LBNL-xxxxx

# Analysis of Casualty Risk per Police-Reported Crash for Model Year 2000 to 2004 Vehicles, using Crash Data from Five States

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

Prepared by

Tom Wenzel
Energy Analysis Department
Environmental Energy Technologies Division
Lawrence Berkeley National Laboratory
Berkeley, CA 94720

March 2011

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

DISCLAIMER

This document was prepared as an account of work sponsored by the United States Government. While this document is believed to contain correct information, neither the United States Government nor any agency thereof, nor The Regents of the University of California, nor any of their employees, makes any warranty, express or implied, or assumes any legal responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by its trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof, or The Regents of the University of California. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof, or The Regents of the University of California.

Ernest Orlando Lawrence Berkeley National Laboratory is an equal opportunity employer.

## Table of Contents

Executive Summary ..... i
1. Introduction ..... 1
2. VIN correction ..... 2
3. Comparison of two measures of risk ..... 6
4. Potential bias using state crash data ..... 30
5. Accounting for driver behavior ..... 37
6. Accounting for crash location (urban vs. rural counties) ..... 42
7. Assessing general quality of vehicle design ..... 45
8. Conclusions ..... 49
9. References ..... 51
Appendix ..... 52

## List of Tables and Figures

Table 2.1. VIN position 10 translations ..... 3
Table 2.2. VIN position 9 shifts ..... 4
Table 2.3. VIN corrections after checking against database of valid, 17-character VINs ..... 4
Table 2.4. Result of all corrections made to VINs ..... 5
Figure 3.1. Five state fatality and casualty risks, per vehicle registration-year and per crash, by vehicle type ..... 7
Figure 3.2. Five state fatality and casualty risks per vehicle registration-year, by vehicle type ..... 8
Figure 3.3. Fatality risks per million vehicle registration-years, by state and vehicle type ..... 8
Figure 3.4. Five state casualty risks, per vehicle registration-year and per crash, by vehicle type 9
Figure 3.5. Five state fatality risks, per vehicle registration-year and per crash, by vehicle type 10
Figure 3.6. Casualty risks per 100,000 vehicle registration-years, by state and vehicle type ..... 10
Figure 3.7. Casualty risks per 100,000 vehicle registration-years, with Pennsylvania on separate scale, by state and vehicle type ..... 11
Figure 3.8. Pennsylvania casualty risks per 100,000 vehicle registration-years, including and excluding moderate injuries, by vehicle type ..... 12
Figure 3.9. Casualty risks per 10,000 crashes, by state and vehicle type ..... 13
Figure 3.10. Casualty risks per 10,000 crashes, indexed to all vehicles in each state, by state and vehicle type ..... 13
Figure 3.11. Five state casualty risk per vehicle registration-year and per crash, by vehicle make and model ..... 14
Table 3.1. Correlations ($R^2$) between casualty risk per crash and per vehicle registration-year, by state and vehicle model ..... 14
Figure 4.7. Crashes per 1,000 vehicle registration-years, by vehicle type and state ..... 35
Figure 4.8. One-vehicle, non-injury crashes and all crashes per 10,000 vehicle registration-years, by vehicle type and state ..... 36
Figure 5.1. Casualty risk in five states, by driver age and gender ..... 37
Figure 5.2. Effect of driver age and gender on driver casualty risks by vehicle type ..... 38
Figure 5.3. Bad driver rating by driver age and gender and vehicle type, US fatal crash data (FARS) ..... 40

Figure 5.4. Bad driver rating by driver age and gender and vehicle type, US police-reported crash data (NASS GES) ....40
Figure 5.5. Casualty risk by vehicle type and driver behavior, NASS GES data ....41
Figure 5.6. Casualty risk by vehicle type and driver behavior, NASS GES data (adjusted scale)41
Figure 6.1. Casualty risk for drivers (excluding young males and elderly drivers), by population density decile and vehicle type ....42
Figure 6.2. Distribution of police-reported crashes, by population density decile and vehicle type ....43
Figure 6.3. Casualty risk per 10,000 crashes, by crash location and vehicle type ....43
Figure 6.4. Five state casualty risks per vehicle registration-year and per crash, in urban I/M areas and rural non-I/M areas, by vehicle type ....44
Figure 6.5. Five state crashes per 1,000 vehicle registration-year, in urban I/M areas and rural non-I/M areas, by vehicle type ....44
Figure 7.1. Relationship between fatality risk to drivers of model year 1997 to 2001 car models and retail price of a model year 1998 car in 2003 ....45
Figure 7.2. Casualty risk by car type and median household income, model year 2000 to 2004 vehicles in NASS GES through 2008 ....46
Figure 7.3. Casualty risk by vehicle model and median household income, selected model year 2000 to 2004 models in NASS GES through 2008 ....47
Figure 7.4. Casualty risk by vehicle model and median household income, selected model year 2000 to 2004 models in NASS GES through 2008 ....47
Figure 7.5. Percent bad drivers by vehicle type and median household income, model year 2000 to 2004 models in NASS GES through 2008 ....48
Figure A.1. Five state fatality and casualty risks to others per vehicle registration-year, by vehicle type ....52
Figure A.2. Five state casualty risks to others per vehicle registration-year and per-crash, by vehicle type ....53
Figure A.3. Five state fatality risks to others per vehicle registration-year and per crash, by vehicle type ....53
Figure A.4. Five state fatality and casualty combined risks per vehicle registration-year, by vehicle type ....54
Figure A.5. Five state casualty combined risks per vehicle registration-year and per-crash, by vehicle type ....54
Figure A.6. Five state fatality combined risks per vehicle registration-year and per crash, by vehicle type ....55

## Executive Summary

In this report we compare two measures of driver risks: fatality risk per vehicle registration-year, and casualty (fatality plus serious injury) risk per police-reported crash. Our analysis is based on three sets of data from five states (Florida, Illinois, Maryland, Missouri, and Pennsylvania): data on all police-reported crashes involving model year 2000 to 2004 vehicles; 2005 county-level vehicle registration data by vehicle model year and make/model; and odometer readings from vehicle emission inspection and maintenance (I/M) programs conducted in urban areas of four of the five states (Florida does not have an I/M program). The two measures of risk could differ for three reasons: casualty risks are different from fatality risk; risks per vehicle registration-year are different from risks per crash; and risks estimated from national data are different from risks from the five states analyzed here. We also examined the effect of driver behavior, crash location, and general vehicle design on risk, as well as sources of potential bias in using the crash data from five states.

### VIN correction

The vehicle identification numbers (VINs) provided in the state crash databases are truncated at 10 to 12 characters; without the full 17-digit VIN we are not able to determine whether these truncated VINs represent legitimate combinations of numerals and characters. We used two pieces of information to determine whether each truncated VIN was valid: the model year reported in the databases independently of the VIN, and a lookup table of over 50 million unique legitimate 17-character VINs, from eight state inspection and maintenance programs throughout the country. After translating VIN position 10 to match the reported model year, and comparing the truncated VINs with the database of full valid VINs to correct VIN position 8 and 11, we increased the number of decodable VINs of model year 2000 to 2004 vehicles by 18%.

### Comparison of two measures of risk

We compared fatality risk per vehicle registration-year and casualty risk per police-reported crash, using the same data: all police reported crashes in five states. For the most part, the trend in casualty risk by vehicle type is quite similar to that of fatality risk, when vehicle registration-years are used as the measure of exposure; however, casualty risks are substantially lower than fatality risks for sports cars and for pickups.

