993.

Impacts on Home Heating Costs of Incentives for Alternative Fuel Vehicles, Technical Report Eight, U.S. Department of Energy, Office of Domestic and International Energy Policy, Washington, DC, May 1992.

with K.G. Duleep, *Costs and Benefits of Automotive Fuel Economy Improvement: A Partial Analysis*, ORNL-6704, Oak Ridge National Laboratory, Oak Ridge, Tennessee, March 1992.

M.A. Deluchi, Q. Wang and D.L. Greene, *Motor Vehicle Fuel Economy: The Forgotten Hydrocarbon Control Strategy?* ORNL-6715, Oak Ridge National Laboratory, Oak Ridge, Tennessee, July 1992.

"Transportation Energy Policy: Back to the Past or Ahead to the Future?" in *Twenty Years of Energy Policy: Looking Toward the Twenty-First Century*, Proceedings of the Twentieth Annual Illinois Energy Conference, Energy Resources Center, University of Illinois at Chicago, Chicago, Illinois, November 23–24, 1992.

with M. Singh, E. Ecklund, R. Bechtold and C. Saricks, "Second Interim Report of the Interagency Commission on Alternative Motor Fuels," Office of Energy Demand, Office of Policy, Planning and Analysis, Washington, DC, 1992.

Q. Wang, D.L. Greene and M.A. Deluchi, "Effects of Increasing Fuel Economy on Gasoline Vehicle HC Emissions," *Proceedings of the 84th Annual Meeting and Exhibition, Air and Waste Management Association*, Vancouver, BC, Canada, June 16–21, 1991.

"Coverage and Quality Problems with Existing Data Resources for Freight Transportation," *Proceedings of the Special Conference on Freight Transportation Data: The Changing Federal Role Since Deregulation*, Transportation Research Circular, No. 367, Transportation Research Board, Washington, DC, 1990.

and M. Singh, *First Interim Report on the Interagency Commission on Alternative Motor Fuels*, Office of Energy Demand Policy, Office of Policy, Planning and Analysis, U.S. Department of Energy, Washington, DC, September 30, 1990.

"Commercial Aircraft Fuel Efficiency Potential Through 2010," *Proceedings 1990 Intersociety Energy Conversion Engineering Conference*, Reno, Nevada, published by the American Institute of Chemical Engineers, August 12–17, 1990.

Energy Efficiency Improvement Potential of Commercial Aircraft to 2010, ORNL-6622, Oak Ridge National Laboratory, Oak Ridge, Tennessee, June 1990.

and Anju Rathi, *Alternative Motor Fuel Use Model: Model Theory and Design, and User's Guide*, ORNL/TM-11448, Oak Ridge National Laboratory, Oak Ridge, Tennessee, April 1990.

with Donald Igo, "Energy Dependence," Chapter 3 in *National Strategic Transportation Study*, U.S. Department of Transportation, Washington, DC, 1989.

"A Context for Estimating Economic and Energy Security Benefits," Technical Report Two, *Assessment of Costs and Benefits of Flexible and Alternative Fuel Use in the U.S. Transportation Sector*, DOE/PE-0091, Office of Policy, Planning and Analysis, U.S. Department of Energy, Washington, DC, September 1989.

and J.C. Liang, *Modeling Operating Weight and Axle Weight Distributions for Highway Vehicles*, ORNL-6437, Oak Ridge National Laboratory, Oak Ridge, Tennessee, July 1988.

et al., "Research Priorities in Transportation and Energy," *Transportation Research Circular*, no. 323, Transportation Research Board, Washington, DC, September 1987.

Simulating the Market for Automotive Fuel Efficiency: The SHRSIM Model, ORNL/TM-10074, Oak Ridge National Laboratory, Oak Ridge, Tennessee, February 1987.

and M.C. Holcomb, *Off-Highway Use of Gasoline in the United States*, U.S. Department of Transportation, Federal Highway Administration, Office of Highway Information Management, Washington, DC, July 1986.

RUMS, A PC-Based FORTRAN Program for Estimating Consumer Surplus Changes Using Multinomial Logit and Hedonic Demand Models, ORNL/TM-10069, Oak Ridge National Laboratory, Oak Ridge, Tennessee, August 1986.

Driver Energy Conservation Awareness Training: Review and Recommendations, ORNL/TM-9897, Oak Ridge National Laboratory, Oak Ridge, Tennessee, May 1986.

R.L. Graves, D.L. Greene, E.W. Gregory, II, *Application of the Adiabatic Diesel to Heavy Trucks: A Technology Assessment*, ORNL/TM-9554, Oak Ridge National Laboratory, Oak Ridge, Tennessee, March 1986.

and R. Kowalski and F. Southworth, *The Transportation Sector in Costa Rica and Opportunities for Energy Conservation*, a report of the Energy Conservation Services Program, U.S. Agency for International Development, Office of Energy, Washington, DC, May 1985.

P.S. Hu, D.L. Greene and L.E. Till, *Motor Vehicle MPG and Market Shares Report: First Six Months of Model Year 1984*, ORNL/TM-9391, Oak Ridge National Laboratory, Oak Ridge, Tennessee, October 1984.

and P.S. Hu and A.B. Rose, *Transportation Energy Use and Efficiency in Tunisia*, ORNL-6066, Oak Ridge National Laboratory, Oak Ridge, Tennessee, August 1984.

"Highway Fuel Use: Trends and Factors," *Proceedings of the Energy Information Administration Symposium on Petroleum Information*, DOE/EIA-0425, Energy Information Administration, Washington, DC, September 1983.

G.F. Roberts and D.L. Greene, *Trends in Heavy Truck Energy Use and Efficiency*, ORNL/TM-8843, Oak Ridge National Laboratory, Oak Ridge, Tennessee, October 1983.

et al., *An Overview of the Wharton School Assessments of the Short-Term Integrated Forecasting System*, DOE/NBB-0046, Energy Information Administration, Office of Statistical Standards, Washington, DC, July 1983.

"Regional Demand Implications for Gasoline Supply Shortages," in T.R. Lakshmanan and P. Nijkamp, eds., *Systems and Models for Energy and Environmental Analysis*, Gower, UK, pp. 206–233, 1983.

and E. Hirst, J. Soderstrom and J. Trimble, *Estimating the Total Impact on Energy Consumption of Department of Energy Conservation Programs*, ORNL-5925, Oak Ridge National Laboratory, Oak Ridge, Tennessee, November 1982.

"Gasoline use in the United States," *Petroleum Supply Monthly*, DOE/EIA-0109 (82-05), U.S. Department of Energy, Washington, DC, 1982.

and G. Kulp, G.H. Walton and D.B. Shonka, *Transportation Energy Use 1973–80: Changes, Trends, and Causes*, ORNL/TM-7953, Oak Ridge National Laboratory, Oak Ridge, Tennessee, April 1982.

G. Samuels, A.B. Rose, D.L. Greene and J.N. Hooker, "Energy Conservation in Transportation," a chapter in Vol. III of *Advances in Energy Systems and Technology*, Peter Auer, ed., Academic Press, New York, 1982.

E. Hirst, D.L. Greene et al., *Energy Use from 1973 to 1980: The Role of Improved Energy Efficiency*, ORNL/CON-79, Oak Ridge National Laboratory, Oak Ridge, Tennessee, December 1981.

The Aggregate Demand for Gasoline and Highway Passenger Vehicles in the United States: A Review of the Literature, 1938–1978, ORNL-5728, Oak Ridge National Laboratory, Oak Ridge, Tennessee, July 1981.

J.N. Hooker, A.B. Rose and D.L. Greene, *End Use Energy Consumption Data Base: Transportation Sector*, DOE/EIA/CR-7405-01, U.S. Department of Energy, Washington, DC, February 1980.

A Statistical Analysis of State VMT Estimates in VMT Statistics, Lifetime VMT, and Current State Methods of Estimating VMT, ORNL/TM-6327, Oak Ridge National Laboratory, Oak Ridge, Tennessee, February 1979.

G. Kulp, D.L. Greene et al., *Regional Analyses of Highway Energy Use*, ORNL-5587, Oak Ridge National Laboratory, Oak Ridge, Tennessee, December 1979.

"A Regional Stock System Model of Highway Gasoline Demand," in *Changing Energy Use Futures, Volume 1*, Proceedings of the Second International Conference on Energy Use Management, Los Angeles, California, October 22–26, 1979.

and T.P. O'Conner, P.D. Patterson, A.B. Rose and D.B. Shonka, *Regional Transportation Energy Conservation Data Book*, ORNL-5435, Oak Ridge National Laboratory, Oak Ridge, Tennessee, September 1978.

Econometric Analysis of the Demand for Gasoline at the State Level, ORNL/TM-6326, Oak Ridge National Laboratory, Oak Ridge, Tennessee, July 1978.

An Investigation of the Variability of Gasoline Consumption among States, ORNL-5391, Oak Ridge National Laboratory, Oak Ridge, Tennessee, April 1978.

R.R. Schmitt and D.L. Greene, "Evaluating Transportation Innovations with the Intervening Opportunities Model," *Proceedings of the Northeast American Institute for Decision Sciences*, April 1977.

M.G. Wolman, D.L. Greene and N.L. Froomer, *An Analysis of the Physical Characteristics of the Shanandoah River Which Affects Its Navigability* and *An Analysis of the Physical Characteristics of the Potomac River Which Affects Its Navigability*, prepared for the Baltimore District, U.S. Army Corps of Engineers, Chesapeake Research Consortium, pub. nos. 37S and 38S, resp., March 1975.

G. Power, M.G. Wolman et al., *Report on the Shenandoah River: An Investigation to Determine Navigability* and *Report on the Potomac River: An Investigation to Determine Navigability*, prepared for the Baltimore District, U.S. Army Corps of Engineers, Chesapeake Research Consortium, pub. nos. 37 and 38, resp., March 1975.

FORTHCOMING PUBLICATIONS

"Rebound 2007: Analysis of National Light-Duty Vehicle Travel Statistics," article in press, *Energy Policy*.

C. Liu and D.L. Greene, "Impacts of Feebates in Combination with Fuel Economy and Emissions Standards on U.S. Light-Duty Vehicles Fuel Use and Greenhouse Gas Emissions," forthcoming, *Transportation Research Record*.

Z. Lin and D.L. Greene, "Significance of Daily VMT Variation over Time and Among Drivers on Assessment of PHEV Energy Impact," forthcoming, *Transportation Research Record*.

Z. Lin and D.L. Greene, "Promoting the Market for Plug-in Hybrid and Battery Electric Vehicles: The Role of Recharge Availability," forthcoming, *Transportation Research Record.*

KARA M. KOCKELMAN

Professor & William J. Murray Jr. Fellow
Department of Civil, Architectural & Environmental Engineering - University of Texas at Austin
Note: Full CV available at www.ce.utexas.edu/prof/kockelman

EDUCATION
University of California at Berkeley   BSCE 1991, MSCE 1996, PhD 1998, MCP 1996

PROFESSIONAL EXPERIENCE – Academic
University of Texas at Austin	Assistant Professor of Civil Engineering	1998-2004
University of Texas at Austin	Associate Professor of Civil Engineering	2004-2009
University of Texas at Austin	Professor of Civil Engineering	2009 – Present

Licensed Professional Engineer, California (Certification #C057380) and Texas (Certification #93443)
Licensed City Planner, AICP 2008 & Member of American Planning Assoc. National & Texas Chapters, 2008

SELECTED RELEVANT REFEREED PUBLICATIONS
1. Musti, S., K. Kortum and K. Kockelman. "Household Energy Use and Travel: Opportunities for Behavioral Change." *Transportation Research* 16D: 49-56, 2011.
2. Kockelman, K., M. Thompson and C. Frei. "Americans' Contributions to Climate Change: Opportunities for Meeting Carbon Targets." *Journal of Urban Planning and Development*, Vol.137 (2): 91-100, 2011.
3. Paul, B., K. Kockelman and S. Musti. "The Light-Duty-Vehicle Fleet's Evolution: Anticipating PHEV Adoption and GHG Emissions across the U.S. Fleet." Forthcoming in *Transp Research Record*, 2011.
4. Kockelman K., M. Bomberg, M. Thompson and C. Whitehead. "GHG Emissions Control Options: Opportunities for Conservation." Report commissioned by National Academy of Science's Committee for the Study on the Relationships among Development Patterns, VMT, and Energy Conservation. June 2009.
5. Khan, M., and K. Kockelman. Predicting the Market Potential of Plug-In Electric Vehicles Using Multiday GPS Data. Proc'gs of the Annual Meeting of the Transportation Research Board and under review for publication in *Transportation Research Record.*

SELECTED OTHER REFEREED PUBLICATIONS
1. Tirumalachetty, S. and K. Kockelman. "Forecasting Greenhouse Gas Emissions from Urban Regions: Microsimulation of Land Use and Transport Patterns in Austin, Texas." Under review for publication in *Journal of Transport Geography*.
2. Tuttle, D. P. and K. Kockelman. "Electrified Vehicle Technology Trends, Infrastructure Implications and Cost Comparisons." Forthcoming in *Journal of the Transportation Research Forum.* Tirumalachetty, S. and K. Kockelman. "The Welfare Implications of Carbon Taxes and Carbon Caps: A Look at U.S. Households." Under review for publication to NOVA Science Publishers as a chapter in *Household Energy: Economics, Consumption and Efficiency.* Bomberg, M., K. Kockelman, M. Thompson. "GHG Emissions Control Options: Assessing Transportation & Electricity Generation Technologies & Policies to Stabilize Climate Change." Under review for publication in the *Journal of the Transportation Research Forum*. Kockelman, K. "Transportation and Land Use Solutions for Low-Carbon Cities." Under review for publication in the *Journal of the Transportation Research Forum*.

Over 95 refereed publications to date, and over 50 technical reports, total.
Over 100 referred conference proceedings and more than 250 additional presentations of research.

SELECTED HONORS AND AWARDS

U.C. Berkeley's University Medal - "Most Distinguished Graduate" of Graduating Class of 5,300 students, 1991

National Science Foundation Faculty Early Career Development (CAREER) Award, 2000-2004

Ford CAREER Award, given by Ford Motor Company's Ford Fund, 2002

One the World's Top 100 Young Innovators, according to MIT's *Technology Review*, May 2002

Named to the National Academy of Engineering's Gallery of Women in Engineering, August 2002

Inaugural recipient of the Annual New Faculty Award, sponsored by the Council of University Transportation Centers and the American Road and Transportation Builders Association, recognizing "outstanding teaching and research contributions to the transportation field" December 2002

Recipient of the Geoffrey J.D. Hewings Award for 2006 presented by the Regional Science Assoc. International

Recipient of ASCE's 2007 Harland Bartholomew Award, in recognition of contributions to the enhancement of the role of the civil engineer in urban planning and development

Recipient of the Women's Transportation Seminar 2007 Heart of Texas Chapter *Woman of the Year* Award

Recipient of ASCE's 2010 Walter L. Huber Research Prize in Transportation Engineering, for contributions in the areas of data acquisition and analysis to facilitate decisions in transport planning and policy-making

RELATED PROFESSIONAL ASSIGNMENTS, SYNERGISTIC TO THIS PROPOSAL

Member, Executive Committee, UT Austin & Texas A&M, NSF-sponsored Plug-in Electric Vehicle Industry-University Research Center (IURC), 2010 – present.

Member, UT Austin's NSF IGERT for Sustainable Grid Integration of Distributed & Renewable Energy Systems, 2010-present, & NSF RCN SEES, for Sustainable Cities: People, Infrastructures & the Energy-Climate-Water Nexus, 2011-present.

Member, Advisory Council on Transportation Statistics (ACTS), Bureau of Transportation Statistics Research and Innovative Technology Administration (RITA), U.S. Department of Transportation, 2010-present.

Member, National Research Council Committee for the Study on Relationships among Development Patterns, VMT, and Energy Conservation, 2007-2009.

Member, RAND Corporation's Panel on Transportation and Climate Change, Washington, DC, June 2008

Member, TRB Committees on Transportation and Land Development, Transport Economics, and Statistical Methods (and Past Chair of TRB's Survey Methods Committee).

Member, Editorial Advisory Boards of *Transportation Research (Part B)*, *Journal of Transport and Land Use, Journal of Regional Science,* and *Papers in Regional Science.*

PAST RELATED ENGAGEMENTS

Member, Transportation Electrification Panel, Sponsored by Indiana University's School of Public and Environmental Affairs, producing report "Plug-in Electric Vehicles: A Practical Plan for Progress" report (available at http://www.indiana.edu/~spea/pubs/TEP_combined.pdf), 2010-2011.

Invited presentation of "Urban Planning, Land Use, and Vehicle Technologies to Reduce GHG Emissions in Cities" for the National Academy of Science's EU-US Frontiers of Engineering Symposium, at UC Irvine, November 3, 2011; and for UT Austin's Energy and Urban Development Conference, November 18, 2011.

Presentation of "PEV Market Potential Using Multi-Day GPS Data" for the NSF EV-TEC Industry-University Research Center workshop on Electric Vehicles - Transportation and Electricity Convergence, Industry Advisory Panel Meeting, Houston, TX, November 2, 2011.

Invited Presenter, US-China NSF Workshop titled "Pathways Toward Low Carbon Cities: Quantifying Baselines and Interventions", Hong Kong, December 13-14, 2010.

Invited Lecture titled "Environmental Impacts of Land Use and Transportation Policies" for Europe's Kuhmo-Nectar Summer School in Transportation Economics, Valencia, Spain, July 2010.

Invited presentation, titled "Incorporating Environment and Equity into Evaluations of Land Use Transportation Systems" for UCLA Symposium titled "Transportation - Land Use -Environment Connection," Lake Arrowhead, CA, October 18-20, 2009.

Invited presentation, titled "Credit Based Congestion Pricing: An Opportunity for Reductions in Greenhouse Gas Emissions?" at the International Colloquium on Transport, Energy and Greenhouse Gases: Will Rationing be Necessary?, 19th Entretiens du Centre Jacques Cartier in Lyon, France, December 4-5, 2006.

Please visit http://www.caee.utexas.edu/prof/kockelman/ for access to pre-prints of all papers and complete vitae (listing all research projects and publications).

CURRICULUM VITAE

R. MICHAEL VAN AUKEN
Dynamic Research Inc.
355 Van Ness Avenue
Torrance, California 92691
(310) 212-5211

EDUCATION

Bachelor of Science in Aerospace Engineering,
University of Michigan, 1980

Master of Science in Aeronautical and Astronautical Engineering,
Stanford University, 1984

Engineer in Aeronautical and Astronautical Engineering,
Stanford University, 1987.

Thesis was on Model Reduction with Generalized Input/Output Weightings, a model reduction method for linearized road vehicle dynamics models with potential application to automotive active suspensions.

PROFESSIONAL EXPERIENCE

1989 to present

Principal Engineer
Dynamic Research Inc., Torrance, California

Involved in various technical activities in the areas of vehicle dynamics and control, crashworthiness and crash avoidance; including ride characteristics, handling, occupant injury assessments and technology effectiveness estimates, for automobiles, motorcycles, and ATVs. This involved mathematical modeling and computer simulation of driver and vehicle systems, data analysis, and interpretation of results; as well as full scale and component testing and model validation. Other activities have included development of tire-road math models; Fourier analysis, sound signal and other types of signal analysis; large scale, computer simulations of multi-body dynamics; finite element analysis; and various types of statistical analyses and experimental design.

Applications in these areas have included studies related to the development of transient ride and noise discomfort metrics for passenger cars, Weibull analysis of automobile transmission failures, and injury and fatality assessment of accident data. The ride and noise discomfort metric development included, for example, a large scale computational effort to analyze, correlate, and interpret driver and vehicle objective and subjective data. Similar studies have

recently been accomplished on the topic of driver attentional workload

1987 to 1989	Sr. Staff Engineer Hughes Aircraft Company, El Segundo, California

Involved in several automotive systems engineering tasks including defining vehicle performance requirements for a chassis subsystem; participating in the design, development, and verification of several vehicle dynamics systems, including developing tire-road contact force and moment math models; and identification of significant driver/vehicle feedback factors for driving simulators.

1980 to 1987	Sr. Research Engineer Lockheed Missiles & Space Company, Inc., Sunnyvale, California

Involved in the design, development, and verification of a attitude determination system based on optimal estimation methods, including allocating requirements to subsystems, simulation based trade studies, identified system anomalies and developed appropriate countermeasures, developed hardware math models, and refined simulations.

Refined and applied simulations to assess dynamic system performance under extreme non-linear disturbance conditions. Refined and applied simulations to assess guidance system performance and mission planning.

1979 to 1980	Engineer Airflow Sciences Corporation, Plymouth, Michigan

Assisted in the development of computational fluid dynamics (CFD) software. Used CFD software to solve specific aerodynamic problems. Assisted in wind tunnel test preparation.

TECHNICAL AND SCIENTIFIC PUBLICATIONS

Van Auken, R.M., Model Reduction with Generalized Input/Output Weighting, Engineer Thesis, Stanford University, Stanford, CA, May, 1987.

Van Auken, M., Zellner, J.W., and Kunkel, D.T., "Correlation of Zwicker's Loudness and Other Noise Metrics with Drivers' Over-the-Road Transient Noise Discomfort," Technical Paper 980585, SAE, 1998.

Kebschull, S.A., Zellner, J.W., Van Auken, M., Rogers, N.M., "Injury Risk/Benefit Analysis of Motorcycle Protective Devices Using Computer Simulation and ISO 13232", Paper Number 98-S10-W-26, Proc. 16th International Technical Conference on the

Enhanced Safety of Vehicles, Windsor, Canada, National Highway Traffic Safety Administration, Washington D.C., 1998, pp. 2357-2374.

Zellner, J.W., Kebschull, Van Auken, R.M., "Analysis of Vehicle Tip Stability in Side Impact Tests," Technical Paper 2000-01-1650, SAE, 2000.

Van Auken, R.M., Zellner, J.W., Kebschull, S.A., Wiley, K.D., Smith, T., Shewchenko, N, Rogers, N.M., "Development of Neck Injury Assessment Criteria for the ISO 13232 Motorcyclist Anthropometric Test Dummy with the Revised Neck," Paper No. 417, Proc. 18th International Technical Conference on the Enhanced Safety of Vehicles, Nagoya, Japan, May 2003.

Van Auken, R.M., and Zellner, J.W., "An Assessment of the Effects of Vehicle Weight and Size on Fatality Risk in 1985 to 1998 Model Year Passenger Cars and 1985 to 1997 Model Year Light Trucks and Vans," Technical Paper 2005-01-1354, SAE Transactions, Vol. 114, Part 6, Detroit, April 11-15, 2005, pp 1607-1622.

Van Auken, R.M., Zellner, J.W., Smith, T., Rogers, N.M., "Development of an Improved Neck Injury Assessment Criteria for the ISO 13232 Motorcyclist Anthropometric Test Dummy," Paper No. 05-0227, Proc. 19th International Technical Conference on the Enhanced Safety of Vehicles, 2005.

Van Auken, R.M., Kebschull, S.A., Broen, P.C., Zellner, J.W., Rogers, N.M., "Development of a Rider Size and Position Model to Determine Motorcycle Protective Device Test Conditions," Paper Number 05-0392, Proc. 19th International Technical Conference on the Enhanced Safety of Vehicles, 2005.

Suzuki, H., Sugimoto, Y., Van Auken, R.M., and Zellner, J.W., "Development of a Prototype Safety Analysis System to Assess and Forecast Vehicle Safety," SAE Technical Paper 2006-01-0718, 2006.

Van Auken, R.M., "Active Suspension Control Synthesis Using Reduced Order Models that Account for the Time Delay Between Road Inputs," ESDA2006-95570, Proc. 8th Biennial ASME Conference on Engineering Systems Design and Analysis, Torino, Italy, July 4-7, 2006.

Zellner, J.W., Van Auken, R.M., Kebschull, S.A., Mun oz, S. Injury Risk-Benefit Analysis of Rollover Protection Systems (ROPS) For All Terrain Vehicles (ATVs) Using Computer Simulation, Full-Scale Testing and ISO 13232, Paper F2008-08-009, Fisita World Congress, Munich, Germany, 2008.

Zellner, J.W., Van Auken, R.M., Chiang, D.P, Broen, P.C., Joseph Kelly, J.K., Sugimoto, Y , "Extension of the Honda-DRI "Safety Impact Methodology" (SIM) for the NHTSA Advanced Crash Avoidance Technology (ACAT) Program and Application to a Prototype Advanced Collision Mitigation Braking System," Technical Paper 2009-01-0781, *SAE Int. J. Passenger. Cars – Mech. Syst.* **2**(1):875-894, 2009.

Van Auken, R.M., "Development and Comparison of Laplace Domain and State-Space Models of a Half-Car with Flexible Body," ESDA2010-24518, Proc. 10th Biennial Conference on Engineering Systems Design and Analysis, Istanbul, Turkey, July 12-14, 2010.

Van Auken, R.M., Zellner, J.W., Silberling, J.Y., Kelly, J., Chiang, D.P., Broen, P., Kirsch, A., Sugimoto, Y., "Extension of the Honda-DRI Safety Impact Methodology for the NHTSA Advanced Crash Avoidance Technology (ACAT) Program and Application to the Evaluation of an Advanced Collision Mitigation Braking System – Final Results of the ACAT-I Program," Technical Paper 2011-01-0581, *SAE Int. J. Passenger Cars – Mech. Syst.* **4**(1):488-508, 2011.

Van Auken, R.M., Smith, T.A., Zellner, J.W., "Development of a Probabilistic Skull Fracture Model for a 50th Percentile Adult Male Motorcyclist ATD Headform," Paper Number 11-0035, Proc. 21st International Technical Conference on the Enhanced Safety of Vehicles, Washington, June 2011.

Van Auken, R.M., Zellner, J.W., Silberling, J.Y., Sugimoto, Y., Urai, Y., "Progress Report on Evaluation of a Pre-Production Head-On Crash Avoidance Assist System Using an Extended "Safety Impact Methodology" (SIM)," Paper Number 11-0270, Proc. 21st International Technical Conference on the Enhanced Safety of Vehicles, Washington, June 2011.

Sugimoto, Y., Sugimoto, Y., Kawakami, K., Hashimoto, T., Zellner, J.W., Van Auken, R.M., "Safety Impact Methodology for Advanced Crash Avoidance Technology (ACAT) Program and Application to Advanced Collision Mitigation Braking System and Head-on Crash Avoidance Assist System," First International Symposium on Future Active Safety Technology Toward Zero-Traffic-Accident, JSAE, Tokyo, September 5-9, 2011.

Van Auken, R.M., Zellner, J.W., "Extension of the Honda-DRI "Safety Impact Methodology" (SIM) for the NHTSA Advanced Crash Avoidance Technology (ACAT) II Program and Application to the Evaluation of a Preproduction Head-On Crash Avoidance Assist System," Technical Paper 2012-01-0291, SAE, forthcoming.

PUBLISHED TECHNICAL REPORTS

VanAuken, R. M. and Zellner, J. W., ATV ROPS Tests and Simulations, DRI-TR-98-2, Dynamic Research, Inc., Torrance, California, October 1998.

Van Auken, R.M., and Zellner, J.W., An Assessment of the Effects of Vehicle Weight and Size Parameters on Fatality Risk in Model Year 1985-98 Passenger Cars and 1985-97 Light Trucks, DRI-TR-02-02, Dynamic Research, Inc., Torrance, California, DOT Docket No. NHTSA-2003-16318-2, February 2002.

Van Auken, R.M., and Zellner, J.W., A Further Assessment of the Effects of Vehicle Weight on Fatality Risk in Model Year 1985-98 Passenger Cars and 1985-97 Light Trucks, DRI-TR-03-01, Dynamic Research, Inc., Torrance, California, DOT Docket No. NHTSA-2003-16318-3, January 2003.

Van Auken, R.M., and Zellner, J.W., A Review of the Results in the 1997 Kahane, 2002 DRI, 2003 DRI, and 2003 Kahane Reports on the Effects of Passenger Car and Light Truck Weight and Size on Fatality Risk, DRI-TR-04-02, Dynamic Research, Inc., Torrance, California, DOT Docket No. NHTSA-2003-16318-7, March 2004.
Kebschull, S.A., Kelly, J., Van Auken, R.M., and Zellner, J.W., An Analysis of the Effects of SUV Weight and Length on SUV Crashworthiness and Compatibility Using Systems Modeling and Risk-Benefit Analysis, DRI-TR-04-04-2, Dynamic Research, Inc., Torrance, California, DOT Docket No. NHTSA-2003-16128-1452, July 2004.

Van Auken, R.M., and Zellner, J.W., Supplemental Results on the Independent Effects of Curb Weight, Wheelbase, and Track on Fatality Risk in 1985-1998 Model Year Passenger Cars and 1985-1997 Model Year LTVs, DRI-TR-05-01, Dynamic Research, Inc., Torrance, California, Federal Docket No. NHTSA-2003-16318-17, 20 May 2005.

Van Auken, R.M., and Zellner, J.W., DRI Comments On Safety Impacts of EPA-NHTSA Proposed Rule To Establish Light-Duty Vehicle Greenhouse Gas Emission Standards and Corporate Average Fuel Economy Standards Dated September 28 2009, DRI-TM-09-86, Dynamic Research, Inc., Torrance, California, Federal Docket No. NHTSA-2009-0059-112.1, 27 November 2009.

Van Auken, R.M., Zellner, J.W., Chiang, D.P., Kelly, J., Silberling, J.Y., Dai, R., Broen, P.C., Kirsch, A.M., Sugimoto, Y., Advanced Crash Avoidance Technologies Program – Final Report of the Honda-DRI Team, Volume I: Executive Summary and Technical Report, DOT HS 811 454, Accomplished by Honda and DRI for the National Highway Traffic Safety Administration, Washington, June, 2011.

Van Auken, R.M., and Zellner, J.W., Updated Analysis of the Effects of Passenger Vehicle Size and Weight On Safety, Phase I: Updated Analysis Based on 1995 to 2000 Calendar Year Data for 1991 to 1999 Model Year Light Passenger Vehicles, DRI-TR-11-01, Dynamic Research Inc., Torrance, California, Federal Docket No. NHTSA-2010-0152-0030, January 2011.

Zellner, J.E., Van Auken, R.M., Kelly, J., Silberling, J.Y., Hagoski, B.K., Sugimoto, Y., Advanced Crash Avoidance Technologies (ACAT) II Program – Final Report of the Honda-DRI Team, Accomplished by Honda and DRI for the National Highway Traffic Safety Administration, Washington, D.C., forthcoming.

Van Auken, R.M., and Zellner, J.W., Updated Analysis of the Effects of Passenger Vehicle Size and Weight On Safety, Phase II: Analysis Based on 2002 to 2008 Calendar Year Data for 2000 to 2007 Model Year Light Passenger Vehicles, DRI-TR-12-01, Dynamic Research Inc., Torrance, California, forthcoming.

PUBLIC PRESENTATIONS

Van Auken, R.M., and Zellner, J.W., "Effects of Vehicle Weight and Size Reductions on Safety," P03-6853, Transportation Research Board 82nd Annual Meeting, Washington, D.C., January 15, 2003.

Van Auken, R.M., and Erwin, S., "Injury Severity in Side Impact Mismatch," CIREN Quarterly Meeting, Washington, D.C., April 3, 2003.

Van Auken, R.M., and Zellner, J.W. "Updated Analysis of the Effects of Passenger Vehicle Size and Weight on Safety," NHTSA Workshop on Vehicle Mass-Size-Safety, Washington, D.C., February 25, 2011.

PEER REVIEWS

Proc. 8th Biennial ASME Conference on Engineering Systems Design and Analysis, Torino, Italy, July 4-7, 2006.

Proc. 10th Biennial Conference on Engineering Systems Design and Analysis, Istanbul, Turkey, July 12-14, 2010.

Appendix B: Conflict of Interest Statements

Conflict of Interest and Bias for Peer Review

Background

Identification and management of potential conflict of interest (COI) and bias issues are vital to the successes and credibility of any peer review consisting of external experts. The questionnaire that follows is consistent with EPA guidance concerning peer reviews.[1]
.

Definitions

Experts in a particular field will, in many cases, have existing opinions concerning the subject of the peer review. These opinions may be considered bias, but are not necessarily conflicts of interest.

Bias: For a peer review, means a predisposition towards the subject matter to be discussed that could influence the candidate's viewpoint.

Examples of bias would be situations in which a candidate:

1. Has previously expressed a position on the subject(s) under consideration by the panel; or

2. Is affiliated with an industry, governmental, public interest, or other group which has expressed a position concerning the subject(s) under consideration by the panel.

Conflict of Interest: For a peer review, as defined by the National Academy of Sciences,[2] includes any of the following:

1. Affiliation with an organization with financial ties directly related to the outcome;

2. Direct personal/financial investments in the sponsoring organization or related to the subject; or

3. Direct involvement in the documents submitted to the peer review panel... that could impair the individual's objectivity or create an unfair competitive advantage for the individual or organization.

---

[1] U.S. EPA (2009). Science Policy Council Peer Review Handbook.

OMB (2004). Final Information Quality Bulletin for Peer Review.

[2] NAS (2003). "Policy and Procedures on Committee Composition and Balance and Conflict or Interest for Committees Used in the Development of Reports" (www.nationalacademies.org/coi).

Policy and Process

- Candidates with COI, as defined above, will not be eligible for membership on those panels where their conflicts apply.

- In general, candidates with bias, as defined above, on a particular issue will be eligible for all panel memberships; however, extreme biases, such as those likely to impair a candidate's ability to contribute to meaningful scientific discourse, will disqualify a candidate.

- Ideally, the composition of each panel will reflect a range of bias for a particular subject, striving for balance.

- Candidates who meet scientific qualifications and other eligibility criteria will be asked to provide written disclosure through a confidential questionnaire of all potential COI and bias issues during the candidate identification and selection process.

- Candidates should be prepared, as necessary, to discuss potential COI and bias issues.

- All bias issues related to selected panelists will be disclosed in writing in the final peer review record.

Conflict of Interest and Bias Questionnaire

Peer Review of LBNL Reports on Vehicle Mass/Footprint Reduction & Safety

Instructions to Candidate Reviewers

1. Please check YES/NO/DON'T KNOW in response to each question.

2. If your answer is YES or DON'T KNOW, please provide a brief explanation of the circumstances.

3. Please make a reasonable effort to answer accurately each question. For example, to the extent a question applies to individuals (or entities) other than you (e.g., spouse, dependents, or their employers), you should make a reasonable inquiry, such as emailing the questions to such individuals/entities in an effort to obtain information necessary to accurately answer the questions.

Questions

1. Are you (or your spouse/partner or dependents) or your current employer, an author, contributor, or an earlier reviewer of the document(s) being reviewed by this panel?

   YES___   NO_X_   DON'T KNOW___

2. Do you (or you spouse/partner or dependents) or your current employer have current plans to conduct or seek work related to the subject of this peer review following the completion of this peer review panel?

   YES_X_   NO___   DON'T KNOW___

3. Do you (or your spouse/partner or dependents) or your current employer have any known financial stake in the outcome of the review *(e.g.,* investment interest in a business related to the subject of peer review)?

   YES___   NO_X_   DON'T KNOW___

4. Have you (or your spouse/partner or dependents) or your current employer commented, reviewed, testified, published, made public statements, or taken positions regarding the subject of this peer review?

   YES___   NO___   DON'T KNOW_X_

5. Do you hold personal values or beliefs that would preclude you from conducting an objective, scientific evaluation of the subject of the review?

   YES___   NO_X_   DON'T KNOW___

6. Do you know of any reason that you might be unable to provide impartial advice or comments on the subject review of the panel?

   YES___   NO_X_   DON'T KNOW___

7. Are you aware of any other factors that may create potential conflict of interest or bias issues for you as a member of the panel?

   YES___   NO_X_   DON'T KNOW___

Acknowledgment

I declare that the disclosed information is true and accurate to the best of my knowledge, and that no real, potential, or apparent conflict of interest or bias is known to me except as disclosed. I further declare that I have made reasonable effort and inquiry to obtain the information needed to answer the questions truthfully, and accurately. I agree to inform SRA promptly of any change in circumstances that would require me to revise the answers that I have provided.

