Title

Assessment of DRI’s Two-Stage Logistic Regression Model Used to Simultaneously Estimate the Relationship between Vehicle Mass or Size Reduction and U.S. Fatality Risk, Crashworthiness/Compatibility, and Crash Avoidance:

Author

Wenzel, Tom, P

Publication Date

2016-01-29

LBNL-1005830

Assessment of DRI's Two-Stage Logistic Regression Model Used to Simultaneously Estimate the Relationship between Vehicle Mass or Size Reduction and U.S. Fatality Risk, Crashworthiness/Compatibility, and Crash Avoidance

Final report prepared for the Office of Energy Efficiency and Renewable Energy,
US Department of Energy

Prepared by

Tom Wenzel
Building Technology and Urban Systems Division
Energy Technologies Area
Lawrence Berkeley National Laboratory
Berkeley, CA 94720

January 2016

This work was supported by the Vehicle Technologies Program,
Office of Energy Efficiency and Renewable Energy of the U.S. Department of Energy
under Contract No. DE-AC02-05CH11231.

Acknowledgements

This report was funded by Carol Schutte of the Vehicle Technologies Program in the Office of Energy Efficiency and Renewable Energy of the U.S. Department of Energy. We are grateful for her support of this research.

This work was supported by the Assistant Secretary for Energy Efficiency and Renewable Energy, of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231.

Executive summary

This report summarizes an effort to replicate the results from a 2-stage regression model developed by Dynamic Research Inc. (DRI) to simultaneously estimate the effect of mass or footprint reduction on the two components of societal fatality risk per vehicle miles of travel, crashes per VMT (crash frequency) and fatality risk once a crash has occurred (crashworthiness/ compatibility).

Lawrence Berkeley National Laboratory (LBNL) was not able to exactly replicate the results from DRI's simultaneous 2-stage regression model. This may be because of discrepancies in how DRI and LBNL classified the state police-reported crash data into crash types. LBNL's analysis of four alternate regression models suggests that the results from DRI's method are sensitive to changes in what data are used in the analysis, or even the particular vehicles included in the decimation sample; in some cases the sign of the estimated relationship from DRI's results changes under an alternate LBNL regression. However, for the most part LBNL's alternate regressions confirm the general results from DRI's simultaneous model, and LBNL's analysis in its Phase 2 report: that mass reduction is associated with an increase in crash frequency (crashes per VMT), but a decrease in fatality risk once a crash has occurred, across all vehicle types. Similar results were obtained after using stopped rather than non-culpable vehicles as the induced exposure records, and replacing footprint with wheelbase and track width.

Table of Contents

Acknowledgements ... i
Executive summary ... iii
1. Introduction ... 1
2. Classification of crash types ... 2
3. Comparison with results in DRI's Phase II report and LBNL sensitivities ... 5
4. Comparison of results in DRI's Supplemental Report ... 11
5. Conclusions ... 14
6. References ... 14
Appendix A. Comparison of DRI and LBNL Two-Stage Results ... 16

List of Tables and Figures

Table 1. Distribution of vehicles in police-reported non-fatal crashes, by state and vehicle type . 3
Table 2. Distribution of PA non-fatal crashes by type and year ... 4
Table 3. Distribution of vehicles in police-reported non-fatal crashes, by vehicle and crash type 5
Table 4. Number of vehicles involved in non-fatal crashes used in regression analyses ... 6
Table 5. Number of vehicles involved in non-fatal crashes used in LBNL regression analyses... 7
Table 6. DRI 2-stage and alternate LBNL 2-stage estimates of the effect of mass or footprint reduction on U.S. societal fatality risk per crash (crashworthiness/compatibility) ... 9
Table 7. DRI 2-stage and alternate LBNL 2-stage estimates of the effect of mass or footprint reduction on U.S. crash frequency per VMT (crash avoidance) ... 10
Table 8. NHTSA 1-stage, DRI 2-stage, and alternate LBNL 2-stage estimates of the effect of mass or footprint reduction on U.S. societal fatality risk per VMT ... 11
Table 9. Crash, driver, and vehicle characteristics of non-culpable and stopped vehicles used for induced exposure ... 12
Table 10. Comparison of DRI and LBNL estimated effect of mass, track width, or wheelbase reduction on U.S. societal fatality risk per crash, crash frequency per VMT, and fatality risk per VMT ... 13
Table A-1. DRI and LBNL 2-stage estimated effect of mass or footprint reduction on US fatality risk per crash, crash frequency per VMT, and fatality risk per VMT ... 16
Table A-2. LBNL 2-stage estimated effect of mass or footprint reduction on US fatality risk per crash, crash frequency per VMT, and fatality risk per VMT, using 13 NHTSA states ... 16

1. Introduction

Reducing vehicle mass is perhaps the easiest and least-costly method to reduce fuel consumption and greenhouse gas emissions from light-duty vehicles. However, the extent to which government regulations should encourage manufacturers to reduce vehicle mass depends on what effect, if any, light-weighting vehicles is expected to have on societal safety. As part of an interagency analysis effort between the National Highway Traffic Safety Administration (NHTSA), the Environmental Protection Agency (EPA), and the Department of Energy (DOE), Lawrence Berkeley National Laboratory (LBNL) has been examining the relationship between vehicle mass and size and U.S. societal fatality and casualty risk, using historical data on recent vehicle designs. This research effort informs the agencies on the extent to which vehicle mass can be reduced in order to meet fuel economy and greenhouse gas emissions standards, without compromising the safety of road users.

In 2012 NHTSA updated its 2003 and 2010 logistic regression analyses of the effect a reduction in light-duty vehicle mass has on US societal fatality risk[1] per vehicle mile of travel (VMT; Kahane 2012); the 2012 analysis is the most thorough investigation of this issue to date. In 2012 LBNL completed two studies that replicated NHTSA's analysis of fatality risk per VMT (Wenzel 2012a) and analyzed the relationship between mass reduction and the two components of risk per VMT, crashes per VMT (or crash frequency) and risk once a crash has occurred (or crashworthiness; Wenzel 2012b).