The trend in casualty risk by vehicle type is similar regardless of whether vehicle registration years or police-reported crashes are used as the measure of exposure. Casualty risks for subcompact and compact cars are relatively lower per crash than per vehicle, while casualty risks for large and import luxury cars, minivans, large SUVs, and pickups are relatively higher per crash than per vehicle.

Casualty risks per vehicle registration-year are remarkably consistent across four of the five states; the exception is Pennsylvania, which has much lower casualty risks than the other states, because it reports a category of injuries, moderate injuries, between serious/incapacitating injuries and minor injuries. On the other hand, the five states have dramatically different fatality risks per vehicle registration-year. These differences could be attributable to small numbers of

fatalities in certain vehicle types and states, distance from trauma centers, or the quality of care that might prevent serious injuries from becoming fatalities.

Casualty risks per police-reported crash are substantially higher in Florida and Maryland than in the other three states, across all vehicle types, in part because of which crashes are required to be reported in each state. When casualty risk per crash for individual vehicle types is indexed to the risk for all vehicles in each state, the relative risks by vehicle type are quite similar in the five states. The exceptions are high risks of sports cars in Pennsylvania and ¾-ton pickups in Pennsylvania and Missouri, and low risks of midsize crossover SUVs in Pennsylvania.

The agreement between casualty risk per vehicle registration year and per crash, by vehicle model, is highest in Pennsylvania ($R^2$=0.72) and lowest in Maryland ($R^2$=0.51). Adjusting casualty risk per vehicle for annual mileage by vehicle model, using state I/M inspection records slightly lowers the correlation with casualty risk per crash in each state. The agreement in casualty risk for 103 popular vehicle models among the five states varies substantially; the correlation among the five states is stronger for casualty risk per vehicle registration-year ($R^2$ between 0.56 and 0.77, depending on state) than for casualty risk per crash ($R^2$ between 0.28 and 0.45, depending on state).

**Accounting for vehicle miles traveled**

We analyzed average odometer readings of model year 2002 vehicles from 2008 and 2009 I/M records in four of the five states, as well as four additional states (California, Colorado, Ohio, and Wisconsin). Average odometer readings are highest in Missouri (87,330) and lowest in Maryland (79,523), a difference of about 10%. When indexed to the average odometer reading for all vehicle types in each state, the average odometer by vehicle type is quite similar across the four states. There is more variability across the states in terms of average miles driven for sports cars, import luxury cars, large SUVs, midsize CSUVs, and pickups than for other vehicle types. Using Illinois as the baseline state, the correlation of average odometer readings for the 103 most prevalent vehicle models is highest for Pennsylvania ($R^2$=0.78) and lowest for Maryland ($R^2$=0.68).

California's I/M program includes vehicles in rural as well as urban counties; vehicles registered in urban California counties are driven 5% fewer miles than vehicles registered in rural counties. Almost all vehicle types are driven less in urban counties than rural counties; the exceptions are 1-ton pickups and fullsize vans. Since average odometer reading varies much more by vehicle type than by driving location, it is probably safe to assume that average mileage taken from vehicles registered in urban areas is representative for vehicles throughout a state.

For most vehicle types, adjusting for miles driven has little to no effect on casualty risk per vehicle registration-year; however, adjusting for miles driven substantially increases casualty risk for sports cars by 30%, which are driven many fewer miles than other vehicles, and slightly reduces casualty risk for fullsize vans and ¾-ton pickups.

**Crash involvement rates**

Crash involvement rates vary consistently by state, with Illinois and Missouri having relatively high crash rates (77 and 66 crashes, respectively, per 1,000 registered vehicle-years) and Florida and Pennsylvania having relatively low crash rates (30 per 1,000 registered vehicle-years). However, when crash frequencies for each vehicle type are indexed to the crash frequency for all vehicles in each state, the trends in indexed crash frequencies among vehicle types are similar across the five states. Some vehicle types have consistently higher crash involvement rates in each state (subcompact, compact cars), while others have consistently low crash rates (large cars, import luxury cars, minivans, large SUVs, midsize CSUVs, ½- and ¾-ton pickups). Sports cars do not have substantially higher crash frequency than other car models, which suggests that their high fatality and casualty risks are due to them being involved in more severe crashes, or to relatively poor crashworthiness compared to other cars. Using Missouri as the baseline state, the correlation ($R^2$) of crash rates for the 103 most prevalent vehicle models are between 0.81 and 0.85 in the four other states.

There is a strong correlation ($R^2$=0.91) between vehicle registration-years and vehicles involved in police-reported crashes for vehicle models, indicating that crashes, rather than registrations, can be safely used as a measure of exposure; the correlation is not improved when one accounts for mileage by vehicle model. However, there is no correlation between average odometer and crash frequency across vehicle types or models.

**Effect on casualty risk**

Calculating risk using all police-reported crashes, rather than vehicle registration-years, as the measure of exposure changes the risk of certain vehicle types relative to that of other vehicle types. For example, subcompact and compact cars have higher crash involvement rates than midsize, large and import luxury cars, so the difference in casualty risk between smaller and larger cars is dramatically reduced when the measure of exposure is changed from registration-years to all police-reported crashes. On the other hand, midsize and large cars, and minivans, have relatively low crash involvement rates, so changing the measure of exposure to crashes increases their casualty risk relative to that of other vehicle types. Similarly, the relatively low crash involvement rates of large SUVs, crossover SUVs, and fullsize pickups results in relatively higher casualty risks when the measure of exposure is changed from vehicle registration-years to all police-reported crashes.

In general, changing the measure of exposure from registration years to all police-reported crashes reduces the sharp decline in casualty risk as car size increases. Casualty risk per crash does decrease as size increases, within each major vehicle type, but for cars and truck-based SUVs the decrease is not as large as when registration-years is used as the measure of exposure.

**Reporting bias by states**

We used NASS GES data to analyze whether casualty risks per crash from the five states, and 12 other states that provide crash data and VINS, are representative of national casualty risks per crash. For most vehicle types, casualty risks are highest in the five VIN states, followed by those

in the twelve other VIN states, and lowest in the remaining states. The twelve other VIN states are most representative of national casualty risk for all vehicle types except pickups; national casualty risks for pickups fall between those in the VIN states and those in the non-VIN states. Based on this analysis of GES data, we would expect that calculating casualty risk using data from all of the SDS VIN states will overstate national casualty risk, for all vehicle types. However, national GES casualty risks are remarkably similar to the casualty risks from all police-reported crashes in the five states, for all vehicle types except pickup trucks, which have substantially lower casualty risks in the national GES data than in the police-reported crash data from the five states.

Another type of reporting bias is the criteria each state requires for a crash to be reported to police, and included in the dataset provided to the NHTSA SDS. As mentioned above, Florida has the highest casualty risk per crash of the five states; indexing the risk by vehicle type to the risk for all vehicles in each state is one method to eliminate the bias. In its 2003 report, NHTSA speculated that the high fatality rate in Florida (from FARS) was in part caused by a large number of high-risk young drivers in that state. However, Florida has an average, if not slightly low, fraction of young male drivers (and a slightly high fraction of elderly drivers) involved in all police-reported crashes, relative to the other states.

**Reporting bias by drivers**

Analysts suspect that one-vehicle, non-rollover, non-injury crashes by pickup trucks are under-reported in state crash databases. If pickup truck owners were not reporting one-vehicle non-injury crashes, we would expect the crash rate for pickup trucks in one-vehicle non-injury crashes to be lower relative to that of other vehicles than in all crashes. However, the data indicate that pickup trucks have <u>higher</u> crash rates, relative to those of other vehicle types, in the types of crashes least likely to be reported (one-vehicle non-injury crashes) than in all crashes combined. This suggests that pickup truck owners are not under-reporting one-vehicle non-injury crashes in the state crash databases.