Tong Donna Chen
Name

_____   11/27/2011
Signature   Date

Conflict of Interest and Bias Questionnaire

Peer Review of LBNL Reports on Vehicle Mass/Footprint Reduction & Safety

Instructions to Candidate Reviewers

1. Please check YES/NO/DON'T KNOW in response to each question.

2. If your answer is YES or DON'T KNOW, please provide a brief explanation of the circumstances.

3. Please make a reasonable effort to answer accurately each question. For example, to the extent a question applies to individuals (or entities) other than you (e.g., spouse, dependents, or their employers), you should make a reasonable inquiry, such as emailing the questions to such individuals/entities in an effort to obtain information necessary to accurately answer the questions.

Questions

1. Are you (or your spouse/partner or dependents) or your current employer, an author, contributor, or an earlier reviewer of the document(s) being reviewed by this panel?

   YES_X__   NO___   DON'T KNOW___

   [I am a peer reviewer for the NHTSA study "Relationships Between Fatality Risk, Mass, and Footprint in Model Year 200-2007 Passenger Cars and LTVs."]

2. Do you (or you spouse/partner or dependents) or your current employer have current plans to conduct or seek work related to the subject of this peer review following the completion of this peer review panel?

   YES_X__   NO___   DON'T KNOW___

   [The Insurance Institute for Highway Safety has in the past and will continue to study the relationship between vehicle size/weight and crash injury risk.]

3. Do you (or your spouse/partner or dependents) or your current employer have any known financial stake in the outcome of the review *(e.g.,* investment interest in a business related to the subject of peer review)?

   YES___   NO_X__   DON'T KNOW___

4. Have you (or your spouse/partner or dependents) or your current employer commented, reviewed, testified, published, made public statements, or taken positions regarding the subject of this peer review?

YES_X__   NO___   DON'T KNOW___

[The Insurance Institute for Highway Safety has in the past and will continue to study the relationship between vehicle size/weight and crash injury risk.]

5. Do you hold personal values or beliefs that would preclude you from conducting an objective, scientific evaluation of the subject of the review?

    YES___   NO_X__   DON'T KNOW___

6. Do you know of any reason that you might be unable to provide impartial advice or comments on the subject review of the panel?

    YES___   NO_X__   DON'T KNOW___

7. Are you aware of any other factors that may create potential conflict of interest or bias issues for you as a member of the panel?

    YES___   NO_X__   DON'T KNOW___

Acknowledgment

I declare that the disclosed information is true and accurate to the best of my knowledge, and that no real, potential, or apparent conflict of interest or bias is known to me except as disclosed. I further declare that I have made reasonable effort and inquiry to obtain the information needed to answer the questions truthfully, and accurately. I agree to inform SRA promptly of any change in circumstances that would require me to revise the answers that I have provided.

Charles M. Farmer, Ph.D.
Director of Statistical Services
Insurance Institute for Highway Safety
Arlington, VA
Name

_____   October 24, 2011____
Signature   Date

Conflict of Interest and Bias Questionnaire

Peer Review of LBNL Reports on Vehicle Mass/Footprint Reduction & Safety

Instructions to Candidate Reviewers

1. Please check YES/NO/DON'T KNOW in response to each question.

2. If your answer is YES or DON'T KNOW, please provide a brief explanation of the circumstances.

3. Please make a reasonable effort to answer accurately each question. For example, to the extent a question applies to individuals (or entities) other than you (e.g., spouse, dependents, or their employers), you should make a reasonable inquiry, such as emailing the questions to such individuals/entities in an effort to obtain information necessary to accurately answer the questions.

Questions

1. Are you (or your spouse/partner or dependents) or your current employer, an author, contributor, or an earlier reviewer of the document(s) being reviewed by this panel?

   YES___   NO_X_   DON'T KNOW___

2. Do you (or you spouse/partner or dependents) or your current employer have current plans to conduct or seek work related to the subject of this peer review following the completion of this peer review panel?

   YES___   NO_X_   DON'T KNOW___

3. Do you (or your spouse/partner or dependents) or your current employer have any known financial stake in the outcome of the review *(e.g.,* investment interest in a business related to the subject of peer review)?

   YES___   NO_X_   DON'T KNOW___

4. Have you (or your spouse/partner or dependents) or your current employer commented, reviewed, testified, published, made public statements, or taken positions regarding the subject of this peer review?

   YES_X_   NO___   DON'T KNOW___

5. Do you hold personal values or beliefs that would preclude you from conducting an objective, scientific evaluation of the subject of the review?

   YES___   NO_X_   DON'T KNOW___

6. Do you know of any reason that you might be unable to provide impartial advice or comments on the subject review of the panel?

   YES___   NO_X_   DON'T KNOW___

7. Are you aware of any other factors that may create potential conflict of interest or bias issues for you as a member of the panel?

   YES___   NO_X_   DON'T KNOW___

Acknowledgment

I declare that the disclosed information is true and accurate to the best of my knowledge, and that no real, potential, or apparent conflict of interest or bias is known to me except as disclosed. I further declare that I have made reasonable effort and inquiry to obtain the information needed to answer the questions truthfully, and accurately. I agree to inform SRA promptly of any change in circumstances that would require me to revise the answers that I have provided.

David L. Greene
Name

_____   10/31/2011
Signature   Date

Conflict of Interest and Bias Questionnaire

Peer Review of LBNL Reports on Vehicle Mass/Footprint Reduction & Safety

Instructions to Candidate Reviewers

1. Please check YES/NO/DON'T KNOW in response to each question.

2. If your answer is YES or DON'T KNOW, please provide a brief explanation of the circumstances.

3. Please make a reasonable effort to answer accurately each question. For example, to the extent a question applies to individuals (or entities) other than you (e.g., spouse, dependents, or their employers), you should make a reasonable inquiry, such as emailing the questions to such individuals/entities in an effort to obtain information necessary to accurately answer the questions.

Questions

1. Are you (or your spouse/partner or dependents) or your current employer, an author, contributor, or an earlier reviewer of the document(s) being reviewed by this panel?

   YES___   NO_X_   DON'T KNOW___

2. Do you (or you spouse/partner or dependents) or your current employer have current plans to conduct or seek work related to the subject of this peer review following the completion of this peer review panel?

   YES___   NO_X_   DON'T KNOW___

3. Do you (or your spouse/partner or dependents) or your current employer have any known financial stake in the outcome of the review *(e.g.,* investment interest in a business related to the subject of peer review)?

   YES___   NO_X_   DON'T KNOW___

4. Have you (or your spouse/partner or dependents) or your current employer commented, reviewed, testified, published, made public statements, or taken positions regarding the subject of this peer review?

   YES___   NO___   DON'T KNOW_X_

   [I have a paper (to be presented at TRB) about crash safety, controlling for vehicle size, fuel economy, etc.]

5. Do you hold personal values or beliefs that would preclude you from conducting an objective, scientific evaluation of the subject of the review?

   YES___   NO_X_   DON'T KNOW___

6. Do you know of any reason that you might be unable to provide impartial advice or comments on the subject review of the panel?

   YES___   NO_X_   DON'T KNOW___

7. Are you aware of any other factors that may create potential conflict of interest or bias issues for you as a member of the panel?

   YES___   NO_X_   DON'T KNOW___

Acknowledgment

I declare that the disclosed information is true and accurate to the best of my knowledge, and that no real, potential, or apparent conflict of interest or bias is known to me except as disclosed. I further declare that I have made reasonable effort and inquiry to obtain the information needed to answer the questions truthfully, and accurately. I agree to inform SRA promptly of any change in circumstances that would require me to revise the answers that I have provided.

Kara Kockelman_____
Name

_____   10/21/11___
Signature   Date

Conflict of Interest and Bias Questionnaire

Peer Review of LBNL Reports on Vehicle Mass/Footprint Reduction & Safety

Instructions to Candidate Reviewers

1. Please check YES/NO/DON'T KNOW in response to each question.

2. If your answer is YES or DON'T KNOW, please provide a brief explanation of the circumstances.

3. Please make a reasonable effort to answer accurately each question. For example, to the extent a question applies to individuals (or entities) other than you (e.g., spouse, dependents, or their employers), you should make a reasonable inquiry, such as emailing the questions to such individuals/entities in an effort to obtain information necessary to accurately answer the questions.

Questions

1. Are you (or your spouse/partner or dependents) or your current employer, an author, contributor, or an earlier reviewer of the document(s) being reviewed by this panel?

   YES___   NO_X_   DON'T KNOW___

2. Do you (or you spouse/partner or dependents) or your current employer have current plans to conduct or seek work related to the subject of this peer review following the completion of this peer review panel?

   YES___   NO_X_   DON'T KNOW___

3. Do you (or your spouse/partner or dependents) or your current employer have any known financial stake in the outcome of the review *(e.g.,* investment interest in a business related to the subject of peer review)?

   YES___   NO_X_   DON'T KNOW___

4. Have you (or your spouse/partner or dependents) or your current employer commented, reviewed, testified, published, made public statements, or taken positions regarding the subject of this peer review?

   YES___   NO_X_   DON'T KNOW___

5. Do you hold personal values or beliefs that would preclude you from conducting an objective, scientific evaluation of the subject of the review?

   YES___   NO__X_   DON'T KNOW___

6. Do you know of any reason that you might be unable to provide impartial advice or comments on the subject review of the panel?

   YES___   NO__X_   DON'T KNOW___

7. Are you aware of any other factors that may create potential conflict of interest or bias issues for you as a member of the panel?

   YES___   NO__X_   DON'T KNOW___

Acknowledgment

I declare that the disclosed information is true and accurate to the best of my knowledge, and that no real, potential, or apparent conflict of interest or bias is known to me except as disclosed. I further declare that I have made reasonable effort and inquiry to obtain the information needed to answer the questions truthfully, and accurately. I agree to inform SRA promptly of any change in circumstances that would require me to revise the answers that I have provided.

R. Michael Van Auken
Name

_____   2011-10-24___
Signature   Date

Appendix C: Peer Review Charge

Charge to Peer Reviewers of LBNL Statistical Analysis of the Effect of Vehicle Mass & Footprint Reduction on Safety

In developing programs to reduce greenhouse gas (GHG) emissions and increase fuel economy of light-duty highway vehicles, the U.S. Environmental Protection Agency (EPA) and the National Highway Transportation Safety Administration (NHTSA) have to evaluate the safety of mass reduction technologies likely to be used to meet future standards. The U.S. Department of Energy (DOE) has contracted with Lawrence Berkeley National Laboratory (LBNL) to perform a statistical analysis of the effect of vehicle mass and footprint reduction on safety. LBNL's analysis of the relationship between vehicle mass, footprint, and societal fatality and casualty risk is comprised of two phases. Phase 1 is an assessment of the NHTSA report *Relationships between Fatality Risk, Mass, and Footprint in Model Year 2000-2007 Passenger Cars and LTVs.* This study uses logistic regression analysis to estimate the relationship of changes in vehicle mass and footprint on US fatality risk per vehicle mile traveled. Phase 2 is an independent logistic regression analysis to estimate the relationship between vehicle mass, footprint and total casualty (fatality plus serious injury) risk, per police-reported crash, using state-level data on all crashes.

The focus of this peer review is to evaluate the assumptions made, data and methods of statistical analysis used, and conclusions from the analysis, for both the LBNL assessment of the NHTSA study and the independent LBNL study using state-level data. A comprehensive peer review by third party experts is an important step for validation of the results of the studies, and how the results of the studies are used in modeling the effect of new fuel economy and greenhouse gas emission standards on vehicle safety. In order to review the LBNL analysis, you should understand and be familiar with the data, assumptions, conclusions, and statistical approach used in the NHTSA 2011 report "Relationship between Fatality Risk, Mass, and Footprint in Model Year 2000-2007 Passenger Cars and LTVs." Also, you may want to review the 2003 NHTSA report http://www.nhtsa.gov/cars/rules/regrev/evaluate/pdf/809662.pdf and the 2010 NHTSA report http://www.nhtsa.gov/staticfiles/rulemaking/pdf/cafe/CAFE_2012-2016_FRIA_04012010.pdf (beginning page 464), depending upon your familiarity with them.

You are asked to review and provide expert comments on the LBNL Phase 1 and LBNL Phase 2 draft reports described above. You are being provided LBNL's Phase 1 draft report and the NHTSA 2011 report, which it addresses. Work is being completed on the Phase 2 study, and a draft report is expected to be available for review by November 22.

EPA is seeking peer reviewers' expert opinions on the statistic methodologies used in the two LBNL studies and whether they are likely to yield realistic estimates of the relationship between vehicle mass, footprint, and total fatality or casualty risk. EPA requests that each reviewer comment on all aspects of the two LBNL studies, with particular emphasis on the methodologies employed, assumptions inherent to the analysis, sources of information employed, methods of calculation and any other key issues the reviewer may identify. Reviewers are encouraged to examine and evaluate the NHTSA study in helping

them to understand the LBNL assessment analysis. Findings of this peer review may be used toward validation and improvement of the statistical analysis conducted by LBNL, and to inform EPA staff on potential use of the regression results for predicting the safety effect of future standards in reducing mass and footprint. No independent data analysis will be required for this review.

Reviewers are asked to orient their comments toward these five general areas: (1) assumptions; (2) control and dependent variables used in the regression models; (3) methodology and statistics; (4) data sets; and (5) recommendations. Possible topics are provided in each area as illustrative examples. Reviewers are expected to identify additional topics or depart from these examples as necessary to best apply their particular set of expertise toward review of the LBNL reports.

Please note that the author intends to make the database and statistical programs available to the public to ensure that the assumptions made and the methods of calculation are transparent and replicable. Thus it will be helpful for reviewers to recommend improvements to the analysis that would utilize publicly available information rather than those that would make use of proprietary information.

Comments should be sufficiently clear and detailed to allow readers to thoroughly understand their relevance to the LBNL studies. ***Please deliver your final written comments to SRA International no later than Thursday, December 22.***

All materials provided to reviewers as well as reviewer comments should be treated as confidential, and should neither be released nor discussed with others outside of the review panel. Once EPA, LBNL, and NHTSA have made their reports and supporting documentation public, EPA will notify reviewers that they may release or discuss the peer review materials and their review comments with others.

Should reviewers have questions about what is required in order to complete this review or need additional background material, please contact Brian Menard at SRA (Brian_Menard@sra.com) or (434-817-4133). If a reviewer has any questions about the EPA peer review process itself, please contact Ms. Ruth Schenk in EPA's Quality Office, National Vehicle and Fuel Emissions Laboratory (schenk.ruth@epa.gov) or (734-214-4017).

Appendix D: Reviews

Donna Chen & Kara Kockelman Review of LBNL Phase 1 Report	
1. ASSUMPTIONS	**COMMENTS**
Please comment on the validity of any assumptions embedded in the LBNL assessment analysis and the independent casualty analysis that could affect the projected relationship between vehicle mass/footprint reductions and fatality/casualty risk. Examples might include assumptions regarding whether recent historical relationships between vehicle weight, size, and safety will continue into the future; potential future improvements in vehicle technology and design may result in compensatory safety benefits; and the annual baseline fatality distribution.	The report does a nice job discussing recent trends in vehicles, such as the increase of ESC, side airbags, and light truck crash compatibility with passenger cars – which will improve safety outcomes for all vehicles, but perhaps most significantly the smaller and lighter vehicles. . It also mentions the phasing out of the lightest and smallest vehicles between model years 2000-2007 (but doesn't mention the makes and models somehow), which were particularly poor safety performers in the past. However, with the introduction of urban commuter vehicles, such as the SmartCar, Mini Cooper, and Fiat 500m, and the growing popularity of smaller, fuel-efficient compact vehicles following gas price increases, this trend does not seem so obvious. Such vehicles should be discussed. The simplistic logistic model employed in this analysis only accounts for two crash outcomes (fatal versus non-fatal) and so neglects the more detailed, and ordered nature of injury severity data, which is unfortunate. The model also assumes error-term homoscedasticity from one crash or individual to the next; in reality certain vehicle types (e.g., pickups) and crash contexts (e.g., high speed crashes) have more uncertainty associated with their severity outcomes. It would be good to point out such limitations for readers.
Please comment on any apparent unstated or implicit assumptions and related caveats or limitations.	The role of driver behavior is briefly addressed in the report but not emphasized sufficiently. Fatality risk is a combination of driver, vehicle, and roadway characteristics. Driver behavioral differences are many and do not solely exist for pickup truck drivers versus car drivers. Socioeconomic data such driver household income, size, and education influence driver attitudes and driving environments. For example, Chen et al. (2010) found that crash risk increases for those living in socioeconomically disadvantaged areas (including households more likely to drive less expensive and older vehicles). Though such data is not typically available in state and national crash

	databases, the importance of these driver and environmental characteristics on crash rates (per mile driven) and fatality risk should be stressed in both reports. It is clearly very difficult to control for, but a major caveat to the NHTSA (& now LBNL) results. We expect that crash severity could be probably be lower for many of the small cars and pickups if they were driven by those who tend to drive more expensive vehicles, under the same settings (e.g., daytime, urban freeway). Similarly, in the second LBNL report (which uses VMT estimates), we expect that crash rates would probably be lower for these types of driver-vehicle-setting combinations.

ADDITIONAL COMMENTS:

Chen, H.Y., Ivers, R.Q., Mariniuk, A.L.C., Boufous, S., Senserrick, T., Woodward, M., Stevenson, M. and Norton R. Socioeconomic status and risk of car crash injury, independent of place of residence and driving exposure: Results from the DRIVE study. Journal of Epidemiology and Community Health 64(10), 2010, pp. 998-1003.

2. CONTROL AND DEPENDENT VARIABLES USED IN THE REGRESSION MODELS	COMMENTS
Please comment on the adequacy of control and dependent variables used in the assessment analysis and independent casualty analysis, and recommend any alternative control or dependent variables that are available for possible inclusion in the analysis. For example, what are the relative merits of the main dependent variables used, fatality risk per estimated VMT, and casualty risk per police-reported crash?	As alluded to above, a primary concern is that the NHTSA analysis (& thus the LBNL analyses) largely neglect the idea that vehicle type (make & model) is very much a proxy for driver type, and a vehicle's crash avoidance may have very little to do with vehicle type. It has a lot to do with the person behind the wheel, and gender & age simply aren't enough to control for such distinctions. Education, risk aversion, ability, wealth, etc., are important covariates. But existing data sets are quite limiting (though the MVOSS & FAR with 3-year driver violation history do offer some valuable insights, not discussed in these reports). In reality, small cars may be less crash prone than Kahane's & Wenzel's results suggest, because they are driven by lower-income, younger, less risk averse people driving in more crash prone settings (e.g., commercial strips rather than pricey residential suburbs). Such key caveats need thoughtful discussion. Four relevant papers on the topics of crash frequency and vehicle size-and-weight implications (by Knipling, Kweon & Kockelman, Wang and Kockelman, and Chen & Kockelman) have been sent to Tom Wenzel. These all include useful literature reviews for further connections to useful findings for citation in the reports, as time allows the contractor.
	The grouping of the vehicles into heavier- and lighter-than-average weight categories essentially splits a "typical" weight vehicle of that type into two categories. The impacts of curb weight and footprint on fatality risk may be easier to interpret if the vehicles were grouped into 3 weight categories (light, average, and heavy) & by type (with the average category representing vehicles within one standard deviation of average weight). Furthermore, the grouping of CUVs and minivans into the same vehicle type category neglects the fact that these vehicles have faced rather different ground clearance requirements (impacting rollover potential), door types (sliding vs. standard), and, perhaps most importantly, can appeal to different types of drivers (as indicated in the market shift of car drivers to CUV drivers).

What additional control variables, such as vehicle make or model, might be included in the regression models?	Vehicle height, a variable which may be more valuable than vehicle type for similarly structured vehicles such as sedans, wagons, CUVs, and minivans, would be a valuable control variable. In addition to a wider track, a lower center of gravity also increases vehicle stability, thereby reducing the risk of rollover. Relevant literature & findings exist, and should be cited. Other variables which have been found in past studies to influence fatality risk such as seat belt use, roadway geometry and division type are not included in this study (which is largely a repeat of the NHTSA study, as specifically contracted by the EPA). To account for driver characteristics that contribute to fatality risk, socioeconomic variables such as household income, education, household size, etc. would be valuable additions. Unfortunately, both state and national crash databases typically do not include such information (outside of MVOSS). Such issues should be flagged for readers. It seems the contractor has done his duty, and the key limitations lie with the original methodology he was to essentially duplicate.
Please comment on any caveats or limitations that these dependent variable or control variables entail with respect to use of the results as the basis for estimating the safety effect of mass reduction.	Please see above comment (in Assumptions section) regarding driver behavior and environment.
ADDITIONAL COMMENTS: Table 2.1 has many indicator variables labeled as "C" for continuous variable (such as ABS, ESC, AWD, DRVMALE, etc). These C's should be removed.	

3. METHODOLOGY AND STATISTICS	COMMENTS
Please comment on the validity and applicability of the methodology LBNL used in assessing the NHTSA 2011 study and its analysis of the relationship between mass, footprint, and risks per police-reported crash.	The report assesses the NHTSA 2011 study in a fair amount of detail and seeks to introduce some additional analyses to better examine the relationship between mass, footprint, and fatality risks. However, due to a lack of control for very specific vehicle differences (which vary by make & sub-model), the exclusion of driver characteristics and crash setting details (which cannot always be controlled for, but are often correlated with vehicle type), the effects of downweighting vehicles and/or shifting vehicle styles and sizes may be overestimated. Simply changing the vehicle on a risky driver in a high-risk setting is unlikely to influence outcomes significantly.
Please review other statistical methods LBNL has used in the analysis, in addition to the logistic regression methodology. Examples include the alternative approaches used by LBNL to assess NHTSA interval estimation results, and LBNL's linear regression analysis of actual, predicted, and residual risk by vehicle model.	In the alternative measures of exposure, the author examines the effect of vehicle manufacturer on fatality risk and treats the luxury models produced by Toyota, Honda, and Nissan as separate manufacturers. However, domestic luxury brands (such as Cadillac & Lincoln) are categorized with their nameplate manufacturers (GM and Ford), which appears inconsistent.

The effect of calendar year variables on fatality risk may be overestimated here, since VMT is tracked by vehicle model and not by calendar year. The trend of greatest fatality risk reductions in light trucks, CUVs and minivans with increasing calendar year may simply be a reflection of rising gas prices in combination with the ailing economy contributing to lower VMT (in these relatively low-fuel-economy vehicles).

It is unclear how the author determined the various percentage replacements of vehicle types in the aggressive vehicle market share shift scenario. (For example, why are 50% of SUVs replaced by CUVs and 60% of small pickups replaced by CUVs? The CUV is a more natural replacement for an SUV, and an SUV a more natural replacement for a pickup.) |

Please comment on caveats or limitations of using non-significant regression estimates to project the safety impact of mass reduction.	First, the t-statistics are not provided in the report which makes it difficult for the reader to assess statistical significance of specific regression estimates (except where noted by the author). Second, inclusion of a statistically insignificant variable can influence the estimates of coefficients associated with related variables. Nevertheless, in general, it is best to keep insignificant estimates if one has a strong defense for their role, since removing such variables (& thus their parameters) will shift the burden of response to a correlated covariate's parameter, thus biasing the latter. We generally keep key covariates in a model up to a pvalue of 0.20 or 0.25 or so, especially in relatively small data sets (e.g., n < 1,000). Covariates for which we have no strong basis can be removed for pvalues > 0.10.
How might the LBNL methodology be strengthened to better represent future vehicle designs and reduce multi-collinearity between mass and footprint in the regression analysis?	Including more vehicle-specific characteristics (such as vehicle height and engine size) reduces the analysis' dependence on vehicle type, since vehicle shapes and structures will continue to evolve. There is also correlation with context (e.g., pickups are driven in more rural locations, with greater hazards [like less lighting, higher speed, & few medians]). Disaggregate data are almost always best, to avoid ecological fallacies & such.
ADDITIONAL COMMENTS: On page 55, it is unclear what is meant by "however; if anything, reduction of this type of fatality will increase detrimental effect of mass reduction in cars."	

4. DATA SETS	COMMENTS
Please comment on the validity and applicability of the datasets used to project changes in risk resulting from reduction in vehicle mass. LBNL's casualty analysis used police-reported crash data from 16 states, while the 2011 NHTSA study used national fatality data, combined with a subset of non-culpable vehicles involved in two-vehicle crashes from police-reported crash data from 13 states.	The acquisition of Polk data for VMT estimates by make & model is valuable, and a contribution to the literature. However, these estimates come from vehicles found in repair shops in non-attainment areas, and so will be biased towards problem-prone vehicles, wealthier households who service their vehicles more regularly, and/or urban (smoggier) areas. Such issues merit careful discussion in the paper, so that readers are well aware of caveats. Related to this, Tom Wenzel indicated (by phone) that he did take a look at CA's extensive odometer reads, which go into some semi-rural locations (not too rural), and he indicated that the VMT values by vehicle type (not controlling for HH attributes & such) are very similar (just 5% longer in rural areas) – except for vans (which are used much more extensively in rural areas). This is interesting to me, and is not that different from what we've seen in the past. For example, Kockelman & Zhao's JTS paper from 2000 (pre-print at http://www.ce.utexas.edu/prof/kockelman/public_html/BTSJournalLDTs.pdf) suggests that, after controlling for various HH attributes & vehicle types, density is/was still very important (tables 1 & 2), but a shift from a density of 1k to 5k persons per sq mile (which is 1.5 vs. 7.8 persons per acre) means an increase of 750 mi/yr/vehicle (which is about 7.5% of annual VMT). Such differences, and their practical significance (or lack thereof) should be discussed in the reports.
Please comment on any apparent, unstated, or implicit impact on estimated risks inherent in the two different approaches, and any related caveats or limitations. For example, what are the strengths and weaknesses of the two measures of vehicle exposure, miles of vehicle travelled scaled up from crash data from 13 states, and number of police-reported crashes?	The use of non-culpable vehicles in two-vehicle crashes as a proxy for vehicles which are "just there" may be distorting the overall distribution of vehicle models. VMT may differ between vehicles that are more prone to run-off-road accidents, at-fault two-car crash vehicles, and non-culpable vehicles.
ADDITIONAL COMMENTS:	

5. RECOMMENDATIONS	COMMENTS
Please comment on whether the LBNL assessment adequately addresses the NHTSA 2011 study and identifies the safety impact from mass reduction. Are the analytic methods and data used to assess the NHTSA study, and estimate the relationship between risk, mass, and footprint, appropriate? Is casualty risk per crash a legitimate measure of vehicle safety? What other methods or data could be used to better predict the effect of future vehicle designs on safety?	While driver fatalities per crash seems a useful measure of vehicle design safety, and examination of fatal crash rates is very valuable (using Polk-based exposure estimates), there are many caveats to work of this type. As noted above: a primary concern remains a neglect of the notion that the type of car is very much a proxy for driver type, and a vehicle's crash avoidance may have very little to do with vehicle type. It has a lot to do with the person behind the wheel. Simply including gender and age variables cannot account for important covariates such as education, risk aversion, driving ability, wealth, etc. In reality, small cars may be less crash prone than Kahane's and Wenzel's results suggest, because they are driven by lower-income, younger, less risk averse people driving in more crash prone settings (e.g., commercial strips rather than pricey residential suburbs). Of course, as noted above, it is very difficult to control for all these variables, and the contractor was asked to rely on the original data. In reality, the best the report authors can do with such data sets is to explain how all the other, relevant attributes may factor in (e.g., quality of driver and typical driving settings), and how they can generate biased estimation (sometimes in either direction). Discussion of relevant literature that looks more deeply at crash outcomes (e.g., Wang or Chen's papers, mentioned above, allowing for heteroscedasticity and individual vehicle attributes, non-driver outcomes, etc.) will also be useful.
Please comment on the overall adequacy of LBNL's assessment of the 2011 NHTSA report and its independent study of casualty risk for predicting the effect of vehicle mass or footprint reduction on safety. Provide any recommended improvements that might reasonably be adopted by the author to improve the analysis.	Overall, the study is a comprehensive assessment of the 2011 NHTSA report and introduces interesting additional analyses to examine the relationship of vehicle mass and footprint reduction on safety. However, as stated previously in the comments here, driver preference for specific car types (including size and mass) is related to driver socioeconomic characteristics and driving behavior. As vehicle, driver, and roadway environment characteristics all contribute to fatality risk, the effects of physical vehicle changes such as mass or footprint reduction on safety should not be overstated when the other two types of characteristics are not sufficiently accounted for.
ADDITIONAL COMMENTS:	
Donna Chen & Kara Kockelman	

Review of LBNL Phase 2 Report	
1. ASSUMPTIONS	**COMMENTS**
Please comment on the validity of any assumptions embedded in the LBNL assessment analysis and the independent casualty analysis that could affect the projected relationship between vehicle mass/footprint reductions and fatality/casualty risk. Examples might include assumptions regarding whether recent historical relationships between vehicle weight, size, and safety will continue into the future; potential future improvements in vehicle technology and design may result in compensatory safety benefits; and the annual baseline fatality distribution.	The Phase 2 report serves as a complimentary document to the Phase 1 report by isolating the effect of vehicle mass and footprint on crashworthiness. Whereas the Phase 1 report analyzes fatality risk per estimated VMT, the Phase 2 report analyzes casualty risk per crash. The parallel structure of the two reports makes it easy for the reader to compare the results of the two analyses.

The binary logistic model employed in this analysis can only account for two injury outcome categories; here it is used to distinguish crashes resulting in serious injury or death from all other crash outcomes. Thus, the model does not account for the ordinal nature of injury severity and neglects the difference between a serious injury and a death.

The report states that "a serious incapacitating injury can be just as traumatic to the victim and her family, and costly from an economic perspective, as a fatality." While serious injuries are very costly to society (and may have similar economic cost implications as deadly crashes), willingness to-pay estimates (which include pain and suffering) price the cost of a fatality at almost 20 times the cost of an incapacitating injury (NSC 2010). Thus, it is difficult to assess the economic cost of the estimates of increases in casualty risk per crash without distinguishing whether that outcome is a serious injury or a death. This limitation of the model should be addressed in the report.

The logistic model also assumes error-term homoscedasticity and cannot account for increases and decreases in the variation of injury outcomes due to vehicle and driver type, for example. Such limitations of the model should be discussed. |
| Please comment on any apparent unstated or implicit assumptions and related caveats or limitations. | The role of driver behavior is briefly addressed in the report but not emphasized sufficiently. Casualty risk is a combination of driver, vehicle, and roadway characteristics. Whereas vehicle characteristics significantly influence crashworthiness, driver behavioral differences play a significant, if not primary, role in determining crash frequency. Socioeconomic data such |

	driver household income, size, and education influence driver attitudes and driving environments. For example, Chen et al. (2010) found that crash risk increases for those living in socioeconomically disadvantaged areas (including households more likely to drive less expensive and older vehicles). Though such data is not typically available in state and national crash databases, the importance of these driver and environmental characteristics on crash rates (per mile driven) and casualty risk should be stressed in both reports. It is clearly very difficult to control for, but is a major caveat to the NHTSA (& now LBNL) results. We expect that crash severity could be probably be lower for many of the small cars and pickups if they were driven by those who tend to drive more expensive vehicles, under the same settings (e.g., daytime, urban freeway). Thus, statements like "a 100-lb reduction in the mass of lighter cars leads to a 1.84% increase in crash frequency" should be accompanied by an explanation of the possibility of the mass variable accounting/proxying for effects of lower income households owning smaller vehicles.

ADDITIONAL COMMENTS:

National Safety Council (2010) Estimating the Costs of Unintentional Injuries. Available online at http://www.nsc.org/news_resources/injury_and_death_statistics/Pages/EstimatingtheCostsofUnintentionalInjuries

Chen, H.Y., Ivers, R.Q., Mariniuk, A.L.C., Boufous, S., Senserrick, T., Woodward, M., Stevenson, M. and Norton R. Socioeconomic status and risk of car crash injury, independent of place of residence and driving exposure: Results from the DRIVE study. *Journal of Epidemiology and Community Health* 64(10), 2010, pp. 998-1003.

2. CONTROL AND DEPENDENT VARIABLES USED IN THE REGRESSION MODELS	COMMENTS
Please comment on the adequacy of control and dependent variables used in the assessment analysis and independent casualty analysis, and recommend any alternative control or dependent variables that are available for possible inclusion in the analysis. For example, what are the relative merits of the main dependent variables used, fatality risk per estimated VMT, and casualty risk per police-reported crash?	As alluded to above, a primary concern is that the NHTSA analysis (& thus the LBNL analyses) largely neglect the idea that vehicle type (make & model) is very much a proxy for driver type, and a vehicle's crash frequency may have very little to do with physical vehicle characteristics. It has a lot to do with the person behind the wheel, and gender and age simply aren't enough to control for such distinctions. Education, risk aversion, ability, wealth, etc., are important covariates. But existing data sets are quite limiting (though the MVOSS & FAR with 3-year driver violation history do offer some valuable insights, not discussed in these reports). In reality, small cars may be less crash prone than Kahane's & Wenzel's results suggest, because they are driven by lower-income, younger, less risk averse people driving in more crash prone settings (e.g., commercial strips rather than pricey residential suburbs). Such key caveats need thoughtful discussion. Four relevant papers on the topics of crash frequency and vehicle size-and-weight implications (by Knipling, Kweon & Kockelman, Wang and Kockelman, and Chen & Kockelman) have been sent to Tom Wenzel. These all include useful literature reviews for further connections to useful findings for citation in the reports, as time allows the contractor.

Independent variables such as vehicle mass and footprint may be accounting for effects of driver socioeconomic factors as discussed in the Assumptions section. Furthermore, vehicle option variables such as AWD and side curtain airbags may be reflecting the effects of driver environment (e.g., those living in areas with icy winters opting for AWD) and attitude (e.g., more risk-averse drivers opting for side curtain airbags) rather than the vehicle technology themselves. While extremely heavy and extremely large vehicles may have significantly different handling and braking characteristics which influence crash frequency and casualty risk, it is unlikely that given the same driver in the same environment, a small change in vehicle mass or footprint would influence the driver's crash proneness. |

What additional control variables, such as vehicle make or model, might be included in the regression models?	Vehicle height, a variable which may be more valuable than vehicle type for similarly structured vehicles such as sedans, wagons, CUVs, and minivans, would be a valuable control variable. In addition to a wider track, a lower center of gravity also increases vehicle stability, thereby reducing the risk of rollover. Relevant literature & findings exist, and should be cited. Other variables which have been found in past studies to influence fatality risk such as seat belt use, roadway geometry and division type are not included in this study. To account for driver characteristics that contribute to casualty risk, socioeconomic variables such as household income, education, household size, etc. would be valuable additions. Unfortunately, both state and national crash databases typically do not include such information (outside of MVOSS). Such issues should be flagged for readers.
Please comment on any caveats or limitations that these dependent variable or control variables entail with respect to use of the results as the basis for estimating the safety effect of mass reduction.	Please see above comment (in Assumptions section) regarding driver behavior and environment.

ADDITIONAL COMMENTS:

Table 2.1 has many indicator variables labeled as "C" for continuous variable (such as ABS, ESC, AWD, DRVMALE, etc). These C's should be removed.