Dynamic Research, Inc. (DRI) released three reports in 2012 analyzing the relationship between vehicle mass and size and societal U.S. fatality risk. The three DRI reports use essentially the same data and methodology developed by NHTSA to estimate the effect of a reduction in vehicle mass on societal fatality risk per VMT, while holding vehicle size (footprint) constant. The Phase I DRI report (DRI 2012a) attempts to identify, and correct, discrepancies in the data and methodology used to replicate the 2003 NHTSA results, for MY91 to MY99 vehicles in calendar years 1995 to 2000; however, DRI was not able to exactly replicate NHTSA's estimated effect of mass or footprint reduction on fatality risk per VMT. In its Phase I report DRI also introduces a 2-stage regression model that simultaneously estimates the effect of mass or footprint reduction on the two components of fatality risk per VMT: crash frequency (crashes per VMT) and crashworthiness/compatibility (fatalities per crash). The estimated effect of mass or footprint reduction on fatality risk per VMT from DRI's 2-stage model is quite similar to its estimates from its 1-stage model, but differences with NHTSA's 1-stage model remain. The Phase II DRI report (DRI 2012b) replicates the 2012 NHTSA results, for MY00 to MY07 vehicles in calendar year 2002 to 2008, and updates DRI's 2-stage regression model. DRI was able to exactly replicate NHTSA's results for fatality risk per VMT in its 1-stage model, most likely because DRI used the same datasets as NHTSA, DRI's estimates for fatality risk per VMT in its 1- and 2-stage models in its Phase II report are quite similar. The Supplemental DRI report (DRI 2012c) assesses the sensitivity of the regression estimates to two changes in the NHTSA methodology: using stopped, instead of non-culpable, vehicles in two-vehicle crashes as the measure of induced exposure, and replacing footprint with vehicle wheelbase and track width as the measure of vehicle size.

---

[1] Societal fatality risk includes the risk to both the occupants of the case vehicle as well as any crash partner or pedestrians.

This report summarizes LBNL's effort to replicate the DRI 2-stage regression model, and its results.

2. Classification of crash types

DRI used the databases of FARS fatality cases and 13-state induced exposure crash data that NHTSA developed for their 2012 report. However, DRI used police-reported crash data from only 10 states (all the NHTSA states except AL, NJ, and WI) to estimate the effect of mass reduction on police-reported non-fatal crashes per VMT and fatalities per police-reported non-fatal crashes. In addition, DRI used its own methodology to classify the state crash data by type of crash. For its 2012 Phase 2 report, LBNL used the same 13 states as NHTSA, but also had to classify crashes by type using its own definitions. In order to replicate DRI's Phase 2 analysis LBNL needs to reconcile the DRI and LBNL distributions of crashes by type, to determine whether the classification schemes are similar.

Table 1 compares the total number of vehicles in non-fatal crashes in the 10 states reported by DRI (in Appendix E.f of DRI's Phase II report, DRI 2012b) and LBNL. The table indicates that LBNL identified almost 4% more cars in non-fatal crashes in the 10 states than DRI; the biggest increases were in PA and MI, while LBNL identified substantially fewer cars in non-fatal crashes than DRI in WA, FL, MO, and NE. The table indicates that these trends are consistent for light trucks and CUVs/minivans, as well as cars. It is not clear why LBNL identified more vehicles involved in non-fatal crashes from the state data than DRI; one potential reason is that DRI excluded vehicles that had a reported model year prior to 1981 in the state crash databases. Another possibility is differences in the programming DRI and LBNL used to decode vehicle identification numbers (VINs).

Table 2 shows the number of police-reported non-fatal crashes in PA, from Appendix E.f in DRI, by crash type and calendar year (2002 crash data for PA are unavailable). Starting in 2006 DRI's count of crashes is one-third of their count in 2005; since this is consistent across all types of crashes Table 2 suggests there is an error in how DRI classified vehicles after 2005 in PA (PA did change the coding of their crash databases in 2006, which may be another source of the problem).

Table 3 compares the total number of vehicles in non-fatal crashes in the 10 states reported by DRI and LBNL, by vehicle and crash type. LBNL identifies substantially more cars involved in rollovers and crashes with a heavy-duty truck and a light-duty truck, but substantially fewer cars involved in a crash with a heavier-than-average car, than DRI. The majority of the discrepancy in Table 3 is for case vehicles involved in crashes with heavier-than-average cars and lighter-than-average light trucks; this suggests that either DRI and LBNL are using different average curb weights to define lighter- and heavier-than-average cars and light trucks, or DRI is classifying certain light truck models as cars (or alternatively LBNL is classifying certain car models as light trucks). These trends are consistent for light trucks and CUVs/minivans as well.

Table 1. Distribution of vehicles in police-reported non-fatal crashes, by state and vehicle type

Vehicle type	State	DRI	LBNL	Difference	Percent difference
Cars	FL	498,430	486,214	-12,216	-2.5%
	KS	109,023	108,039	-984	-0.9%
	KY	268,576	266,731	-1,845	-0.7%
	MD	214,641	215,246	605	0.3%
	MI	474,921	485,384	10,463	2.2%
	MO	325,618	314,076	-11,542	-3.5%
	NE	77,786	71,630	-6,156	-7.9%
	PA	116,391	246,054	129,663	111.4%
	WA	201,910	182,052	-19,858	-9.8%
	WY	13,839	14,114	275	2.0%
	Total	2,301,135	2,389,540	88,405	3.8%
Light trucks	FL	280,020	270,673	-9,347	-3.3%
	KS	75,908	74,143	-1,765	-2.3%
	KY	171,910	169,073	-2,837	-1.7%
	MD	88,628	92,170	3,542	4.0%
	MI	359,487	366,443	6,956	1.9%
	MO	194,313	186,977	-7,336	-3.8%
	NE	52,845	47,211	-5,634	-10.7%
	PA	58,802	119,996	61,194	104.1%
	WA	114,969	105,267	-9,702	-8.4%
	WY	19,829	19,965	136	0.7%
	Total	1,416,711	1,451,918	35,207	2.5%
CUVs/ minivans	FL	110,801	107,778	-3,023	-2.7%
	KS	28,747	28,384	-363	-1.3%
	KY	64,549	63,562	-987	-1.5%
	MD	48,530	47,780	-750	-1.5%
	MI	171,976	175,052	3,076	1.8%
	MO	76,563	73,519	-3,044	-4.0%
	NE	21,546	19,858	-1,688	-7.8%
	PA	28,266	65,320	37,054	131.1%
	WA	56,489	50,155	-6,334	-11.2%
	WY	3,562	3,623	61	1.7%
	Total	611,029	635,031	24,002	3.9%