**Accounting for driver age and gender**

Elderly drivers (over 65 years old) have higher casualty risk than other drivers, in virtually all vehicle types. For most vehicle types, young drivers (under 25 years old) have the same casualty risk as middle age drivers, and all drivers. Apparently calculating risk per police-reported crash accounts for much of the effect of poor driving behavior of young drivers that is observed in calculating risk per vehicle or vehicle-mile.

**Accounting for bad driving behavior**

Driver age and gender is a crude measure of driving behavior; a more accurate measure would be the circumstances of the current crash, and past driving record. We analyzed NHTSA's "bad driver rating" variable (based on whether alcohol, drugs, or reckless driving was involved in the current crash, as well as the driver's driving record over the last three years) by vehicle type, and found that young males have the highest, and elderly drivers the lowest, bad driver rating, across all vehicle types, using FARS data for model years 2003 through 2007. Young females have

much lower bad driver ratings than young males; this suggests that the higher casualty risks for women compared to men are the result of their relative frailty rather than risky driving behaviors on their part. Similarly, elderly drivers are very safe drivers, but they face high casualty risks because of their frailty.

The casualty risks for bad drivers are substantially higher than those for not bad drivers, for all vehicle types; this suggests that bad drivers are involved in a greater number of serious crashes, perhaps involving higher speeds, than not bad drivers. Bad drivers account for a greater increase in casualty risk or certain vehicle types (import luxury cars, SUVs, crossover SUVs, and pickups) than they do for most types of cars (including sports cars), minivans and fullsize vans.

**Accounting for crash location**

We used the population density of the county in which a crash occurred to assess the relationship between casualty risk and crash location. Casualty risk decreases as county population density increases, with the highest casualty risks in the most rural counties, and the lowest risks in the most urban counties, for all vehicles types. However, because a large fraction of pickup trucks are driven in rural areas (22%, as opposed to 10% for other vehicle types), accounting for the population density of the county in which a crash occurs reduces the casualty risk in pickup trucks by about 15% relative to that of other vehicle types. Because crash frequency in rural counties is four times higher than in urban counties, the effect of driving in rural counties on casualty risk per crash is smaller (twice as high as in urban counties) than the effect on casualty risks per vehicle registration-year (ten times as high in rural counties as in urban counties).

**Accounting for general quality of vehicle design**

Using NASS GES, which reports the zip code on the driver's license, there clearly is a relationship between driver income (the median household income of the driver's zip code) and casualty risk per crash for cars, with casualty risk decreasing as driver income increases for each car type. In addition, risk consistently decreases as income increases for five of the ten most prevalent vehicle models; however, there is no similar relationship for the other five prevalent vehicle models. Therefore it is not clear how strong the relationship between driver income and risk is when one accounts for vehicle model. The decrease in casualty risk as income increases is not the result of better driving behavior by drivers with higher incomes, as the percent of bad drivers does not increase with increasing income, with the possible exception of drivers with the highest incomes.

## 1. Introduction

For its updated analysis of the relationship between vehicle weight/footprint and safety, NHTSA and LBNL plan to use the same database of vehicle attributes by vehicle model year, make, model, and trim. The attributes to be analyzed include curb weight, footprint (wheelbase times track width), other vehicle measurements thought to influence safety (door sill height, bumper height, average height of force, etc.), and other attributes that vary by vehicle model (presence of electronic stability controls, side airbags, primary or secondary energy absorbing system, etc.).

However, NHTSA intends to replicate its 2003 analysis, which estimates fatality risk per registered vehicle-mile (Kahane, 2003). Fatalities are taken from FARS, vehicle registrations from RL Polk Company, and annual vehicle mileage from Carfax, or state vehicle emission inspection and maintenance (I/M) test results, or GES. NHTSA intends to use its previous approach to derive non-fatal crashes for its statistical analysis: using non-culpable vehicles involved in all police-reported crashes from about 15 states, to allocate national vehicle registration years to combinations of crashes, drivers and vehicles. The crash data will come from 15 states that provide vehicle identification number (VIN) in the data they provide to NHTSA's State Data System.

LBNL intends to analyze casualty (fatality and serious/incapacitating injury) risk per police-reported crash, using all crashes from the same 15 states. This is an extension of LBNL's earlier analysis of casualty risk using crash data from five states (Wenzel, 2010), by obtaining data through 2007 to include model years 2005 through 2007, and by obtaining data for additional states that report VIN. Because the two agencies are planning to use different data and methodologies to assess the influence of vehicle weight and footprint on safety, it is important to understand how the different approaches compare.

This report summarizes an analysis of the differences in the two approaches to estimating risk, and examines several other aspects of casualty risk per police-reported crash. Section 2 provides details on the corrections made to the truncated VINs in the police-reported crash data to maximize the number of vehicle observations that can be used in the analysis. Section 3 analyzes in detail the differences in the two measures of risk, fatality risk per vehicle registration-year and casualty risk per crash. Section 4 examines two sources of potential bias in the state crash data that could affect the estimates of casualty risk. Sections 5 and 6 discuss how accounting for driver behavior and crash location, respectively, affects casualty risk per crash. Finally, Section 7 examines the extent to which the general quality of vehicle design, expressed as vehicle resale value, as well as driver income independent of vehicle design, explains casualty risk.

## 2. VIN correction

There are over 28 million vehicle records from the five states, including medium- and heavy-duty vehicles and motorcycles; VIN decoding reduces the number of records to 17 million light-duty vehicles. The vehicle model year is missing from 1.6 million of these records (9%), mostly from Illinois (22% of all Illinois crashes) and Pennsylvania (9% of all Pennsylvania crashes). Of the vehicles with reported model years between 1981 and 2006, the reported model year matches that in the decoded VIN for 90%.

Because a VIN is a 17-character combination of numerals and letters, VINs often are entered incorrectly in databases, with characters omitted, entered incorrectly, or transposed. Position 9 of the VIN is a check digit that allows the determination if a full 17-digit VIN is a legitimate combination of numerals and letters. However, the VINs provided in the state crash databases are truncated at 10 to 12 characters; without the full 17-digit VIN we are not able to determine, via the check digit, whether these truncated VINs represent legitimate combinations of numerals and characters. The VIN decoder developed by NHTSA corrects the characters in VIN positions 1 through 7, based on common transcription errors. However, the decoder does not make changes to VIN positions 8 through 11.

We used two pieces of information to determine whether each truncated VIN was valid: the model year reported in the databases independently of the VIN, and a lookup table of over 50 million unique legitimate 17-character VINs, from eight state inspection and maintenance programs throughout the country (including four of the five states that provided crash data: IL, MD, MO, and PA). The VIN correction process described here was used for the dataset used in all analysis discussed elsewhere in this report.

In the first step we corrected a subset of VINs so that the character in position 10 matched the reported model year. First we translated VIN position 10 for characters that are commonly entered incorrectly. For example, if the reported model year was 1981 and “8” was entered in VIN position 10, the “8” was translated into a “B”, which is the correct character for model year 1981 vehicles. As shown in Table 2.1, a total of nearly 320,000 translations were made; over 70% of the translations were for model year 1995 vehicles incorrectly identified as 2005 vehicles (i.e., the character in VIN position 10 was incorrectly entered as an “S” rather than as a “5”).