3. METHODOLOGY AND STATISTICS	COMMENTS
Please comment on the validity and applicability of the methodology LBNL used in assessing the NHTSA 2011 study and its analysis of the relationship between mass, footprint, and risks per police-reported crash.	The Phase 2 report enhances the findings in the Phase 1 report by isolating the effect of vehicle mass and footprint on crashworthiness. Like the Phase 1 report, this analysis goes into a fair amount of detail and seeks to introduce additional analyses to better examine the relationship between mass, footprint, and casualty risks. However, due to a lack of control for very specific vehicle differences (which vary by make & sub-model), the exclusion of driver characteristics and crash setting details (which cannot always be controlled for, but are often correlated with vehicle type), the effects of downweighting vehicles and/or shifting vehicle styles and sizes may be overestimated. Simply changing the vehicle mass or footprint on a risky driver in a high-risk setting is unlikely to influence crash outcomes significantly.
Please review other statistical methods LBNL has used in the analysis, in addition to the logistic regression methodology. Examples include the alternative approaches used by LBNL to assess NHTSA interval estimation results, and LBNL's linear regression analysis of actual, predicted, and residual risk by vehicle model.	In the alternative measures of exposure, the author examines the effect of vehicle manufacturer on fatality risk and treats the luxury models produced by Toyota, Honda, and Nissan as separate manufacturers. However, domestic luxury brands (such as Cadillac & Lincoln) are categorized with their nameplate manufacturers (GM and Ford), which appears inconsistent.
Please comment on caveats or limitations of using non-significant regression estimates to project the safety impact of mass reduction.	First, the t-statistics are not provided in the report which makes it difficult for the reader to assess statistical significance of specific regression estimates (except where noted by the author). Second, inclusion of a statistically insignificant variable can influence the estimates of coefficients associated with related variables. Nevertheless, in general, it is best to keep insignificant estimates if one has a strong defense for their role, since removing such variables (& thus their parameters) will shift the burden of response to a correlated covariate's parameter, thus biasing the latter. We generally keep key covariates in a model up to a pvalue of 0.20 or 0.25 or so, especially in relatively small data sets (e.g., n < 1,000). Covariates for which we have no strong basis can be removed for pvalues > 0.10.
How might the LBNL methodology be strengthened to better represent future vehicle designs and reduce multi-collinearity between mass and footprint in the regression analysis?	Including more vehicle-specific characteristics (such as vehicle height and engine size) reduces the analysis' dependence on vehicle type, since vehicle shapes and structures will continue to evolve. There is also correlation with context (e.g., pickups are driven in more rural locations, with greater hazards [like less lighting, higher speed, & few medians]). Disaggregate data are

	almost always best, to avoid ecological fallacies & such.
ADDITIONAL COMMENTS:	

4. DATA SETS	COMMENTS
Please comment on the validity and applicability of the datasets used to project changes in risk resulting from reduction in vehicle mass. LBNL's casualty analysis used police-reported crash data from 16 states, while the 2011 NHTSA study used national fatality data, combined with a subset of non-culpable vehicles involved in two-vehicle crashes from police-reported crash data from 13 states.	The Phase 2 report uses an unusually extensive data set of police-reported crash data from 13 states which the author compares in detail to national data sets to illustrate similarities and differences. The author is very thorough in addressing the difference in definitions of "serious" and "incapacitating" injuries across different states and the effects of such inconsistency on the regression results. Since casualty risk in the report accounts for serious injuries but not minor injuries, the author should note that police-reported injury levels may also be poor indicators of the actual or Modified Abbreviated Injury Scale (MAIS) level, following medical evaluation. Farmer (2003) found that 41% of injuries reported by U.S. police as incapacitating received MAIS ratings of "minor injury" by health care professionals using NASS Crashworthiness Data System (CDS). Thus, the results of the estimated casualty risk increases and decreases rely heavily on the assumption that police errors in reporting actual MAIS ratings are consistent across states.
Please comment on any apparent, unstated, or implicit impact on estimated risks inherent in the two different approaches, and any related caveats or limitations. For example, what are the strengths and weaknesses of the two measures of vehicle exposure, miles of vehicle travelled scaled up from crash data from 13 states, and number of police-reported crashes?	The Phase 1 analysis used non-culpable vehicles in two-vehicle crashes as a proxy for induced exposure crashes. In contrast, Phase 2 analysis uses data from vehicles involved in one-car crashes and the responsible vehicle in two-car crashes. The exclusion of the not-at-fault vehicle in two-car crashes may be distorting the distribution of crash frequency and casualty risk across different vehicle makes and models if crash-prone drivers are more likely to drive certain types of vehicles.
ADDITIONAL COMMENTS: Farmer, C.M. Reliability of police-reported information for determining crash and injury severity. *Traffic Injury Prevention* 4(1), 2003, pp.38-44.	

5. RECOMMENDATIONS	COMMENTS
Please comment on whether the LBNL assessment adequately addresses the NHTSA 2011 study and identifies the safety impact from mass reduction. Are the analytic methods and data used to assess the NHTSA study, and estimate the relationship between risk, mass, and footprint, appropriate? Is casualty risk per crash a legitimate measure of vehicle safety? What other methods or data could be used to better predict the effect of future vehicle designs on safety?	As noted above, a primary concern remains a neglect of the notion that the type of car is very much a proxy for driver type, and a vehicle's crash avoidance may have very little to do with vehicle type. It has a lot to do with the person behind the wheel. Simply including gender and age variables cannot account for important covariates such as education, risk aversion, driving ability, wealth, etc. In reality, small cars may be less crash prone than Kahane's and Wenzel's results suggest, because they are driven by lower-income, younger, less risk averse people driving in more crash prone settings (e.g., commercial strips rather than pricey residential suburbs). Alas, it is very difficult to control for all these variables, since they are not readily available in data sets. In reality, the best the report authors can do with such data sets is to explain how all the other, relevant attributes may factor in (e.g., quality of driver and typical driving settings), and how they can generate biased estimation (sometimes in either direction). Discussion of relevant literature that looks more deeply at crash outcomes (e.g., Wang or Chen's papers, mentioned above, allowing for heteroscedasticity and individual vehicle attributes, non-driver outcomes, etc.) will also be useful.
Please comment on the overall adequacy of LBNL's assessment of the 2011 NHTSA report and its independent study of casualty risk for predicting the effect of vehicle mass or footprint reduction on safety. Provide any recommended improvements that might reasonably be adopted by the author to improve the analysis.	Overall, the study is an enriching complementary document to the Phase 1 assessment of the 2011 NHTSA report. The parallel structure of the two reports allows the reader to easily compare and contrast the various additional analyses which examine the relationship of vehicle mass and footprint reduction on safety. However, as stated previously in the comments here, driver preference for specific car types (including size and mass) is related to driver socioeconomic characteristics and driving behavior. As vehicle, driver, and roadway environment characteristics all contribute to fatality risk, the effects of physical vehicle changes such as mass or footprint reduction on safety should not be overstated when the other two types of characteristics are not sufficiently accounted for.
ADDITIONAL COMMENTS:	

Charles Farmer **Review of LBNL Phase 1 and 2 Reports**	
1. ASSUMPTIONS	COMMENTS
Please comment on the validity of any assumptions embedded in the LBNL assessment analysis and the independent casualty analysis that could affect the projected relationship between vehicle mass/footprint reductions and fatality/casualty risk. Examples might include assumptions regarding whether recent historical relationships between vehicle weight, size, and safety will continue into the future; potential future improvements in vehicle technology and design may result in compensatory safety benefits; and the annual baseline fatality distribution.	The report concludes "that much of the detrimental effect of mass or footprint reduction on risk can be attributed to the tendency for mass or footprint reduction to increase crash frequency, rather than to reduce vehicle crashworthiness (risk once a crash has occurred)." However, the interpretation of casualties per crash as inversely proportional to crashworthiness ignores the possibility that injury severity also depends upon the circumstances of the crash. Casualties per crash must be divided into casualties per *severe* crash and *severe* crashes per crash, where a *severe* crash would be one involving more energy, e.g., high-speed or rollover. It could be that weight reduction <u>increases</u> casualties per severe crash (i.e., reduces crashworthiness), but reduces the likelihood that a crash is severe.
Please comment on any apparent unstated or implicit assumptions and related caveats or limitations.	The statistical models assume no interaction between the vehicle size/weight measures and any of the numerous covariates, but this may not be true. For example, size/weight reductions may differently affect vehicles with and without ESC if they affect vehicle handling. It is risky to make statements such as that on p. 11 of the Phase I report: *Therefore, the mass of a lighter car could be reduced by 800 lbs while adding ESC, without increasing fatality risk.*
ADDITIONAL COMMENTS: NHTSA's fatality analysis covered calendar years 2001-08, but the casualty analysis excludes 2008. Such exclusion is understandable given that 2008 data were at the time unavailable for a majority of the states (I think they are available now). However, 2008 was an unusual year and may have affected the size and weight effect estimates. The footnote on p. 6 of the Phase II report states that an analysis including the available 2008 data will be summarized in Appendix A. I don't see Appendix A. Is an analysis planned including 2008 data?	

2. CONTROL AND DEPENDENT VARIABLES USED IN THE REGRESSION MODELS	COMMENTS
Please comment on the adequacy of control and dependent variables used in the assessment analysis and independent casualty analysis, and recommend any alternative control or dependent variables that are available for possible inclusion in the analysis. For example, what are the relative merits of the main dependent variables used, fatality risk per estimated VMT, and casualty risk per police-reported crash?	One needs to restrict control variables to those that are available and reliable. A problem when combining state databases is that the states often are not consistent as to the variables coded and the definitions of those variables. This severely limits the list of possible control variables.
What additional control variables, such as vehicle make or model, might be included in the regression models?	I think that already there are too many control variables in the regression models. Instead I would consider defining different classifications of crash types. Table 2.2 of the Phase II report shows that the distribution of crash types for casualty crashes is very different from that for fatal crashes.
Please comment on any caveats or limitations that these dependent variable or control variables entail with respect to use of the results as the basis for estimating the safety effect of mass reduction.	Model overspecification could be the reason for results that are non-intuitive, especially in the Phase II analyses of police-reported crashes. Control variables may be correlated with each other or with the size and weight variables. For example, Figure 2.9 of the Phase I report implies that torso side airbags increase fatality risk in CUVs.
ADDITIONAL COMMENTS: The sensitivity results of Chapter 5 (Phase II) point out the extreme differences in results when changing the control variables. For example, including vehicle make changes the effect of a 100-lb reduction in heavier cars from -0.91% to +0.55% (see Fig 5.3).	

3. METHODOLOGY AND STATISTICS	COMMENTS
Please comment on the validity and applicability of the methodology LBNL used in assessing the NHTSA 2011 study and its analysis of the relationship between mass, footprint, and risks per police-reported crash.	Figure 2.11 of the Phase II report implies that NHTSA's fitting of a separate regression model for each of the 9 crash types was unnecessary, at least for the analysis of casualty risk per crash. I don't recall seeing a similar analysis for fatality risk per VMT. Is it possible to get essentially the same results as the NHTSA study using a single regression model?
Please review other statistical methods LBNL has used in the analysis, in addition to the logistic regression methodology. Examples include the alternative approaches used by LBNL to assess NHTSA interval estimation results, and LBNL's linear regression analysis of actual, predicted, and residual risk by vehicle model.	The graphs in LBNL's analysis of risk by vehicle model seem to indicate different trends for light and heavy vehicles (e.g., Figure 4.1 of Phase I). However, only simple linear relationships are examined, unlike the NHTSA analyses, which modeled piecewise linear relationships. Also, it's not clear why the 3rd and last rows of Table 4.1 list different numbers (both are labeled CUVs/minivans), and some entries in Table 4.1 disagree with Figures 4.6-4.11.
Please comment on caveats or limitations of using non-significant regression estimates to project the safety impact of mass reduction.	Making projections from non-significant regression estimates is proper so long as the resulting confidence intervals are constructed conservatively (to account for the accumulated imprecision). In that sense, I prefer NHTSA's jackknife approach to the standard errors produced by SAS (see p. 13 of Phase II).
How might the LBNL methodology be strengthened to better represent future vehicle designs and reduce multi-collinearity between mass and footprint in the regression analysis?	
ADDITIONAL COMMENTS:	

4. DATA SETS	COMMENTS
Please comment on the validity and applicability of the datasets used to project changes in risk resulting from reduction in vehicle mass. LBNL's casualty analysis used police-reported crash data from 16 states, while the 2011 NHTSA study used national fatality data, combined with a subset of non-culpable vehicles involved in two-vehicle crashes from police-reported crash data from 13 states.	A major limitation of the Phase II analysis is a bias that may be due to the patterns of missing data. In particular, the vehicle identification number (VIN) is missing or mistyped for many crash records. High-severity crashes (especially fatal) are more likely to have detailed police investigation, so VINs (and other variables) in these crashes may be more complete. State crash files are therefore much less reliable than FARS.
Please comment on any apparent, unstated, or implicit impact on estimated risks inherent in the two different approaches, and any related caveats or limitations. For example, what are the strengths and weaknesses of the two measures of vehicle exposure, miles of vehicle traveled scaled up from crash data from 13 states, and number of police-reported crashes?	The VMT weights provided by NHTSA were scaled to represent the entire US. Comments on pp. 9 and 18 of the Phase II report seem to acknowledge this deficiency, promising to adjust these to the 13 states in the future. Was any adjustment made, such as multiplying the weights by the proportion of annual US VMT accounted for by each of these states? The accuracy of the VMT weights is critical is we are to believe the somewhat surprising results concerning crashes per VMT.

ADDITIONAL COMMENTS:

Statements above Figure 2.7 in the Phase II report imply that the effects of weight reduction on crashes per VMT and fatalities per crash should add up to the effect on fatalities per VMT. This is not the case. For example, a 1.43% increase in crashes per VMT and a 0.76% decrease in fatalities per crash would imply a 2.16% decrease in fatalities per VMT (i.e., 1 - 0.9924/1.0143). The fact that the model on fatalities per VMT yields an estimated 1.08% increase should be a cause for concern. Either the VMT weights are inaccurate or the control variables have different effects on crash frequency and crashworthiness.

5. RECOMMENDATIONS	COMMENTS
Please comment on whether the LBNL assessment adequately addresses the NHTSA 2011 study and identifies the safety impact from mass reduction. Are the analytic methods and data used to assess the NHTSA study, and estimate the relationship between risk, mass, and footprint, appropriate? Is casualty risk per crash a legitimate measure of vehicle safety? What other methods or data could be used to better predict the effect of future vehicle designs on safety?	Casualty risk per crash does not fully measure the effects of vehicle size and weight reductions on society. Casualty risk per VMT best coincides with the NHTSA analysis of fatalities per VMT. The breakdown of casualty risk per VMT into the crash frequency and crashworthiness components is of interest. However, the surprising results reported here make everything suspect. For example, the Phase II report concludes that "the detrimental effect of male drivers has to do with their higher tendency of getting into a serious crash rather than their sensitivity to injury once a serious crash has occurred" (p. 24). A few pages later it concludes that "male drivers have essentially no effect on crash frequency, but cause a statistically significant increase in fatality risk once a crash occurs" (p. 28).
Please comment on the overall adequacy of LBNL's assessment of the 2011 NHTSA report and its independent study of casualty risk for predicting the effect of vehicle mass or footprint reduction on safety. Provide any recommended improvements that might reasonably be adopted by the author to improve the analysis.	Overall these are reasonably good studies. The Phase I report does a very good job of assessing the NHTSA report of fatality risk. However, the Phase II report should be more cautious in its conclusions concerning casualty risk. The casualty analysis is based solely on police-reported data from 13 states, which: 1. May not be representative of the US as a whole. 2. Are inconsistent in the information given and the way in which it is coded. 3. Suffer from information that is missing, inaccurate, or unclear.

ADDITIONAL COMMENTS:

Column G of Table 6.1 in the Phase II report provides the most appropriate comparison to results from the NHTSA report (Column A). For both fatalities and casualties per VMT, a 100-lb weight reduction is most harmful in lighter cars, less harmful in heavier cars and lighter light trucks, and slightly beneficial in heavier light trucks, minivans, and crossovers.

Review of "An Analysis of the Relationship between Casualty Risk per Crash and Vehicle Mass and Footprint for Model Year 2000-2007 Light-Duty Vehicles"

David L. Greene

January 10, 2012

Summary

The Phase I and Phase II analyses by Tom Wenzel of LBNL have been executed diligently and consistently in accord with the methods and data used in the original NHTSA analysis. The studies contain many valuable, new insights. The phase I study highlights the weakening relationship between vehicle mass and highway fatalities. This is not only seen in decreasing coefficient estimates but in the very large number of results that are not statistically significant. When regressions were done separately by footprint deciles, vehicle mass was statistically significantly positively related to fatalities only for light-duty trucks in rollovers, there were almost as many cases in which mass was negatively related to fatalities (9 vs. 13 out of 27) and there were more instances of statistically significant negative relationships than positive relationships. Given that so many tests are being jointly conducted, it is quite possible that when joint probabilities are considered, there is no significant relationship between mass and fatalities (more on joint probabilities later). Showing the weakness and inconsistency of these results is an important contribution.

Another meaningful contribution of the phase I study, and one that deserves more emphasis, is a logical inference from the following findings: 1) much of the variance in risk remains unexplained even by the most complete models, 2) control variables explain 1 to 2 orders of magnitude more of the variance than the variables of interest (mass and size), 3) when key control variables are removed or changed it strongly influences the coefficients of mass and size. These results have very important implications for the robustness of the results and the likelihood that some or all of the apparently statistically significant relationships are due to spurious correlations with omitted or imperfectly controlled factors. Noting that exposure measures are control variables with constrained coefficients, the following observation from the phase I study is especially perceptive.

> "Calculating risk as total fatalities per induced exposure crash, rather than per vehicle mile traveled, reverses the sign of mass reductions on risk in cars and the lighter light trucks, with mass reduction leading to a <u>reduction</u> in risk in all vehicle types."

Finally, the phase I report notes that if only the control variables are included in the regression and not size or mass, the resulting residuals from the regression are uncorrelated with size or mass. Given these findings (as well as those of phase II) the conclusions that,

> "The 2011 NHTSA study, and this report, conclude that the effect of mass reduction on US fatality risk is small."

should be revised with the following emendation, "… and probably non-existent."

Both studies, like the NHTSA analysis, have shortcomings in terms of interpreting the results and the language used to describe the results, and acknowledging the limitations of the data and methodology. The limitations are extensive. The interpretation of the results of the LBNL studies commits two important, related errors. The first is to attribute inferred coefficients of mass and size as representing only the effects of vehicle mass and size when, as the phase I and II study results indicate, there is a virtual certainty that aliasing effects are present due to a combination of omitted variables, errors in variables and correlations among variables. Given that estimated driver and environmental factors tend to have 1-2 orders of magnitude larger impacts on safety outcomes than vehicle factors, the almost certain presence of aliasing effects must be explicitly acknowledged as severely limiting the ability to draw inferences about the effects of vehicle attributes. Second, the language used in interpreting results fails to acknowledge that the analysis does not address the effects of down-weighting or down-sizing specific vehicles or vehicle designs, but instead relies on correlations between vehicle weight and size in existing vehicle designs. In existing vehicles, weight and size are correlated with each other and many other vehicle attributes (and driver and environmental attributes, as well). Thus, the study is not actually measuring the effects of down-weighting via the material substitution and design changes likely to occur as a consequence of fuel economy and emissions standards. An early example of the kind of misleading language referred to here can be found on page iv.

> "For example, a 100-lb reduction in the mass of lighter cars leads to a 1.84% increase in crash frequency (columns B), while mass reduction leads to a 0.76% decrease in the number of fatalities per crash (column C);"

This statement is misleading in that it implies causality rather than correlation, and it is additionally misleading in that it implies that the inference applies to removing weight from specific vehicles. Neither is correct. A better statement would be the following.

> "For example, vehicles in the lighter class that are 100 lbs. lighter are correlated with a 1.84% increase in crash frequency.…"

There are so many examples of this misleading language that it is not feasible to list them all. All should be corrected, however. Failure to correct them could lead to serious misinterpretation of the studies' findings.

Following in the footsteps of the seminal study by DRI, the NHTSA and LBNL studies contribute to the literature in three important ways: 1) the LBNL and NHTSA studies recognize that the societal safety perspective is the correct perspective to when assessing the impacts of fuel economy and emissions regulations, 2) they recognize that vehicle dimensions and vehicle mass may have separate and potentially different impacts on both the likelihood of a crash and the outcomes of the crash and, 3) the LBNL phase two analysis makes an additional contribution by attempting to disentangle factors affecting the likelihood of a crash and factors affecting the outcomes of a crash.

Speaking of the DRI study, I am puzzled about why there are no references cited in the phase I study and only a handful all by Kahane and Wenzel, in the phase II study. This is perhaps due to the scope of work defined for the two studies but there are highly relevant studies in the literature that could have been cited, those by DRI foremost among them. Making use of the insights from these studies would have been helpful in interpreting the results of both phase I and phase II.

Lack of a Theory or Model of the Phenomenon

Both the NHTSA and LBNL studies lack a rigorous theory of the process by which down-weighting at constant size or down-sizing at constant mass affect societal safety either through crash avoidance or crashworthiness. This is not a trivial shortcoming because it affects the ability to formulate hypotheses and interpret results. Prior to the dissenting report on safety of the NRC 2002 CAFÉ report, the physics of elastic collisions between objects was typically cited as the underlying physical model. That report showed how taking the societal perspective renders that model inappropriate. What remains appears to be far more complex, involving the quantity of kinetic energy, the ability of vehicle designs to absorb that energy so as to minimize maximum deceleration rates, stability, maneuverability, safety technologies, and more.

The consequences of the lack of a rigorous theory are that it is not known, a priori, what the signs of coefficients are expected to be, let alone what their quantitative relationships should be. Hypotheses must be formulated based on intuition and the interpretation of results is likewise ad hoc. One implication of this is that results that suggest that lower fatalities are associated with lower vehicle mass have equal standing, a priori, with results that indicate that higher fatalities are associated with lower vehicle mass, and similarly for vehicle size. There are no surprising or unsurprising results, in theory.

This also makes it difficult to develop a plan for statistically testing the model or theory and its implications. It would have been helpful to the reader to have been presented early on in the report with such a plan of analysis.

On the Virtual Certainty of Aliasing

The LBNL report typically attributes causal effects to correlations between mass or size and safety. In fact, most or all of the observed correlations are almost certainly affected by aliasing effects. There is ample evidence for this inference in the results presented in the LBNL phase II report.

The coefficients of mass and size change in important ways when different model formulations are estimated. Removing and adding control variables changes the magnitudes and sometimes the signs of the mass and size variables. This means that, at a minimum, the mass and size variables alias the effects of the omitted control variables. The question is whether the aliasing is eliminated entirely by the inclusion of the control variables available or whether some aliasing remains either because not all relevant and correlated control variables have been included or because the included control variables are imperfect measures of the factors they are intended to represent.

The latter seems highly likely for the following reasons. First, the overall explanatory power of the full models (including control variables), as measured by their R^2 is low. Most of the variance in casualties and fatalities remains unexplained. Second, at least some of the important included control variables are only crude approximations of the factors they are intended to represent. For example, dummy variables represent differences in state reporting practices, age and gender represent risky driving behavior differences among owners of different sizes and masses of vehicles, the presence or absence of a kind of safety equipment represents both its performance and use in a particular vehicle, and calendar year dummy variables represent unknown factors associated with the respective calendar years. Such practices are common and their use is appropriate. Third, the control variables generally account for 1-2 orders of magnitude more variance in the casualty and fatality variables than do vehicle weight and size. To recap, the amount of unexplained variability in the dependent variables is larger relative to the variance statistically explained by the most complete models. Control variables are correlated with size and mass, and they account for 1-2 orders of magnitude more variability in the dependent variables than the variables of interest, mass and size. Therefore, even small correlations of size and mass with omitted variables or with errors (imprecision) of the control variables could easily result in biased estimates for the effects of size and mass on the dependent variables.

Tom Wenzel is to be commended for providing the results that definitively demonstrate the three key points made above. The above is not a criticism of the analysis nor of the results, per se. It is a criticism of their interpretation. In light of the above, the results should be interpreted in light of the virtual certainty that many of the estimated coefficients are likely to be biased in ways that make their interpretation highly uncertain. The implication is that phrases such as "down-weighting or down-sizing caused" to "mass (or size) and unobserved correlated factors are associated with…"

On Joint Probabilities

The NHTSA and LBNL studies do not correctly interpret their results as joint statistical tests. When testing a hypothesis on, for example, 5 vehicle classes simultaneously, a result for one equation that might be statistically significant on its own may not be statistically significant as one of five related tests.

Statistical analyses comprised of multiple regressions too often overlook the fact that tests of statistical significance designed for individual regressions may not apply in the case of multiple regressions. That is the case here. NHTSA conducted 5 analyses to infer relationships between mass differences among vehicles holding footprint constant for 5 classes of cars. The results showed one relationship out of five was statistically significant. As table 1 illustrates, using a simple example, if one conducts 5 trials, each with a 0.05 probability of given result, there is a 22.6% probability of finding at least one such result in the five trials. Thus, the joint significance level of the overall result (1 statistically significant regression out of 5) is 0.226, rather than 0.05.

Table 1. Simplified Illustration of the Joint Probability of Inferences in Multiple Regressions

# Significant Regressions	Combinations	0.95	0.05	Joint Probability
0	1	0.773780938	1	77.37809%
1	5	0.81450625	0.05	20.36266%
2	10	0.857375	0.0025	2.14344%
3	10	0.9025	0.000125	0.11281%
4	5	0.95	0.00000625	0.00297%
5	1	1	3.125E-07	0.00003%
				100.0%
				22.6%

So there is between a 1:4 and a 1:5 chance of getting one statistically significant result by pure chance. In fact, the actual significance level of the results is more complicated to calculate, and probably a bit smaller than 0.226. Thus, it is very appropriate for Dr. Kahane to add the qualifier "if any" to his conclusions about the relationship between the societal highway fatalities and mass reduction, holding footprint constant. Had appropriate tests of joint statistical significance been used to evaluate the results in the NHTSA and LBNL studies, the significance levels very likely would not meet accepted criteria for statistical significance. This could change the conclusions of the studies from the inference that mass is correlated with fatalities or casualties in some case but not others to the lack of statistically significant evidence that mass is correlated with fatalities or injuries on the highway. This is an important difference.

Page-by-Page Comments

I will make page by page comments on the phase II study only, since that contains the overwhelming share of original contributions and the key findings of the phase I study are recapitulated there.

p. iii   Paragraph 3. This would be a very good place to acknowledge the importance of driver behavior and environment on crash avoidance especially.

p. iv   Para. 3. This would be a good place to discuss probability inference in joint tests.

Para 4. The statement about a 100-lb reduction in the mass, etc., is a good example of misleading language.

p. v   Para. 1. Again, it is misleading to say that mass reduction increases crash frequency, for reasons stated above.

Para. 2. It is more accurate to describe the association of lower vehicle mass with casualty risk than the "effect of mass reduction on…" casualty risk.
Para. 3. Would benefit greatly from joint probability inferences.

Para. 4. As noted above, this shows how much more important the control variables are than the variables of interest.

Para 5. Again, these are correlations not necessarily effects.

p. vi   Para. 1. Again, mass reduction is misleading terminology and you do not know if it increases casualty risk or not, you know only a correlation. Why is this so important? It is the virtual certainty of spurious correlations, or aliasing, as noted above.

Para. 4. (1st bullet) This is clear evidence of aliasing. Take variables out of the regression and the coefficients of interest change in important ways. Are there no important factors still missing? Are the variables included perfect measures of the factors of interest? Of course not. Thus, there must be remaining aliasing. How bad is it? We don't know.

The third bullet shows the same effect with a different set of variables.

p. vii   Para. 4. No, your analysis does not indicate "…that much of the detrimental effect of mass or footprint reduction on risk can be attributed to the tendency for mass or footprint reduction to increase crash frequency." Again, you have correlation, not causation and you have good reason to believe that what you are seeing is affected by spurious correlations.

Para. 5. The "effect" is small, 1-2 orders of magnitude smaller than correlations with other control variables, and IS strongly affected by which variables are in the equation, as stated on the previous page, and there is a great deal of unexplained variance. Please reconsider the meaning of these results in light of the comments above.

Finally, as the last paragraph of the ES implies, it would be far better not to speak in terms of "reducing" mass or size. That is not what is happening in your data set.

p. 1   Para. 4. Risk per VMT includes the effects of how well vehicles are driven as well as how well they can be driven. I think there is no chance that you have fully accounted for how well vehicles are driven.

p. 2   Para. 2. Exposure measures are explanatory variables whose elasticity is constrained to 1. That is, it is assumed that an increase in vehicle use of 1 vehicle mile produces a 1 unit increase in the chance of a fatality (or casualty as the case may be). This is actually a maintained hypothesis. If this hypothesis is incorrect, it can bias the other coefficients in the equation. Thus, the change from fatalities/VMT to fatalities/registration-year, to fatalities per crash not only changes the meaning of the analysis, it also may bias coefficients in the event that the true relationship between fatalities and VMT is not an elasticity of 1.

p. 3   Line 1. Please acknowledge that your "accounting for differences in driver characteristics, crash locations, and other vehicle attributes" is incomplete and that this could affect your inferences about size and mass.

p. 3   Para. 2. NHTSA's use of "non-culpable" vehicles involved in two vehicle crashes as an exposure measure raises its own issues. How non-culpable was the non-culpable vehicle. Often this is a matter of degree, rather than black or white. Driver behavior may also be involved. It seems to

me this is just another potential measure of exposure that may or may not be better than any other measure and may introduce new sources of bias in the analysis.

p. 3   Para. 4. Induced exposure needs to be defined. What is it intended to mean? This needs to be explicit.

Also, I am startled that there are no equations in these reports. Equations can provide an unequivocal explanation of the assumed relationships that cannot be adequately accomplished by words, in many cases. Why no equations?

p. 7   Para 3. CUVs and minivans are involved in fewer crashes with stationary objects than cars. Why? Is it the drivers, the vehicles, or the passengers? How well can you control for such differences? Not well. What does this mean for your analysis?

p. 9   Para 1. Here an equation showing how the weighting was done would be very helpful.

p. 9   Bullet 1. Excluding these vehicle types implies that the control variables in the model are not adequate to account for whatever makes these vehicle types different from the vehicles included in the analysis. First, this is an admission that the model is not adequate to explain the fatalities associated with these vehicles. Second, it is an admission that if they were included the coefficients on the variables of interest would likely be biased by spurious correlations. Clearly, it would not even be sufficient to include the vehicles along with a control variable (e.g., X = 1 if vehicle is a police car, 0 otherwise). This is yet another indication that the model suffers generally from omitted variables, errors in variables and correlation among right-hand side variables.

p. 10   Line 3. Sentence does not make sense. Please correct.

p. 12   Sect. 2.3. Please provide an equation.

p. 13   Were the confidence intervals calculated using e^x-1? Please state explicitly or, better, show an equation.

Para. 2. Here another instance where you say "mass reduction increases societal fatality risk" but you really are not entitled to say that. It is misleading. Also, the NHTSA CI's are larger, as they should be in a joint test.

Para. 3. These results require an underlying theory for interpretation. The lack of one makes it seem like there is just no consistency in the results.

Para. 4. The fact that the results for fatalities per crash differ substantially from fatality per VMT may be very important. Taken at face value, it would imply that any negative effect of reduced mass is due to its effect on crash avoidance (crash probability) rather than crashworthiness. This is where the lack of a theory is most troubling. Why would that be? Are lighter vehicles less easily controlled, etc.? Or, as seems much more likely, is there a spurious correlation between mass and other omitted or imperfectly measured factors (including driver behavior) that lead to an increased probability of a crash? Consider, for example, driver age. Driver age is related to crash involvement. Driver age is a control variable. But are all young drivers the

same? Is it possible that young drivers more prone to risky behavior tend to drive lighter vehicles? If so, this could partly explain the result observed. Of course, this is just speculative, but the point is that correlations with imperfectly measured and omitted factors are highly likely to be present in the data and, if there, could easily affect the statistical inferences.

p. 14   Here we see that changing the exposure measure influence the effect of mass and size on fatalities and casualties, which is more evidence that spurious correlations are likely to be biasing estimated coefficients for mass and size.

The bar graphs with confidence intervals are well done and convey a great deal of information effectively. The patterns of magnitude and statistical significance are difficult to interpret, partly because there is no explicit theory of what should happen and partly, perhaps, because the relationships are actually not real.

p. 16   Para. 1. Reduction in the mass of lighter cars increases crash frequency but reduces fatalities per crash. This is contradictory to the previously maintained theory that mass protects due to the physics of velocity changes in elastic collisions. Indeed, there is no theoretical explanation for these results, only speculation.

p. 18   Para. 1. Here is a good example of such casual speculation.

p. 19   Para. 1. Developing VMT weights for the 13 states is a good idea, given the effect of exposure measures on inferences. Still the results would not be definitive.

p. 20   Para. 2. Mass reduction leads to a large reduction in risk only in crashes with objects for heavier cars?   There is only one type of crash in which the simple physics of collisions leads to an unambiguous benefit for increased mass, and that is collisions with moveable or breakable objects. This finding contradicts even that.

Para. 3. More speculation, this time about rollovers.

p. 21   Para. 3. "Curiously,…" Curiouser and curiouser.

Figure 2.15 printed without labels. Could be my computer but the other graphs were fine.

p. 24   Para. 2. This is probably a very important finding that needs further investigation and explanation. As figure 2.16 illustrates well, the correlations with mass and size are orders of magnitude smaller than the correlations with driver and environmental factors. This is why even small correlations with omitted or imperfectly measured control factors could be, are even likely to be, predominantly responsible for the estimated coefficients of mass and size.

p. 25   Para. 2. The results for minivans discussed here could be due to what is going on inside the vehicle as much or more than the vehicle itself. How could these results be explained in terms of the vehicle itself. Figure 2.17 shows this again. The effects of calendar year dummies, which can only be considered rough approximations to unknown and various time-related factors, have much large effects than size or mass. Again in figure 2.20.

p. 28   Para. 1. "Surprisingly,…" How can side airbags, which deploy only in a crash, reduce crash frequency but not fatality risk in a crash? Only if the real effect is a reflection of who buys a CUV/minivan with side airbags and how and where they drive. There are more surprising inferences in paragraph 2 about male and female drivers. Surprising relative to what theory?

p. 32   The problem here is not numerical multicollinearity (numerical difficulties inverting the cross-product matrix) but the more complex problem of correlations among right-hand side variables, omitted variables, errors in variables, and correlations of included variables with omitted and imperfectly measured variables. This leads to biased estimates.

p. 36   Table 3.1 cries out for inference based on joint probabilities. What is the probability of observing "success" in at least 3 out of 27 trials when the true probability of success is only 0.05. See discussion above. The probability is certainly much higher than 0.05 and probably closer to 0.5. The implication is that, taken together, these results do not show any statistically significant relationship between mass or size and risk per crash. If there were a rigorous underlying theory, the interpretation might be different (patterns of significance could matter) but there is none. Again, good graphs on succeeding pages.

p. 41   The statistical significance of such a relationship should be the same whether bins are used or not. Is it?

p. 42   Para. 3. R-squared is not the correct measure of statistical significance. Is the coefficient of weight significantly different from zero?

p. 43   Para. 2. This is perhaps the key finding of the phase I and II analyses. Control variables explain 1-2 orders of magnitude more variance than size or mass. Still, most of the variance remains unexplained and is uncorrelated with mass or size. It is very likely there is nothing going on here.

p. 47   Para. 2. More evidence of correlation of mass and size with control variables and how changing definitions or excluding control variables results in important changes in the coefficient of mass and size. Such results are considered unstable.

p. 49   Para. 1. The rationale for the grouping of manufacturers is not obvious. Can you explain it?