Table 2. Distribution of PA non-fatal crashes by type and year

	Crash type	2003	2004	2005	2006	2007	2008	Total
DRI	Rollover	153	169	192	81	102	116	813
	Fixed object	3,760	4,855	4,837	2,531	3,444	4,150	23,577
	Ped etc.	570	813	792	151	195	13	2,534
	HDT	692	737	872	272	366	345	3,284
	Lgt car	4,766	6,448	5,729	1,717	1,809	1,866	22,335
	Hvy car	3,598	5,373	5,390	1,599	2,030	2,030	20,020
	Lgt LT	1,754	2,363	1,966	642	704	642	8,071
	Hvy LT	1,211	1,845	1,891	589	794	784	7,114
	Other	5,629	8,162	8,039	2,087	2,303	2,423	28,643
	Total	22,133	30,765	29,708	9,669	11,747	12,369	116,391
LBNL	Rollover	619	690	920	965	1,163	1,274	5,631
	Fixed object	4,735	5,483	6,518	6,976	8,693	9,486	41,891
	Ped etc.	633	728	842	925	1,048	1,110	5,286
	HDT	764	942	1,102	1,273	1,565	1,411	7,057
	Lgt car	5,821	7,034	7,374	8,119	8,437	7,736	44,521
	Hvy car	3,527	4,493	5,076	5,880	6,722	6,157	31,855
	Lgt LT	3,325	4,108	4,340	5,239	5,849	5,153	28,014
	Hvy LT	1,621	2,217	2,798	3,485	4,313	3,958	18,392
	Other	7,660	9,221	10,304	11,206	12,019	12,997	63,407
	Total	28,705	34,916	39,274	44,068	49,809	49,282	246,054

Table 3. Distribution of vehicles in police-reported non-fatal crashes, by vehicle and crash type

Vehicle type	Crash type	DRI	LBNL	Difference	Percent difference
Cars	Rollover	17,967	21,424	3,457	19.2%
	Fixed object	213,292	215,567	2,275	1.1%
	Ped etc.	42,685	41,049	-1,636	-3.8%
	HDT	65,509	88,528	23,019	35.1%
	Lgt car	389,037	397,003	7,966	2.0%
	Hvy car	437,351	341,056	-96,295	-22.0%
	Lgt LT	191,132	299,889	108,757	56.9%
	Hvy LT	188,499	214,763	26,264	13.9%
	Other	755,663	770,261	14,598	1.9%
	Total	2,301,135	2,389,540	88,405	3.8%
Light trucks	Rollover	33,518	37,837	4,319	12.9%
	Fixed object	148,346	144,124	-4,222	-2.8%
	Ped etc.	23,774	22,217	-1,557	-6.5%
	HDT	36,459	48,587	12,128	33.3%
	Lgt car	232,596	233,226	630	0.3%
	Hvy car	265,810	202,715	-63,095	-23.7%
	Lgt LT	114,601	177,621	63,020	55.0%
	Hvy LT	114,056	127,443	13,387	11.7%
	Other	447,551	458,148	10,597	2.4%
	Total	1,416,711	1,451,918	35,207	2.5%
CUVs/ minivans	Rollover	5,048	5,744	696	13.8%
	Fixed object	37,270	37,024	-246	-0.7%
	Ped etc.	11,453	11,003	-450	-3.9%
	HDT	15,330	21,352	6,022	39.3%
	Lgt car	104,989	107,891	2,902	2.8%
	Hvy car	122,373	93,400	-28,973	-23.7%
	Lgt LT	50,933	82,329	31,396	61.6%
	Hvy LT	51,772	59,914	8,142	15.7%
	Other	211,861	216,374	4,513	2.1%
	Total	611,029	635,031	24,002	3.9%

3. Comparison with results in DRI's Phase II report and LBNL sensitivities

DRI decimated the state police-reported crash data "due to memory limitations in the Fortran software." To accomplish this DRI randomly sampled 15 police-reported non-fatal crashes for each fatal FARS crash in the 10 states, for each combination of state, calendar year, and crash type. (For CUVs/minivans, DRI sampled 50 police-reported non-fatal crashes for each fatal FARS crash.) This resulted in a random sample of 9%, 12%, and 24% of cars, light trucks, and CUVs/minivans, respectively, involved in non-fatal crashes in the state databases, with the same distribution by state, calendar year, and crash type as the number of FARS cases. While this approach attempts to correct for differences in reporting thresholds, and reporting bias, across states, it may result in biased regression estimates if the samples are not randomly selected from all available vehicles.

In order to test the sensitivity of the regression estimates to which vehicles happened to be randomly sampled, LBNL replicated DRI's approach and sampled a similar number of vehicles. However, since LBNL could not replicate DRI's classification of vehicles by crash type, the total number of vehicles LBNL sampled differs somewhat from the number DRI sampled. Table 4 shows the number of cases DRI used in their 2012 Phase II report, the target sample based on DRI's method and LBNL's counts of vehicles, and the actual cases LBNL used in their first sensitivity analysis. Table 4 also shows the number of cases LBNL used in a second sensitivity, which it sampled a second set of vehicles. To the extent possible, the vehicles included in the alternative LBNL decimation case were different from those used in the initial decimation case. If the actual number of non-fatal police-reported crashes for a given state, calendar year, and crash type was more than twice the target number desired, there is no overlap in the vehicles actually sampled under the two cases. However, where the target sample exceeded the number available, the vehicles sampled in each case are exactly the same.