**Table 2.1. VIN position 10 translations**

| Reported model year | Original (incorrect) VIN character | Translated (correct) VIN character | Vehicles | Percent distribution |
|---|---|---|---|---|
| 1981 | 8 | B | 1,944 | 1% |
| 1986 | 6 | G | 31,378 | 10% |
| 1990 | 1 | L | 18,174 | 6% |
| 1995 | 5 | S | 230,731 | 72% |
| 2000 | 4 | Y | 31,574 | 10% |
| 2001 | L | 1 | 1,718 | 1% |
| 2004 | Y | 4 | 1,374 | 0% |
| 2005 | S | 5 | 1,564 | 0% |
| 2006 | G | 6 | 34 | 0% |
| Total | | | 318,491 | 100% |

We noticed that a large number of records with incorrect characters in VIN position 10, based on the reported model year, had the correct character for that model year entered in VIN position 9, the check digit. That is, the value for position 10 was erroneously entered in position 9 instead. It appears that this shifting occurred for 24% of all VINs with the incorrect value in position 10, and close to 30% for vehicles with VIN position 10 representing model years 2000 to 2004. Therefore we shifted the character entered into position 9 into position 10, as well as all subsequent characters in the VIN, and entered an X for position 9 (a legitimate value for the check digit). As shown in Table 2.2, a total of 410,000 shifts were made, 52,000 for model years 2000 to 2004.

The second step in VIN correction involved comparing the truncated VINs against a database of legitimate 17-character VINs, after applying the NHTSA VIN decoder (which made additional translations of VIN positions). Decoded truncated VINs from the state crash data were first compared against the first 11 digits of the full valid VINs; any truncated VINs that did not match the database of valid VINs were then compared against the first 10 digits of the valid VINs (that is, ignoring VIN position 11, the plant at which the vehicle was manufactured). We then examined the remaining truncated VINs that did not match any of the VINs in the valid VIN database, and represented more than 20 vehicles in the state crash data. We found two VINs that were legitimate yet were not included in the valid VIN database, both for 2001 PT Cruisers; we added those to the database of valid VINs.[1] Of the remaining unmatched VINs, many had transcription errors in VIN position 8, the vehicle engine code; for example, a 5 was entered instead of an S. After checking against the full-VIN database, we translated these VINs to correct them, so they would match a VIN in the full-VIN database. These corrections were made to 237,000 vehicles between model years 2000 and 2004, as shown in Table 2.3.

[1] The NHTSA decoder translates V7 from B to 8, even though B is the correct V7 for MY01 PT Cruisers. That is, the decoder properly decodes MY01 PT Cruisers, but the VIN will not appear in the valid VIN database.

**Table 2.2. VIN position 9 shifts**

| Reported model year | Vehicles | Percent distribution | Percent of all incorrect model year |
|---|---|---|---|
| 1981 | 2,783 | 0.7% | 12% |
| 1982 | 3,176 | 0.8% | 11% |
| 1983 | 4,252 | 1.0% | 24% |
| 1984 | 7,229 | 1.8% | 22% |
| 1985 | 11,990 | 2.9% | 25% |
| 1986 | 13,231 | 3.2% | 25% |
| 1987 | 16,176 | 3.9% | 23% |
| 1988 | 19,724 | 4.8% | 21% |
| 1989 | 23,761 | 5.8% | 28% |
| 1990 | 19,491 | 4.8% | 22% |
| 1991 | 23,311 | 5.7% | 30% |
| 1992 | 23,000 | 5.6% | 26% |
| 1993 | 29,203 | 7.1% | 29% |
| 1994 | 31,579 | 7.7% | 27% |
| 1995 | 25,490 | 6.2% | 20% |
| 1996 | 27,114 | 6.6% | 22% |
| 1997 | 25,402 | 6.2% | 23% |
| 1998 | 23,445 | 5.7% | 26% |
| 1999 | 24,928 | 6.1% | 29% |
| 2000 | 16,310 | 4.0% | 19% |
| 2001 | 14,175 | 3.5% | 30% |
| 2002 | 10,415 | 2.5% | 26% |
| 2003 | 7,305 | 1.8% | 29% |
| 2004 | 3,945 | 1.0% | 27% |
| 2005 | 1,932 | 0.5% | 27% |
| 2006 | 243 | 0.1% | 23% |
| 2007 | 5 | 0.0% | 18% |
| Total | 409,615 | 100% | 24% |

**Table 2.3. VIN corrections after checking against database of valid, 17-character VINs**

| Reported model year | Vehicles | Percent distribution |
|---|---|---|
| 2000 | 80,777 | 34% |
| 2001 | 62,373 | 26% |
| 2002 | 48,225 | 20% |
| 2003 | 29,234 | 12% |
| 2004 | 16,723 | 7% |
| Total | 237,332 | 100% |

The net result of these corrections is show in Table 2.4. The combination of VIN corrections increased the number of decodable VINs, and therefore the number of model year 2000 to 2004 vehicles for analysis, by 18%.

**Table 2.4. Result of all corrections made to VINs**

| | All MY00-04 vehicles (based on model year reported in state crash databases) | | Where reported model year matches VIN model year | | Where model year matches and truncated VIN matches full VIN database | |
|---|---|---|---|---|---|---|
| VIN correction | Vehicles | Percent increase | Vehicles | Percent increase | Vehicles | Percent increase |
| Uncorrected | 1,856,502 | | 1,722,584 | | 1,449,322 | |
| Translate V10 | 1,891,168 | 1.9% | 1,757,250 | 2.0% | 1,473,569 | 1.7% |
| Shift V9 | 1,943,318 | 2.8% | 1,809,400 | 3.0% | 1,483,454 | 0.7% |
| Translate V8 | 2,185,715 | 12.5% | 2,046,732 | 13.1% | 1,711,502 | 15.4% |

## 3. Comparison of two measures of risk

NHTSA intends to examine the relationship between vehicle weight and fatality risk per registered vehicle-mile, while LBNL plans to analyze the relationship between vehicle weight and casualty (fatality and serious/incapacitating injury) risk per police-reported crash. There are several reasons why the two approaches might result in different estimates of risk; we address each of these issues in this section:

- Fatality risks are different from casualty risks. A specific vehicle design may reduce fatalities, while increasing serious injuries, relative to other vehicle designs.
- Risks calculated per vehicle registration-year, or mileage-adjusted registration year, are different from risks calculated per police-reported crash.
- Risks in 5, or 15, states are different from national risks.
- Different threshholds individual states have for reporting crashes, and casualties, resulting in biased estimates of risk in some states.
- Drivers of certain vehicle types, such as pickup trucks, may tend to under-report crashes, which can bias risk per crash.

### 3.1. Fatality risks are different from casualty risks

LBNL purchased vehicle registration data from R.L. Polk for the five states for which we have police-reported crash data, in order to analyze risks per vehicle registration-years. Registration data for model year 2000 through 2004 vehicles as of 2005, by vehicle model year, make, model, and registration county, were purchased. We calculated fatality and casualty risks pre vehicle registration-year, using casualties from the crash data from each state. The vehicle registrations for each model year were multiplied by the number of years vehicles of that model year had been on the road by 2005; i.e., registrations of model year 2000 vehicles were multiplied by 5, registrations of model year 2001 vehicles were multiplied by 4, etc. This method does not account for different retirement rates by vehicle type and model, and could be improved by using actual registrations by model year in each calendar year. Because Pennsylvania did not provide crash data for 2002, we multiplied model year 2000 vehicle registrations in Pennsylvania by 4, model year 2001 registrations by 3, etc.