         Para. 2. Yet more evidence for the instability of the model and likelihood that variables still missing from the model, plus errors in measuring the included control variables are likely biasing the inferences. The results described in this paragraph do not make sense to me. How can they be explained other than random results?

p. 50   Para. 1. More casual interpretation of results. OK, maybe, maybe not. Same for paragraph 2. The economy faltered in 2008 but the big negative effect was in 2007. The downward trend started in 2004. This correlates neither with vehicle mass changes over time nor economic growth as measured by real GDP. Idle speculation.

p. 51   Para. 1. Good discussion of the gratuitous speculation by NHTSA about the meaning of the observed correlations. This is more a Rorschach test than statistical analysis.

p. 51   Para. 2. "We have no explanation for why…" More of this kind of honest appraisal is needed in studies like these.

p. 52   More results for which there is no explanation.

p. 53   Why the interaction between calendar year dummy variables and safety equipment? The presence of the safety equipment on a particular vehicle is established. What has calendar year (not model year) to do with it? Again, one suspects spurious correlations.

p. 55   One needs to think carefully about the reasons why vehicles would be excluded. It does not appear that NHTSA did that. First line of first paragraph "used" rather than "sused".

p. 57   Well reasoned. It is interesting that NHTSA resisted including footprint or size in previous analyses on the grounds of correlation with mass. These results show that assertion was groundless.

p. 59   Para. 3. Rather than say risk per VMT accounts for two effects, it is better to say it includes or comprises two effects. But this statement also ignores the important influences of drivers and environment and their potential correlations with other factors. Yes, it includes how well a vehicle can be driven, but more importantly it includes how well a vehicle IS driven. That is in there too and is very likely to be correlated with make, model, and other vehicle attributes.

p. 60   Para. 3. Here again, the conclusions are misstated. It is not a genuine "reduction" in mass, but an association with the mass of vehicles. And how does it "lead" to and increase in crash frequency? What is the theory or model that predicts this? Driver and environment are very likely mixed up in these results to an unknown but likely substantial degree. So what can we really conclude? Not this.

As I read the conclusions and inferences I find myself asking, why?, why?, repeatedly without any sound explanations. Page 61, paragraph 3 contains more "surprising" results. Surprising because they are contrary to theory? Surprising because they are contrary to intuition? Surprising because they are random? To what can we attribute so many "surprising" results, and how many must there be before one concludes that the analysis is not revealing what we had hoped it would.

Para. 4. Again, this cries out for joint statistical inference. Three statistically significant results out of 27 is probably nothing statistically significant at all.

p. 62   Para. 1. What this shows, again, is that the coefficients of mass and size are strongly influenced by which control variables are included in the model and how they are defined. These results and their implications need explanation. The bottom line is that the effects of mass and size are likely to be (after the necessary joint significance calculations are done) not statistically significant, not consistent, and not robust.

p. 63   How do mass and footprint reduction (again, it's not really reduction in the sense of designing lighter vehicles to increase fuel economy or reduce GHG emissions, the issue at hand) increase crash frequency. What is the theory? I don't find a theory in either the NHTSA report or the phase I and phase II studies. Absent a theory, these results seem sufficiently unstable and

inconsistent to be highly questionable as evidence of any relationship between mass or size and crashworthiness or crash avoidance. I think joint estimation of significance levels would provide additional support for this view.

Michael Van Auken Review of LBNL Phase 1 Report	
1. ASSUMPTIONS	**COMMENTS**
Please comment on the validity of any assumptions embedded in the LBNL assessment analysis and the independent casualty analysis that could affect the projected relationship between vehicle mass/footprint reductions and fatality/casualty risk. Examples might include assumptions regarding whether recent historical relationships between vehicle weight, size, and safety will continue into the future; potential future improvements in vehicle technology and design may result in compensatory safety benefits; and the annual baseline fatality distribution.	The basic assumptions, methodology, and data are primarily the same as in the Kahane (2011) report. These include the following: 1) The probability of a crash fatality is proportional to the vehicle miles travelled (VMT), except as noted in Section 5.1 2) The logarithm of probability of fatality per VMT for a given curb weight, footprint, and control variable values varies as a linear combination of the curb weight, footprint, and control variables within the domain of the data. 3) The logistic regression methods determine a maximum likelihood estimate of model coefficients. 4) It is assumed that the above relationships remain constant in the recent past (i.e., 2000-2007 model year vehicles in the 2002-2008 calendar years), present, and near future (i.e., 2017-2025 model year vehicles). The first assumption that crash fatalities are proportional to VMT rather than the number of vehicle registration years (VRY) is appropriate because the fatalities cannot occur if the vehicles are not driven on the road (i.e., VMT = 0). This assumption is qualified however because VMT is more difficult to measure than VRY and therefore may be less accurate. On the other hand the probability of a fatal crash or the number of fatalities in a crash may also depend on the vehicle occupancy. The analysis in Section 5.1 is a commendable attempt to explore the sensitivity to this assumption, however the Kahane (1997) and DRI (2003-2005) reports have shown that some driver, vehicle, and environmental factors may be underrepresented or overrepresented in unweighted induced-exposure data. VRY could have also been considered as a measure of exposure.

	The second and third assumptions are appropriate provided that it is recognized that it is essentially impossible with currently available knowledge and information to model all of the factors that could affect the probability of fatality in a crash, and that the objective of the analysis is to identify overall trends versus vehicle weight and footprint. In general the probability of fatality depends on other many other factors which have not been modelled (e.g., driver behavior factors, vehicle design factors, roadway design factors, EMT factors) and these unmodeled factors are assumed to be uncorrelated with vehicle weight and footprint, and/or are represented by the other control variables. The latter assumption might or might not be valid. The fourth assumption is perhaps the weakest because it assumes that future vehicles will have the same design characteristics as past vehicles, and that the characteristics of the vehicle population (e.g., collision partner weight, size, type) will also remain the same. A commendable attempt to partially address this effect is described in Section 6. These effects can be perhaps better addressed by the "Volpe model" described in Kahane (2011) of the Honda-DRI fleet systems model described in Refs (2324), which can be used to forecast the effects of mass reductions of individual makes and models on a year-by-year basis.
Please comment on any apparent unstated or implicit assumptions and related caveats or limitations.	The induced-exposure data set provided by NHTSA is based on the "non-culpable" vehicle in two-vehicle crashes. It is assumed that the dataset is a reprehensive sample of the driver and environmental exposure factors for vehicle use. However, since these cases include moving vehicles, some vehicle-driver-environmental conditions may be under or over represented in this data depending on how they affect the ability of a non-culpable vehicle to avoid a crash. Results in Ref (17) indicate that the estimated effect of weight and size reduction are sensitive to whether the induced-exposure data are based on the Kahane (2003) non-culpable vehicle definition of the Kahane (1996) stopped vehicle definition.

	Unfortunately it is not currently possible to test this sensitivity with the NHTSA-provided induced-exposure data.
ADDITIONAL COMMENTS: None come to mind.	

2. CONTROL AND DEPENDENT VARIABLES USED IN THE REGRESSION MODELS	COMMENTS
Please comment on the adequacy of control and dependent variables used in the assessment analysis and independent casualty analysis, and recommend any alternative control or dependent variables that are available for possible inclusion in the analysis. For example, what are the relative merits of the main dependent variables used, fatality risk per estimated VMT, and casualty risk per police-reported crash?	The main metric used in both the Kahane (2011) and Wenzel (2011a) reports is the total number of fatalities (except as noted). Reducing the total number of fatalities, which includes both subject vehicle occupants and collision partner fatalities, is desirable from a societal viewpoint. Fatal crash occurrence is related to the total number of fatalities, which has been used by Kahane (2003, 2011) to address concerns about double counting.

VMT is a good measure of accident exposure provided that it can be accurately determined.

[Note: This peer review does not address the Wenzel (2011b) companion report (Ref 3) which examines the risks per police-reported crash. See Ref 4 for comments on the companion report.] |
What additional control variables, such as vehicle make or model, might be included in the regression models?	None come to mind.
Please comment on any caveats or limitations that these dependent variable or control variables entail with respect to use of the results as the basis for estimating the safety effect of mass reduction.	None come to mind.
ADDITIONAL COMMENTS: The underlying reasons for some of the estimated effects are unknown at this time, but presumably involve driver, vehicles, environment or accident factors that have not been controlled for in the Kahane (2011) and Wenzel analyses. See, for example, Refs 17 and 25.	

3. METHODOLOGY AND STATISTICS	COMMENTS
Please comment on the validity and applicability of the methodology LBNL used in assessing the NHTSA 2011 study and its analysis of the relationship between mass, footprint, and risks per police-reported crash.	The logistic regression methods seem to be appropriate. The confidence intervals are based on the logistic regression Wald Chi-Square statistic, which as Kahane (2003, 2011) has demonstrated does not include all sources of variation. However, these confidence intervals are useful because they do provide some indication of the uncertainty in the results. [Note: This peer review does not address the Wenzel (2011b) companion report, which examines the risks per police-reported crash. See Ref 4 for comments on the companion report.]
Please review other statistical methods LBNL has used in the analysis, in addition to the logistic regression methodology. Examples include the alternative approaches used by LBNL to assess NHTSA interval estimation results, and LBNL's linear regression analysis of actual, predicted, and residual risk by vehicle model.	The correlations in Section 3 appear to be assessed using the Coefficient of Multiple Determination (R^2) based on a linear fit to the data (e.g., the correlation between footprint versus curb weight in Figure 3.1 on p. 14). The linear regression model attributes the differences between the dependent variable (vertical axis) and the linear fit to the independent variable (horizontal axis) to random effects. If there is no preference as to the choice of independent and dependent variables (e.g., footprint versus curb weight, or curb weight versus footprint), then the linear trend and R^2 result would be different if the two variables were interchanged, and having two different yet equally valid results would be undesirable. If the variation in the data can be attributed to both variables (e.g., footprint and curb weight), then it would be better to report the square of the sample correlation coefficient r^2, where r is computed according to Eqn (1). The trend lines in these correlation figures should not be computed using a linear regression. Instead, the trend line should pass through the sample means (i.e., (x, y)), and have a slope equal to the ratio of the sample standard deviations in the data (i.e., s_y / s_x). Therefore, the reported correlation results do not depend on the ordering of the data variables.

	Note this comments does not apply to linear trends indicated in Section 4, for which the Coefficient of Multiple Determination (R^2) seems appropriate.
	The Coefficient of Multiple Determination (R^2) is frequently used in the Wenzel (2011a) report as an indicator of the statistical importance of a linear trend (e.g., R^2 values in Tables 4.1 and 4.2 that are greater than 0.3 are shown in blue font). It would be better to report the standard error, confidence interval, and/or probability value as measures of the statistical significance of a linear trend.
Please comment on caveats or limitations of using non-significant regression estimates to project the safety impact of mass reduction.	Regression estimates are random numbers which have an unknown expected value and variance, and known sample value and standard error. If the sample value can be explained by a zero expected value and known standard error then the result is considered not statistically significantly different than zero and therefore the result is not considered to be statistically significant. However, If we can combine this estimate with other estimates then the unknown expected values and variances can also be combined using the same transformation, and the statistical significance of the combined result can be tested. Therefore, depending on the sample values and inter-correlation, the combined result may be statistically significant even if the individual estimates are not statistically significant.
	For example, the results from each of the nine different crash types can be combined into an overall estimate and the standard error calculated assuming that the results for each crash type are independent of each other. Then the statistical significance of the combined effect can be determined.
	However, and Kahane (2011) points out there are two sources of uncertainty in the regression results. The first is the FARS based sampling error which is uncorrelated across crash types because they are based on different fatal cases (Kahane 2011, p. 77). The second is

	the state based induced-exposure sampling error which is correlated across crash types because they are based on the same induced-exposure cases. Therefore a confidence interval estimated using the "jackknife" method described by Kahane (2011) and accounting for correlation of these two error sources would be more accurate than a simple estimate based on the Wald Chi-Square statistic and assumed independence.
How might the LBNL methodology be strengthened to better represent future vehicle designs and reduce multi-collinearity between mass and footprint in the regression analysis?	The effects of multi-collinearity can be mitigated by 1) obtaining more data, 2) pooling data from different crash type or vehicle types, or 3) reducing the number of regression variables. The first option would require more calendar years and/or model years, which would involve added newer data as it becomes available (or using older data). The second option might be to recombine the CUVs and minivans with truck based vans and adding a control variable to compensate for the differences in the vehicles types. The third option might involve removing statistically insignificant control variables or removing control variables that would not be expected to have an effect on the probability of fatality in the crash (e.g., the side airbag variable is not included in pedestrian crashes because it is not expected to affect pedestrian fatality risk). The number of driver age control variables might be reduced from eight to three (as in the Kahane (1997) and DRI (2002-2005) studies). Finally, a linear curb weight model instead of a two-piece linear model may help to better elucidate the general trend.

The Variance Inflation Factor (VIF) has been suggested as a measure of multi-collinearity in the Kahane (2010 and 2011) reports, however this diagnostic metric does not account for differences in database size (i.e., Options 1 and 2 above). The Wenzel (2011a) report does not discuss the Variance Inflation Factor or report any VIF results. |
| ADDITIONAL COMMENTS:

None come to mind. | |

4. DATA SETS	COMMENTS
Please comment on the validity and applicability of the datasets used to project changes in risk resulting from reduction in vehicle mass. LBNL's casualty analysis used police-reported crash data from 16 states, while the 2011 NHTSA study used national fatality data, combined with a subset of non-culpable vehicles involved in two-vehicle crashes from police-reported crash data from 13 states.	The induced-exposure data set provided by NHTSA is based on the non-culpable vehicles in two-vehicle crashes. See the comments in Table 1 on the limitations of this data. In addition, there are also many differences in the coding variables and values used by the different states, which tend to make the recoding to a common data set imprecise. [Note: This peer review does not address the Wenzel (2011b) companion report, which examines the risks per police-reported crash. See Ref 4 for comments on the companion report.]
Please comment on any apparent, unstated, or implicit impact on estimated risks inherent in the two different approaches, and any related caveats or limitations. For example, what are the strengths and weaknesses of the two measures of vehicle exposure, miles of vehicle travelled scaled up from crash data from 13 states, and number of police-reported crashes?	The number of fatal cases tends to be much less than the number of induced-exposure cases. Therefore the effective numbers of degrees-of-freedom in the statistical estimates tend to be limited by the available number of fatal cases. For example, it would not be possible to estimate the effects of two variables (e.g., just curb weight and footprint) if we had data for only one fatal case even if we had thousands of induced-exposure cases. Therefore it is desirable to use data for the entire US in order to get a large sample of fatal cases for the logistic regressions. This then requires the available induced-exposure data (i.e., from 13 states) to be "scaled up" the US level using the method described in Kahane (2003 and 2011). The result is the best currently available estimate of vehicle exposure. There may be some concerns about the accuracy of the vehicle miles-travelled data because the difficulty estimating the number of vehicle miles travelled at the make-model-year-state level of detail. [Note: This peer review does not address the Wenzel (2011b) companion report, which examines the risks per police-reported crash. See Ref 4 for comments on the companion report.]

ADDITIONAL COMMENTS:

None come to mind.

5. RECOMMENDATIONS	COMMENTS
Please comment on whether the LBNL assessment adequately addresses the NHTSA 2011 study and identifies the safety impact from mass reduction. Are the analytic methods and data used to assess the NHTSA study, and estimate the relationship between risk, mass, and footprint, appropriate? Is casualty risk per crash a legitimate measure of vehicle safety? What other methods or data could be used to better predict the effect of future vehicle designs on safety?	The basic methodology described by Kahane (2011) seems appropriate; however some results using this method and data are not well understood and need further diagnosis. The induced-exposure data set provided by NHTSA is based on the non-culpable vehicles in two-vehicle crashes. See the Table 1 comments on the limitations of this data. [Note: This peer review does not address the Wenzel (2011b) companion report, which examines the risks per police-reported crash.]
Please comment on the overall adequacy of LBNL's assessment of the 2011 NHTSA report and its independent study of casualty risk for predicting the effect of vehicle mass or footprint reduction on safety. Provide any recommended improvements that might reasonably be adopted by the author to improve the analysis.	The Wenzel (2011a) report provides a valuable supplement to the analysis and results in the Kahane (2011) report. [Note: This peer review does not address the Wenzel (2011b) companion report, which examines the risks per police-reported crash.]
ADDITIONAL COMMENTS: See additional comments and recommendations in Tables 6 and 7.	

Table 6. Additional General Comments and Recommendations

Section	COMMENTS AND RECOMMENDATIONS
3	Use of R^2 is confusing. Suggest using lower case "r" when referring to the sample correlation coefficient (Box, Hunter, Hunter, 1978, P. 61); or upper case R when referring to the regression coefficient of multiple determination (Draper and Smith, 1981, p. 90).
All	In most cases the reported results are just estimates, but are not described as such. For example, "The effect of mass reduction on heavier cars and CUVs and minivans are not statistically significant" on p. iii should say "The ***estimated*** effect of mass reduction on heavier cars and CUVs and minivans are not statistically significant." This distinction is important when comparing results based on different models and assumptions because the different models and assumptions do not change the effect itself, but rather the estimate of the effect. For example, the statement "The first sensitivity, in dark purple, includes the weight variables in the regression model but excludes the footprint variable; this model tests the effect of mass reduction while allowing footprint to vary with vehicle mass. This sensitivity increases the risk from a 100-lb mass reduction in cars (from 1.43% to 2.64% for lighter cars, and from 0.48% to 1.94% for heavier cars) and CUVs/minivans (from a 0.47% decrease in risk to a 0.52% increase in risk); however, there is no change in fatality risk in light-duty trucks" on page 15 is misleading. It would be better to state that "The first sensitivity, in dark purple, includes the weight variables in the regression model but excludes the footprint variable; this model tests the ***estimated*** effect of mass reduction while allowing footprint to vary with vehicle mass. ~~This sensitivity~~***Removing the footprint variable from the regression model*** increases the ***estimated*** risk from a 100-lb mass reduction in cars (from 1.43% to 2.64% for lighter cars, and from 0.48% to 1.94% for heavier cars) and CUVs/minivans (from a 0.47% decrease in risk to a 0.52% increase in risk); however, ~~there is no~~***the*** change in ***the estimated*** fatality risk in light-duty trucks ***is very small and not statistically significant***." This also applies to table and figure captions. For example, "Table ES.1. Effect of mass and footprint reduction on fatality risk, under alternative regression model specifications" should say "Table ES.1. ***Estimated effects*** of mass and footprint reduction on fatality risk, under alternative regression model specifications." "Figure 3.3 Effect of reduction in mass or footprint on US fatality risk per VMT, by vehicle type: mass only, footprint only, and both" should say "Figure 3.3 ***Estimated effects*** of reduction in mass or footprint on US fatality risk per VMT, by vehicle type: mass only, footprint only, and both." Overall the word "effect" appears over 200 times in this report with the "estimated" or other qualifier. In some cases this may be appropriate and in other cases it is not appropriate. It is recommended that the author review each instance and revise as appropriate.

4	Figures 4.1 through 4.17 do not control for the effect of vehicle size (e.g. footprint), which has been shown to be correlated with vehicle weight (e.g., Figure 3.1), and therefore these figures may be misleading. It is strongly suggested that the horizontal axis label be changed to "Curb weight (lbs) and corresponding changes in size," and/or a note such as the following be added to each figure: "Note these results do not control for the effect of vehicle size on fatality risk. Therefore the horizontal axis represents changes to both vehicle weight and vehicle size."
4	The statistical significance of the linear trends in Figures 4.1 through 4.17 are not reported. It would be helpful if the confidence intervals or statistical significance of the linear trends were reported, either in addition to or instead of R^2. The confidence intervals for the estimated slopes should be added to the results in Tables 4.1 and 4.2.

Table 7. Additional Specific Comments and Recommendations

Section	Page	COMMENTS AND RECOMMENDATIONS
Executive Summary, 7	iii, 65	2nd paragraph refers to "our analysis," however the results are the same as the NHTSA analysis. The author should clarify who or what "our analysis" refers to and how it relates to the NHTSA analysis. Perhaps the statement "LBNL was able to reproduce the NHTSA analysis, which finds that…" would be more appropriate.
Executive Summary	iv	Last bullet – suggest changing the statement that "Logistic regression does not allow a statistic" to "Logistic regression *methods do not have* a statistic."
Executive Summary, 4	Iv, v, 22, 66	Suggest changing "variance in risk" to *variation* in risk" throughout.
Executive Summary, 7	viii, 69	The numerical results for the NHTSA preferred model in Tables ES.1 and 7.1 are slightly different than the results reported in the NHTSA report. For example 1.43%/0.48%/0.52%/-0.40%/-0.47% should be *1.44*%/*0.47*%/0.52%/-*0.39*%/-*0.46*%
4	31-32	Figures 4.12 through 4.14 have the results for small and heavy-duty pickups combined, which is inconsistent with the results in Table 4.1
4	35	The R^2 values in Table 4.1 are different than the values in Figures 4.6, 4.8, 4.9, 4.11.
5.1	37	The subsection title should be "Alternative measures of exposure *and outcome*" because fatal crashes and fatalities are measures of the crash outcome, not exposure.
5.2	39	It would be helpful to list the 18 manufacturer dummy variables in a table.
5.2	39	It is unclear why Lexus, Acura, and Infinity are treated as separate manufacturers, but Cadillac and Lincoln are not.
5.2	39	It is unclear why AM General is considered a Chrysler brand. The AM General Hummer was sold by GM beginning with the 2001 model year.
5.3	44	It would be helpful if the figures include error bars or shading to indicate the confidence intervals.
6.4	63	Table 6.3 – Suggest adding a note that the 72,316 total includes fatalities that are counted more than once in crashes involving more than one vehicle type.

Michael Van Auken Review of LBNL Phase 2 Report	
1. ASSUMPTIONS	**COMMENTS**
Please comment on the validity of any assumptions embedded in the LBNL assessment analysis and the independent casualty analysis that could affect the projected relationship between vehicle mass/footprint reductions and fatality/casualty risk. Examples might include assumptions regarding whether recent historical relationships between vehicle weight, size, and safety will continue into the future; potential future improvements in vehicle technology and design may result in compensatory safety benefits; and the annual baseline fatality distribution.	The basic assumptions, methodology, and data are primarily the same as in the Kahane (2011) report, but have been extended to include serious injuries as well as fatalities, and also address crash involvement (i.e., fatalities and serious injuries per accident, and also accidents per VMT).. These include the following: 1) The probability of a crash fatality or serious injury is proportional to the number of accidents (provided the crash conditions remain the same); and the probability of an accident is proportional to the vehicle miles travelled (VMT). 2) The logarithm of probabilities of fatality or serious injury per accident, and accidents per VMT for a given curb weight, footprint, and control variable values varies as a linear combination of the curb weight, footprint, and control variables within the domain of the data. 3) The logistic regression methods determine a maximum likelihood estimate of model coefficients. 4) It is assumed that the above relationships remain constant in the recent past (i.e., 2000-2007 model year vehicles in the 2002-2008 calendar years), present, and near future (i.e., 2017-2025 model year vehicles). The first assumption that crash fatalities and serious injuries are proportional to the number of accidents provided the crash conditions remain the same seems self-evident (e.g., if two fatal crashes had exactly the same conditions, then the expected number of fatalities for the two crashes would be twice the value for just one of the crashes). The assumption that the number of accidents are proportional to VMT rather than the number of vehicle registration years (VRY) is also appropriate because accidents cannot occur if the vehicles are not

| | driven on the road (i.e., VMT = 0). This assumption is qualified however because VMT is more difficult to measure than VRY and therefore may be less accurate. On the other hand the probability of a fatal crash or the number of fatalities in a crash may also depend on the vehicle occupancy.

The second and third assumptions are appropriate provided that it is recognized that it is essentially impossible with currently available knowledge and information to model all of the factors that could affect the probability of fatality in a crash, and that the objective of the analysis is to identify overall trends versus vehicle weight and footprint. In general the probability of fatality depends on other many other factors which have not been modelled (e.g., driver behavior factors, vehicle design factors, roadway design factors, EMT factors), and these unmodeled factors are assumed to be uncorrelated with vehicle weight and footprint, and/or are represented by the other control variables. The latter assumption might or might not be valid.

The fourth assumption is perhaps the weakest because it assumes that future vehicles will have the same design characteristics as past vehicles, and that the characteristics of the vehicle population (e.g., collision partner weight, size, type) will also remain the same. |
|---|---|
| Please comment on any apparent unstated or implicit assumptions and related caveats or limitations. | The induced-exposure data set provided by NHTSA is based on the "non-culpable" vehicle in two-vehicle crashes. It is assumed that the dataset is a reprehensive sample of the driver and environmental exposure factors for vehicle use. However, since these cases include moving vehicles, some vehicle-driver-environmental conditions may be under or over represented in this data depending on how they affect the ability of a non-culpable vehicle to avoid a crash. Results in Ref (17) indicated that the estimated effect of weight and size reduction are sensitive to whether the induced-exposure data are based on the Kahane (2003) non-culpable vehicle definition of the Kahane (1997) stopped vehicle definition. |

| | Unfortunately it is not currently possible to test this sensitivity with the NHTSA-provided induced-exposure data.

It is also assumed that the accident data from the 13 or 16 states are representative of all US states. Figure 2.1 in Wenzel (2011b) provides a useful comparison of the distribution of fatalities in the US and 13 states by the nine different crash types. |
|---|---|
| ADDITIONAL COMMENTS:

None come to mind. | |

2. CONTROL AND DEPENDENT VARIABLES USED IN THE REGRESSION MODELS	COMMENTS
Please comment on the adequacy of control and dependent variables used in the assessment analysis and independent casualty analysis, and recommend any alternative control or dependent variables that are available for possible inclusion in the analysis. For example, what are the relative merits of the main dependent variables used, fatality risk per estimated VMT, and casualty risk per police-reported crash?	Reducing the total number of fatalities and serious injuries is desirable from a societal viewpoint. This includes both subject vehicle occupant and collision partner (e.g., other vehicle occupant, pedestrian) fatalities and serious injuries. VMT is a good measure of accident exposure provided that it can be accurately determined. The number of fatalities and serious injuries per accident is a measure of vehicle crashworthiness (i.e., effect of a crash on the subject vehicle occupants) and crash compatibility (i.e., effect of a crash on the other vehicle occupants or vulnerable road users). Subject vehicle occupant fatalities and serious injuries per accident are a measure of the subject vehicle crashworthiness. Collision partner fatalities and serious injuries per accident are a measure of vehicle crash compatibility. The number of accidents per VMT is a measure of the crash avoidance capabilities of a given vehicle.
What additional control variables, such as vehicle make or model, might be included in the regression models?	None come to mind.
Please comment on any caveats or limitations that these dependent variable or control variables entail with respect to use of the results as the basis for estimating the safety effect of mass reduction.	None come to mind.
ADDITIONAL COMMENTS: The underlying reasons for some of the estimated effects are unknown at this time, but presumably involve driver, vehicle, environment or accident factors than have not been controlled for in the Kahane (2011) and Wenzel (2011b) analyses. See, for example, Refs 17 and 23.	

3. METHODOLOGY AND STATISTICS	COMMENTS
Please comment on the validity and applicability of the methodology LBNL used in assessing the NHTSA 2011 study and its analysis of the relationship between mass, footprint, and risks per police-reported crash.	The logistic regression methods seem to be appropriate. The confidence intervals are based on the logistic regression Wald Chi-Square statistic, which as Kahane (2003, 2011) has demonstrated does not include all sources of variation. However, these confidence intervals are useful because they do provide some indication of the uncertainty in the results.

The two-stage results for the 13-state fatalities per crash and 13-state crashes per VMT, and the one-stage result for 13-state fatalities per VMT reported in Tables ES.1 and 6.1 and Figure 2.7 were computed using independent logistic regressions. The differences between the two-stage results and the one-stage results for fatalities per crash could have been eliminated by using the "simultaneous three-way" logistic regression method described in DRI (2003). This method imposes the constraint that the combined two-stage estimated and the one-stage estimated are equal. |
| Please review other statistical methods LBNL has used in the analysis, in addition to the logistic regression methodology. Examples include the alternative approaches used by LBNL to assess NHTSA interval estimation results, and LBNL's linear regression analysis of actual, predicted, and residual risk by vehicle model. | The correlations in Section 3 appear to be assessed using the Coefficient of Multiple Determination (R^2) based on a linear fit to the data (e.g., the correlation between footprint versus curb weight in Figure 3.1 on p. 32). The linear regression model attributes the differences between the dependent variable (vertical axis) and the linear fit to the independent variable (horizontal axis) to random effects. If there is no preference as to the choice of independent and dependent variables (e.g., footprint versus curb weight, or curb weight versus footprint), then the linear trend and R^2 result would be different if the two variables were interchanged, and having two different yet equally valid results would be undesirable.

If the variation in the data can be attributed to both variables (e.g., footprint and curb weight), then it would be better to report the square of the sample correlation coefficient r^2, where r is computed according to Eqn (1). The trend lines in these correlation figures should |

	not be computed using a linear regression. Instead, the trend line should pass through the sample means (i.e., (x, y)), and have a slope equal to the ratio of the sample standard deviations in the data (i.e., s_y / s_x). Therefore, the reported correlation results do not depend on the ordering of the data variables. Note this comments does not apply to linear trends indicated in Section 4, for which the Coefficient of Multiple Determination (R^2) seems appropriate. The Coefficient of Multiple Determination (R^2) is frequently used in the Wenzel (2011b) report as an indicator of the statistical importance of a linear trend (e.g., R^2 values in Tables 4.1 and 4.2 were compared to 0.3). It would be better to report the standard error, confidence interval, and/or probability value as measures of the statistical significance of a linear trend.
Please comment on caveats or limitations of using non-significant regression estimates to project the safety impact of mass reduction.	Regression estimates are random numbers which have an unknown expected value and variance, and known sample value and standard error. If the sample value can be explained by a zero expected value and known standard error then the result is considered not statistically significantly different than zero and therefore the result is not considered to be statistically significant. However, If we can combine this estimate with other estimates then the unknown expected values and variances can also be combined using the same transformation, and the statistical significance of the combined result can be tested. Therefore, depending on the sample values and inter-correlation, the combined result may be statistically significant even if the individual estimates are not statistically significant. For example, the results from each of the nine different crash types can be combined into an overall estimate and the standard error calculated assuming that the results for each crash type are independent of each other. Then the statistical significance of the combined effect can be

	determined.
	However, and Kahane (2011) points out there are two sources of uncertainty in the regression results. The first is the FARS based sampling error which is uncorrelated across crash types because they are based on different fatal cases (Kahane 2011, p. 77). The second is the state based induced-exposure sampling error which is correlated across crash types because they are based on the same induced-exposure cases. Therefore a confidence interval estimated using the "jackknife" method described by Kahane (2011) and accounting for correlation of these two error sources would be more accurate than a simple estimate based on the Wald Chi-Square statistic and assumed independence.
How might the LBNL methodology be strengthened to better represent future vehicle designs and reduce multi-collinearity between mass and footprint in the regression analysis?	The effects of multi-collinearity can be mitigated by 1) obtaining more data, 2) pooling data from different crash type or vehicle types, or 3) reducing the number of regression variables. The first option would require more states (for serious injuries and police-reported accidents), calendar years and/or model years, which would involve added newer data as it becomes available (or using older data). The second option might be to recombine the CUVs and minivans with truck based vans and adding a control variable to compensate for the differences in the vehicles types. The third option might involve removing statistically insignificant control variables or removing control variables that would not be expected to have an effect on the probability of crash or crash outcome (e.g., the side airbag variable is not included in pedestrian crashes because it is not expected to affect pedestrian fatality risk). The number of driver age control variables might be reduced from eight to three (as in the Kahane (1997) and DRI (2002-2005) studies). Finally, a linear curb weight model instead of a two-piece linear model may help to better elucidate the general trend.
	The Variance Inflation Factor (VIF) has been suggested as a measure of multi-collinearity in the Kahane (2010 and 2011) reports, however this

	diagnostic metric does not account for differences in database size (i.e., Options 1 and 2 above). The Wenzel (2011b) report does not discuss the Variance Inflation Factor or report any VIF results.
ADDITIONAL COMMENTS: None come to mind.	

4. DATA SETS	COMMENTS
Please comment on the validity and applicability of the datasets used to project changes in risk resulting from reduction in vehicle mass. LBNL's casualty analysis used police-reported crash data from 16 states, while the 2011 NHTSA study used national fatality data, combined with a subset of non-culpable vehicles involved in two-vehicle crashes from police-reported crash data from 13 states.	The induced-exposure data set provided by NHTSA is based on the non-culpable vehicles in two-vehicle crashes. See the comments in Table 1 on the limitations of this data. The use of property damage accident data and cases with serious injury from the 13 states seems appropriate (with the noted qualification that the different states may have different accident reporting thresholds and injury reporting criteria). The concerns about the use of data for the 3 additional states (Georgia, Illinois, and New Mexico) have also been noted. In addition, there are also many differences in the coding variables and values used by the different states, which tend to make the recoding to a common data set (either induced-exposure, police-reported accident, or severe injury) imprecise.
Please comment on any apparent, unstated, or implicit impact on estimated risks inherent in the two different approaches, and any related caveats or limitations. For example, what are the strengths and weaknesses of the two measures of vehicle exposure, miles of vehicle travelled scaled up from crash data from 13 states, and number of police-reported crashes?	The number of fatal or serious injury cases tends to be much less than the number of induced-exposure cases (and the number of police-reported accidents). Therefore the effective numbers of degrees-of-freedom in the statistical estimates tend to be limited by the available number of fatal or serious injury cases. For example, it would not be possible to estimate the effects of two variables (e.g., just curb weight and footprint) if we had data for only one fatal of serious injury case even if we had thousands of induced-exposure cases. Therefore it is desirable to use data for the entire US in order to get a large sample of fatal cases for the logistic regressions. This then requires the available induced-exposure data (i.e., from 13 states) to be "scaled up" the US level using the method described in Kahane (2003 and 2011). The result is the best currently available estimate of vehicle exposure. There may be some concerns about the accuracy of the vehicle miles-travelled data because the difficulty estimating the number of vehicle miles travelled at the make-model-year-state level of detail.

ADDITIONAL COMMENTS:

None come to mind.

5. RECOMMENDATIONS	COMMENTS
Please comment on whether the LBNL assessment adequately addresses the NHTSA 2011 study and identifies the safety impact from mass reduction. Are the analytic methods and data used to assess the NHTSA study, and estimate the relationship between risk, mass, and footprint, appropriate? Is casualty risk per crash a legitimate measure of vehicle safety? What other methods or data could be used to better predict the effect of future vehicle designs on safety?	The basic methodology described by Kahane (2011) seems appropriate; and the extension by Wenzel (2011b) are also appropriate. However some results using these methods and data are not well understood and need further diagnosis. The induced-exposure data set provided by NHTSA is based on the non-culpable vehicles in two-vehicle crashes. See the Table 1 comments on the limitations of this data. The state accident data files tend to have different database variable and coding definitions and criteria, which could confound the results.
Please comment on the overall adequacy of LBNL's assessment of the 2011 NHTSA report and its independent study of casualty risk for predicting the effect of vehicle mass or footprint reduction on safety. Provide any recommended improvements that might reasonably be adopted by the author to improve the analysis.	The Wenzel (2011b) report provides a valuable supplement to the analysis and results in the Kahane (2011) report.
ADDITIONAL COMMENTS: See additional comments and recommendations in Tables 6 and 7.	

Table 6. Additional General Comments and Recommendations

Section	COMMENTS AND RECOMMENDATIONS
4	Use of R^2 is confusing. Suggest using lower case "r" when referring to the sample correlation coefficient (Box, Hunter, Hunter, 1978, P. 61); or upper case R when referring to the regression coefficient of multiple determination (Draper and Smith, 1981, p. 90).
All	In most cases the reported results are just estimates, but are not described as such. The word "effect" appears several hundred times in this report with the "estimated" or other qualifier. In some cases this may be appropriate and in other cases it is not appropriate. It is recommended that the author review each instance and revise as appropriate.
All	"Crashworthiness" in most instances should be changed to "crashworthiness and crash compatibility" because the fatalities and/or serious injuries may either be in the subject vehicle (crashworthiness effect) or collision partner (crash compatibility effect).
4	The statistical significance of the linear trends in Figures 4.1 through 4.9 are not reported. It would be helpful if the confidence intervals or statistical significance of the linear trends were reported, either in addition to or instead of R^2. The confidence intervals for the estimated slopes should be added to the results in Tables 4.1 and 4.2.