Table 4. Number of vehicles involved in non-fatal crashes used in regression analyses

Vehicle type	Actual cases DRI used	Target sample based on DRI method, LBNL counts	Actual cases LBNL used (model 1)	LBNL model 2: Alternative decimation
Car	201,735	203,799	206,005	204,018
Light truck	172,384	172,005	172,847	172,467
CUV/minivan	146,358	152,576	153,111	152,676
Total	520,477	528,380	531,963	529,161

If possible, sampling should be avoided, as it may introduce bias in the estimates and reduce the explanatory power of the regression model. In addition, DRI's decimation method assumes that fatality rates (i.e. number of fatalities per crash) are constant across crash types. Some types of crashes, especially rollovers and crashes with pedestrians/cycles/motorcycles, have higher societal fatality rates per crash than others. This approach also does not allow for differences in fatality rates per crash among states that may not be accounted for in the other control variables NHTSA used (driving at night, on high-speed roads, in rural areas, and in high-fatality states).

On the other hand, when using all of the police-reported crash data one must account for differences in reporting thresholds, and reporting bias, among the states. A better approach than DRI's decimation method may be to use all of the non-fatal police-reported crash data, while including internal control variables for each state to control for differences in reporting thresholds, and reporting bias, among states.

Table 5 shows the number of vehicles LBNL used in its three alternative models 3 through 5. Model 3 applies DRI's decimation and sampling method to all 13 states included in the NHTSA analysis (as mentioned above, DRI was not able to obtain police-reported crash data for AL, NJ, and WI). This provides a total of over 670,000 vehicles for analysis, a 25 percent increase from the sample DRI used from the 10 states. Model 4 uses all crash data from the 13 states, and includes an internal control variable for each of the states except Florida. This method increases the number of vehicles used by an order of magnitude, to over 6.4 million. In model 5 LBNL replicates the data that NHTSA used in its analysis, by using "synthetic" data for the four state/CY combinations where police-reported crash data were not available (PA in 2002, MI in 2002 and 2003, and WY in 2008). Following NHTSA's example, LBNL duplicated the crash data from the most recent year of available data (PA in 2003, MI in 2004, and WY in 2007). In addition, for model 5 LBNL excluded all vehicles whose reported model year in the state crash data did not match the model year from the decoded VIN, as NHTSA did in their analysis (all data from WA were used in model 5, as WA does not report the vehicle model year in its police-reported crash database). The combination of these two changes results in slightly higher number of vehicles used in LBNL model 5 than in model 4, as shown in Table 5.

Table 5. Number of vehicles involved in non-fatal crashes used in LBNL regression analyses

Vehicle type	LBNL model 3: Decimation using 13 states	LBNL model 4: All crash data from 13 states	LBNL model 5: All crash data from 13 states, duplicate data for missing states/CY, only vehicles whose reported MY matches VIN
Car	262,628	3,480,677	3,485,912
Light truck	216,662	2,020,045	2,043,807
CUV/minivan	191,104	925,558	934,785
Total	670,394	6,426,280	6,464,504

Tables 6 through 8 show compare the estimated effect of mass or footprint reductions on U.S. societal fatality risk per crash, crash frequency per VMT, and societal fatality risk per VMT, respectively, from the NHTSA 2012 and DRI 2012 reports with LBNL's five alternate regression models. The DRI estimates for footprint reduction in Tables 6 through 8 are the average estimates for footprint reduction in lighter- and heavier-than-average cars and light trucks, weighted by the number of fatalities, as reported in DRI 2012.

Table 6 compares the DRI and LBNL 2-stage estimates on the effect on fatality risk per crash (crashworthiness and crash compatibility). LBNL model 1 predicts results similar to those for DRI only for mass reduction in heavier-than-average light trucks and footprint reductions in CUVs/minivans. LBNL model 1 predicts a higher, and statistically significant, increase in fatality risk per crash in heavier-than-average cars, but substantially higher decreases in fatality risk per crash in lighter-than-average light trucks and CUVs/minivans, than DRI. In addition LBNL model 3 estimates a smaller decrease than DRI in fatality risk per crash from footprint reduction in light trucks.

LBNL model 2 shows the sensitivity of the estimates to the particular vehicles that are included in the decimated sample. LBNL's estimates are nearly identical for mass reduction in heavier-than-average cars CUVs/minivans, and footprint reduction in CUVs/minivans; however, the alternate data decimation results in substantially different estimates for the effect of mass or footprint reduction in light trucks. In some cases estimates in LBNL model 2 are more similar to DRI's estimates than those in LBNL model 1. The comparison of LBNL models 1 and 2 indicates that the particular vehicles included in the decimation sample may bias the regression model estimates.

LBNL model 3 shows the sensitivity of the estimates to including the three additional states used in NHTSA's analysis (AL, NJ, and WI). Including the three additional states increases the estimated beneficial effect of mass reduction in lighter-than-average cars on fatality risk per crash, to a statistically-significant 0.55%. LBNL model 3 reduces the estimated beneficial effect of mass reduction on lighter-than-average light trucks and CUVs/minivans; these estimates are closer to DRI's estimates. In terms of footprint reduction, LBNL model 3 estimates a small (0.57%), but statistically-significant, increase in fatality risk per crash in cars, and a slightly smaller increase in fatality risk per crash in CUVs/minivans than in LBNL models 1 and 2 (1.49% vs. 1.65% and 1.60%).

LBNL model 4 shows the sensitivity of the estimates to including all of the state crash data, rather than a decimation sample. The regression model includes an internal control variable for each state, except Florida, to account for differences in reporting thresholds and bias across the states. Estimates from LBNL model 4 are comparable to those from model 3 only for mass reductions in lighter-than average cars and light trucks. LBNL model 4 estimates a substantially higher beneficial effect on risk per crash from mass reduction in heavier-than-average light trucks and CUVs/minivans, a smaller beneficial effect on risk per crash from footprint reduction in light trucks, and substantially higher detrimental effects on risk per crash from footprint reduction in cars and CUVs/minivans.