Figure 3.1 shows fatality and casualty risks by vehicle type, using data from the police-reported crashes in five states. Each type of risk is calculated using two measures of exposure: vehicle registration-years, as described above, using Polk registration data for each state, and all police-reported crashes, from the state data. The units for each of the four risks are different, as noted in the figure legend; risks per vehicle registration-year (in blue) are plotted on the left-hand axis, while risks per crash are plotted on the right-hand axis (in red). Fatality risks are represented by solid symbols and lines, while casualty risks are represented by open symbols and dashed lines.

Figure 3.1 allows comparison of risks based on fatalities as opposed to casualties, and whether using a different measure of exposure (police-reported crashes rather than vehicle registration years) has on relative risk by vehicle type. Figure 3.2 compares fatality and casualty risk, both per registered vehicle-years. Although the absolute risks under each measure are different, for the most part, the trend in casualty risk by vehicle type is quite similar to that for fatality risk;

however, casualty risks per 100,000 registration-years (open blue) are substantially lower than fatality risks per million registration-years (solid blue) for sports cars and for pickups.

Figure 3.2 indicates that large cars have roughly the same fatality risk as midsize cars, which is surprising. In Florida and Maryland fatality risk is 9% and 14% higher, respectively, in large cars than in midsize cars, while in Pennsylvania and Illinois fatality risk in large cars is 11% and 18% lower, respectively, than the risk in midsize cars. The fatality risks in large cars in Figure 3.2 may be overstated, as these vehicles tend to be driven by the elderly who are less likely to survive a serious crash.

Because fatality risks are based on data from only 5 states, risks for the least popular vehicle types (midsize CSUVs, 1-ton pickups, fullsize vans, small SUVs, and small CSUVs) are the least reliable due to the small number of fatalities in those vehicles (from 33 to 92). Figure 3.2 indicates that, with the exception of sports cars and pickups, the relationship of casualty risk per vehicle registration-year by vehicle type is quite similar to that of fatality risk.

As noted in Kahane 2003, there is a wide range in fatality risk per mileage-adjusted vehicle-years across states. Figure 3.3 compares fatality risk per million registered vehicle-years in each of the five states, by vehicle type. The figure indicates that fatality risks in Illinois and Maryland tend to be lower than in Florida, Missouri, and Pennsylvania. For the most part, the trends in risk by vehicle type across states are similar; notable exceptions are the consistently increasing risk in Maryland as pickup size increases. These anomalies may be explained by relatively small numbers of fatalities in a given vehicle type and state.