Table 7. Additional Specific Comments and Recommendations

Section	Page	COMMENTS AND RECOMMENDATIONS
Executive Summary, 4	iv, v, 22, 66	Suggest changing "variance in risk" to "***variation*** in risk" throughout.
Executive Summary, 6	vii, 63	The statement "In conclusion, casualty risk per crash is not necessarily a better metric than fatality risk per VMT for evaluating the effect of mass or footprint reduction on risk; rather, it provides a different perspective in assessing the benefits or drawbacks of mass and footprint reduction on safety in vehicles. However, it does allow the separation of risk per VMT to be separated into its two components, crash frequency and risk per crash" suggests that the casualty risk per crash metric was needed in order to assess the crash frequency and risk per crash, which is incorrect. The DRI (2003-2012) methods have also estimated the effects of weight and size on crash frequency (A/E) and risk per crash (F/E) in terms of fatalities.
5.3	49	It would be helpful to list the 18 manufacturer dummy variables in a table.
5.3	48-49	It is unclear why Lexus, Acura, and Infinity are treated as separate manufacturers, but Cadillac and Lincoln are not.
5.3	49	It is unclear why AM General is considered a Chrysler brand. The AM General Hummer was sold by GM beginning with the 2003 model year.

August 9, 2012

<u>MEMORANDUM</u>

SUBJECT: EPA Response to Comments on the peer review of *LBNL Statistical Analysis of the Effect of Vehicle Mass & Footprint Reduction on Safety) (LBNL Phase 1 and 2 Reports)*, prepared by Tom Wenzel, Lawrence Berkeley National Laboratory.

FROM: Cheryl Caffrey, Assessment and Standards Division
                     Office of Transportation and Air Quality, U.S. Environmental Protection Agency

The *LBNL Phase 1 and 2 Reports* were reviewed by Donna Chen and Kara Kockelman (University of Texas at Austin), Charles Farmer (Insurance Institute for Highway Safety, David Greene (Oak Ridge National Laboratory), and Michael Van Auken (Dynamic Research, Inc.).

This memo includes a compilation of comments prepared by SRA International; LBNL's responses to, and actions in response to, those comments are interspersed with the comments and noted by [LBNL].

Assessment of NHTSA's Report "Relationships Between Fatality Risk, Mass, and Footprint in Model Year 2000-2007 Passenger Cars and LTVs" (LBNL Phase 1 Report)	
1. ASSUMPTIONS	**COMMENTS**
Please comment on the validity of any assumptions embedded in the LBNL assessment analysis and the independent casualty analysis that could affect the projected relationship between vehicle mass/footprint reductions and fatality/casualty risk. Examples might include assumptions regarding whether recent historical relationships between vehicle weight, size, and safety will continue into the future; potential future improvements in vehicle technology and design may result in compensatory safety benefits; and the annual baseline fatality distribution.	**[Chen and Kockelman]** [1] The report does a nice job discussing recent trends in vehicles, such as the increase of ESC, side airbags, and light truck crash compatibility with passenger cars – which will improve safety outcomes for all vehicles, but perhaps most significantly the smaller and lighter vehicles. . It also mentions the phasing out of the lightest and smallest vehicles between model years 2000-2007 (but doesn't mention the makes and models somehow), which were particularly poor safety performers in the past. However, with the introduction of urban commuter vehicles, such as the SmartCar, Mini Cooper, and Fiat 500m, and the growing popularity of smaller, fuel-efficient compact vehicles following gas price increases, this trend does not seem so obvious. Such vehicles should be discussed.

[LBNL] The vehicles mentioned were particular models of small light vehicles that had poor on-road safety records. However as Figures 4.7 and 4.10 indicate, there are several recent car models weighing less than 2,500 lbs that have fatality risks similar to models weighing 1,000 lbs more, even after accounting for all differences in vehicle, driver and crash characteristics except mass and footprint.

[2] The simplistic logistic model employed in this analysis only accounts for two crash outcomes (fatal versus non-fatal) and so neglects the more detailed, and ordered nature of injury severity data, which is unfortunate. The model also assumes error-term homoscedasticity from one crash or individual to the next; in reality certain vehicle types (e.g., pickups) and crash contexts (e.g., high speed crashes) have more uncertainty associated with their severity outcomes. It would be good to point out such limitations for readers.

[LBNL] The LBNL Phase 1 report followed the methodology of the NHTSA 2011 report, which analyzes the estimated relationships between vehicle mass, footprint and US fatalities per VMT. The LBNL Phase 2 report analyzes the relationships between vehicle mass, footprint, and casualties, defined as fatalities plus serious or incapacitating injuries. The effects on other injury severities (i.e. minor injuries) were not analyzed because of the differences in reporting of injury severity in the thirteen states.

[Van Auken] The basic assumptions, methodology, and data are primarily the same as in the Kahane (2011) |

report. These include the following:

1) The probability of a crash fatality is proportional to the vehicle miles travelled (VMT), except as noted in Section 5.1
2) The logarithm of probability of fatality per VMT for a given curb weight, footprint, and control variable values varies as a linear combination of the curb weight, footprint, and control variables within the domain of the data.
3) The logistic regression methods determine a maximum likelihood estimate of model coefficients.
4) It is assumed that the above relationships remain constant in the recent past (i.e., 2000-2007 model year vehicles in the 2002-2008 calendar years), present, and near future (i.e., 2017-2025 model year vehicles).

The first assumption that crash fatalities are proportional to VMT rather than the number of vehicle registration years (VRY) is appropriate because the fatalities cannot occur if the vehicles are not driven on the road (i.e., VMT = 0). This assumption is qualified however because VMT is more difficult to measure than VRY and therefore may be less accurate. On the other hand the probability of a fatal crash or the number of fatalities in a crash may also depend on the vehicle occupancy. The analysis in Section 5.1 is a commendable attempt to explore the sensitivity to this assumption, however the Kahane (1997) and DRI (2003-2005) reports have shown that some driver, vehicle, and environmental factors may be underrepresented or overrepresented in unweighted induced-exposure data. VRY could have also been considered as a measure of exposure.

[LBNL] NHTSA and LBNL used vehicle miles traveled, rather than vehicle registration-years, as the measure of exposure because a vehicle that is not driven has zero risk. Nonetheless, LBNL conducted a sensitivity using vehicle registration years rather than VMT as the measure of exposure. This alternative resulted in lower estimated effects of mass reduction on risk in lighter cars and light trucks, no change in CUVs/minivans, but a substantially higher estimated effect of mass reduction in heavier cars (from an estimated 0.51% increase in risk to an estimated 2.40% increase in risk, as shown in Alternative 6 in new Table ES.1).

The second and third assumptions are appropriate provided that it is recognized that it is essentially impossible with currently available knowledge and information to model all of the factors that could affect the probability of fatality in a crash, and that the objective of the analysis is to identify overall trends versus vehicle weight and footprint. In general the probability of fatality depends on other many other factors which have not been modeled (e.g., driver behavior factors, vehicle design factors, roadway design factors, EMT factors) and these unmodeled factors are assumed to be uncorrelated with vehicle weight and footprint,

	and/or are represented by the other control variables. The latter assumption might or might not be valid. The fourth assumption is perhaps the weakest because it assumes that future vehicles will have the same design characteristics as past vehicles, and that the characteristics of the vehicle population (e.g., collision partner weight, size, type) will also remain the same. A commendable attempt to partially address this effect is described in Section 6. These effects can be perhaps better addressed by the "Volpe model" described in Kahane (2011) of the Honda-DRI fleet systems model described in Refs (2324), which can be used to forecast the effects of mass reductions of individual makes and models on a year-by-year basis.
Please comment on any apparent unstated or implicit assumptions and related caveats or limitations.	**[Chen and Kockelman]** The role of driver behavior is briefly addressed in the report but not emphasized sufficiently. Fatality risk is a combination of driver, vehicle, and roadway characteristics. Driver behavioral differences are many and do not solely exist for pickup truck drivers versus car drivers. Socioeconomic data such driver household income, size, and education influence driver attitudes and driving environments. For example, Chen et al. (2010) found that crash risk increases for those living in socioeconomically disadvantaged areas (including households more likely to drive less expensive and older vehicles). Though such data is not typically available in state and national crash databases, the importance of these driver and environmental characteristics on crash rates (per mile driven) and fatality risk should be stressed in both reports. It is clearly very difficult to control for, but a major caveat to the NHTSA (& now LBNL) results. We expect that crash severity could be probably be lower for many of the small cars and pickups if they were driven by those who tend to drive more expensive vehicles, under the same settings (e.g., daytime, urban freeway). Similarly, in the second LBNL report (which uses VMT estimates), we expect that crash rates would probably be lower for these types of driver-vehicle-setting combinations. *[LBNL] We agree that it would be preferable to include additional variables that account for driver behavior rather than just driver age and gender. However, variables such as driver age, gender, or other socio-economic characteristics probably have a less direct effect on risk than the actual driving skill or behavior of individual drivers. The FARS data includes information on whether alcohol or drug use, or speeding or reckless driving, was a factor in the current crash, as well as the driver's record for the last three years. NHTSA 2003 found that including a "bad driver" rating variable comprised of 8 of these variables did not have an appreciable effect on the relationship between vehicle mass and risk. Although these variables are included in the FARS fatal crash records, they are not included in the state crash databases, so a similar variable cannot be used in the regression models. However, Alternative 12 in new Table ES.2, and the third column in new Figure 5.14, indicates that excluding bad drivers (after also excluding reported driver alcohol or drug use) results in additional estimated increases in fatality risk as vehicle mass decreases, and estimated decreases in risk as*

vehicle footprint decreases. LBNL believes that driver behavior in the current crash and their recent driving record, and not general socio-economic information, is the best available measure to account for the effect driver skill and behavior has on risk in individual vehicles.

Nonetheless, LBNL explored the effect of adding a control variable for one additional driver socio-economic characteristic, household income. Unfortunately, crash data do not include household income of vehicle drivers. FARS includes the driver's zip code, from his or her license, which could be merged with US Census data by zip code to obtain the average median income of households in a particular zip code. However, very few states include driver zip code in their databases of police-reported crashes.

Kweon and Kockelman (2003) merged police-reported crash data from GES with household demographic data, including household income, from the National Household Travel Survey (NHTS), to examine crash rates per annual VMT by driver age, gender, and vehicle type. Each dataset includes population weights to derive national estimates of crashes and households, which the authors use to estimate national crash rates for driver age, gender and vehicle type. However, there are several limitations with the NHTS data. First, the survey includes only 25,000 national households, supplemented by an additional 124,000 households from "add-on" state and municipal jurisdictions. The national weights provided in the NHTS most likely do not represent the national distribution of vehicle registrations by make and model. Second, 30% of the households in the 2009 NHTS reported income "over $100,000"; using the income bins provided in the NHTS to estimate household income will likely understate the actual income of the households participating in the survey.

Chen el al (2010) made use of an Australian survey of 20,000 young adult drivers that asked questions regarding driving behavior. The injury severity of 127 of these participants who were admitted to hospitals after a vehicle crash was analyzed as a function of socio-economic status. The socio-economic status of the drivers was based on average or median values from the zip code where they (or their parents, since most were at University) lived.

Using 2010 California vehicle registration data, LBNL calculated the median household income by registered owner's zip code, using 2000 census data, then took the average household income by vehicle make and model. Predicted US fatality risk is correlated with this measure of average household income by vehicle model, with risk decreasing as income increases; the correlation is highest for cars and CUVs, with an R^2 of 0.60 (or a correlation coefficient r of 0.77) or higher, as shown in new Figure 4.20 and the first column of new Table 4.7. However, while income tends to increase as mass increases, the correlation is not strong (R^2 is highest for 4-door cars, at 0.16, and CUVs, at 0.19, as shown in new Figure 4.21 and the middle columns of new Table 4.7).

As with predicted fatality risk, there is a wide range in the average income of households owning vehicle models with similar mass. This analysis is summarized at the end of Section 4 of the Final Phase 1 report.

Using average household income by vehicle model, rather than driver alcohol or drug use, or reckless driving, as an alternative indicator of driver behavior substantially reduces the estimated detrimental effect of mass reduction on risk in cars, and slightly increases the estimated detrimental effect of mass reduction on risk in light trucks. This analysis is summarized at the end of Section 5.4 of the Final Phase 1 report.

[Farmer] The statistical models assume no interaction between the vehicle size/weight measures and any of the numerous covariates, but this may not be true. For example, size/weight reductions may differently affect vehicles with and without ESC if they affect vehicle handling. It is risky to make statements such as that on p. 11 of the Phase I report: *Therefore, the mass of a lighter car could be reduced by 800 lbs while adding ESC, without increasing fatality risk.*

[LBNL] T*he sentence has been revised to: "For instance, a 100-lb reduction in curb weight for an underweight car is estimated to increase risk by 1.4%, while installing ESC would reduce risk by 11.4%; the models estimate that the beneficial effect of adding ESC is nearly ten times that of reducing mass by 100 lbs. "*

[Van Auken] The induced-exposure data set provided by NHTSA is based on the "non-culpable" vehicle in two-vehicle crashes. It is assumed that the dataset is a reprehensive sample of the driver and environmental exposure factors for vehicle use. However, since these cases include moving vehicles, some vehicle-driver-environmental conditions may be under or over represented in this data depending on how they affect the ability of a non-culpable vehicle to avoid a crash. Results in Ref (17) indicate that the estimated effect of weight and size reduction are sensitive to whether the induced-exposure data are based on the Kahane (2003) non-culpable vehicle definition of the Kahane (1996) stopped vehicle definition.
Unfortunately it is not currently possible to test this sensitivity with the NHTSA-provided induced-exposure data.

[LBNL] *NHTSA developed vehicle miles traveled weighting factors for induced exposure cases based on stopped, rather than non-culpable, vehicles. LBNL investigated what effect this alternative method of exposure has on the estimated effect of mass or footprint reduction on risk, in Alternative 15 in Table ES.3 and Section 5.6 in the Final Phase 1 report.*

164

ADDITIONAL COMMENTS:

[Chen and Kockelman] Chen, H.Y., Ivers, R.Q., Mariniuk, A.L.C., Boufous, S., Senserrick, T., Woodward, M., Stevenson, M. and Norton R. Socioeconomic status and risk of car crash injury, independent of place of residence and driving exposure: Results from the DRIVE study. Journal of Epidemiology and Community Health 64(10), 2010, pp. 998-1003.

2. CONTROL AND DEPENDENT VARIABLES USED IN THE REGRESSION MODELS	COMMENTS
Please comment on the adequacy of control and dependent variables used in the assessment analysis and independent casualty analysis, and recommend any alternative control or dependent variables that are available for possible inclusion in the analysis. For example, what are the relative merits of the main dependent variables used, fatality risk per estimated VMT, and casualty risk per police-reported crash?	**[Chen and Kockelman]** [1] As alluded to above, a primary concern is that the NHTSA analysis (& thus the LBNL analyses) largely neglect the idea that vehicle type (make & model) is very much a proxy for driver type, and a vehicle's crash avoidance may have very little to do with vehicle type. It has a lot to do with the person behind the wheel, and gender & age simply aren't enough to control for such distinctions. Education, risk aversion, ability, wealth, etc., are important covariates. But existing data sets are quite limiting (though the MVOSS & FAR with 3-year driver violation history do offer some valuable insights, not discussed in these reports). In reality, small cars may be less crash prone than Kahane's & Wenzel's results suggest, because they are driven by lower-income, younger, less risk averse people driving in more crash prone settings (e.g., commercial strips rather than pricey residential suburbs). Such key caveats need thoughtful discussion. Four relevant papers on the topics of crash frequency and vehicle size-and-weight implications (by Knipling, Kweon & Kockelman, Wang and Kockelman, and Chen & Kockelman) have been sent to Tom Wenzel. These all include useful literature reviews for further connections to useful findings for citation in the reports, as time allows the contractor. *[LBNL] As discussed above, LBNL believes that driver behavior in the current crash and their recent driving record, and not general socio-economic information, is the best available measure to account for the effect driver skill and behavior have on risk in individual vehicles. As described in Section 5.4, LBNL used this measure of driver behavior in one of its sensitivity analyses in the draft final report. Nonetheless, LBNL examined the effect of including a measure of driver household income on risk in Section 5.4 in the Final Phase 1 report.* [2] The grouping of the vehicles into heavier- and lighter-than-average weight categories essentially splits a "typical" weight vehicle of that type into two categories. The impacts of curb weight and footprint on fatality risk may be easier to interpret if the vehicles were grouped into 3 weight categories (light, average, and heavy) & by type (with the average category representing vehicles within one standard deviation of average weight). Furthermore, the grouping of CUVs and minivans into the same vehicle type category neglects the fact that these vehicles have faced rather different ground clearance requirements (impacting rollover potential), door types (sliding vs. standard), and, perhaps most importantly, can appeal to different types of drivers (as indicated in the market shift of car drivers to CUV drivers). *[LBNL] As described in NHTSA 2011, NHTSA included CUVs and minivans in the same vehicle category in part because in most respects they are more "car-like" than conventional light-duty trucks (i.e. pickups, SUVs, and*

fullsize vans). A MINIVAN indicator variable is included in the regressions to estimate the separate effect of minivan vs. CUV mass reduction on risk. The relatively small number of CUVs and minivans during the analysis period precludes them from being separated into two distinct groups for analysis.

[Farmer] One needs to restrict control variables to those that are available and reliable. A problem when combining state databases is that the states often are not consistent as to the variables coded and the definitions of those variables. This severely limits the list of possible control variables.

[LBNL] *The control variables used from the state crash databases are either derived from the VIN, or based on the driver age and gender, or the time or location of the crash. LBNL believes that these variables are consistently coded among the thirteen states.*

[Van Auken] [1] The main metric used in both the Kahane (2011) and Wenzel (2011a) reports is the total number of fatalities (except as noted). Reducing the total number of fatalities, which includes both subject vehicle occupants and collision partner fatalities, is desirable from a societal viewpoint. Fatal crash occurrence is related to the total number of fatalities, which has been used by Kahane (2003, 2011) to address concerns about double counting.

[2] VMT is a good measure of accident exposure provided that it can be accurately determined. [Note: This peer review does not address the Wenzel (2011b) companion report (Ref 3) which examines the risks per police-reported crash. See Ref 4 for comments on the companion report.]

What additional control variables, such as vehicle make or model, might be included in the regression models?	**[Chen and Kockelman]** [1] Vehicle height, a variable which may be more valuable than vehicle type for similarly structured vehicles such as sedans, wagons, CUVs, and minivans, would be a valuable control variable. In addition to a wider track, a lower center of gravity also increases vehicle stability, thereby reducing the risk of rollover. Relevant literature & findings exist, and should be cited.
	[LBNL] *The propensity for a vehicle to roll over increases as its center of gravity increases (and static stability factor decreases), while risk in frontal crashes may decline as bumper height increases (and the degree of overlap with the bumper of a crash partner increases). However, overall vehicle height is only a crude proxy for these two other vehicle dimensions that are thought to correlate with safety, and therefore was not included in the regression models.*
	[2] Other variables which have been found in past studies to influence fatality risk such as seat belt use,

roadway geometry and division type are not included in this study (which is largely a repeat of the NHTSA study, as specifically contracted by the EPA).

[LBNL] Seat belt use was not included as a control variable because it is notoriously under-reported for non-fatal injuries. Roadway geometry and roadway division were not included because they are not consistently reported or coded in the state databases of police-reported crashes used to develop the induced exposure cases.

[3] To account for driver characteristics that contribute to fatality risk, socioeconomic variables such as household income, education, household size, etc. would be valuable additions. Unfortunately, both state and national crash databases typically do not include such information (outside of MVOSS). Such issues should be flagged for readers. It seems the contractor has done his duty, and the key limitations lie with the original methodology he was to essentially duplicate.

[LBNL] As discussed above, LBNL believes that driver behavior in the current crash and their recent driving record, and not general socio-economic information, is the best available measure to account for the effect driver skill and behavior have on risk in individual vehicles. As described in Section 5.4, LBNL used this measure of driver behavior in one of its sensitivity analyses in the draft final report. Nonetheless, LBNL examined the effect of including a measure of driver household income on risk in Section 5.4 in the Final Phase 1 report.

Please comment on any caveats or limitations that these dependent variable or control variables entail with respect to use of the results as the basis for estimating the safety effect of mass reduction.	**[Chen and Kockelman]** Please see above comment (in Assumptions section) regarding driver behavior and environment. *[LBNL] As discussed above, an additional sensitivity analysis was run using a control variable for household income, based on California vehicle registration data in 2010.* **[Farmer]** Model overspecification could be the reason for results that are non-intuitive, especially in the Phase II analyses of police-reported crashes. Control variables may be correlated with each other or with the size and weight variables. For example, Figure 2.9 of the Phase I report implies that torso side airbags increase fatality risk in CUVs. *[LBNL] It is possible that inclusion of too many control variables leads to non-intuitive results in certain cases. This problem might be resolved by removing control variables (other than vehicle weight or footprint) that are*

	not statistically significant. The example cited suggests that torso side airbags increase fatality risk in CUVs/minivans by an estimated 0.7%, is a small effect, and not statistically-significant. NHTSA and LBNL ran a sensitivity where non-significant control variables are removed; this sensitivity has only a small effect on the original estimates, as shown in Alternative 19 in Table ES.3 and in Section 5.6.

ADDITIONAL COMMENTS:

[Chen and Kockelman] Table 2.1 has many indicator variables labeled as "C" for continuous variable (such as ABS, ESC, AWD, DRVMALE, etc). These C's should be removed.

[LBNL] These vehicle variables are continuous because, for some models, the VIN does not indicate whether a particular vehicle is equipped with that option or not. In these cases the fraction of that model that is equipped with the particular feature is used. On the other hand, the DRVMALE variable is always a discrete variable, and is now coded as such in the table.

[Van Auken] The underlying reasons for some of the estimated effects are unknown at this time, but presumably involve driver, vehicles, environment or accident factors that have not been controlled for in the Kahane (2011) and Wenzel analyses. See, for example, Refs 17 and 25.

[LBNL] The report notes that other vehicle, driver, or crash characteristics may account for the remaining residual risk not explained by the control variables included in the regression models. As discussed above, additional sensitivity analyses were run using an alternative control variable for vehicle characteristics, initial vehicle purchase price (using values in Polk's VIN decoding software), and an alternative control variable for driver behavior, household income (based on California vehicle registration data in 2010). The results of these sensitivities are included in Sections 5.2 and 5.4 in the Final Phase 1 report.

3. METHODOLOGY AND STATISTICS	COMMENTS
Please comment on the validity and applicability of the methodology LBNL used in assessing the NHTSA 2011 study and its analysis of the relationship between mass, footprint, and risks per police-reported crash.	**[Chen and Kockelman]** The report assesses the NHTSA 2011 study in a fair amount of detail and seeks to introduce some additional analyses to better examine the relationship between mass, footprint, and fatality risks. However, due to a lack of control for very specific vehicle differences (which vary by make & sub-model), the exclusion of driver characteristics and crash setting details (which cannot always be controlled for, but are often correlated with vehicle type), the effects of downweighting vehicles and/or shifting vehicle styles and sizes may be overestimated. Simply changing the vehicle on a risky driver in a high-risk setting is unlikely to influence outcomes significantly.

[LBNL] As discussed above, additional sensitivity analyses were run using an alternative control variable for vehicle characteristics, initial vehicle purchase price (using values in Polk's VIN decoding software), and an alternative control variable for driver behavior, household income (based on California vehicle registration data in 2010). The results of these sensitivities are included in Sections 5.2 and 5.4 in the Final Phase 1 report.

[Van Auken] The logistic regression methods seem to be appropriate. The confidence intervals are based on the logistic regression Wald Chi-Square statistic, which as Kahane (2003, 2011) has demonstrated does not include all sources of variation. However, these confidence intervals are useful because they do provide some indication of the uncertainty in the results. [Note: This peer review does not address the Wenzel (2011b) companion report, which examines the risks per police-reported crash. See Ref 4 for comments on the companion report.] |
| Please review other statistical methods LBNL has used in the analysis, in addition to the logistic regression methodology. Examples include the alternative approaches used by LBNL to assess NHTSA interval estimation results, and LBNL's linear regression analysis of actual, predicted, and residual risk by vehicle model. | **[Chen and Kockelman]** [1] In the alternative measures of exposure, the author examines the effect of vehicle manufacturer on fatality risk and treats the luxury models produced by Toyota, Honda, and Nissan as separate manufacturers. However, domestic luxury brands (such as Cadillac & Lincoln) are categorized with their nameplate manufacturers (GM and Ford), which appears inconsistent.

[LBNL] The analysis has been revised to include the luxury models Lexus, Acura, and Infiniti with their nameplate manufacturers Toyota, Honda, and Nissan. An additional sensitivity was conducted using separate indicator variables for the five luxury brands Lexus, Acura, Infiniti, Cadillac, and Lincoln. Adding variables for the five luxury brands increases the estimated detrimental effect of mass reduction on risk for lighter cars (from 1.55% to 2.04 %), and dramatically increases the estimated detrimental effect of mass reduction on risk for heaver cars (from 0.51% in NHTSA's preferred model, to 0.75% after including 14 major vehicle makes, to 1.80% after including 14 makes plus the five luxury brands), as shown in Alternative 8 in Table ES.2. The vehicle manufacturer variables have little effect on the estimated effect of lighter light truck mass reduction, but both |

analyses reduce the slightly beneficial estimated effect of mass reduction in heavier light trucks. Including the manufacturer variables changes the estimated effect of mass reduction in CUVs/minivans from an estimated 0.38% reduction in risk to an estimated 1.62% or 1.28% <u>increase</u> in risk).

LBNL also examined the effect of replacing the vehicle make indicator variables with a single continuous variable for vehicle initial purchase price, taken from the Polk VIN decoder. Initial purchase price can be seen as a proxy for the general quality of vehicle design; we expect that more expensive vehicles will have lower risk than less expensive vehicles. Replacing the vehicle make variables with the variable for purchase price slightly reduces the estimated effect of mass reduction on lighter cars and light trucks, increases the estimated effect of mass reduction in heavier cars (from 0.51% to 0.84%), and substantially decreases the estimated effect of mass reduction in CUVs/minivans (from a 0.38% reduction in risk to a 0.92% reduction in risk). This analysis is summarized in Section 5.2 of the Final Phase 1 report.

[2] The effect of calendar year variables on fatality risk may be overestimated here, since VMT is tracked by vehicle model and not by calendar year. The trend of greatest fatality risk reductions in light trucks, CUVs and minivans with increasing calendar year may simply be a reflection of rising gas prices in combination with the ailing economy contributing to lower VMT (in these relatively low-fuel-economy vehicles).

[LBNL] As discussed in Section 5.3 of the LBNL Final Phase 1 report, the VMT weights NHTSA developed do not reflect the reduction in driving that occurred in 2008 (as shown in Figure 5.9). Therefore, the large reduction in risk estimated by the CY08 variables likely is the result of less driving in that year, and not the result of widespread changes in the fleet of vehicles on the road as potential crash partners, as NHTSA has surmised/postulated. However, the lower VMT in CY08 does not explain the consistent trend in reduced estimated fatality risk over the other calendar years.

[3] It is unclear how the author determined the various percentage replacements of vehicle types in the aggressive vehicle market share shift scenario. (For example, why are 50% of SUVs replaced by CUVs and 60% of small pickups replaced by CUVs? The CUV is a more natural replacement for an SUV, and an SUV a more natural replacement for a pickup.)

The aggressive shift in market share scenario was selected to maximize replacement of truck-based vehicles, such as pickups and SUVs, with car-based vehicles, such as CUVs and minivans, that tend to be safer and have

higher fuel economy than truck-based vehicles.

[Van Auken] [1] The correlations in Section 3 appear to be assessed using the Coefficient of Multiple Determination (R^2) based on a linear fit to the data (e.g., the correlation between footprint versus curb weight in Figure 3.1 on p. 14). The linear regression model attributes the differences between the dependent variable (vertical axis_ and the linear fit to the independent variable (horizontal axis) to random effects. If there is no preference as to the choice of independent and dependent variables (e.g., footprint versus curb weight, or curb weight versus footprint), then the linear trend and R^2 result would be different if the two variables were interchanged, and having two different yet equally valid results would be undesirable.

[LBNL] *The R^2 values are the same if the two variables are interchanged.*

[2] If the variation in the data can be attributed to both variables (e.g., footprint and curb weight), then it would be better to report the square of the sample correlation coefficient r^2, , where r is computed according to Eqn (1). The trend lines in these correlation figures should not be computed using a linear regression. Instead, the trend line should pass through the sample means (I.e. (x, y)), and have a slope equal to the ratio of the sample standard deviations in the data (i.e., s_y / s_x). Therefore, the reported correlation results do not depend on the ordering of the data variables.

[LBNL] *Again, the R^2 values are the same if the two variables are interchanged. The R^2 values in Figures 3.1 and 3.2, and Figures 4.1 through 4.5, have been replaced with the correlation coefficient r, and a new Table 3.1 has been added to Section 3 that lists r and VIF for curb weight and footprint by vehicle type.*

Note this comments does not apply to linear trends indicated in Section 4, for which the Coefficient of Multiple Determination (R^2) seems appropriate.

[3] The Coefficient of Multiple Determination (R^2) is frequently used in the Wenzel (2011a) report as an indicator of the statistical importance of a linear trend (e.g., R^2 values in Tables 4.1 and 4.2 that are greater than 0.3 are shown in blue font). It would be better to report the standard error, confidence interval, and/or probability value as measures of the statistical significance of a linear trend.

[LBNL] *Whether the estimated relationships shown in Tables 4.1 and 4.2 in the Final Draft report are statistically significant has been added to the tables (see new Tables 4.2 and 4.4 in the Final Phase 1 report). However, the point of the section is to indicate that, while the relationship may be statistically significant on*

	average, as indicated by the confidence level around the estimated effect, there remains a large range in risk for individual vehicle models with similar mass, as indicated by the relatively low R^2 values.
Please comment on caveats or limitations of using non-significant regression estimates to project the safety impact of mass reduction.	**[Chen and Kockelman]** First, the t-statistics are not provided in the report which makes it difficult for the reader to assess statistical significance of specific regression estimates (except where noted by the author). Second, inclusion of a statistically insignificant variable can influence the estimates of coefficients associated with related variables. Nevertheless, in general, it is best to keep insignificant estimates if one has a strong defense for their role, since removing such variables (& thus their parameters) will shift the burden of response to a correlated covariate's parameter, thus biasing the latter. We generally keep key covariates in a model up to a pvalue of 0.20 or 0.25 or so, especially in relatively small data sets (e.g., n < 1,000). Covariates for which we have no strong basis can be removed for p values > 0.10. *[LBNL] The 95% confidence intervals are included in every figure in the document, to indicate whether the estimated variable is statistically different from zero.* *NHTSA is planning to run a sensitivity where non-significant control variables are removed; LBNL will run similar sensitivities for its revisions to its Phase 1 and Phase 2 reports.* **[Van Auken]** Regression estimates are random numbers which have an unknown expected value and variance, and known sample value and standard error. If the sample value can be explained by a zero expected value and known standard error then the result is considered not statistically significantly different than zero and therefore the result is not considered to be statistically significant. However, If we can combine this estimate with other estimates then the unknown expected values and variances can also be combined using the same transformation, and the statistical significance of the combined result can be tested. Therefore, depending on the sample values and inter-correlation, the combined result may be statistically significant even if the individual estimates are not statistically significant. For example, the results from each of the nine different crash types can be combined into an overall estimate and the standard error calculated assuming that the results for each crash type are independent of each other. Then the statistical significance of the combined effect can be determined. However, and Kahane (2011) points out there are two sources of uncertainty in the regression results. The first is the FARS based sampling error which is uncorrelated across crash types because they are based on

	different fatal cases (Kahane 2011, p. 77). The second is the state based include-exposure sampling error which is correlated across crash types because they are based on the same included-exposure cases. Therefore a confidence interval estimated using the jackknife method described by Kahane (2011) and accounting for correlation of these two error sources would be more accurate than a simple estimate based on the Wald Chi-Square statistic and assumed independence. *[LBNL] LBNL agrees that NHTSA's jack-knife method is a preferable method of estimating confidence intervals; however, the jack-knife method obtains even larger estimates of uncertainty than using the standard errors output from the regression models.*
How might the LBNL methodology be strengthened to better represent future vehicle designs and reduce multi-collinearity between mass and footprint in the regression analysis?	**[Chen and Kockelman]** Including more vehicle-specific characteristics (such as vehicle height and engine size) reduces the analysis' dependence on vehicle type, since vehicle shapes and structures will continue to evolve. There is also correlation with context (e.g., pickups are driven in more rural locations, with greater hazards [like less lighting, higher speed, & few medians]). Disaggregate data are almost always best, to avoid ecological fallacies & such. *[LBNL] As discussed above, overall vehicle height is only a crude proxy for two vehicle dimensions, center of gravity and bumper height, that are thought to correlate with safety in rollover and two-vehicle crashes, respectively, and therefore was not included in the regression models. Other vehicle attributes that might affect risk are engine power-to-weight ratio, braking distance, and handling capabilities; LBNL may examine the effect of these accounting for these vehicle attributes in future analyses.* *LBNL's analysis does account for whether a vehicle is driven in a rural area, defined as a county in which the population density is less than 250 residents per square mile of land area.* **[Van Auken]** [1] The effects of multi-collinearity can be mitigated by 1) obtaining more data, 2) pooling data from different crash type or vehicle types, or 3) reducing the number of regression variables. The first option would require more calendar years and/or model years, which would involve added newer data as it becomes available (or using older data). The second option might be to recombine the CUVs and minivans with truck based vans and adding a control variable to compensate for the differences in the vehicles types. The third option might involve removing statistically insignificant control variables or removing control variables that would not be expected to have an effect on the probability of fatality in the crash (e.g., the side airbag variable

	is not included in pedestrian crashes because it is not expected to affect pedestrian fatality risk). The number of driver age control variables might be reduced from eight to three (as in the Kahane (1997) and DRI (2002-2005) studies). Finally, a linear curb weight model instead of a two-piece linear model may help to better elucidate the general trend. [2] The Variance Inflation Factor (VIF) has been suggested as a measure of multi-collinearity in the Kahane (2010 and 2011) reports, however this diagnostic metric does not account for differences in database size (i.e., Options 1 and 2 above). The Wenzel (2011a) report does not discuss the Variance Inflation Factor or report any VIF results. *[LBNL] A table of VIF results, Table 3.1, has been added to the Final Phase 1 report.*

ADDITIONAL COMMENTS:

[Chen and Kockelman] On page 55, it is unclear what is meant by "however; if anything, reduction of this type of fatality will increase detrimental effect of mass reduction in cars."

[LBNL] Full adoption of ESC is expected to reduce the number of, and fatalities in, rollovers and one-vehicle crashes. Because rollovers and one-vehicle crashes are the only types of crashes in which mass reduction is estimated to reduce fatality risk, reducing the number of these fatalities by full adoption of ESC increases the net estimated effect of mass reduction over all types of crashes. Similarly, full adoption of side airbags is estimated to reduce the number of fatalities in side impact crashes, as least for cars (Figure 6.4). However, Figure 6.7 indicates that the estimated effect of mass reduction on risk when a car is struck in the side is the same or lower than when a car is involved in a fatal crash. Therefore the reduction in the number of side impact fatalities expected from full adoption of side airbags will likely increase the estimated effect of mass reduction on risk across all types of crashes.