LBNL model 5 includes duplicated data for the four state/CY combinations for which data were not available, and excludes vehicles where the model year reported in the state crash data did not match the decoded VIN. In most cases, these changes result in only small differences from the estimates in model 4; however, model 5 now estimates a relatively large (1.12%), and statistically-significant, reduction in fatality risk per crash from mass reduction in heavier-than-average cars, a large increases in risk per crash from footprint reduction in cars (1.25%) and CUVs/minivans (2.57%).

Table 6. DRI 2-stage and alternate LBNL 2-stage estimates of the effect of mass or footprint reduction on U.S. societal fatality risk per crash (crashworthiness/compatibility)

Variable	Case vehicle type	A. NHTSA 1-stage estimate (fatalities per crash)	B. DRI 2-stage estimate (fatalities per crash)	C. Alternate LBNL 2-stage estimates (fatalities per crash)				
				1. 10 states, decimated crash data	2. 10 states, alternate decimated crash data	3.13 states, decimated crash data	4. 13 states, all crash data	5. 13 states, all crash data, duplicated data, where MY matches
Mass reduction	Cars < 3106 lbs	—	-0.51%	-0.35%	-0.29%	-0.55%	-0.53%	-0.65%
	Cars > 3106 lbs	—	0.42%	0.65%	0.65%	-0.16%	-0.51%	-1.12%
	LTs < 4594 lbs	—	-0.80%	-1.46%	-1.11%	-0.97%	-0.96%	-1.03%
	LTs > 4594 lbs	—	-1.17%	-1.15%	-1.01%	-1.11%	-1.56%	-1.59%
	CUV/minivan	—	-0.96%	-1.34%	-1.35%	-1.02%	-1.42%	-1.53%
Footprint reduction	Cars	—	0.20%	-0.45%	-0.48%	0.57%	0.95%	1.25%
	LTs	—	-1.81%	-1.57%	-1.84%	-1.63%	-1.46%	-1.37%
	CUV/minivan	—	1.64%	1.65%	1.60%	1.49%	2.41%	2.57%

Estimates that are statistically significant at the 95% level are shown in red.
* DRI estimates using decimated data for 10 states. Estimates for track width and wheelbase reductions in cars and light trucks are the weighted average DRI estimates by crash partner weight based on the number of fatalities.

Table 7 compares the DRI and LBNL 2-stage estimates on the effect of mass or footprint reduction on crash frequency (crashes per VMT). LBNL model 1 predicts results similar to those for DRI only for mass reduction in heavier-than-average light trucks. LBNL model 1 predicts smaller increases in crash frequency per VMT than DRI from mass reduction in lighter-than-average cars and from footprint reduction in light trucks, while predicting larger increases in crash frequency than DRI from mass reduction in lighter-than-average light trucks and CUVs/minivans, and especially from footprint reduction in cars.

LBNL model 2 shows that the estimates for crash frequency from mass or footprint reduction in light trucks are particularly sensitive to the vehicles that are included in the decimation sample. As shown in model 3, adding the three additional states increases the estimated effect of mass reduction on crash frequency in heavier-than-average cars to a statistically-significant 0.75%, but slightly reduces the increase in crash frequency from mass reduction in lighter-than-average light trucks and CUVs/minivans, and substantially reduces the increase in crash frequency from footprint reduction in cars.

LBNL model 4, which uses all of the 13-state crash data, has little effect on the estimates on crash frequency from mass reduction in cars, lighter-than-average light trucks, and CUVs/minivans, but substantially increases the estimate of crash frequency from mass reduction in heavier-than-average light trucks, and decreases the estimated increase in crash frequency from footprint reduction in cars and light trucks. Model 4 estimates a statistically-significant, but small, decrease in crash frequency from footprint reduction in CUVs/minivans.

LBNL model 5 estimates larger increases in crash frequency from mass reduction, particularly for heavier-than-average cars (from 0.77% to 1.41%), but smaller increases in crash frequency from footprint reduction in cars and light trucks, and a larger decrease in crash frequency from footprint reduction in CUVs/minivans.

Table 7. DRI 2-stage and alternate LBNL 2-stage estimates of the effect of mass or footprint reduction on U.S. crash frequency per VMT (crash avoidance)

Variable	Case vehicle type	A. NHTSA 1-stage estimate (crashes per VMT)	B. DRI 2-stage estimate (crashes per VMT)	C. Alternate LBNL 2-stage estimates (crashes per VMT)				
				1. 10 states, decimated crash data	2. 10 states, alternate decimated crash data	3.13 states, decimated crash data	4. 13 states, all crash data	5. 13 states, all crash data, duplicated data, where MY matches
Mass reduction	Cars < 3106 lbs	—	2.10%	1.83%	1.86%	1.95%	1.84%	1.96%
	Cars > 3106 lbs	—	0.16%	-0.10%	-0.05%	0.75%	0.77%	1.41%
	LTs < 4594 lbs	—	1.41%	1.81%	1.50%	1.30%	1.29%	1.41%
	LTs > 4594 lbs	—	0.83%	0.76%	0.67%	0.71%	1.04%	1.08%
	CUV/minivan	—	0.52%	0.86%	0.83%	0.62%	0.80%	0.89%
Footprint reduction	Cars	—	1.20%	2.26%	2.25%	1.27%	1.08%	0.81%
	LTs	—	1.59%	1.28%	1.45%	1.35%	1.20%	1.09%
	CUV/minivan	—	0.31%	0.15%	0.29%	0.25%	-0.44%	-0.58%

Estimates that are statistically significant at the 95% level are shown in red.
* DRI estimates using decimated data for 10 states. Estimates for track width and wheelbase reductions in cars and light trucks are the weighted average DRI estimates by crash partner weight based on the number of fatalities.

Table 8 compares the DRI and LBNL 2-stage estimates with the NHTSA 1-stage estimates on the effect of mass or footprint reduction on U.S. societal fatality risk per VMT. LBNL model 1 predicts results more similar to NHTSA's results than DRI's results for mass reduction in heavier-than-average cars, and footprint reduction in cars and CUVs/minivans. However, DRI's 2-stage results for fatality risk per VMT remain more similar to NHTSA's 1-stage results than LBNL's 2-stage results for mass reduction in lighter-than-average cars, light trucks, and CUVs/minivans, and for footprint reduction in light trucks. LBNL model 2 improves the agreement with NHTSA's 1-stage results for mass reduction in lighter-than-average cars and heavier-than-average light trucks. LBNL model 3 substantially reduces the estimated effect on fatality risk per VMT from mass reduction in lighter-than-average cars (to 1.40%) and from footprint reduction in CUV/minivans (to 1.74%).