**Figure 3.1. Five state fatality and casualty risks, per vehicle registration-year and per crash, by vehicle type**

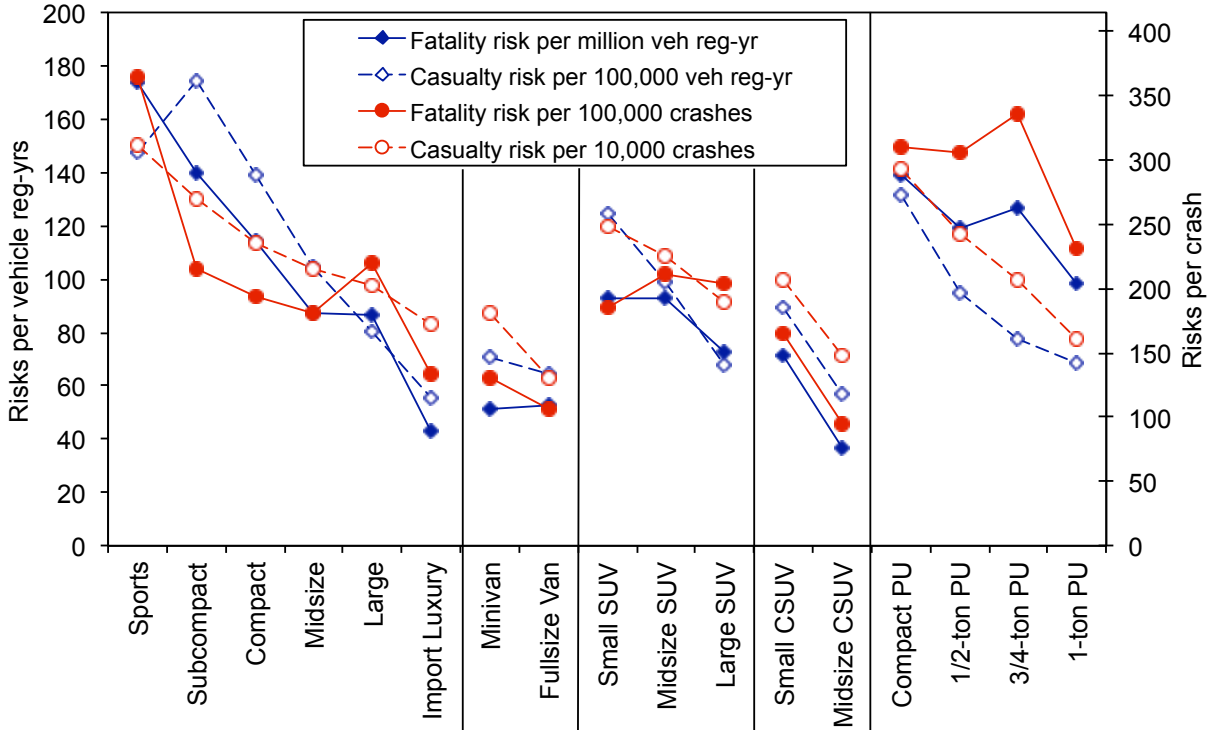


**Figure 3.2. Five state fatality and casualty risks per vehicle registration-year, by vehicle type**

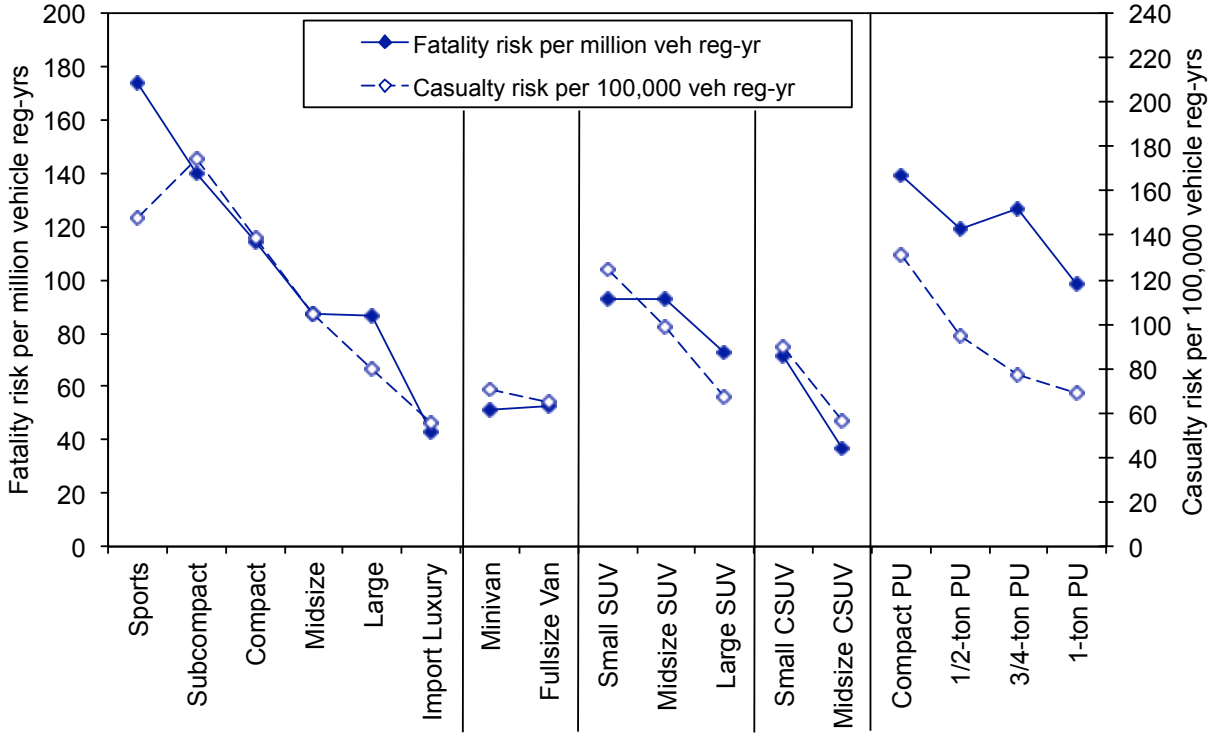
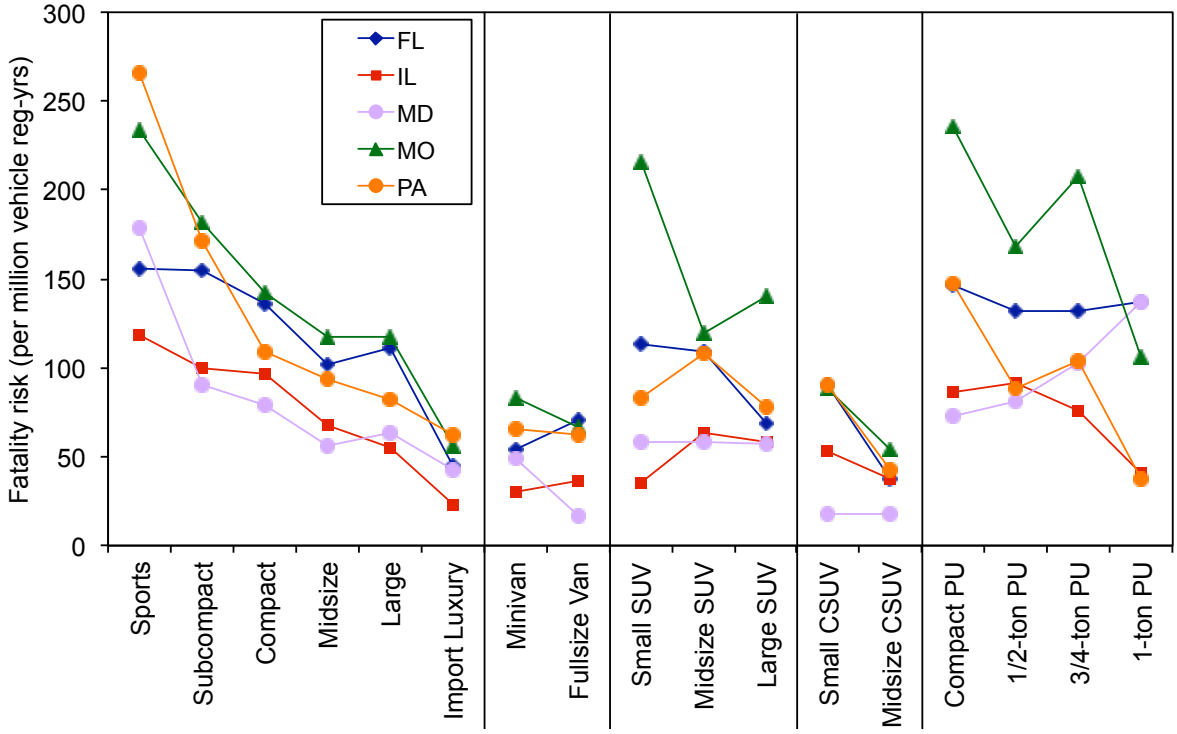



**Figure 3.3. Fatality risks per million vehicle registration-years, by state and vehicle type**



### 3.2. Risks per vehicle are different from risks per crash

Figure 3.4 plots the casualty risks using two different measures of exposure, vehicle registration years (open blue diamonds) and crashes (open red circles), from Figure 3.1; the right-hand scale is adjusted so that the two risks for midsize cars overlap. Here we see that casualty risk for subcompact and compact cars are relatively lower per crash than per vehicle, while casualty risks for large and import luxury cars, minivans, large SUVs, and pickups are relatively higher per crash than per vehicle. The trends in casualty risk by vehicle type are similar regardless of whether vehicle registration years or police-reported crashes are used as the measure of exposure.

Finally Figure 3.5 compares fatality risks per registration-year and per crash. Again the trends in fatality risk by vehicle type, when measured per registration-year and per crash, are quite similar.

Figure 3.6 compares casualty risks per 100,000 vehicle registration-years by vehicle type, across the five states. Here we see remarkable consistency across four of the five states, both in terms of absolute casualty risk levels and trends in risk by vehicle type. The exception is Pennsylvania, which has much lower casualty risks than the other states; for example, midsize cars have a fatality risk of only 31 in Pennsylvania, but between 105 (in Missouri) and 129 (in Florida) in the other states.