4. DATA SETS	COMMENTS
Please comment on the validity and applicability of the datasets used to project changes in risk resulting from reduction in vehicle mass. LBNL's casualty analysis used police-reported crash data from 16 states, while the 2011 NHTSA study used national fatality data, combined with a subset of non-culpable vehicles involved in two-vehicle crashes from police-reported crash data from 13 states.	**[Chen and Kockelman]** [1] The acquisition of Polk data for VMT estimates by make & model is valuable, and a contribution to the literature. However, these estimates come from vehicles found in repair shops in non-attainment areas, and so will be biased towards problem-prone vehicles, wealthier households who service their vehicles more regularly, and/or urban (smoggier) areas. Such issues merit careful discussion in the paper, so that readers are well aware of caveats. *[LBNL] The largest source of vehicles in the Polk odometer data are vehicles that report for a regularly-scheduled vehicle emission inspection and maintenance (I/M) program inspection. In urban areas with poor air quality, all eligible vehicles are required to report for inspection every one or two years. While the Polk VMT data may be skewed towards urban areas (which typically include suburban areas, and sometimes rural areas, depending on the state I/M program), they are not biased towards problem-prone vehicles or those owned by wealthier households.* [2] Related to this, Tom Wenzel indicated (by phone) that he did take a look at CA's extensive odometer reads, which go into some semi-rural locations (not too rural), and he indicated that the VMT values by vehicle type (not controlling for HH attributes & such) are very similar (just 5% longer in rural areas) – except for vans (which are used much more extensively in rural areas). This is interesting to me, and is not that different from what we've seen in the past. For example, Kockelman & Zhao's JTS paper from 2000 (pre- print at http://www.ce.utexas.edu/prof/kockelman/public_html/ BTSJournalLDTs.pdf) suggests that, after controlling for various HH attributes & vehicle types, density is/was still very important (tables 1 & 2), but a shift from a density of 1k to 5k persons per sq mile (which is 1.5 vs. 7.8 persons per acre) means an increase of 750 mi/yr/vehicle (which is about 7.5% of annual VMT). Such differences, and their practical significance (or lack thereof) should be discussed in the reports. *[LBNL] The vehicle mile traveled weights NHTSA assigns to each vehicle in the induced exposure dataset are based on the average odometer reading by vehicle make, model, and year from a database of odometer readings provided by Polk. Many of these odometer readings were obtained through state-run emissions inspection/maintenance programs, which tend to operate in urban areas of states with poor air quality. In a previous analysis LBNL (LBNL 2011a) found that vehicles registered in rural counties (<200 population per square mile) are driven between 2% and 8% more miles than vehicles registered in urban counties (>200 population per square mile), depending on vehicle type (full size vans registered in urban counties are driven 3%*

	more miles than those registered in rural counties; however this difference is not statistically significant). Therefore the Polk database of odometer readings is likely skewed towards vehicles driven in urban areas, which tend to be driven fewer miles than comparable vehicles registered in rural areas. *An additional limitation of the VMT weights developed by NHTSA is that they do not account for the reduction in miles driven in response to higher gas prices and the economic recession in 2008, as noted in Section 5.3 of LBNL's Final Phase 1 report. However, despite these limitations, the average VMT weights NHTSA has developed are an improvement over the averages used in the 2003 analysis.* **[Van Auken]** The inducted-exposure data set provided by NHTSA is based on the non-culpable vehicles in two-vehicle crashes. See the comments in Table 1 on the limitations of this data. In addition, there are also many differences in the coding variables and values used by the different states, which tend to make the recoding to a common data set imprecise. [Note: This peer review does not address the Wenzel (2011b) companion report, which examines the risks per police-reported crash. See Ref 4 for comments on the companion report.] **[LBNL]** *As discussed above, NHTSA has developed vehicle miles traveled weighting factors for induced exposure cases based on stopped, rather than non-culpable, vehicles. LBNL summarized the effect this alternative method of exposure has on the estimated effect of mass or footprint reduction on risk in Alternative 15 in Table ES.3, and Section 5.6, in the Final Phase 1 report.*
Please comment on any apparent, unstated, or implicit impact on estimated risks inherent in the two different approaches, and any related caveats or limitations. For example, what are the strengths and weaknesses of the two measures of vehicle exposure, miles of vehicle traveled scaled up from crash data from 13 states, and number of police-reported crashes?	**[Chen and Kockelman]** The use of non-culpable vehicles in two-vehicle crashes as a proxy for vehicles which are "just there" may be distorting the overall distribution of vehicle models. VMT may differ between vehicles that are more prone to run-off-road accidents, at-fault two-car crash vehicles, and non-culpable vehicles. **[LBNL]** *As discussed above, NHTSA has developed vehicle miles traveled weighting factors for induced exposure cases based on stopped, rather than non-culpable, vehicles. LBNL summarized the effect this alternative method of exposure has on the estimated effect of mass or footprint reduction on risk in Alternative 15 in Table ES.3, and Section 5.6, in the Final Phase 1 report.* **[Van Auken]** [1] The number of fatal cases tends to be much less than the number of induced-exposure cases. Therefore the effective numbers of degrees-of-freedom in the statistical estimates tend to be limited by the

| | available number of fatal cases. For example, it would not be possible to estimate the effects of two variables (e.g., just curb weight and footprint) if we had data for only one fatal case even if we had thousands of induced-exposure cases. Therefore it is desirable to use data for the entire US in order to get a large sample of fatal cases for the logistic regressions. This then requires the available induced-exposure data (i.e., from 13 states) to be "scaled up" the US level using the method described in Kahane (2003 and 2011). The result is the best currently available estimate of vehicle exposure.

[2] There may be some concerns about the accuracy of the vehicle miles-traveled data because the difficulty estimating the number of vehicle miles travelled at the make-model-year=state level of detail. [Note: This peer review does not address the Wenzel (2011b) companion report, which examines the risks per police-reported crash. See Ref 4 for comments on the companion report.] |
| ADDITIONAL COMMENTS: | |

5. RECOMMENDATIONS	COMMENTS
Please comment on whether the LBNL assessment adequately addresses the NHTSA 2011 study and identifies the safety impact from mass reduction. Are the analytic methods and data used to assess the NHTSA study, and estimate the relationship between risk, mass, and footprint, appropriate? Is casualty risk per crash a legitimate measure of vehicle safety? What other methods or data could be used to better predict the effect of future vehicle designs on safety?	**[Chen and Kockelman]** While driver fatalities per crash seems a useful measure of vehicle design safety, and examination of fatal crash rates is very valuable (using Polk- based exposure estimates), there are many caveats to work of this type. As noted above: a primary concern remains a neglect of the notion that the type of car is very much a proxy for driver type, and a vehicle's crash avoidance may have very little to do with vehicle type. It has a lot to do with the person behind the wheel. Simply including gender and age variables cannot account for important covariates such as education, risk aversion, driving ability, wealth, etc. In reality, small cars may be less crash prone than Kahane's and Wenzel's results suggest, because they are driven by lower-income, younger, less risk averse people driving in more crash prone settings (e.g., commercial strips rather than pricey residential suburbs). Of course, as noted above, it is very difficult to control for all these variables, and the contractor was asked to rely on the original data. In reality, the best the report authors can do with such data sets is to explain how all the other, relevant attributes may factor in (e.g., quality of driver and typical driving settings), and how they can generate biased estimation (sometimes in either direction). Discussion of relevant literature that looks more deeply at crash outcomes (e.g., Wang or Chen's papers, mentioned above, allowing for heteroscedasticity and individual vehicle attributes, non-driver outcomes, etc.) will also be useful.

[LBNL] As discussed above, additional sensitivity analyses were run using a control variable for household income, based on California vehicle registration data in 2010 (see Section 5.4 in the Final Phase 1 report). And sensitivity analyses were also run using a control variable for initial vehicle purchase price as a proxy for quality of vehicle design, using values in Polk's VIN decoding software (see Section 5.2 in the Final Phase 1 report).

[Van Auken] [1] The basic methodology described by Kahane (2011) seems appropriate; however some results using this method and data are not well understood and need further diagnosis.

[2] The induced-exposure data set provided by NHTSA is based on the non-culpable vehicles in two-vehicle crashes. See the Table 1 comments on the limitations of this data. [Note: This peer review does not address the Wenzel (2011b) companion report, which examines the risks per police-reported crash.]

[LBNL] As discussed above, NHTSA has developed vehicle miles traveled weighting factors for induced exposure cases based on stopped, rather than non-culpable, vehicles. LBNL summarized the effect this alternative method of exposure has on the estimated effect of mass or footprint reduction on risk in Alternative 15 in Table ES.3, and Section 5.6, in the Final Phase 1 report. |

Please comment on the overall adequacy of LBNL's assessment of the 2011 NHTSA report and its independent study of casualty risk for predicting the effect of vehicle mass or footprint reduction on safety. Provide any recommended improvements that might reasonably be adopted by the author to improve the analysis.	**[Chen and Kockelman]** Overall, the study is a comprehensive assessment of the 2011 NHTSA report and introduces interesting additional analyses to examine the relationship of vehicle mass and footprint reduction on safety. However, as stated previously in the comments here, driver preference for specific car types (including size and mass) is related to driver socioeconomic characteristics and driving behavior. As vehicle, driver, and roadway environment characteristics all contribute to fatality risk, the effects of physical vehicle changes such as mass or footprint reduction on safety should not be overstated when the other two types of characteristics are not sufficiently accounted for. **[Farmer]** Overall these are reasonably good studies. The Phase I report does a very good job of assessing the NHTSA report of fatality risk. **[Van Auken]** The Wenzel (2011a) report provides a valuable supplement to the analysis and results in the Kahane (2011) report. [Note: This peer review does not address the Wenzel (2011b) companion report, which examines the risks per police-reported crash.]
ADDITIONAL COMMENTS: **[Van Auken]** See attached tables 6 and 7 below.	

Table 6. Additional General Comments and Recommendations Mike Van Auken	
Section	**COMMENTS AND RECOMMENDATIONS**
3	Use of R^2 is confusing. Suggest using lower case "r" when referring to the sample correlation coefficient (Box, Hunter, Hunter, 1978, P. 61); or upper case R when referring to the regression coefficient of multiple determination (Draper and Smith, 1981, p. 90). *[LBNL] Figures 3.1 and 3.2 and 4.1 through 4.5 report r and not R^2 in the Final Phase 1 report. In addition, a new Table 3.1 with r and VIF, and a new Table 4.1 with r and R^2, have been added.*
All	In most cases the reported results are just estimates, but are not described as such. For example, "The effect of mass reduction on heavier cars and CUVs and minivans are not statistically significant" on p. iii should say "The ***estimated*** effect of mass reduction on heavier cars and CUVs and minivans are not statistically significant." This distinction is important when comparing results based on different models and assumptions because the different models and assumptions do not change the effect itself, but rather the estimate of the effect. For example, the statement "The first sensitivity, in dark purple, includes the weight variables in the regression model but excludes the footprint variable; this model tests the effect of mass reduction while allowing footprint to vary with vehicle mass. This sensitivity increases the risk from a 100-lb mass reduction in cars (from 1.43% to 2.64% for lighter cars, and from 0.48% to 1.94% for heavier cars) and CUVs/minivans (from a 0.47% decrease in risk to a 0.52% increase in risk); however, there is no change in fatality risk in light-duty trucks" on page 15 is misleading. It would be better to state that "The first sensitivity, in dark purple, includes the weight variables in the regression model but excludes the footprint variable; this model tests the ***estimated*** effect of mass reduction while allowing footprint to vary with vehicle mass. ~~This sensitivity~~***Removing the footprint variable from the regression model*** increases the ***estimated*** risk from a 100-lb mass reduction in cars (from 1.43% to 2.64% for lighter cars, and from 0.48% to 1.94% for heavier cars) and CUVs/minivans (from a 0.47% decrease in risk to a 0.52% increase in risk); however, ~~there is no~~***the*** change in ***the estimated*** fatality risk in light-duty trucks ***is very small and not statistically significant***." This also applies to table and figure captions. For example, "Table ES.1. Effect of mass and footprint reduction on fatality risk, under alternative regression model specifications" should say "Table ES.1. ***Estimated effects*** of mass and footprint reduction on fatality risk, under alternative regression model specifications." "Figure 3.3 Effect of reduction in mass or footprint on US fatality risk per VMT, by vehicle type: mass only, footprint only, and both" should say "Figure 3.3 ***Estimated effects*** of reduction in mass or footprint on US fatality risk per VMT, by vehicle type: mass only, footprint only, and both." Overall the word "effect" appears over 200 times in this report with the "estimated" or other qualifier. In some cases this may be

	appropriate and in other cases it is not appropriate. It is recommended that the author review each instance and revise as appropriate.
	[LBNL] The word "estimated" will be used extensively in the final report. In addition the following text was added in the Executive Summary, Section 1, and Section 7:
	Although the purpose of the NHTSA and LBNL reports is to estimate the effect of vehicle mass reduction on societal risk, this is not exactly what the regression models are estimating. Rather, they are estimating the recent historical relationship between mass and risk, after accounting for most measurable differences between vehicles, drivers, and crash times and locations. In essence, the regression models are comparing the risk of a 2600-lb Dodge Neon with that of a 2500-lb Honda Civic, after attempting to account for all other differences between the two vehicles. The models are not estimating the effect of literally removing 100 lbs from the Neon, leaving everything else unchanged.
	In addition, the analyses are based on the relationship of vehicle mass and footprint on risk for recent vehicle designs (model year 2000 to 2007). These relationships may or may not continue into the future as manufacturers utilize new vehicle designs and incorporate new technologies, such as more extensive use of strong lightweight materials and specific safety technologies. Therefore, throughout this report we use the phrase "the estimated effect of mass (or footprint) reduction on risk" as shorthand for "the estimated change in risk as a function of its relationship to mass (or footprint) for vehicle models of recent design."
4	Figures 4.1 through 4.17 do not control for the effect of vehicle size (e.g. footprint), which has been shown to be correlated with vehicle weight (e.g., Figure 3.1), and therefore these figures may be misleading. It is strongly suggested that the horizontal axis label be changed to "Curb weight (lbs) and corresponding changes in size," and/or a note such as the following be added to each figure: "Note these results do not control for the effect of vehicle size on fatality risk. Therefore the horizontal axis represents changes to both vehicle weight and vehicle size."
	[LBNL] The figures, as well as Tables 4.2 and 4.4, show the relationship between risk and vehicle mass, before and after accounting for all modelled differences across vehicle models except mass and footprint, and are labelled as such. A new Table 4.3 has been added to the final report that compares predicted risk after accounting for all modelled differences across vehicle models except mass (but including footprint), and a new Table 4.5 shows the relationship between predicted risk and footprint after accounting for all differences except footprint (but including mass); this increases the estimated risks in some cases, makes some statistically significant, but does not appreciably reduce the range in risk by vehicle model.
4	The statistical significance of the linear trends in Figures 4.1 through 4.17 are not reported. It would be helpful if the confidence

intervals or statistical significance of the linear trends were reported, either in addition to or instead of R^2.

The confidence intervals for the estimated slopes should be added to the results in Tables 4.1 and 4.2.

[LBNL] Tables 4.2 through 4.5 in the Final Phase 1 report summarize the linear trends and statistical significance of Figures 4.1 through 4.17.

Table 7. Additional Specific Comments and Recommendations Mike Van Auken		
Section	**Page**	**COMMENTS AND RECOMMENDATIONS**
Executive Summary, 7	iii, 65	2nd paragraph refers to "our analysis," however the results are the same as the NHTSA analysis. The author should clarify who or what "our analysis" refers to and how it relates to the NHTSA analysis. Perhaps the statement "LBNL was able to reproduce the NHTSA analysis, which finds that…" would be more appropriate. **[LBNL]** *The suggested change has been made.*
Executive Summary	iv	Last bullet – suggest changing the statement that "Logistic regression does not allow a statistic" to "Logistic regression ***methods do not have*** a statistic." **[LBNL]** *The suggested change has been made.*
Executive Summary, 4	iv, v, 22, 66	Suggest changing "variance in risk" to ***variation*** in risk" throughout. **[LBNL]** *The suggested change has been made.*
Executive Summary, 7	viii, 69	The numerical results for the NHTSA preferred model in Tables ES.1 and 7.1 are slightly different than the results reported in the NHTSA report. For example 1.43%/0.48%/0.52%/-0.40%/-0.47% should be ***1.44***%/***0.47***%/0.52%/***-0.39***%/***-0.46***% **[LBNL]** *We are not sure why LBNL's results are slightly different from NHTSA's; perhaps due to rounding differences, or the factor LBNL used to convert log-odds ratios to probabilities, as mentioned in the Final Phase 1 report.*
4	31-32	Figures 4.12 through 4.14 have the results for small and heavy-duty pickups combined, which is inconsistent with the results in Table 4.1 **[LBNL]** *The figures in the Final Phase 1 report present the data for small and heavy-duty pickups separately.*
4	35	The R^2 values in Table 4.1 are different than the values in Figures 4.6, 4.8, 4.9, 4.11. **[LBNL]** *The R^2 values in the bottom four rows of the table are incorrect; these have been corrected in the Final Phase 1 report.*

5.1	37	The subsection title should be "Alternative measures of exposure **and outcome**" because fatal crashes and fatalities are measures of the crash outcome, not exposure.
		[LBNL] The title has been changed to "Alternative measures of risk"
5.2	39	It would be helpful to list the 18 manufacturer dummy variables in a table.
		[LBNL] The indicator variables for the vehicle manufacturers have been added to the final report.
5.2	39	It is unclear why Lexus, Acura, and Infinity are treated as separate manufacturers, but Cadillac and Lincoln are not.
		[LBNL] The results of two regressions have been included in the Final Phase 1 report: one including the five luxury brands in their parent manufacturers, and one accounting for each of the five luxury brands (both in Figure 5.2). Accounting for the five luxury brands substantially increases the estimated detrimental effect of car mass reduction on risk, as shown in Section 5.3 in the Final Phase 1 report.
5.2	39	It is unclear why AM General is considered a Chrysler brand. The AM General Hummer was sold by GM beginning with the 2001 model year.
		[LBNL] AM General has been removed from the Chrysler brand and included in the Other category.
5.3	44	It would be helpful if the figures include error bars or shading to indicate the confidence intervals.
6.4	63	Table 6.3 – Suggest adding a note that the 72,316 total includes fatalities that are counted more than once in crashes involving more than one vehicle type.
		[LBNL] Table 6.3 is mis-labelled; it represents the total number of fatalities between 2004 and 2008 for model year 2000 to 2007 light duty vehicles, not the average annual fatalities. The text in the paragraph above explains that the number of fatalities in mulit-vehicle crashes is divided by the number of case vehicles, to avoid double-counting of fatalities. The title of Table 6.3 has been corrected in the Final Phase 1 report, and VMT for each vehicle type added.

An Analysis of the Relationship between Casualty Risk Per Crash and Vehicle Mass and Footprint for Model Year 2000-2007 Light-Duty Vehicles (LBNL Phase 2 Report)	
1. ASSUMPTIONS	**COMMENTS**
Please comment on the validity of any assumptions embedded in the LBNL assessment analysis and the independent casualty analysis that could affect the projected relationship between vehicle mass/footprint reductions and fatality/casualty risk. Examples might include assumptions regarding whether recent historical relationships between vehicle weight, size, and safety will continue into the future; potential future improvements in vehicle technology and design may result in compensatory safety benefits; and the annual baseline fatality distribution.	**[Chen and Kockelman]** [1] The Phase 2 report serves as a complimentary document to the Phase 1 report by isolating the effect of vehicle mass and footprint on crashworthiness. Whereas the Phase 1 report analyzes fatality risk per estimated VMT, the Phase 2 report analyzes casualty risk per crash. The parallel structure of the two reports makes it easy for the reader to compare the results of the two analyses. [2] The binary logistic model employed in this analysis can only account for two injury outcome categories; here it is used to distinguish crashes resulting in serious injury or death from all other crash outcomes. Thus, the model does not account for the ordinal nature of injury severity and neglects the difference between a serious injury and a death. *[LBNL] The LBNL Phase 2 report analyzes the relationships between vehicle mass, footprint, and casualties, defined as fatalities plus serious or incapacitating injuries. The effects on other injury severities (i.e. minor injuries) were not analyzed because of the differences in reporting of injury severity in the thirteen states.* [3] The report states that "a serious incapacitating injury can be just as traumatic to the victim and her family, and costly from an economic perspective, as a fatality." While serious injuries are very costly to society (and may have similar economic cost implications as deadly crashes), willingness to-pay estimates (which include pain and suffering) price the cost of a fatality at almost 20 times the cost of an incapacitating injury (NSC 2010). Thus, it is difficult to assess the economic cost of the estimates of increases in casualty risk per crash without distinguishing whether that outcome is a serious injury or a death. This limitation of the model should be addressed in the report. *[LBNL] The focus of the analysis is to estimate the relationship between changes in vehicle mass or footprint on fatality or injury risk. Estimating the economic cost of any estimated increase in risk is beyond the scope of the analysis.* [4] The logistic model also assumes error-term homoscedasticity and cannot account for increases and

decreases in the variation of injury outcomes due to vehicle and driver type, for example. Such limitations of the model should be discussed.

[LBNL] As mentioned above, the effects on other injury severities (i.e. minor injuries) were not analyzed because of the differences in reporting of injury severity in the thirteen states.

[Farmer] The report concludes "that much of the detrimental effect of mass or footprint reduction on risk can be attributed to the tendency for mass or footprint reduction to increase crash frequency, rather than to reduce vehicle crashworthiness (risk once a crash has occurred)." However, the interpretation of casualties per crash as inversely proportional to crashworthiness ignores the possibility that injury severity also depends upon the circumstances of the crash. Casualties per crash must be divided into casualties per *severe* crash and *severe* crashes per crash, where a *severe* crash would be one involving more energy, e.g., high-speed or rollover. It could be that weight reduction <u>increases</u> casualties per severe crash (i.e., reduces crashworthiness), but reduces the likelihood that a crash is severe.

[LBNL] The state crash data do not consistently provide a measure of crash severity, such as an accurate estimate of travel speed, to change the measure of exposure from all police-reported crashes to all severe police-reported crashes. Figure 2.10 in the Final Phase 2 report shows that using casualty, as opposed to all, crashes as the measure of exposure does change the estimated detrimental effect of mass reduction on fatality risk per crash

[Van Auken] The basic assumptions, methodology, and data are primarily the same as in the Kahane (2011) report, but have been extended to include serious injuries as well as fatalities, and also address crash involvement (i.e., fatalities and serious injuries per accident, and also accidents per VMT).. These include the following:

1) The probability of a crash fatality or serious injury is proportional to the number of accidents (provided the crash conditions remain the same); and the probability of an accident is proportional to the vehicle miles travelled (VMT).
2) The logarithm of probabilities of fatality or serious injury per accident, and accidents per VMT for a given curb weight, footprint, and control variable values varies as a linear combination of the curb weight, footprint, and control variables within the domain of the data.
3) The logistic regression methods determine a maximum likelihood estimate of model coefficients.
4) It is assumed that the above relationships remain constant in the recent past (i.e., 2000-2007 model year

vehicles in the 2002-2008 calendar years), present, and near future (i.e., 2017-2025 model year vehicles).

The first assumption that crash fatalities and serious injuries are proportional to the number of accidents provided the crash conditions remain the same seems self-evident (e.g., if two fatal crashes had exactly the same conditions, then the expected number of fatalities for the two crashes would be twice the value for just one of the crashes). The assumption that the number of accidents are proportional to VMT rather than the number of vehicle registration years (VRY) is also appropriate because accidents cannot occur if the vehicles are not driven on the road (i.e., VMT = 0). This assumption is qualified however because VMT is more difficult to measure than VRY and therefore may be less accurate. On the other hand the probability of a fatal crash or the number of fatalities in a crash may also depend on the vehicle occupancy.

The second and third assumptions are appropriate provided that it is recognized that it is essentially impossible with currently available knowledge and information to model all of the factors that could affect the probability of fatality in a crash, and that the objective of the analysis is to identify overall trends versus vehicle weight and footprint. In general the probability of fatality depends on other many other factors which have not been modeled (e.g., driver behavior factors, vehicle design factors, roadway design factors, EMT factors), and these unmodeled factors are assumed to be uncorrelated with vehicle weight and footprint, and/or are represented by the other control variables. The latter assumption might or might not be valid.

The fourth assumption is perhaps the weakest because it assumes that future vehicles will have the same design characteristics as past vehicles, and that the characteristics of the vehicle population (e.g., collision partner weight, size, type) will also remain the same.

| Please comment on any apparent unstated or implicit assumptions and related caveats or limitations. | [Chen and Kockelman] The role of driver behavior is briefly addressed in the report but not emphasized sufficiently. Casualty risk is a combination of driver, vehicle, and roadway characteristics. Whereas vehicle characteristics significantly influence crashworthiness, driver behavioral differences play a significant, if not primary, role in determining crash frequency. Socioeconomic data such driver household income, size, and education influence driver attitudes and driving environments. For example, Chen et al. (2010) found that crash risk increases for those living in socioeconomically disadvantaged areas (including households more likely to drive less expensive and older vehicles). Though such data is not typically available in state and national crash databases, the importance of these driver and environmental characteristics on crash rates (per mile driven) and casualty risk should be stressed in both reports. It is clearly very difficult to control for, but is a major caveat to the NHTSA (& now LBNL) results. We expect that crash severity could be probably be lower for many of the small cars and pickups if they were driven by those who tend to drive more expensive vehicles, |

under the same settings (e.g., daytime, urban freeway). Thus, statements like "a 100-lb reduction in the mass of lighter cars leads to a 1.84% increase in crash frequency" should be accompanied by an explanation of the possibility of the mass variable accounting/proxying for effects of lower income households owning smaller vehicles.

[LBNL] We agree that it would be preferable to include additional variables that account for driver behavior rather than just driver age and gender. However, variables such as driver age, gender, or other socio-economic characteristics probably have a less direct effect on risk than the actual driving skill or behavior of individual drivers. The FARS data includes information on whether alcohol or drug use, or speeding or reckless driving, was a factor in the current crash, as well as the driver's record for the last three years. NHTSA 2003 found that including a "bad driver" rating variable comprised of 8 of these variables did not have an appreciable effect on the relationship between vehicle mass and risk. Although these variables are included in the FARS fatal crash records, they are not included in the state crash databases, so a similar variable cannot be used in the regression models. However, Alternative 12 in Tables ES.2 and 7.1, and the third column in Figure 5.14, in the Final Phase 1 report indicates that excluding bad drivers (after also excluding reported driver alcohol or drug use) results in additional estimated increases in fatality risk as vehicle mass decreases, and estimated decreases in risk as vehicle footprint decreases. LBNL believes that driver behavior in the current crash and their recent driving record, and not general socio-economic information, is the best available measure to account for the effect driver skill and behavior has on risk in individual vehicles.

Nonetheless, LBNL explored the effect of adding a control variable for one additional driver socio-economic characteristic, household income. Unfortunately, crash data do not include household income of vehicle drivers. FARS includes the driver's zip code, from his or her license, which could be merged with US Census data by zip code to obtain the average median income of households in a particular zip code. However, very few states include driver zip code in their databases of police-reported crashes.

Kweon and Kockelman (2003) merge police-reported crash data from GES with household demographic data, including household income, from the National Household Travel Survey (NHTS), to examine crash rates per annual VMT by driver age, gender, and vehicle type. Each dataset includes population weights to derive national estimates of crashes and households, which the authors use to estimate national crash rates for driver age, gender and vehicle type. However, there are several limitations with the NHTS data. First, the survey includes only 25,000 national households, supplemented by an additional 124,000 households from "add-on" state and municipal jurisdictions. The national weights provided in the NHTS most likely do not represent the national distribution of vehicle registrations by make and model. Second, 30% of the households in the 2009

NHTS reported income "over $100,000"; using the income bins provided in the NHTS to estimate household income will likely understate the actual income of the households participating in the survey.

Chen el al (2010) makes use of an Australian survey of 20,000 young adult drivers that asked questions regarding driving behavior. The injury severity of 127 of these participants who were admitted to hospitals after a vehicle crash was analyzed as a function of socio-economic status. The socio-economic status of the drivers was based on average or median values from the zip code where they (or their parents, since most were at University) lived.

Using 2010 California vehicle registration data, LBNL calculated the median household income by registered owner's zip code, using 2000 census data, then took the average household income by vehicle make and model. As discussed in the Final Phase 1 report, predicted US fatality risk is correlated with this measure of average household income by vehicle model, with risk decreasing as income increases; the correlation is highest for cars and CUVs, with an R^2 of 0.60 (or a correlation coefficient r of 0.77) or higher, as shown in new Figure 4.20 and the first column of new Table 4.7. However, while income tends to increase as mass increases, the correlation is not strong (R^2 is highest for 4-door cars, at 0.16, and CUVs, at 0.19, as shown in new Figure 4.21 and the middle columns of new Table 4.7). As with predicted fatality risk, there is a wide range in the average income of households owning vehicle models with similar mass. This analysis is summarized at the end of Section 4 of the Final Phase 1 report.

Using average household income by vehicle model, rather than driver alcohol or drug use, or reckless driving, as an alternative indicator of driver behavior substantially reduces the estimated detrimental effect of mass reduction on risk in cars and CUVs/minivans, and has little effect on the relationship in light trucks. This analysis is summarized in Section 5.5 of the Final Phase 2 report.

[Van Auken] [1] The induced-exposure data set provided by NHTSA is based on the "non-culpable" vehicle in two-vehicle crashes. It is assumed that the dataset is a reprehensive sample of the driver and environmental exposure factors for vehicle use. However, since these cases include moving vehicles, some vehicle-driver-environmental conditions may be under or over represented in this data depending on how they affect the ability of a non-culpable vehicle to avoid a crash. Results in Ref (17) indicated that the estimated effect of weight and size reduction are sensitive to whether the induced-exposure data are based on the Kahane (2003) non-culpable vehicle definition of the Kahane (1997) stopped vehicle definition. Unfortunately it is not

	currently possible to test this sensitivity with the NHTSA-provided induced-exposure data.
	[LBNL] NHTSA has developed vehicle miles traveled weighting factors for induced exposure cases based on stopped, rather than non-culpable, vehicles. LBNL investigated what effect this alternative method of exposure has on the estimated effect of mass or footprint reduction on risk, in Alternative 15 in Table ES.3 and Section 5.6 in the Final Phase 1 report.
	[2] It is also assumed that the accident data from the 13 or 16 states are representative of all US states. Figure 2.1 in Wenzel (2011b) provides a useful comparison of the distribution of fatalities in the US and 13 states by the nine different crash types.

ADDITIONAL COMMENTS:

[Chen and Kockelman] National Safety Council (2010) Estimating the Costs of Unintentional Injuries. Available online at http://www.nsc.org/news_resources/injury_and_death_statistics/Pages/ EstimatingtheCostsofUnintentionalInjuries

Chen, H.Y., Ivers, R.Q., Mariniuk, A.L.C., Boufous, S., Senserrick, T., Woodward, M., Stevenson, M. and Norton R. Socioeconomic status and risk of car crash injury, independent of place of residence and driving exposure: Results from the DRIVE study. *Journal of Epidemiology and Community Health* 64(10), 2010, pp. 998-1003.

[Farmer] NHTSA's fatality analysis covered calendar years 2001-08, but the casualty analysis excludes 2008. Such exclusion is understandable given that 2008 data were at the time unavailable for a majority of the states (I think they are available now). However, 2008 was an unusual year and may have affected the size and weight effect estimates. The footnote on p. 6 of the Phase II report states that an analysis including the available 2008 data will be summarized in Appendix A. I don't see Appendix A. Is an analysis planned including 2008 data?

[LBNL] The footnote is incorrect, there was no Appendix A in the final draft report. For the Final Phase 2 report LBNL included the 12 states that have provided crash data for 2008 (all but Wyoming).

2. CONTROL AND DEPENDENT VARIABLES USED IN THE REGRESSION MODELS	COMMENTS
Please comment on the adequacy of control and dependent variables used in the assessment analysis and independent casualty analysis, and recommend any alternative control or dependent variables that are available for possible inclusion in the analysis. For example, what are the relative merits of the main dependent variables used, fatality risk per estimated VMT, and casualty risk per police-reported crash?	**[Chen and Kockelman]** [1] As alluded to above, a primary concern is that the NHTSA analysis (& thus the LBNL analyses) largely neglect the idea that vehicle type (make & model) is very much a proxy for driver type, and a vehicle's crash frequency may have very little to do with physical vehicle characteristics. It has a lot to do with the person behind the wheel, and gender and age simply aren't enough to control for such distinctions. Education, risk aversion, ability, wealth, etc., are important covariates. But existing data sets are quite limiting (though the MVOSS & FAR with 3-year driver violation history do offer some valuable insights, not discussed in these reports). In reality, small cars may be less crash prone than Kahane's & Wenzel's results suggest, because they are driven by lower-income, younger, less risk averse people driving in more crash prone settings (e.g., commercial strips rather than pricey residential suburbs). Such key caveats need thoughtful discussion. Four relevant papers on the topics of crash frequency and vehicle size-and-weight implications (by Knipling, Kweon & Kockelman, Wang and Kockelman, and Chen & Kockelman) have been sent to Tom Wenzel. These all include useful literature reviews for further connections to useful findings for citation in the reports, as time allows the contractor. *[LBNL] As discussed above, LBNL believes that driver behavior in the current crash and their recent driving record, and not general socio-economic information, is the best available measure to account for the effect driver skill and behavior have on risk in individual vehicles. As described in Section 5.4 of the Final Phase 1 report, LBNL used these measures of driver behavior in one of its sensitivity analyses. Unfortunately, the states do not consistently report this information in their crash databases. LBNL examined the effect of including a measure of driver household income on risk in Section 5.5 in the Final Phase 2 report.* [2] Independent variables such as vehicle mass and footprint may be accounting for effects of driver socioeconomic factors as discussed in the Assumptions section. Furthermore, vehicle option variables such as AWD and side curtain airbags may be reflecting the effects of driver environment (e.g., those living in areas with icy winters opting for AWD) and attitude (e.g., more risk-averse drivers opting for side curtain airbags) rather than the vehicle technology themselves. While extremely heavy and extremely large vehicles may have significantly different handling and braking characteristics which influence crash frequency and casualty risk, it is unlikely that given the same driver in the same environment, a small change in vehicle mass or footprint would influence the driver's crash proneness.