Using all crash data reduces the estimated detrimental effect of mass reduction, while increasing the estimated detrimental effect of footprint reduction, on fatality risk per VMT for all vehicle types, as shown in LBNL model 4. LBNL model 5 estimates very similar effects of mass or footprint reduction on U.S. societal fatality risk per VMT as LBNL model 4.

Table 8. NHTSA 1-stage, DRI 2-stage, and alternate LBNL 2-stage estimates of the effect of mass or footprint reduction on U.S. societal fatality risk per VMT

Variable	Case vehicle type	A. NHTSA 1-stage estimate (fatalities per VMT)	B. DRI 2-stage estimate (fatalities per VMT)	C. Alternate LBNL 2-stage estimates (fatalities per VMT)				
				1. 10 states, decimated crash data	2. 10 states, alternate decimated crash data	3.13 states, decimated crash data	4. 13 states, all crash data	5. 13 states, all crash data, duplicated data, where MY matches
Mass reduction	Cars < 3106 lbs	1.56%	1.58%	1.48%	1.57%	1.40%	1.31%	1.30%
	Cars > 3106 lbs	0.52%	0.58%	0.55%	0.59%	0.59%	0.27%	0.29%
	LTs < 4594 lbs	0.52%	0.61%	0.36%	0.39%	0.33%	0.33%	0.38%
	LTs > 4594 lbs	-0.34%	-0.34%	-0.39%	-0.34%	-0.40%	-0.52%	-0.51%
	CUV/minivan	-0.37%	-0.45%	-0.48%	-0.52%	-0.40%	-0.62%	-0.64%
Footprint reduction	Cars	1.89%	1.40%	1.80%	1.78%	1.84%	2.03%	2.06%
	LTs	-0.07%	-0.23%	-0.29%	-0.38%	-0.28%	-0.25%	-0.28%
	CUV/minivan	1.73%	1.96%	1.80%	1.89%	1.74%	1.96%	1.99%

Estimates that are statistically significant at the 95% level are shown in red.
* DRI estimates using decimated data for 10 states. Estimates for track width and wheelbase reductions in cars and light trucks are the weighted average DRI estimates by crash partner weight based on the number of fatalities.

4. Comparison of results in DRI's Supplemental Report

In its 2012 Supplemental report (DRI 2012c) DRI points out that replacing footprint with two variables, wheelbase and track width, reduces the maximum VIF of the regression models from 6.2 for cars, 10.9 for light trucks, and 8.7 for CUVs/minivans to 5.6 for cars, 6.6 for light trucks, and 7.2 for CUVs/minivans (Tables 2 and 1). Although DRI argues that replacing footprint with wheelbase and track width reduces the multi-collinearity in the regression model, the maximum VIF is still greater than 5.

DRI argued that using non-culpable vehicles in two-vehicle crashes as a proxy for all vehicles travelling on road may understate the exposure or VMT of vehicle/driver combinations that could have avoided a two-vehicle crash. DRI proposed using stopped vehicles struck by another vehicle, rather than the vehicle determined not to be at fault, as a proxy for the distribution of vehicle/driver combinations on the road to be used for induced exposure. While this proposal makes sense, it must be recognized that using stopped rather than non-culpable vehicles dramatically reduces the number of vehicles in the police-reported crash data available for the induced exposure data: there are 2,242,871 non-culpable vehicles in the NHTSA dataset, while there are only 610,689 (73% fewer) stopped vehicles in the alternate dataset NHTSA developed on DRI's recommendation.

As shown in Table 9, there are slightly smaller fractions of crashes at night, in rural counties, on high speed roads, and with male, young, or old drivers in stopped vehicles than in non-culpable vehicles. This suggests that stopped vehicles are less influenced by these risky crash or driver characteristics than non-culpable vehicles. Note that there are much larger fractions of vehicles characterized by these risk factors in the entire database of police-reported crashes than in the two subsets used for the induced exposure. However, there are slightly higher fractions of crashes involving risky sporty or AWD cars using stopped rather than non-culpable vehicles. On balance, stopped vehicles, and their drivers, are slightly less risky than non-culpable vehicles/drivers, and therefore may better represent the actual population of on-road vehicles. However, using stopped rather than non-culpable vehicles substantially reduces the sample of vehicles available for the induced exposure dataset.

Table 9. Crash, driver, and vehicle characteristics of non-culpable and stopped vehicles used for induced exposure

Variable		Vehicles involved in two-vehicle crashes used for induced exposure: Non-culpable	Vehicles involved in two-vehicle crashes used for induced exposure: Stopped	Vehicles in all crashes
NITE		16.1%	13.5%	21.7%
RURAL		24.0%	22.6%	26.6%
SPDLIM55		17.0%	12.4%	24.1%
DRVMALE		46.5%	45.1%	55.4%
DRVAGE	14 to 30	30.3%	27.2%	37.0%
	30 to 50	42.8%	45.9%	40.7%
	50 to 70	22.6%	23.4%	21.1%
	70 to 96	4.3%	3.4%	5.2%
VEHTYPE	2-dr car	7.8%	7.5%	8.2%
	4-dr car	43.0%	39.9%	42.9%
	Sporty car	1.4%	1.6%	1.5%
	Police car	0.8%	0.7%	1.1%
	AWD car	1.4%	2.0%	1.3%
	Sm pickup	10.1%	10.2%	11.0%
	Lg pickup	2.5%	2.3%	3.1%
	SUV	15.9%	17.5%	15.3%
	CUV	8.4%	9.3%	7.1%
	Minivan	6.8%	6.8%	6.1%
	Full van	2.0%	2.3%	2.5%

Table 10 shows the sensitivity of the LBNL estimates to the combination of these two changes suggested by DRI, replacing footprint with wheelbase and track width, and using stopped rather than non-culpable vehicles as the basis for induced exposure. The table compares NHTSA's 1-stage fatality risk per VMT estimates with the DRI and LBNL 2-stage estimates for fatality risk per VMT, as well as for fatality risk per crash and crash frequency per VMT. The LBNL estimates are from model 5, which uses all crash data from 13 states (including an internal control variable for each state except Florida), as well as duplicate data for four state/CY combinations with missing data, while excluding vehicles where the model year reported in the state crash data did not match the decoded VIN; this dataset most closely resembles the methodology NHTSA used in its analyses.