**Figure 3.4. Five state casualty risks, per vehicle registration-year and per crash, by vehicle type**

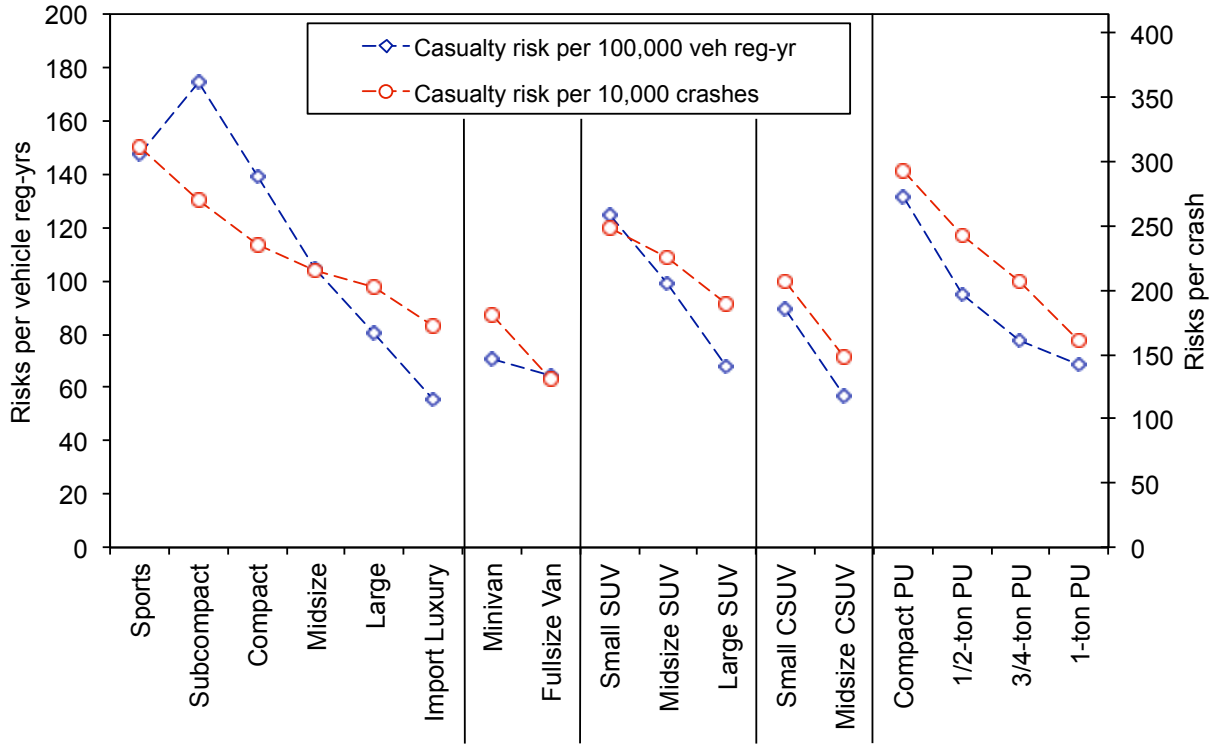


**Figure 3.5. Five state fatality risks, per vehicle registration-year and per crash, by vehicle type**

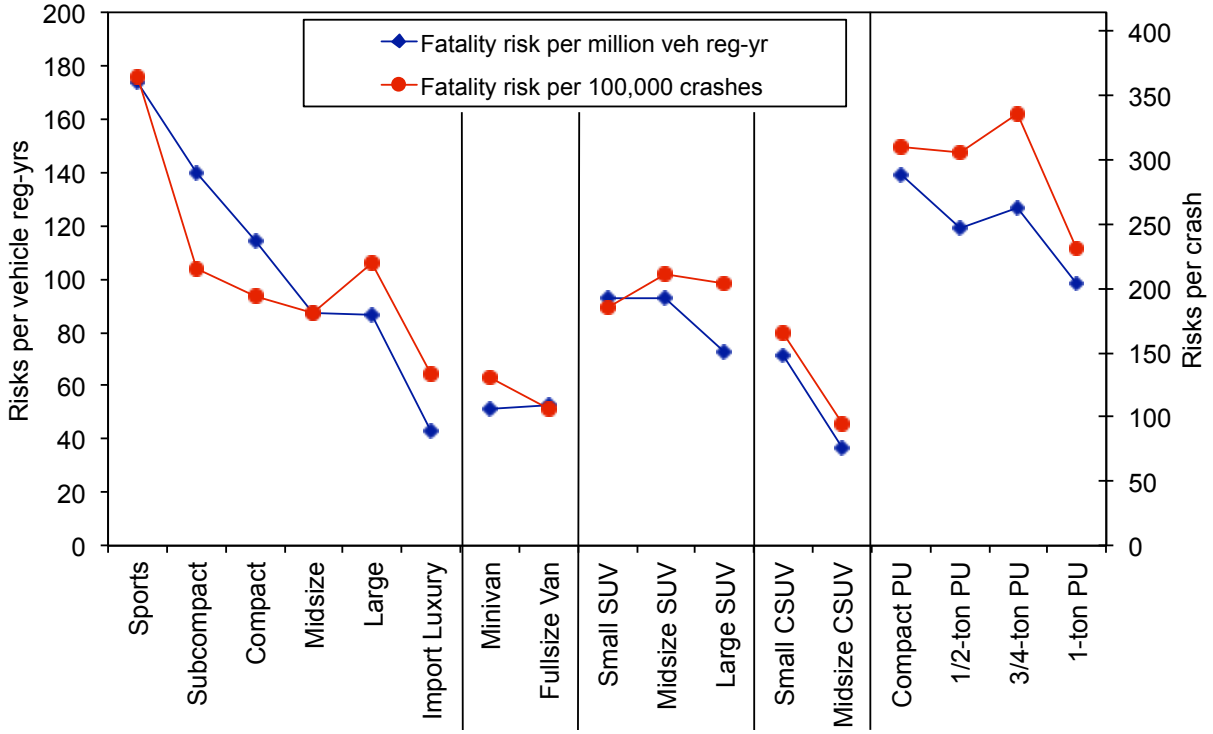
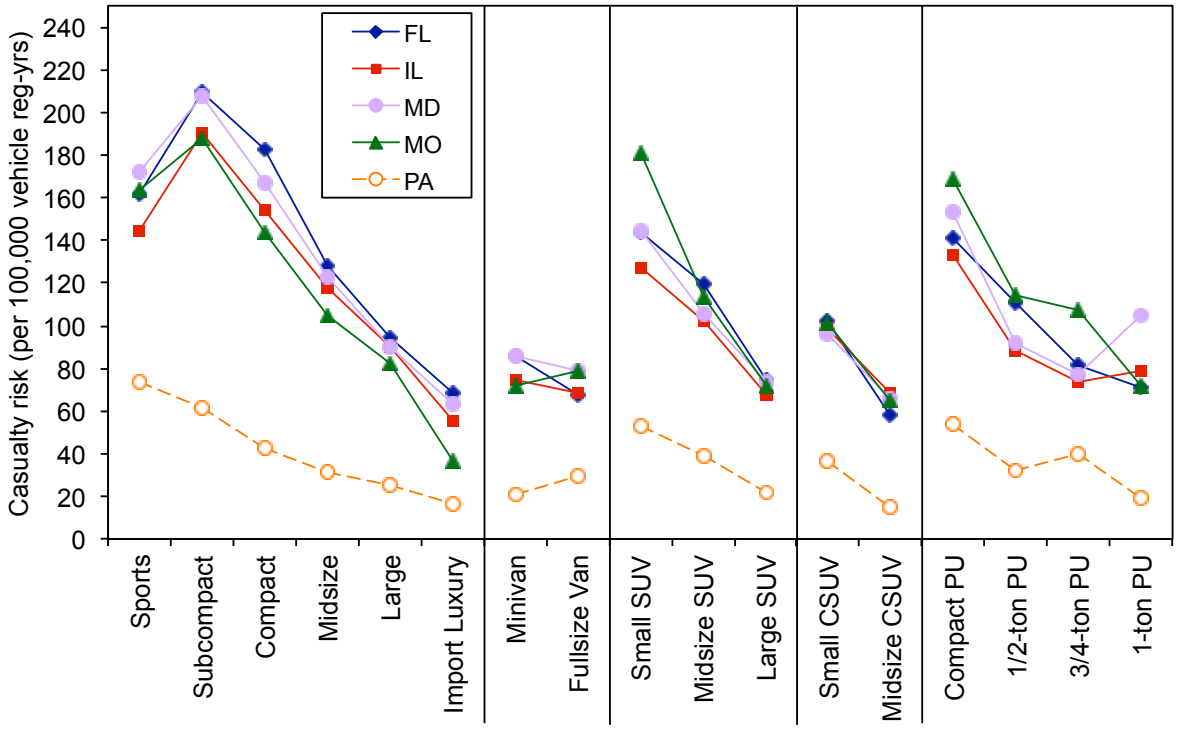



**Figure 3.6. Casualty risks per 100,000 vehicle registration-years, by state and vehicle type**



Pennsylvania is unique among the five states studied in that it reports a category of injuries, moderate injuries, between serious/incapacitating injuries and minor injuries. Figure 3.7 shows the same data as Figure 3.6, but plots Pennsylvania casualty risks excluding moderate injuries, using a different (right hand) scale. The figure demonstrates that, for the most part, the trends in casualty risk by vehicle type in Pennsylvania are quite similar to those in the other four states. The exceptions are that sports cars and ¾-ton pickups have higher casualty risks relative to other vehicles types in Pennsylvania than in the other states.