	[LBNL] We agree that it would be preferable to include other variables to control for individual driver behavior and skill; unfortunately such data are not available.
	[Farmer] One needs to restrict control variables to those that are available and reliable. A problem when combining state databases is that the states often are not consistent as to the variables coded and the definitions of those variables. This severely limits the list of possible control variables.
	[LBNL] The control variables used from the state crash databases are either derived from the VIN, or based on the driver age and gender, or the time or location of the crash. LBNL believes that these variables are consistently coded among the thirteen states.
	[Van Auken] [1] Reducing the total number of fatalities and serious injuries is desirable from a societal viewpoint. This includes both subject vehicle occupant and collision partner (e.g., other vehicle occupant, pedestrian) fatalities and serious injuries.
	[2] VMT is a good measure of accident exposure provided that it can be accurately determined.
	[3] The number of fatalities and serious injuries per accident is a measure of vehicle crashworthiness (i.e., effect of a crash on the subject vehicle occupants) and crash compatibility (i.e., effect of a crash on the other vehicle occupants or vulnerable road users). Subject vehicle occupant fatalities and serious injuries per accident are a measure of the subject vehicle crashworthiness. Collision partner fatalities and serious injuries per accident are a measure of vehicle crash compatibility.
	[4] The number of accidents per VMT is a measure of the crash avoidance capabilities of a given vehicle.
What additional control variables, such as vehicle make or model, might be included in the regression models?	**[Chen and Kockelman]** [1] Vehicle height, a variable which may be more valuable than vehicle type for similarly structured vehicles such as sedans, wagons, CUVs, and minivans, would be a valuable control variable. In addition to a wider track, a lower center of gravity also increases vehicle stability, thereby reducing the risk of rollover. Relevant literature & findings exist, and should be cited.
	[LBNL] The propensity for a vehicle to roll over increases as its center of gravity increases (and static stability factor decreases), while risk in frontal crashes may decline as bumper height increases (and the degree of overlap with the bumper of a crash partner increases). However, overall vehicle height is only a crude proxy for these two other vehicle dimensions that are thought to correlate with safety, and therefore was not included in

	the regression models.
	[2] Other variables which have been found in past studies to influence fatality risk such as seat belt use, roadway geometry and division type are not included in this study.
	[LBNL] Seat belt use was not included as a control variable because it is notoriously under-reported for non-fatal injuries. Roadway geometry and roadway division were not included because they are not consistently reported or coded in the state databases of police-reported crashes.
	[3] To account for driver characteristics that contribute to casualty risk, socioeconomic variables such as household income, education, household size, etc. would be valuable additions. Unfortunately, both state and national crash databases typically do not include such information (outside of MVOSS). Such issues should be flagged for readers.
	[LBNL] As discussed above, an additional sensitivity analysis was run using a control variable for household income, based on California vehicle registration data in 2010; see Section 5.5 in the Final Phase 2 report.
	[Farmer] I think that already there are too many control variables in the regression models. Instead I would consider defining different classifications of crash types. Table 2.2 of the Phase II report shows that the distribution of crash types for casualty crashes is very different from that for fatal crashes.
	[LBNL] It is possible that inclusion of too many control variables leads to non-intuitive results in certain cases. This problem might be resolved by removing control variables (other than vehicle weight or footprint) that are not statistically significant. NHTSA and LBNL ran a sensitivity where non-significant control variables were removed; see Section 5.8 in the Final Phase 2 report.
Please comment on any caveats or limitations that these dependent variable or control variables entail with respect to use of the results as the basis for estimating the safety effect of mass reduction.	**[Chen and Kockelman]** Please see above comment (in Assumptions section) regarding driver behavior and environment. *[LBNL] As discussed above, an additional sensitivity analysis was run using a control variable for household income, based on California vehicle registration data in 2010; see Section 5.5 in the Final Phase 2 report.*

	[Farmer] Model overspecification could be the reason for results that are non-intuitive, especially in the Phase II analyses of police-reported crashes. Control variables may be correlated with each other or with the size and weight variables. For example, Figure 2.9 of the Phase I report implies that torso side airbags increase fatality risk in CUVs. *[LBNL] It is possible that inclusion of too many control variables leads to non-intuitive results in certain cases. This problem might be resolved by removing control variables (other than vehicle weight or footprint) that are not statistically significant. The example cited suggests that torso side airbags increase fatality risk in CUVs/minivans by an estimated 0.7%, is a small effect, and not statistically-significant. LBNL ran a sensitivity where non-significant control variables were removed; see Section 5.8 in the Final Phase 2 report.*

ADDITIONAL COMMENTS:

[Chen and Kockelman] Table 2.1 has many indicator variables labeled as "C" for continuous variable (such as ABS, ESC, AWD, DRVMALE, etc). These C's should be removed.

[LBNL] These vehicle variables are continuous because, for some models, the VIN does not indicate whether a particular vehicle is equipped with that option or not. In these cases the fraction of that model that is equipped with the particular feature is used. On the other hand, the DRVMALE variable is always a discrete variable, and is be coded as such in the table in the Final Phase 2 report.

[Farmer] The sensitivity results of Chapter 5 (Phase II) point out the extreme differences in results when changing the control variables. For example, including vehicle make changes the effect of a 100-lb reduction in heavier cars from -0.91% to +0.55% (see Fig 5.3).

[Van Auken] The underlying reasons for some of the estimated effects are unknown at this time, but presumably involve driver, vehicle, environment or accident factors than have not been controlled for in the Kahane (2011) and Wenzel (2011b) analyses. See, for example, Refs 17 and 23.

[LBNL] The report notes that other vehicle, driver, or crash characteristics may account for the residual risk not explained by the control variables included in the regression models. As discussed above, additional sensitivity analyses were run using an alternative control variable for vehicle characteristics, initial vehicle purchase price (using values in Polk's VIN decoding software), and a control variable for driver behavior, household income (based on California vehicle registration data in 2010). The results of these sensitivities are shown in Sections 5.3 and 5.5 in the Final Phase 2 report.

3. METHODOLOGY AND STATISTICS	COMMENTS
Please comment on the validity and applicability of the methodology LBNL used in assessing the NHTSA 2011 study and its analysis of the relationship between mass, footprint, and risks per police-reported crash.	**[Chen and Kockelman]** The Phase 2 report enhances the findings in the Phase 1 report by isolating the effect of vehicle mass and footprint on crashworthiness. Like the Phase 1 report, this analysis goes into a fair amount of detail and seeks to introduce additional analyses to better examine the relationship between mass, footprint, and casualty risks. However, due to a lack of control for very specific vehicle differences (which vary by make & sub-model), the exclusion of driver characteristics and crash setting details (which cannot always be controlled for, but are often correlated with vehicle type), the effects of downweighting vehicles and/or shifting vehicle styles and sizes may be overestimated. Simply changing the vehicle mass or footprint on a risky driver in a high-risk setting is unlikely to influence crash outcomes significantly.

[LBNL] As discussed above, additional sensitivity analyses were run using an alternative control variable for vehicle characteristics, initial vehicle purchase price (using values in Polk's VIN decoding software), and an alternative control variable for driver behavior, household income (based on California vehicle registration data in 2010). The results of these sensitivities are shown in Sections 5.3 and 5.5 in the Final Phase 2 report.

[Farmer] Figure 2.11 of the Phase II report implies that NHTSA's fitting of a separate regression model for each of the 9 crash types was unnecessary, at least for the analysis of casualty risk per crash. I don't recall seeing a similar analysis for fatality risk per VMT. Is it possible to get essentially the same results as the NHTSA study using a single regression model?

[LBNL] Figure 2.1 of the Final Phase 1 report has the same analysis (although the second and third columns are reversed). NHTSA ran a separate regression model for each of nine crash types in order to reweight the estimated coefficients to reflect an estimated reduction in the number of fatal rollovers and crashes with a stationary object from widespread adoption of ESC in the vehicle fleet. Running a single regression model across all crash types obtains essentially the same result as running nine regression models.

[Van Auken] [1] The logistic regression methods seem to be appropriate. The confidence intervals are based on the logistic regression Wald Chi-Square statistic, which as Kahane (2003, 2011) has demonstrated does not include all sources of variation. However, these confidence intervals are useful because they do provide some indication of the uncertainty in the results.

[2] The two-stage results for the 13-state fatalities per crash and 13-state crashes per VMT, and the one-stage |

	result for 13-state fatalities per VMT reported in Tables ES.1 and 6.1 and Figure 2.7 were computed using independent logistic regressions. The differences between the two-stage results and the one-stage results for fatalities per crash could have been eliminated by using the "simultaneous three-way" logistic regression method described in DRI (2003). This method imposes the constraint that the combined two-stage estimated and the one-stage estimated are equal.
Please review other statistical methods LBNL has used in the analysis, in addition to the logistic regression methodology. Examples include the alternative approaches used by LBNL to assess NHTSA interval estimation results, and LBNL's linear regression analysis of actual, predicted, and residual risk by vehicle model.	**[Chen and Kockelman]** In the alternative measures of exposure, the author examines the effect of vehicle manufacturer on fatality risk and treats the luxury models produced by Toyota, Honda, and Nissan as separate manufacturers. However, domestic luxury brands (such as Cadillac & Lincoln) are categorized with their nameplate manufacturers (GM and Ford), which appears inconsistent. *[LBNL] The analysis has been revised to include the luxury models Lexus, Acura, and Infiniti with their nameplate manufacturers Toyota, Honda, and Nissan. An additional sensitivity was conducted using separate indicator variables for the five luxury brands Lexus, Acura, Infiniti, Cadillac, and Lincoln. The results of these analyses are summarized in Section 5.3 in the Final Phase 2 report.* *LBNL also examined the effect of replacing the vehicle make indicator variables with a single continuous variable for vehicle initial purchase price, taken from the Polk VIN decoder. Initial purchase price can be seen as a proxy for the general quality of vehicle design; we expect that more expensive vehicles will have lower risk than less expensive vehicles. This analysis is summarized in Section 5.5 in the Final Phase 2 report.* **[Van Auken]** [1] The correlations in Section 3 appear to be assessed using the Coefficient of Multiple Determination (R^2) based on a linear fit to the data (e.g., the correlation between footprint versus curb weight in Figure 3.1 on p. 32). The linear regression model attributes the differences between the dependent variable (vertical axis) and the linear fit to the independent variable (horizontal axis) to random effects. If there is no preference as to the choice of independent and dependent variables (e.g., footprint versus curb weight, or curb weight versus footprint), then the linear trend and R^2 result would be different if the two variables were interchanged, and having two different yet equally valid results would be undesirable. *[LBNL] The R^2 values are the same if the two variables are interchanged.*

	[2] If the variation in the data can be attributed to both variables (e.g., footprint and curb weight), then it would be better to report the square of the sample correlation coefficient r^2, where r is computed according to Eqn (1). The trend lines in these correlation figures should not be computed using a linear regression. Instead, the trend line should pass through the sample means (i.e., (x, y)), and have a slope equal to the ratio of the sample standard deviations in the data (i.e., s_y / s_x). Therefore, the reported correlation results do not depend on the ordering of the data variables. *[LBNL] Again, the R^2 values are the same if the two variables are interchanged. The R^2 values in Figures 3.1 and 3.2, and Figures 4.1 through 4.6, have been replaced with the correlation coefficient r, and a new Table 3.1 has been added to Section 3 that lists r and VIF for curb weight and footprint by vehicle type, in the Final Phase 2 report.* Note this comments does not apply to linear trends indicated in Section 4, for which the Coefficient of Multiple Determination (R^2) seems appropriate. [3] The Coefficient of Multiple Determination (R^2) is frequently used in the Wenzel (2011b) report as an indicator of the statistical importance of a linear trend (e.g., R^2 values in Tables 4.1 and 4.2 were compared to 0.3). It would be better to report the standard error, confidence interval, and/or probability value as measures of the statistical significance of a linear trend. *[LBNL] The significance of the estimated relationships have been added to the tables in Section 4 of the Final Phase 2 report. However, the point of the section is to indicate that while the relationship may be statistically significant, on average (as indicated by the confidence level around the estimated effect), there remains a large range in risk for individual vehicle models with similar mass (as indicated by the relatively low R^2 values).*
Please comment on caveats or limitations of using non-significant regression estimates to project the safety impact of mass reduction.	**[Chen and Kockelman]** First, the t-statistics are not provided in the report which makes it difficult for the reader to assess statistical significance of specific regression estimates (except where noted by the author). Second, inclusion of a statistically insignificant variable can influence the estimates of coefficients associated with related variables. Nevertheless, in general, it is best to keep insignificant estimates if one has a strong defense for their role, since removing such variables (& thus their parameters) will shift the burden of response to a correlated covariate's parameter, thus biasing the latter. We generally keep key covariates in a model up to a pvalue of 0.20 or 0.25 or so, especially in relatively small data sets (e.g., n < 1,000). Covariates for which we have no strong basis can be removed for pvalues > 0.10.

[LBNL] The 95% confidence intervals are included in every figure in the document, to indicate whether the estimated variable is statistically different from zero. LBNL ran a sensitivity where non-significant control variables are removed; see Section 5.8 of the Final Phase 2 report.

[Farmer] Making projections from non-significant regression estimates is proper so long as the resulting confidence intervals are constructed conservatively (to account for the accumulated imprecision). In that sense, I prefer NHTSA's jackknife approach to the standard errors produced by SAS (see p. 13 of Phase II).

[LBNL] LBNL agrees that NHTSA's jack-knife method is a preferable method of estimating confidence intervals; however, the jack-knife method obtains even larger estimates of uncertainty than using the standard errors output from the regression models.

[Van Auken] [1] Regression estimates are random numbers which have an unknown expected value and variance, and known sample value and standard error. If the sample value can be explained by a zero expected value and known standard error then the result is considered not statistically significantly different than zero and therefore the result is not considered to be statistically significant. However, If we can combine this estimate with other estimates then the unknown expected values and variances can also be combined using the same transformation, and the statistical significance of the combined result can be tested. Therefore, depending on the sample values and inter-correlation, the combined result may be statistically significant even if the individual estimates are not statistically significant.

For example, the results from each of the nine different crash types can be combined into an overall estimate and the standard error calculated assuming that the results for each crash type are independent of each other. Then the statistical significance of the combined effect can be determined.

[2] However, and Kahane (2011) points out there are two sources of uncertainty in the regression results. The first is the FARS based sampling error which is uncorrelated across crash types because they are based on different fatal cases (Kahane 2011, p. 77). The second is the state based induced-exposure sampling error which is correlated across crash types because they are based on the same induced-exposure cases. Therefore a confidence interval estimated using the "jackknife" method described by Kahane (2011) and accounting for correlation of these two error sources would be more accurate than a simple estimate based on the Wald Chi-Square statistic and assumed independence.

[LBNL] LBNL agrees that NHTSA's jack-knife method is a preferable method of estimating confidence intervals;

	however, the jack-knife method obtains even larger estimates of uncertainty than using the standard errors output from the regression models.
How might the LBNL methodology be strengthened to better represent future vehicle designs and reduce multi-collinearity between mass and footprint in the regression analysis?	**[Chen and Kockelman]** Including more vehicle-specific characteristics (such as vehicle height and engine size) reduces the analysis' dependence on vehicle type, since vehicle shapes and structures will continue to evolve. There is also correlation with context (e.g., pickups are driven in more rural locations, with greater hazards [like less lighting, higher speed, & few medians]). Disaggregate data are almost always best, to avoid ecological fallacies & such. *[LBNL] As discussed above, overall vehicle height is only a crude proxy for two vehicle dimensions, center of gravity and bumper height, that are thought to correlate with safety in rollover and two-vehicle crashes, respectively, and therefore was not included in the regression models. Other vehicle attributes that might affect risk are engine power-to-weight ratio, braking distance, and handling capabilities; LBNL may examine the effect of these accounting for these vehicle attributes in future analyses.* *LBNL's analysis does account for whether a vehicle is driven in a rural area, defined as a county in which the population density is less than 250 residents per square mile of land area.* **[Van Auken]** [1] The effects of multi-collinearity can be mitigated by 1) obtaining more data, 2) pooling data from different crash type or vehicle types, or 3) reducing the number of regression variables. The first option would require more states (for serious injuries and police-reported accidents), calendar years and/or model years, which would involve added newer data as it becomes available (or using older data). The second option might be to recombine the CUVs and minivans with truck based vans and adding a control variable to compensate for the differences in the vehicles types. The third option might involve removing statistically insignificant control variables or removing control variables that would not be expected to have an effect on the probability of crash or crash outcome (e.g., the side airbag variable is not included in pedestrian crashes because it is not expected to affect pedestrian fatality risk). The number of driver age control variables might be reduced from eight to three (as in the Kahane (1997) and DRI (2002-2005) studies). Finally, a linear curb weight model instead of a two-piece linear model may help to better elucidate the general trend. [2] The Variance Inflation Factor (VIF) has been suggested as a measure of multi-collinearity in the Kahane

	(2010 and 2011) reports, however this diagnostic metric does not account for differences in database size (i.e., Options 1 and 2 above). The Wenzel (2011b) report does not discuss the Variance Inflation Factor or report any VIF results.
	[LBNL] A table of VIF results, Table 3.1, has been added to the Final Phase 2 report.
ADDITIONAL COMMENTS:	

4. DATA SETS	COMMENTS
Please comment on the validity and applicability of the datasets used to project changes in risk resulting from reduction in vehicle mass. LBNL's casualty analysis used police-reported crash data from 16 states, while the 2011 NHTSA study used national fatality data, combined with a subset of non-culpable vehicles involved in two-vehicle crashes from police-reported crash data from 13 states.	**[Chen and Kockelman]** [1] The Phase 2 report uses an unusually extensive data set of police-reported crash data from 13 states which the author compares in detail to national data sets to illustrate similarities and differences. The author is very thorough in addressing the difference in definitions of "serious" and "incapacitating" injuries across different states and the effects of such inconsistency on the regression results. [2] Since casualty risk in the report accounts for serious injuries but not minor injuries, the author should note that police-reported injury levels may also be poor indicators of the actual or Modified Abbreviated Injury Scale (MAIS) level, following medical evaluation. Farmer (2003) found that 41% of injuries reported by U.S. police as incapacitating received MAIS ratings of "minor injury" by health care professionals using NASS Crashworthiness Data System (CDS). Thus, the results of the estimated casualty risk increases and decreases rely heavily on the assumption that police errors in reporting actual MAIS ratings are consistent across states. *[LBNL] Inconsistencies in injury reporting across states are another reason to include a control variable for each state. The following text has been added to Section 2.3 of the Final Phase 2 report:* *"Another type of bias is inaccurate reporting of injury outcomes by police officers at the scene of a crash. Using detailed NASS CDS records, in which a crash investigator tracks hospital records of victims in a small sample of police-reported crashes, Farmer (2003) found that 41% of injuries that police responders coded as serious or incapacitating received Modified Abbreviated Injury Scale (MAIS) ratings of "minor injury" by health care professionals. The possibility that these injury reporting errors are not consistent across states is another reason to include a control variable for the state in which the crash occurred."* **[Farmer]** A major limitation of the Phase II analysis is a bias that may be due to the patterns of missing data. In particular, the vehicle identification number (VIN) is missing or mistyped for many crash records. High-severity crashes (especially fatal) are more likely to have detailed police investigation, so VINs (and other variables) in these crashes may be more complete. State crash files are therefore much less reliable than FARS. *[LBNL] LBNL assumes that missing or erroneous VINs are equally distributed among vehicle year, make and model. In its primary analysis LBNL used only vehicles whose model year reported in the crash database matched the model year from the decoded VIN (as well as all vehicles in certain years in Washington where model year was not recorded in the state database). In addition, in Section 5.7 of the Final Phase 2 report LBNL*

	corrected obviously erroneous VINs, by translating VIN position 10, shifting the VIN one position to the right starting at position 9 (both based on the reported model year), and translating VIN position 8, the vehicle engine code (based on a large database of known valid VINS, as described in Wenzel 2011a). Combined these changes increased the number of available records by only 1%.
	[Van Auken] [1] The induced-exposure data set provided by NHTSA is based on the non-culpable vehicles in two-vehicle crashes. See the comments in Table 1 on the limitations of this data.
	[2] The use of property damage accident data and cases with serious injury from the 13 states seems appropriate (with the noted qualification that the different states may have different accident reporting thresholds and injury reporting criteria). The concerns about the use of data for the 3 additional states (Georgia, Illinois, and New Mexico) have also been noted.
	[LBNL] *Because of the differences in crash or injury reporting thresholds and injury severity reporting across states, LBNL used a control variable for each state in which the crash occurred.*
	[3] In addition, there are also many differences in the coding variables and values used by the different states, which tend to make the recoding to a common data set (either induced-exposure, police-reported accident, or severe injury) imprecise.
	[LBNL] *The coding of the variable for type of crash may vary across states. However, the remaining control variables used from the state crash databases are either derived from the VIN, or based on the driver age and gender, or the time or location of the crash. LBNL believes that these variables are consistently coded among the thirteen states.*

Please comment on any apparent, unstated, or implicit impact on estimated risks inherent in the two different approaches, and any related caveats or limitations. For example, what are the strengths and weaknesses of the two measures of vehicle exposure, miles of vehicle traveled scaled up from crash data from 13 states, and number of police-reported crashes?	**[Chen and Kockelman]** The Phase 1 analysis used non-culpable vehicles in two-vehicle crashes as a proxy for induced exposure crashes. In contrast, Phase 2 analysis uses data from vehicles involved in one-car crashes and the responsible vehicle in two-car crashes. The exclusion of the not-at-fault vehicle in two-car crashes may be distorting the distribution of crash frequency and casualty risk across different vehicle makes and models if crash-prone drivers are more likely to drive certain types of vehicles.
	[LBNL] The Phase 1 analysis used a subset of non-culpable vehicles in two-vehicle crashes to assign vehicle, driver, and crash characteristics to the total number of US vehicle registrations and miles driven. The measure of exposure in the Phase 1 analysis is national miles driven. The Phase 2 analysis used all non-fatality (or non-casualty) crashes in the 13 states as the measure of exposure for estimating fatalities (or casualties) per crash; the same VMT from the Phase 1 study was used to estimate crashes per VMT, and fatalities (or casualties) per VMT.
	[Farmer] The VMT weights provided by NHTSA were scaled to represent the entire US. Comments on pp. 9 and 18 of the Phase II report seem to acknowledge this deficiency, promising to adjust these to the 13 states in the future. Was any adjustment made, such as multiplying the weights by the proportion of annual US VMT accounted for by each of these states? The accuracy of the VMT weights is critical is we are to believe the somewhat surprising results concerning crashes per VMT.
	[LBNL] LBNL requested that NHTSA recalculate the VMT weights to represent the total number of vehicles registered, and the estimated total number of miles driven, in the thirteen states, rather than for the entire US. NHTSA declined to provide LBNL VMT weights for the 13 states.
	In their previous analyses DRI similarly found that crash frequency (per VMT) increases as mass decreases. And in Section 1.6 of its 2011 report NHTSA notes that small and light vehicles historically have had higher crash and insurance claim frequency per vehicle mile traveled, despite their theoretical advantage in terms of handling, braking, and accelerating.
	[Van Auken] [1] The number of fatal or serious injury cases tends to be much less than the number of induced-exposure cases (and the number of police-reported accidents). Therefore the effective numbers of degrees-of-freedom in the statistical estimates tend to be limited by the available number of fatal or serious injury cases. For example, it would not be possible to estimate the effects of two variables (e.g., just curb weight and footprint) if we had data for only one fatal of serious injury case even if we had thousands of induced-exposure cases. Therefore it is desirable to use data for the entire US in order to get a large sample of

	fatal cases for the logistic regressions. This then requires the available induced-exposure data (i.e., from 13 states) to be "scaled up" the US level using the method described in Kahane (2003 and 2011). The result is the best currently available estimate of vehicle exposure. [2] There may be some concerns about the accuracy of the vehicle miles-travelled data because the difficulty estimating the number of vehicle miles travelled at the make-model-year-state level of detail. *[LBNL] RL Polk provided NHTSA a database of average odometer reading by vehicle year, make and model based on hundreds of thousands of odometer readings. Because many of these records came from state emission inspection and maintenance programs, it is likely that they are skewed towards vehicles driven in urban areas, which tend to be driven fewer miles than comparable vehicles registered in rural areas.* *An additional limitation of the VMT weights developed by NHTSA is that they do not account for the reduction in miles driven in response to higher gas prices and the economic recession in 2008, as noted in Section 5.3 of LBNL's Phase 1 report. Despite these limitations, the average VMT weights NHTSA has developed are an improvement over the averages used in the 2003 analysis.*

ADDITIONAL COMMENTS:

[Chen and Kockelman] Farmer, C.M. Reliability of police-reported information for determining crash and injury severity. *Traffic Injury Prevention* 4(1), 2003, pp.38-44.

[Farmer] Statements above Figure 2.7 in the Phase II report imply that the effects of weight reduction on crashes per VMT and fatalities per crash should add up to the effect on fatalities per VMT. This is not the case. For example, a 1.43% increase in crashes per VMT and a 0.76% decrease in fatalities per crash would imply a 2.16% <u>decrease</u> in fatalities per VMT (i.e., 1 - 0.9924/1.0143). The fact that the model on fatalities per VMT yields an estimated 1.08% <u>increase</u> should be a cause for concern. Either the VMT weights are inaccurate or the control variables have different effects on crash frequency and crashworthiness.

[LBNL] Table ES.1 reports for lighter-than-average cars a 2.00% increase in crash frequency, a 0.54% decrease in fatalities per crash, and a 1.42% increase in fatalities per VMT; in this case the first two estimates sum almost exactly to the third (2.00% - 0.54% = 1.46%). In other cases the first two estimates do not sum to the third. In its previous studies DRI solved the three equations, crashes per VMT, fatalities per crash, and fatalities per VMT, simultaneously, which forces the estimated effects on fatalities per VMT to equal the sum of the estimated effects on crashes per VMT and fatalities per crash.

5. RECOMMENDATIONS	COMMENTS
Please comment on whether the LBNL assessment adequately addresses the NHTSA 2011 study and identifies the safety impact from mass reduction. Are the analytic methods and data used to assess the NHTSA study, and estimate the relationship between risk, mass, and footprint, appropriate? Is casualty risk per crash a legitimate measure of vehicle safety? What other methods or data could be used to better predict the effect of future vehicle designs on safety?	**[Chen and Kockelman]** As noted above, a primary concern remains a neglect of the notion that the type of car is very much a proxy for driver type, and a vehicle's crash avoidance may have very little to do with vehicle type. It has a lot to do with the person behind the wheel. Simply including gender and age variables cannot account for important covariates such as education, risk aversion, driving ability, wealth, etc. In reality, small cars may be less crash prone than Kahane's and Wenzel's results suggest, because they are driven by lower-income, younger, less risk averse people driving in more crash prone settings (e.g., commercial strips rather than pricey residential suburbs). Alas, it is very difficult to control for all these variables, since they are not readily available in data sets. In reality, the best the report authors can do with such data sets is to explain how all the other, relevant attributes may factor in (e.g., quality of driver and typical driving settings), and how they can generate biased estimation (sometimes in either direction). Discussion of relevant literature that looks more deeply at crash outcomes (e.g., Wang or Chen's papers, mentioned above, allowing for heteroscedasticity and individual vehicle attributes, non-driver outcomes, etc.) will also be useful.

[LBNL] As discussed above, additional sensitivity analyses were run using an alternative control variable for vehicle characteristics, initial vehicle purchase price (using values in Polk's VIN decoding software), and an alternative control variable for driver behavior, household income (based on California vehicle registration data in 2010). The results of these sensitivities are shown in Sections 5.3 and 5.5 in the Final Phase 2 report.

[Farmer] Casualty risk per crash does not fully measure the effects of vehicle size and weight reductions on society. Casualty risk per VMT best coincides with the NHTSA analysis of fatalities per VMT. The breakdown of casualty risk per VMT into the crash frequency and crashworthiness components is of interest. However, the surprising results reported here make everything suspect. For example, the Phase II report concludes that "the detrimental effect of male drivers has to do with their higher tendency of getting into a serious crash rather than their sensitivity to injury once a serious crash has occurred" (p. 24). A few pages later it concludes that "male drivers have essentially no effect on crash frequency, but cause a statistically significant increase in fatality risk once a crash occurs" (p. 28).

[LBNL] A summary discussing unexpected results has been added to the end of Section 2 in the Final Phase 2 report.

[Van Auken] [1] The basic methodology described by Kahane (2011) seems appropriate; and the extension by Wenzel (2011b) are also appropriate. However some results using these methods and data are not well |

	understood and need further diagnosis.
	[2] The induced-exposure data set provided by NHTSA is based on the non-culpable vehicles in two-vehicle crashes. See the Table 1 comments on the limitations of this data.
	[3] The state accident data files tend to have different database variable and coding definitions and criteria, which could confound the results.
Please comment on the overall adequacy of LBNL's assessment of the 2011 NHTSA report and its independent study of casualty risk for predicting the effect of vehicle mass or footprint reduction on safety. Provide any recommended improvements that might reasonably be adopted by the author to improve the analysis.	**[Chen and Kockelman]** Overall, the study is an enriching complementary document to the Phase 1 assessment of the 2011 NHTSA report. The parallel structure of the two reports allows the reader to easily compare and contrast the various additional analyses which examine the relationship of vehicle mass and footprint reduction on safety. However, as stated previously in the comments here, driver preference for specific car types (including size and mass) is related to driver socioeconomic characteristics and driving behavior. As vehicle, driver, and roadway environment characteristics all contribute to fatality risk, the effects of physical vehicle changes such as mass or footprint reduction on safety should not be overstated when the other two types of characteristics are not sufficiently accounted for.
	[LBNL] As discussed above, additional sensitivity analyses were run using an alternative control variable for vehicle characteristics, initial vehicle purchase price (using values in Polk's VIN decoding software), and an alternative control variable for driver behavior, household income (based on California vehicle registration data in 2010). The results of these sensitivities are shown in Sections 5.3 and 5.5 in the Final Phase 2 report.
	[Farmer] Overall these are reasonably good studies. The Phase I report does a very good job of assessing the NHTSA report of fatality risk. However, the Phase II report should be more cautious in its conclusions concerning casualty risk. The casualty analysis is based solely on police-reported data from 13 states, which:
	1. May not be representative of the US as a whole. 2. Are inconsistent in the information given and the way in which it is coded. 3. Suffer from information that is missing, inaccurate, or unclear.
	[Van Auken] The Wenzel (2011b) report provides a valuable supplement to the analysis and results in the Kahane (2011) report.
ADDITIONAL COMMENTS:	

[Farmer] Column G of Table 6.1 in the Phase II report provides the most appropriate comparison to results from the NHTSA report (Column A). For both fatalities and casualties per VMT, a 100-lb weight reduction is most harmful in lighter cars, less harmful in heavier cars and lighter light trucks, and slightly beneficial in heavier light trucks, minivans, and crossovers.

[Van Auken] See attached tables 6 and 7 below.

Table 6. Additional General Comments and Recommendations Mike Van Auken	
Section	**COMMENTS AND RECOMMENDATIONS**
4	Use of R^2 is confusing. Suggest using lower case "r" when referring to the sample correlation coefficient (Box, Hunter, Hunter, 1978, P. 61); or upper case R when referring to the regression coefficient of multiple determination (Draper and Smith, 1981, p. 90). *[LBNL] Figures 3.1 and 3.2, and 4.1 through 4.5, report r and not R^2 in the Final Phase 2 report. In addition, a new Table 3.1 with r and VIF, and a new Table 4.1 with r and R^2, have been added.*
All	In most cases the reported results are just estimates, but are not described as such. The word "effect" appears several hundred times in this report with the "estimated" or other qualifier. In some cases this may be appropriate and in other cases it is not appropriate. It is recommended that the author review each instance and revise as appropriate. *[LBNL] The word "estimated" will be used extensively in the final report. In addition the following text will be included in the Executive Summary, Section 1, and Section 6:* *Although the purpose of the NHTSA and LBNL reports is to estimate the effect of vehicle mass reduction on societal risk, this is not how the regression models should be interpreted. Rather, they are estimating the recent historical relationship between mass and risk, after accounting for most measurable differences between vehicles, drivers, and crash times and locations. In essence, the regression models are comparing the risk of a 2600-lb Dodge Neon with that of a 2500-lb Honda Civic, after attempting to account for all other differences between the two vehicles. The models are not estimating the effect of literally removing 100 lbs from the Neon, leaving everything else unchanged.* *In addition, the analyses are based on the relationship of vehicle mass and footprint on risk for recent vehicle designs (model year 2000 to 2007). These relationships may or may not continue into the future as manufacturers utilize new vehicle designs and incorporate new technologies, such as more extensive use of strong lightweight materials and specific safety technologies. Therefore, throughout this report we use the phrase "the estimated effect of mass (or footprint) reduction on risk" as shorthand for "the estimated change in risk as a function of its relationship to mass (or footprint) for vehicle models of recent design."*
All	"Crashworthiness" in most instances should be changed to "crashworthiness and crash compatibility" because the fatalities and/or serious injuries may either be in the subject vehicle (crashworthiness effect) or collision partner (crash compatibility

	effect).
	[LBNL] *Text has been changed*
4	The statistical significance of the linear trends in Figures 4.1 through 4.9 are not reported. It would be helpful if the confidence intervals or statistical significance of the linear trends were reported, either in addition to or instead of R^2. The confidence intervals for the estimated slopes should be added to the results in Tables 4.1 and 4.2.
	[LBNL] *New Tables 4.1 through 4.5 summarize the linear trends and statistical significance of Figures 4.1 through 4.17 in the Final Phase 2 report.*

Table 7. Additional Specific Comments and Recommendations Mike Van Auken		
Section	**Page**	**COMMENTS AND RECOMMENDATIONS**
Executive Summary, 4	iv, v, 22, 66	Suggest changing "variance in risk" to "***variation*** in risk" throughout. *[LBNL] The suggested change has been made.*
Executive Summary, 6	vii, 63	The statement "In conclusion, casualty risk per crash is not necessarily a better metric than fatality risk per VMT for evaluating the effect of mass or footprint reduction on risk; rather, it provides a different perspective in assessing the benefits or drawbacks of mass and footprint reduction on safety in vehicles. However, it does allow the separation of risk per VMT to be separated into its two components, crash frequency and risk per crash" suggests that the casualty risk per crash metric was needed in order to assess the crash frequency and risk per crash, which is incorrect. The DRI (2003-2012) methods have also estimated the effects of weight and size on crash frequency (A/E) and risk per crash (F/E) in terms of fatalities. *[LBNL] Text has been changed.*
5.3	49	It would be helpful to list the 18 manufacturer dummy variables in a table. *[LBNL] The indicator variables for the vehicle manufacturers has been added to the final report.*
5.3	48-49	It is unclear why Lexus, Acura, and Infinity are treated as separate manufacturers, but Cadillac and Lincoln are not. *[LBNL] The results of two regressions has been included in the final report: one including the five luxury brands in their parent manufacturers, and one accounting for each of the five luxury brands. Accounting for the five luxury brands has little effect on the estimated detrimental effect of car mass reduction on risk, as shown in Figure 5.3 in the Final Phase 2 report.*
5.3	49	It is unclear why AM General is considered a Chrysler brand. The AM General Hummer was sold by GM beginning with the 2003 model year. *[LBNL] AM General has been removed from the Chrysler brand and included in the Other category.*

Review of "An Analysis of the Relationship between Casualty Risk per Crash and Vehicle Mass and Footprint for Model Year 2000-2007 Light-Duty Vehicles"

David L. Greene

January 10, 2012

Summary

The Phase I and Phase II analyses by Tom Wenzel of LBNL have been executed diligently and consistently in accord with the methods and data used in the original NHTSA analysis. The studies contain many valuable, new insights. The phase I study highlights the weakening relationship between vehicle mass and highway fatalities. This is not only seen in decreasing coefficient estimates but in the very large number of results that are not statistically significant. When regressions were done separately by footprint deciles, vehicle mass was statistically significantly positively related to fatalities only for light-duty trucks in rollovers, there were almost as many cases in which mass was negatively related to fatalities (9 vs. 13 out of 27) and there were more instances of statistically significant negative relationships than positive relationships. Given that so many tests are being jointly conducted, it is quite possible that when joint probabilities are considered, there is no significant relationship between mass and fatalities (more on joint probabilities later). Showing the weakness and inconsistency of these results is an important contribution.

Another meaningful contribution of the phase I study, and one that deserves more emphasis, is a logical inference from the following findings: 1) much of the variance in risk remains unexplained even by the most complete models, 2) control variables explain 1 to 2 orders of magnitude more of the variance than the variables of interest (mass and size), 3) when key control variables are removed or changed it strongly influences the coefficients of mass and size. These results have very important implications for the robustness of the results and the likelihood that some or all of the apparently statistically significant relationships are due to spurious correlations with omitted or imperfectly controlled factors. Noting that exposure measures are control variables with constrained coefficients, the following observation from the phase I study is especially perceptive.

> "Calculating risk as total fatalities per induced exposure crash, rather than per vehicle mile traveled, reverses the sign of mass reductions on risk in cars and the lighter light trucks, with mass reduction leading to a reduction in risk in all vehicle types."

Finally, the phase I report notes that if only the control variables are included in the regression and not size or mass, the resulting residuals from the regression are uncorrelated with size or mass. Given these findings (as well as those of phase II) the conclusions that,

> "The 2011 NHTSA study, and this report, conclude that the effect of mass reduction *while maintaining footprint* on *societal* US fatality risk is small, *and may be non-existent*."

should be revised with the following emendation, "… and probably non-existent."

[LBNL] Text will be revised as indicated in bold above.