Comparing columns A, D, and G, the three estimates for fatality risk per VMT, the DRI estimates more closely match the NHTSA estimates for mass reduction in cars and CUVs/minivans, track width reduction in CUVs/minivans, and wheelbase reduction in cars; all of these estimates are small, and none are statistically-significant. On the other hand, the LBNL estimates more closely match the NHTSA estimates for mass reduction in light trucks, track width reduction in cars and light trucks, and wheelbase reduction in light trucks and CUVs/minivans; five of these six estimates are statistically-significant.

In terms of fatalities per crash (comparing columns B and E), LBNL estimates a more beneficial effect than DRI from mass reduction in cars, heavier-than-average light trucks, and CUVs/minivans; from track width reduction in light trucks; and from wheelbase reduction in CUVs/minivans. In terms of crash frequency per VMT (comparing columns C and F), LBNL estimates a more beneficial effect than DRI from mass reduction only in lighter-than-average light trucks, from track width reduction only in CUVs/minivans, but from wheelbase reduction in all three types of vehicles. Note that LBNL estimates a fairly large (1.37%), statistically-significant reduction in crash frequency per VMT from track width reduction in CUVs/minivans.

Table 10. Comparison of DRI and LBNL estimated effect of mass, track width, or wheelbase reduction on U.S. societal fatality risk per crash, crash frequency per VMT, and fatality risk per VMT

			DRI 2-stage estimate*			LBNL 2-stage estimate**		
Variable	Case vehicle type	A. NHTSA 1-stage (fatalities per VMT)	B. Crashworthiness/ compatibility (fatalities per crash)	C. Crash avoidance (crashes per VMT)	D. Combined (fatalities per VMT)	E. Crashworthiness/ compatibility (fatalities per crash)	F. Crash avoidance (crashes per VMT)	G. Combined (fatalities per VMT)
Mass reduction	Cars < 3106 lbs	0.26%	-1.05%	1.42%	0.36%	-1.26%	1.71%	0.45%
	Cars > 3106 lbs	-0.90%	-0.36%	-0.12%	-0.48%	-2.19%	1.87%	-0.32%
	LTs < 4594 lbs	-0.10%	-0.70%	0.75%	0.06%	-0.54%	0.42%	-0.12%
	LTs > 4594 lbs	-0.97%	-1.34%	0.50%	-0.84%	-1.68%	0.83%	-0.85%
	CUV/minivan	-0.14%	-0.61%	0.39%	-0.22%	-1.39%	0.92%	-0.47%
Track width reduction	Cars	6.04%	1.34%	1.03%	2.37%	3.66%	1.11%	4.78%
	LTs	0.90%	-0.82%	1.23%	0.41%	-2.14%	2.81%	0.68%
	CUV/minivan	-0.55%	-0.44%	0.13%	-0.31%	0.33%	-1.37%	-1.04%
Wheelbase reduction	Cars	0.38%	-0.43%	0.81%	0.38%	0.30%	0.18%	0.47%
	LTs	-0.09%	-1.12%	0.77%	-0.35%	-0.44%	0.21%	-0.23%
	CUV/minivan	1.45%	2.20%	0.45%	2.65%	1.82%	0.04%	1.86%

Estimates that are statistically significant at the 95% level are shown in red.

* DRI estimates using decimated data for 10 states. Estimates for track width and wheelbase reductions in cars and light trucks are the weighted average DRI estimates by crash partner weight based on the number of fatalities.

** LBNL estimates using all non-fatal crash data for 13 states (with 12 internal control variables), duplicated data for missing state/CY combinations, and where the MY reported in state crash data matches the decoded VIN.

5. Conclusions

LBNL was not able to exactly replicate the results from DRI's simultaneous 2-stage regression model. This may be because of discrepancies in how DRI and LBNL classified the state police-reported crash data into crash types. LBNL's analysis of four alternate regression models suggests that the results from DRI's method are sensitive to changes in what data are used in the analysis, or even the particular vehicles included in the decimation sample; in some cases the sign of the estimated relationship from DRI's results changes under an alternate LBNL regression. However, for the most part LBNL's alternate regressions confirm the general results from DRI's simultaneous model, and LBNL's analysis in its Phase 2 report: that mass reduction is associated with an increase in crash frequency (crashes per VMT), but a decrease in fatality risk once a crash has occurred, across all vehicle types. Similarly, while LBNL was not able to replicate DRI's results from its simultaneous model using stopped rather than non-culpable vehicles as the induced exposure records, and replacing footprint with wheelbase and track width, LBNL's general results after these changes to the baseline model were comparable to DRI's, with mass reduction associated with an increase in crash frequency, and a decrease in fatalities per crash.

6. References

Dynamic Research, Inc. 2012a. *Updated Analysis of the Effects of Passenger Vehicle Size and Weight on Safety, Phase I: Updated Analysis Base on 1995 to 2000 Calendar Year Data for 1991 to 1999 Model Year Light Passenger Vehicles*. Report prepared for the International Council on Clean Transportation, the Energy Foundation, and American Honda Motor Company. January. DRI-TR-11-01.

Dynamic Research, Inc. 2012b. *Updated Analysis of the Effects of Passenger Vehicle Size and Weight on Safety, Phase II: Preliminary Analysis Based on 2002 to 2008 Calendar Year Data for 2000 to 2007 Model Year Light Passenger Vehicles*. Report prepared for the International Council on Clean Transportation, the Energy Foundation, and American Honda Motor Company. January. DRI-TR-12-01.