Both studies, like the NHTSA analysis, have shortcomings in terms of interpreting the results and the language used to describe the results, and acknowledging the limitations of the data and methodology. The limitations are extensive. The interpretation of the results of the LBNL studies commits two important, related errors. The first is to attribute inferred coefficients of mass and size as representing only the effects of vehicle mass and size when, as the phase I and II study results indicate, there is a virtual certainty that aliasing effects are present due to a combination of omitted variables, errors in variables and correlations among variables. Given that estimated driver and environmental factors tend to have 1-2 orders of magnitude larger impacts on safety outcomes than vehicle factors, the almost certain presence of aliasing effects must be explicitly acknowledged as severely limiting the ability to draw inferences about the effects of vehicle attributes. Second, the language used in interpreting results fails to acknowledge that the analysis does not address the effects of down-weighting or down-sizing specific vehicles or vehicle designs, but instead relies on correlations between vehicle weight and size in existing vehicle designs. In existing vehicles, weight and size are correlated with each other and many other vehicle attributes (and driver and environmental attributes, as well). Thus, the study is not actually measuring the effects of down-weighting via the material substitution and design changes likely to occur as a consequence of fuel economy and emissions standards. An early example of the kind of misleading language referred to here can be found on page iv.

> "For example, a 100-lb reduction in the mass of lighter cars leads to a 1.84% increase in crash frequency (columns B), while mass reduction leads to a 0.76% decrease in the number of fatalities per crash (column C);"

This statement is misleading in that it implies causality rather than correlation, and it is additionally misleading in that it implies that the inference applies to removing weight from specific vehicles. Neither is correct. A better statement would be the following.

> "For example, vehicles in the lighter class that are 100 lbs. lighter are correlated with a 1.84% increase in crash frequency.…"

There are so many examples of this misleading language that it is not feasible to list them all. All should be corrected, however. Failure to correct them could lead to serious misinterpretation of the studies' findings.

[LBNL] The word "estimated" will be used extensively in the final report. For example, the above sentence will be rewritten as:

"For example, the models estimate that 100-lb lower mass in lighter-than-average cars is associated with a 2.00% increase in crash frequency (column B), while lower mass is associated with a 0.54% decrease in the number of fatalities per crash (column C);"

In addition, the following text will be included in the Executive Summary, Section 1 and Section 6 of the reports:

"Although the purpose of the NHTSA and LBNL reports is to estimate the effect of vehicle mass reduction on societal risk, this is not exactly what the regression models are estimating. Rather, they are estimating the recent historical relationship between mass and risk, after accounting for most measurable differences between vehicles, drivers, and crash times and locations. In essence, the regression models are comparing the risk of a 2600-lb Dodge Neon with that of a 2500-lb Honda Civic, after attempting to account for all other differences between the two vehicles. The models are <u>not</u> estimating the effect of literally removing 100 lbs from the Neon, leaving everything else unchanged.

In addition, the analyses are based on the relationship of vehicle mass and footprint on risk for recent vehicle designs (model year 2000 to 2007). These relationships may or may not continue into the future as manufacturers utilize new vehicle designs and incorporate new technologies, such as more extensive use of strong lightweight materials and specific safety technologies. Therefore, throughout this report we use the phrase "the estimated effect of mass (or footprint) reduction on risk" as shorthand for "the estimated change in risk as a function of its relationship to mass (or footprint) for vehicle models of recent design.""

Following in the footsteps of the seminal study by DRI, the NHTSA and LBNL studies contribute to the literature in three important ways: 1) the LBNL and NHTSA studies recognize that the societal safety perspective is the correct perspective to when assessing the impacts of fuel economy and emissions regulations, 2) they recognize that vehicle dimensions and vehicle mass may have separate and potentially different impacts on both the likelihood of a crash and the outcomes of the crash and, 3) the LBNL phase two analysis makes an additional contribution by attempting to disentangle factors affecting the likelihood of a crash and factors affecting the outcomes of a crash.

Speaking of the DRI study, I am puzzled about why there are no references cited in the phase I study and only a handful all by Kahane and Wenzel, in the phase II study. This is perhaps due to the scope of work defined for the two studies but there are highly relevant studies in the literature that could have been cited, those by DRI foremost among them.   Making use of the insights from these studies would have been helpful in interpreting the results of both phase I and phase II.

[LBNL] LBNL has summarized the 2003 DRI studies, as well as updated results published by DRI in 2011; see Sections 2 and 3 in the Final Phase 2 report.

Lack of a Theory or Model of the Phenomenon

Both the NHTSA and LBNL studies lack a rigorous theory of the process by which down-weighting at constant size or down-sizing at constant mass affect societal safety either through crash avoidance or crashworthiness. This is not a trivial shortcoming because it affects the ability to formulate hypotheses and interpret results. Prior to the dissenting report on safety of the NRC 2002 CAFÉ report, the physics of elastic collisions between objects was typically cited as the underlying physical model.   That report showed how taking the societal perspective renders that model inappropriate. What remains appears

to be far more complex, involving the quantity of kinetic energy, the ability of vehicle designs to absorb that energy so as to minimize maximum deceleration rates, stability, maneuverability, safety technologies, and more.

The consequences of the lack of a rigorous theory are that it is not known, a priori, what the signs of coefficients are expected to be, let alone what their quantitative relationships should be. Hypotheses must be formulated based on intuition and the interpretation of results is likewise ad hoc. One implication of this is that results that suggest that lower fatalities are associated with lower vehicle mass have equal standing, a priori, with results that indicate that higher fatalities are associated with lower vehicle mass, and similarly for vehicle size. There are no surprising or unsurprising results, in theory.

This also makes it difficult to develop a plan for statistically testing the model or theory and its implications. It would have been helpful to the reader to have been presented early on in the report with such a plan of analysis.

[LBNL] Section 1.5 of NHTSA's 2011 report summarizes the hypothetical physical relationships between vehicle mass, footprint, and societal fatality risk. LBNL has added a section summarizing this discussion at the end of Section 1 of the Final Phase 1 and Phase 2 reports.

On the Virtual Certainty of Aliasing

The LBNL report typically attributes causal effects to correlations between mass or size and safety. In fact, most or all of the observed correlations are almost certainly affected by aliasing effects. There is ample evidence for this inference in the results presented in the LBNL phase II report.

The coefficients of mass and size change in important ways when different model formulations are estimated. Removing and adding control variables changes the magnitudes and sometimes the signs of the mass and size variables. This means that, at a minimum, the mass and size variables alias the effects of the omitted control variables. The question is whether the aliasing is eliminated entirely by the inclusion of the control variables available or whether some aliasing remains either because not all relevant and correlated control variables have been included or because the included control variables are imperfect measures of the factors they are intended to represent.

The latter seems highly likely for the following reasons. First, the overall explanatory power of the full models (including control variables), as measured by their R^2 is low. Most of the variance in casualties and fatalities remains unexplained. Second, at least some of the important included control variables are only crude approximations of the factors they are intended to represent. For example, dummy variables represent differences in state reporting practices, age and gender represent risky driving behavior differences among owners of different sizes and masses of vehicles, the presence or absence of a kind of safety equipment represents both its performance and use in a particular vehicle, and calendar year dummy variables represent unknown factors associated with the respective calendar years. Such practices are common and their use is appropriate. Third, the control variables generally account for 1-2 orders of magnitude more variance in the casualty and fatality variables than do vehicle weight and size. To recap, the amount of unexplained variability in the dependent variables is larger

relative to the variance statistically explained by the most complete models. Control variables are correlated with size and mass, and they account for 1-2 orders of magnitude more variability in the dependent variables than the variables of interest, mass and size. Therefore, even small correlations of size and mass with omitted variables or with errors (imprecision) of the control variables could easily result in biased estimates for the effects of size and mass on the dependent variables.

[LBNL] The only control variables used in the NHTSA regression models that are correlated with size or mass are the HD_PKP variable (r of 0.65 on OVERWT00 and r of 0.54 on FOOTPRNT) and the SUV variable (r of 0.62 on FOOTPRNT). The control variable LBNL has added to the sensitivity analyses for initial vehicle purchase price, PRICE000, also is correlated with mass and footprint. However it is possible that other controls not included in the regressions may be correlated with mass or footprint.

Tom Wenzel is to be commended for providing the results that definitively demonstrate the three key points made above. The above is not a criticism of the analysis nor of the results, per se. It is a criticism of their interpretation. In light of the above, the results should be interpreted in light of the virtual certainty that many of the estimated coefficients are likely to be biased in ways that make their interpretation highly uncertain. The implication is that phrases such as "down-weighting or down-sizing caused" to "mass (or size) and unobserved correlated factors are associated with…"

[LBNL] The following text will be added to the executive summary:

"It is unclear why lower vehicle mass is associated with higher crash frequency, but lower risk per crash, in the regression models. It is possible that including variables that more accurately account for important differences among vehicles and driver behavior would reverse this relationship. On the other hand, it is also possible that over thirty years of improvements in vehicle design to achieve high crash test ratings have enabled manufacturers to use clever vehicle design to mitigate the hypothetical safety penalty of low mass vehicles."

"The large remaining unexplained variance in risk by vehicle model could be attributable to other differences in vehicle design, or how drivers who select certain vehicles drive them. It is possible that including variables that account for these factors in the regression models would change the estimated relationship between mass or footprint and risk."

Additional text, discussing unexpected results, has been added to the end of Section 2 of the Final Phase 2 report.

On Joint Probabilities

The NHTSA and LBNL studies do not correctly interpret their results as joint statistical tests. When testing a hypothesis on, for example, 5 vehicle classes simultaneously, a result for one equation that might be statistically significant on its own may not be statistically significant as one of five related tests.

Statistical analyses comprised of multiple regressions too often overlook the fact that tests of statistical significance designed for individual regressions may not apply in the case of multiple regressions. That is the case here. NHTSA conducted 5 analyses to infer relationships between mass differences among

vehicles holding footprint constant for 5 classes of cars. The results showed one relationship out of five was statistically significant. As table 1 illustrates, using a simple example, if one conducts 5 trials, each with a 0.05 probability of given result, there is a 22.6% probability of finding at least one such result in the five trials. Thus, the joint significance level of the overall result (1 statistically significant regression out of 5) is 0.226, rather than 0.05.

Table 1. Simplified Illustration of the Joint Probability of Inferences in Multiple Regressions

# Significant Regressions	Combinations	0.95	0.05	Joint Probability
0	1	0.773780938	1	77.37809%
1	5	0.81450625	0.05	20.36266%
2	10	0.857375	0.0025	2.14344%
3	10	0.9025	0.000125	0.11281%
4	5	0.95	0.00000625	0.00297%
5	1	1	3.125E-07	0.00003%
				100.0%
				22.6%

So there is between a 1:4 and a 1:5 chance of getting one statistically significant result by pure chance. In fact, the actual significance level of the results is more complicated to calculate, and probably a bit smaller than 0.226. Thus, it is very appropriate for Dr. Kahane to add the qualifier "if any" to his conclusions about the relationship between the societal highway fatalities and mass reduction, holding footprint constant. Had appropriate tests of joint statistical significance been used to evaluate the results in the NHTSA and LBNL studies, the significance levels very likely would not meet accepted criteria for statistical significance. This could change the conclusions of the studies from the inference that mass is correlated with fatalities or casualties in some case but not others to the lack of statistically significant evidence that mass is correlated with fatalities or injuries on the highway. This is an important difference.

[LBNL] The comment implies that the significance of the estimate on mass for one of five types of vehicles, or for a handful of nine crash types, does not translate into whether the estimate of mass on risk across all vehicle types, or across all crash types, is statistically significant. For this reason we compared the results from NHTSA's nine regressions for nine types of crashes, in which the estimated coefficients were weighted by the number of fatalities in each type of crash, with results from a single regression model for all crash types combined (without controlling for crash type). The second and third columns in Figure 2.1 of the Final Phase 1 report present this comparison. The figure indicates that the estimated effects of mass reduction on risk, and the statistical uncertainty of those estimates, are virtually identical, whether one takes the weighted average estimates from the nine regression models by crash type or whether one runs a single regression model including all types of crashes.

Figure 6.13 in the Final Phase 1 report similarly shows the estimated effect of mass reduction from nine regression models, each including all vehicle types (with control variables for vehicle type) in each regression model. However, this analysis does not run a single regression model for all types of crashes. We ran a single regression model, across nine crash types and three vehicle types, including control variables for vehicle types but not for crash types. The estimated effect of a 100-lb reduction in mass is a 0.15% (+/- 0.23%) increase in US fatality risk per VMT holding footprint constant, while the estimated effect of a 1-square foot reduction in footprint is 0.43% (+/- 0.23%) increase in risk, holding mass constant.

These results suggest that, across the entire vehicle fleet, the estimated effect of mass reduction on fatality risk is very small, and not statistically significant, even though mass is associated with risk for certain vehicle types in certain types of crashes. However, we believe NHTSA is justified in examining the relationship between mass and risk for certain classes of vehicles.

NHTSA found that there is a statistically-significant, albeit small, estimated increase in fatality risk as mass decreases in one of five vehicle classes studied, lighter-than-average cars. Although it is possible that this result was obtained by pure chance, as expected by joint probabilities or the multiple-comparison fallacy, the fact that this result was obtained for the lightest cars, the vehicle class we expect to be most sensitive to mass reduction, does not appear to be a purely random result.

Page-by-Page Comments

I will make page by page comments on the phase II study only, since that contains the overwhelming share of original contributions and the key findings of the phase I study are recapitulated there.

p. iii   Paragraph 3. This would be a very good place to acknowledge the importance of driver behavior and environment on crash avoidance especially.

p. iv   Para. 3. This would be a good place to discuss probability inference in joint tests.

           Para 4. The statement about a 100-lb reduction in the mass, etc., is a good example of misleading language.

p. v   Para. 1. Again, it is misleading to say that mass reduction increases crash frequency, for reasons stated above.

           Para. 2. It is more accurate to describe the association of lower vehicle mass with casualty risk than the "effect of mass reduction on…" casualty risk.
           Para. 3. Would benefit greatly from joint probability inferences.

           Para. 4. As noted above, this shows how much more important the control variables are than the variables of interest.

           Para 5. Again, these are correlations not necessarily effects.

p. vi   Para. 1. Again, mass reduction is misleading terminology and you do not know if it increases casualty risk or not, you know only a correlation. Why is this so important? It is the virtual certainty of spurious correlations, or aliasing, as noted above.

Para. 4. (1st bullet) This is clear evidence of aliasing. Take variables out of the regression and the coefficients of interest change in important ways. Are there no important factors still missing? Are the variables included perfect measures of the factors of interest? Of course not. Thus, there must be remaining aliasing. How bad is it? We don't know.

The third bullet shows the same effect with a different set of variables.

p. vii   Para. 4. No, your analysis does not indicate "…that much of the detrimental effect of mass or footprint reduction on risk can be attributed to the tendency for mass or footprint reduction to increase crash frequency." Again, you have correlation, not causation and you have good reason to believe that what you are seeing is affected by spurious correlations.

Para. 5. The "effect" is small, 1-2 orders of magnitude smaller than correlations with other control variables, and IS strongly affected by which variables are in the equation, as stated on the previous page, and there is a great deal of unexplained variance. Please reconsider the meaning of these results in light of the comments above.

Finally, as the last paragraph of the ES implies, it would be far better not to speak in terms of "reducing" mass or size. That is not what is happening in your data set.

p. 1   Para. 4. Risk per VMT includes the effects of how well vehicles are driven as well as how well they can be driven. I think there is no chance that you have fully accounted for how well vehicles are driven.

[LBNL] Including driver age does account for how well vehicles are driven, to some extent. In addition, LBNL ran a sensitivity where vehicles with "bad" drivers (i.e. crashes involving drug or alcohol use, or the driver was cited for speeding, another traffic violation, or had been cited in the previous three years) were excluded from the analysis. Figure 5.14 of the Final Phase 1 report shows that this sensitivity found that the estimated detrimental effect of mass reduction on risk per VMT increased substantially for cars and lighter light trucks, and the estimated beneficial effect of mass reduction on risk decreased substantially for heavier light trucks and CUVs/minivans. Unfortunately, this detailed information on the driver is not available in the state crash data, so a similar analysis cannot be done to estimate the effect of removing bad drivers on crash frequency or on risk per crash. However, LBNL did run a sensitivity including average household income; see Section 5.5 in the Final Phase 2 report.

p. 2   Para. 2. Exposure measures are explanatory variables whose elasticity is constrained to 1. That is, it is assumed that an increase in vehicle use of 1 vehicle mile produces a 1 unit increase in the chance of a fatality (or casualty as the case may be). This is actually a maintained hypothesis. If this hypothesis is incorrect, it can bias the other coefficients in the equation. Thus, the change from fatalities/VMT to fatalities/registration-year, to fatalities per crash not only changes the meaning of the analysis, it also may bias coefficients in the event that the true relationship between fatalities and VMT is not an elasticity of 1.

[LBNL] We expect that the regressions using a different measure of exposure or of risk would estimate a different relationship between vehicle mass and risk.

p. 3   Line 1. Please acknowledge that your "accounting for differences in driver characteristics, crash locations, and other vehicle attributes" is incomplete and that this could affect your inferences about size and mass.

[LBNL] The phrase "…included in the NHTSA regression models" will be added to the end of this sentence.

p. 3   Para. 2. NHTSA's use of "non-culpable" vehicles involved in two vehicle crashes as an exposure measure raises its own issues. How non-culpable was the non-culpable vehicle. Often this is a matter of degree, rather than black or white. Driver behavior may also be involved. It seems to me this is just another potential measure of exposure that may or may not be better than any other measure and may introduce new sources of bias in the analysis.

[LBNL] DRI suggested using stopped vehicles, rather than non-culpable vehicles, in two vehicle crashes as the measure of exposure. NHTSA and LBNL ran sensitivity analyses using this alternative measure of exposure; see Section 5.6 in the Final Phase 1 report.

p. 3   Para. 4. Induced exposure needs to be defined. What is it intended to mean? This needs to be explicit.

Also, I am startled that there are no equations in these reports. Equations can provide an unequivocal explanation of the assumed relationships that cannot be adequately accomplished by words, in many cases. Why no equations?

[LBNL] The method that NHTSA used to develop the vehicle registration and vehicle miles traveled weights using the non-culpable vehicles from the crash data from 13 states, national and state vehicle registration data from R.L. Polk, and average vehicle odometer reading by vehicle year, make, and model is complicated, and thoroughly described in Sections 2.3 through 2.6 of the 2012 NHTSA final report.

p. 7   Para 3. CUVs and minivans are involved in fewer crashes with stationary objects than cars. Why? Is it the drivers, the vehicles, or the passengers? How well can you control for such differences? Not well. What does this mean for your analysis?

p. 9   Para 1. Here an equation showing how the weighting was done would be very helpful.

[LBNL] Please refer to Sections 2.3 through 2.6 of the 2012 NHTSA final report for a detailed description of how NHTSA created the VMT weights.

p. 9   Bullet 1. Excluding these vehicle types implies that the control variables in the model are not adequate to account for whatever makes these vehicle types different from the vehicles included in the analysis. First, this is an admission that the model is not adequate to explain the fatalities associated with these vehicles. Second, it is an admission that if they were included the coefficients on the variables of interest would likely be biased by spurious correlations. Clearly, it would not even be sufficient to include the vehicles along with a control variable (e.g.,

X = 1 if vehicle is a police car, 0 otherwise). This is yet another indication that the model suffers generally from omitted variables, errors in variables and correlation among right-hand side variables.

[LBNL] LBNL agrees that excluding these vehicle types is a tacit admission by NHTSA that the control variables used in their regression models do not fully account for differences in vehicle types or the behavior of their drivers. However, LBNL ran a sensitivity in both Phase 1 and Phase 2 draft final reports where these four vehicle types were included in the regression models, with a control variable for each vehicle type. Figure 5.15 in the Final Phase 1 report indicates that including these vehicle types increases the estimated detrimental effect of mass reduction on US fatality risk per VMT for lighter-than-average cars, but increases the estimated beneficial effect for heavier-than-average light trucks. Figure 5.12 in the Final Phase 2 report shows that including these vehicle types has little effect on the estimated effect of mass reduction on casualty risk per crash.

p. 10   Line 3. Sentence does not make sense. Please correct. *[LBNL] The sentence fragment has been deleted*.

p. 12   Sect. 2.3. Please provide an equation. *[LBNL] The term "equation" has been replaced with the term "conversion factor."*

p. 13   Were the confidence intervals calculated using $e^x - 1$? Please state explicitly or, better, show an equation. *[LBNL] Yes, the end of this paragraph states that the confidence intervals were calculated the same way.*

Para. 2. Here another instance where you say "mass reduction increases societal fatality risk" but you really are not entitled to say that. It is misleading. Also, the NHTSA CI's are larger, as they should be in a joint test.

Para. 3. These results require an underlying theory for interpretation. The lack of one makes it seem like there is just no consistency in the results.

[LBNL] A section summarizing the physical theory of vehicles and fatality risk has been added after Section 1 in both the Final Phase 1 and Phase 2 reports.

Para. 4. The fact that the results for fatalities per crash differ substantially from fatality per VMT may be very important. Taken at face value, it would imply that any negative effect of reduced mass is due to its effect on crash avoidance (crash probability) rather than crashworthiness. This is where the lack of a theory is most troubling. Why would that be? Are lighter vehicles less easily controlled, etc.? Or, as seems much more likely, is there a spurious correlation between mass and other omitted or imperfectly measured factors (including driver behavior) that lead to an increased probability of a crash? Consider, for example, driver age. Driver age is related to crash involvement. Driver age is a control variable. But are all young drivers the same? Is it possible that young drivers more prone to risky behavior tend to drive lighter vehicles? If so, this could partly explain the result observed. Of course, this is just speculative, but the point is that correlations with imperfectly measured and omitted factors are highly likely to be present in the data and, if there, could easily affect the statistical inferences.

[LBNL] We expect that the regressions using a different measure of exposure would estimate a different relationship between vehicle mass and risk, although the differences could also be explained by other factors not included in the regression models.

p. 14 Here we see that changing the exposure measure influence the effect of mass and size on fatalities and casualties, which is more evidence that spurious correlations are likely to be biasing estimated coefficients for mass and size.

[LBNL] Again, we expect that the regressions using a different measure of exposure would estimate a different relationship between vehicle mass and risk. Although the differences could also be explained by other factors not included in the regression models.

The bar graphs with confidence intervals are well done and convey a great deal of information effectively. The patterns of magnitude and statistical significance are difficult to interpret, partly because there is no explicit theory of what should happen and partly, perhaps, because the relationships are actually not real.

p. 16 Para. 1. Reduction in the mass of lighter cars increases crash frequency but reduces fatalities per crash. This is contradictory to the previously maintained theory that mass protects due to the physics of velocity changes in elastic collisions. Indeed, there is no theoretical explanation for these results, only speculation.

[LBNL] It is unclear why lower vehicle mass is associated with higher crash frequency, but lower risk per crash, in the regression models. It is possible that including variables that more accurately account for important differences among vehicles and driver behavior would reverse this relationship.

Adding vehicle purchase price substantially reduces the estimated increase in crash frequency as vehicle mass decreases for all vehicle types; mass reduction is now estimated to slightly decrease crash frequency in the case of heavier-than-average cars, but is still estimated to increase crash frequency for the other four types of vehicles. See Figure 4.10 in the Final Phase 2 report.

On the other hand, it is also possible that over thirty years of improvements in vehicle design to achieve high crash test ratings have enabled manufacturers to use clever vehicle design to mitigate the hypothetical safety penalty of low mass vehicles.

Text to this effect has been added to the Final Phase 2 report.

p. 18 Para. 1. Here is a good example of such casual speculation.

p. 19 Para. 1. Developing VMT weights for the 13 states is a good idea, given the effect of exposure measures on inferences. Still the results would not be definitive.

p. 20 Para. 2. "Mass reduction leads to a large reduction in risk only in crashes with objects for heavier cars?" There is only one type of crash in which the simple physics of collisions leads to an unambiguous benefit for increased mass, and that is collisions with moveable or breakable objects. This finding contradicts even that.

[LBNL] As summarized in the new Section 1.1 in the Final reports, additional mass may be beneficial if it enables the vehicle to knock down the object and continue moving, thereby reducing its delta V than if the vehicle was stopped by the object. In a previous study NHTSA estimated that the object is knocked down in about 25% of frontal collisions with stationary objects (Partyka, 1995). However, if the object is immovable, reducing vehicle mass would lower the kinetic energy of the crash, thereby reducing the amount of energy for the vehicle's structure to absorb, and likely reducing risk. Additional crush space or structural strength would increase the amount of crash energy the vehicle can absorb, tending to reduce risk to occupants.

Figures 2.12 through 2.14, and Table 2.3, in the Final Phase 2 report indicates that lower mass is associated with a reduction in 13-state casualty risk per crash in crashes with a stationary object, for four of the five vehicle classes, although the estimated reduction is large enough to become statistically-significant only for heavier-than-average cars (3.75%) and CUVs/minivans (2.60%). Note that in the same table lower footprint is associated with a significant 3.55% increase in casualty risk in car crashes with an object; smaller size substantially increases risk in crashes with objects for light trucks (0.99%) and CUVs/minivans (5.56%) as well.

The relationship between lower mass and fatality risk per crash in crashes with stationary objects is even stronger (not shown in the Final Phase 2 report, but shown in the table below); for fatality risk per crash, mass reduction is estimated to reduce fatality risk per crash in crashes with stationary objects in all five vehicle classes, and the estimated reductions are statistically significant for cars and the lighter light trucks. As in casualty risk per crash, footprint reduction is estimated to increase risk in all three types of vehicles.

Estimated effect of mass or footprint reduction on risk per crash, in crashes with a stationary object

Variable	Vehicle type	Casualty risk per crash	Fatality risk per crash
Weight	Cars < 3106	-0.80%	-3.96%
	Cars > 3106	-3.75%	-5.14%
	LTs < 4594	-0.78%	-3.98%
	LTs > 4594	0.04%	-0.81%
	CUVs/minivans	-2.60%	-5.77%
Footprint	Cars	3.55%	4.41%
	LTs	0.99%	2.11%
	CUVs/minivans	5.56%	5.93%

Para. 3. More speculation, this time about rollovers.

[LBNL] A section summarizing the physical theory of vehicles and fatality risk has been added after Section 1 in both final reports. As noted in that section, reducing a vehicle's mass without changing its roof structure would reduce the force applied on the roof once a vehicle rolls over, and thus reducing risk.

p. 21   Para. 3. "Curiously,…" Curiouser and curiouser.

Figure 2.15 printed without labels. Could be my computer but the other graphs were fine.

p. 24   Para. 2. This is probably a very important finding that needs further investigation and explanation. As figure 2.16 illustrates well, the correlations with mass and size are orders of magnitude smaller than the correlations with driver and environmental factors. This is why even small correlations with omitted or imperfectly measured control factors could be, are even likely to be, predominantly responsible for the estimated coefficients of mass and size.

p. 25   Para. 2. The results for minivans discussed here could be due to what is going on inside the vehicle as much or more than the vehicle itself. How could these results be explained in terms of the vehicle itself. Figure 2.17 shows this again. The effects of calendar year dummies, which can only be considered rough approximations to unknown and various time-related factors, have much large effects than size or mass. Again in figure 2.20.

[LBNL] Text discussing unexpected results such as this one will be added to the end of Section 2 in the Final Phase 2 report.

p. 28   Para. 1. "Surprisingly,…" How can side airbags, which deploy only in a crash, reduce crash frequency but not fatality risk in a crash? Only if the real effect is a reflection of who buys a CUV/minivan with side airbags and how and where they drive. There are more surprising inferences in paragraph 2 about male and female drivers. Surprising relative to what theory?

[LBNL] The coefficients for side airbags in cars have the expected effect; airbags have a small effect on crash frequency, but a large beneficial effect on risk per crash. The unexpected results for CUVs/minivans could be attributable to the relatively small number of vehicles included in the analysis. Text discussing unexpected results such as this one will be added to the end of Section 2 in the Final Phase 2 report.

p. 32   The problem here is not numerical multicollinearity (numerical difficulties inverting the cross-product matrix) but the more complex problem of correlations among right-hand side variables, omitted variables, errors in variables, and correlations of included variables with omitted and imperfectly measured variables. This leads to biased estimates.

p. 36   Table 3.1 cries out for inference based on joint probabilities. What is the probability of observing "success" in at least 3 out of 27 trials when the true probability of success is only 0.05. See discussion above. The probability is certainly much higher than 0.05 and probably closer to 0.5. The implication is that, taken together, these results do not show any statistically significant relationship between mass or size and risk per crash. If there were a rigorous underlying theory, the interpretation might be different (patterns of significance could matter) but there is none. Again, good graphs on succeeding pages.

[LBNL] We believe that the results in Table 3.1 (Table 3.2 in the Final Phase 2 report) speak for themselves, without having to calculate joint probabilities. Casualty risk increases with decreasing mass in more than half of the footprint deciles in only 9 of the 27 vehicle/crash combinations, and the increase is statistically significant in fewer than half of the footprint deciles in each of these 9 cases.

p. 41   The statistical significance of such a relationship should be the same whether bins are used or not. Is it?

[LBNL] Because the risk of a casualty is a binary outcome for an individual vehicle, binning is necessary, either by increments of weight, make and model, or some other factor, to present the relationship between risk per VMT and weight in a plot. The relationship between risk per VMT and curb weight is statistically significant when the values for each vehicle weight bin are weighted by the number of observations in each bin.

The statistical significance of the relationship between casualty risk and weight bin has been included in a new Table 4.1 in the Final Phase 2 report.

p. 42   Para. 3. R-squared is not the correct measure of statistical significance. Is the coefficient of weight significantly different from zero?

[LBNL] A new Table 4.1 has been added to both final reports which includes the statistical significance of the relationship. However, the R^2 value is shown to indicate that, when the data are binned by vehicle weight the relationship appears stronger than when the data are plotted by vehicle make and model, in the first column of new Table 4.2. For example, the relationship between US fatality risk per VMT and weight for individual car models in new Table 4.2 in the Final Phase 1 report (4.0%) is statistically significant; however, the R^2 is only 0.17. Therefore, on average the relationship is statistically significant, but there is a wide range in risk for individual car models.

p. 43   Para. 2. This is perhaps the key finding of the phase I and II analyses. Control variables explain 1-2 orders of magnitude more variance than size or mass. Still, most of the variance remains unexplained and is uncorrelated with mass or size. It is very likely there is nothing going on here.

p. 47   Para. 2. More evidence of correlation of mass and size with control variables and how changing definitions or excluding control variables results in important changes in the coefficient of mass and size. Such results are considered unstable.

[LBNL] Section 5.1 in the Final Phase 2 report discusses the importance of including control variables for each of the thirteen states, to account for differences among the states in reporting rates for non-injury crashes.

p. 49   Para. 1. The rationale for the grouping of manufacturers is not obvious. Can you explain it?

[LBNL] The analysis has been revised to include the luxury models Lexus, Acura, and Infiniti with their nameplate manufacturers Toyota, Honda, and Nissan. An additional sensitivity was conducted using separate indicator variables for the five luxury brands Lexus, Acura, Infiniti, Cadillac, and Lincoln; see Section 5.3 in the Final Phase 2 report.

Para. 2. Yet more evidence for the instability of the model and likelihood that variables still missing from the model, plus errors in measuring the included control variables are likely biasing the inferences. The results described in this paragraph do not make sense to me. How can they be explained other than random results?

p. 50   Para. 1. More casual interpretation of results. OK, maybe, maybe not. Same for paragraph 2. The economy faltered in 2008 but the big negative effect was in 2007. The downward trend started in 2004. This correlates neither with vehicle mass changes over time nor economic growth as measured by real GDP. Idle speculation.

p. 51   Para. 1. Good discussion of the gratuitous speculation by NHTSA about the meaning of the observed correlations. This is more a Rorschach test than statistical analysis.

p. 51   Para. 2. "We have no explanation for why…" More of this kind of honest appraisal is needed in studies like these.

p. 52   More results for which there is no explanation.

p. 53   Why the interaction between calendar year dummy variables and safety equipment? The presence of the safety equipment on a particular vehicle is established. What has calendar year (not model year) to do with it? Again, one suspects spurious correlations.

p. 55   One needs to think carefully about the reasons why vehicles would be excluded. It does not appear that NHTSA did that. First line of first paragraph "used" rather than "sused". *[LBNL] Text has been changed.*

p. 57   Well reasoned. It is interesting that NHTSA resisted including footprint or size in previous analyses on the grounds of correlation with mass. These results show that assertion was groundless.

p. 59   Para. 3. Rather than say risk per VMT accounts for two effects, it is better to say it includes or comprises two effects. *[LBNL] Text has been changed.* But this statement also ignores the important influences of drivers and environment and their potential correlations with other factors. Yes, it includes how well a vehicle can be driven, but more importantly it includes how well a vehicle IS driven. That is in there too and is very likely to be correlated with make, model, and other vehicle attributes.

p. 60   Para. 3. Here again, the conclusions are misstated. It is not a genuine "reduction" in mass, but an association with the mass of vehicles. And how does it "lead" to and increase in crash frequency? What is the theory or model that predicts this? Driver and environment are very likely mixed up in these results to an unknown but likely substantial degree. So what can we really conclude? Not this*. [LBNL] Text has been changed.*

As I read the conclusions and inferences I find myself asking, why?, why?, repeatedly without any sound explanations. Page 61, paragraph 3 contains more "surprising" results. Surprising because they are contrary to theory? Surprising because they are contrary to intuition? Surprising because they are random? To what can we attribute so many "surprising" results, and how many must there be before one concludes that the analysis is not revealing what we had hoped it would.

Para. 4. Again, this cries out for joint statistical inference. Three statistically significant results out of 27 is probably nothing statistically significant at all.

p. 62   Para. 1. What this shows, again, is that the coefficients of mass and size are strongly influenced by which control variables are included in the model and how they are defined. These results and their implications need explanation. The bottom line is that the effects of mass and size are likely to be (after the necessary joint significance calculations are done) not statistically significant, not consistent, and not robust.

p. 63   How do mass and footprint reduction (again, it's not really reduction in the sense of designing lighter vehicles to increase fuel economy or reduce GHG emissions, the issue at hand) increase crash frequency. What is the theory? I don't find a theory in either the NHTSA report or the phase I and phase II studies. Absent a theory, these results seem sufficiently unstable and inconsistent to be highly questionable as evidence of any relationship between mass or size and crashworthiness or crash avoidance. I think joint estimation of significance levels would provide additional support for this view.

[LBNL] Section 1.5 of NHTSA's 2011 report summarizes the hypothetical physical relationships between vehicle mass, footprint, and societal fatality risk. LBNL has added a section summarizing this discussion at the end of Section 1 in each of the final reports.

4.   References

Chen, H.Y., Ivers, R.Q., Mariniuk, A.L.C., Boufous, S., Senserrick, T., Woodward, M., Stevenson, M. and Norton R. 2010. "Socioeconomic status and risk of car crash injury, independent of place of residence and driving exposure: Results from the DRIVE study." *Journal of Epidemiology and Community Health* 64(10), pp. 998-1003.

Kweon, Y-J. and Kockelman, K.M. 2003. "Overall Injury Risk to Different Drivers: Combining Exposure, Frequency, and Severity Models." Accident Analysis and Prevention 35(4), pp. 441-450.

Wenzel, Tom P. 2011a. *Assessment of NHTSA's Report "Relationships Between Fatality Risk, Mass, and Footprint in Model Year 2000-2007 Passenger Cars and LTVs"*. Preliminary report prepared for the Office of Energy Efficiency and Renewable Energy, US Department of Energy, Berkeley, California. August. LBNL-.

Wenzel, Tom P. 2011b. *An Analysis of the Relationship between Casualty Risk Per Crash and Vehicle Mass and Footprint for Model Year 2000-2007 Light-Duty Vehicles.* Preliminary report prepared for the Office of Energy Efficiency and Renewable Energy, US Department of Energy, Berkeley, California. December. LBNL-.

Wenzel, Tom P. 2012a. *Assessment of NHTSA's Report "Relationships Between Fatality Risk, Mass, and Footprint in Model Year 2000-2007 Passenger Cars and LTVs".* Final report prepared for the Office of Energy Efficiency and Renewable Energy, US Department of Energy, Berkeley, California. August. LBNL-.

Wenzel, Tom P. 2012b. *An Analysis of the Relationship between Casualty Risk Per Crash and Vehicle Mass and Footprint for Model Year 2000-2007 Light-Duty Vehicles.* Final report prepared