Dynamic Research, Inc. 2012c. *Updated Analysis of the Effects of Passenger Vehicle Size and Weight on Safety: Supplemental Results on the Sensitivity of the Estimates for 2002 to 2008 Calendar Year Data for 2000 to 2007 Model Year Light Passenger Vehicles to Induced-Exposure Data and Vehicle Size Variables*. Report prepared for the International Council on Clean Transportation, the Energy Foundation, and American Honda Motor Company. February. DRI-TR-12-03-01.

Kahane, C.J. 2012. *Relationships Between Fatality Risk, Mass, and Footprint in Model Year 2000-2007 Passenger Cars and LTVs*. Final report prepared for the National Center for Statistics and Analysis, National Highway Traffic Safety Administration, DOT HS 811 665, 2012; August.

Wenzel, Tom. 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.

Lawrence Berkeley National Laboratory. August. LBNL-5698E.

Wenzel, Tom. 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 for the Office of Energy Efficiency and Renewable Energy, US Department of Energy. Lawrence Berkeley National Laboratory. August. LBNL-5697E.

Appendix A. Comparison of DRI and LBNL Two-Stage Results

Table A-1. DRI and LBNL 2-stage estimated effect of mass or footprint reduction on US fatality risk per crash, crash frequency per VMT, and fatality risk per VMT

			DRI 2-stage estimate*			LBNL 2-stage estimates: 1. 10 states, decimated crash data			LBNL 2-stage estimates: 2. 10 states, alternate decimated crash data		
Variable	Case vehicle type	A. NHTSA 1-stage (fatalities per VMT)	B. Crashworthiness/ compatibility (fatalities per crash)	C. Crash avoidance (crashes per VMT)	D. Combined (fatalities per VMT)	E. Crashworthiness/ compatibility (fatalities per crash)	F. Crash avoidance (crashes per VMT)	G. Combined (fatalities per VMT)	H. Crashworthiness/ compatibility (fatalities per crash)	I. Crash avoidance (crashes per VMT)	J. Combined (fatalities per VMT)
Mass reduction	Cars < 3106 lbs	1.56%	-0.51%	2.10%	1.58%	-0.35%	1.83%	1.48%	-0.29%	1.86%	1.57%
	Cars > 3106 lbs	0.52%	0.42%	0.16%	0.58%	0.65%	-0.10%	0.55%	0.65%	-0.05%	0.59%
	LTs < 4594 lbs	0.52%	-0.80%	1.41%	0.61%	-1.46%	1.81%	0.36%	-1.11%	1.50%	0.39%
	LTs > 4594 lbs	-0.34%	-1.17%	0.83%	-0.34%	-1.15%	0.76%	-0.39%	-1.01%	0.67%	-0.34%
	CUV/minivan	-0.37%	-0.96%	0.52%	-0.45%	-1.34%	0.86%	-0.48%	-1.35%	0.83%	-0.52%
Footprint reduction	Cars	1.89%	0.20%	1.20%	1.40%	-0.45%	2.26%	1.80%	-0.48%	2.25%	1.78%
	LTs	-0.07%	-1.81%	1.59%	-0.23%	-1.57%	1.28%	-0.29%	-1.84%	1.45%	-0.38%
	CUV/minivan	1.73%	1.64%	0.31%	1.96%	1.65%	0.15%	1.80%	1.60%	0.29%	1.89%

Estimates that are statistically significant at the 95% level are shown in red.
* DRI estimates using decimated data for 10 states. Estimates for track width and wheelbase reductions in cars and light trucks are the weighted average DRI estimates by crash partner weight based on the number of fatalities.

Table A-2. LBNL 2-stage estimated effect of mass or footprint reduction on US fatality risk per crash, crash frequency per VMT, and fatality risk per VMT, using 13 NHTSA states

			LBNL 2-stage estimates: 3. 13 states, decimated crash data			LBNL 2-stage estimates: 4. 13 states, all crash data			LBNL 2-stage estimates: 5. 13 states, all crash data, duplicated data, where MY matches		
Variable	Case vehicle type	A. NHTSA 1-stage (fatalities per VMT)	B. Crashworthiness/ compatibility (fatalities per crash)	C. Crash avoidance (crashes per VMT)	D. Combined (fatalities per VMT)	E. Crashworthiness/ compatibility (fatalities per crash)	F. Crash avoidance (crashes per VMT)	G. Combined (fatalities per VMT)	H. Crashworthiness/ compatibility (fatalities per crash)	I. Crash avoidance (crashes per VMT)	J. Combined (fatalities per VMT)
Mass reduction	Cars < 3106 lbs	1.56%	-0.55%	1.95%	1.40%	-0.53%	1.84%	1.31%	-0.65%	1.96%	1.30%
	Cars > 3106 lbs	0.52%	-0.16%	0.75%	0.59%	-0.51%	0.77%	0.27%	-1.12%	1.41%	0.29%
	LTs < 4594 lbs	0.52%	-0.97%	1.30%	0.33%	-0.96%	1.29%	0.33%	-1.03%	1.41%	0.38%
	LTs > 4594 lbs	-0.34%	-1.11%	0.71%	-0.40%	-1.56%	1.04%	-0.52%	-1.59%	1.08%	-0.51%
	CUV/minivan	-0.37%	-1.02%	0.62%	-0.40%	-1.42%	0.80%	-0.62%	-1.53%	0.89%	-0.64%
Footprint reduction	Cars	1.89%	0.57%	1.27%	1.84%	0.95%	1.08%	2.03%	1.25%	0.81%	2.06%
	LTs	-0.07%	-1.63%	1.35%	-0.28%	-1.46%	1.20%	-0.25%	-1.37%	1.09%	-0.28%
	CUV/minivan	1.73%	1.49%	0.25%	1.74%	2.41%	-0.44%	1.96%	2.57%	-0.58%	1.99%

Estimates that are statistically significant at the 95% level are shown in red.
* DRI estimates using decimated data for 10 states. Estimates for track width and wheelbase reductions in cars and light trucks are the weighted average DRI estimates by crash partner weight based on the number of fatalities.