**Figure 3.7. Casualty risks per 100,000 vehicle registration-years, with Pennsylvania on separate scale, by state and vehicle type**

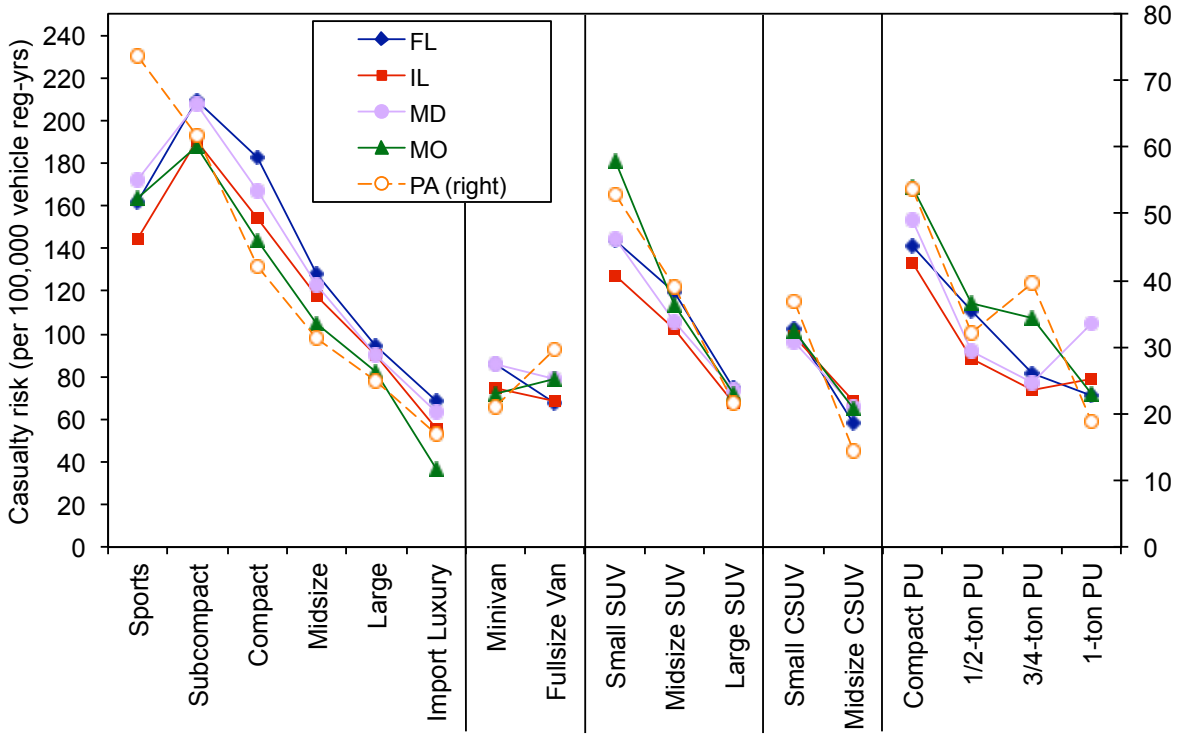
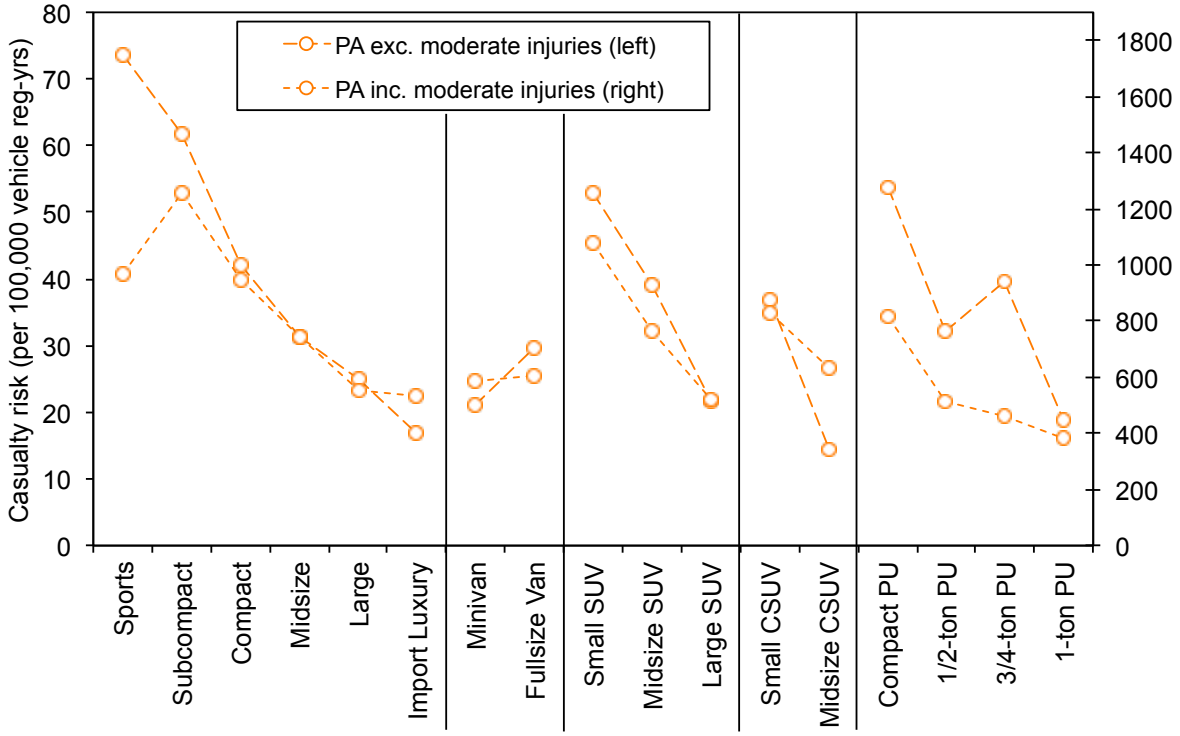



Figure 3.8 compares the trends in Pennsylvania casualty risks by vehicle type, when moderate injuries are included and excluded as casualties. Including moderate injuries as casualties results in much higher casualty risk in Pennsylvania than in the other four states. Figure 3.8 indicates that, when moderate injuries are included, the casualty risk for sports cars is substantially lower, and follows the trend of the other four states (that is, is lower than the casualty risk of subcompact cars). When moderate injuries are included as casualties, casualty risks in pickups are substantially lower than when moderate injuries are not included, and are more similar to the trend in the other four states (that is, casualty risk for ¾-ton pickups is comparable to that of ½-ton pickups).

**Figure 3.8. Pennsylvania casualty risks per 100,000 vehicle registration-years, including and excluding moderate injuries, by vehicle type**




Figures 3.6 and 3.7 suggest that casualty risks are quite similar across states, when vehicle registration-years are used as the measure of exposure. The small differences in casualty risks among the four states (excluding Pennsylvania) may be explained by differences in miles driven among vehicle types, or other factors.

On the other hand, Figure 3.3 suggests that differences among the states can result in dramatically different absolute fatality risks per vehicle registration-year. These differences could be attributable to distance from trauma centers, or the quality of care that might prevent serious injuries from becoming fatalities. Some of these differences could be attributable to small numbers of fatalities in certain vehicle types and states.

Figure 3.9 compares the casualty risks per 10,000 crashes by vehicle type among the five states. Casualty risks per crash are substantially higher in Florida and Maryland than in the other three states, across all vehicle types. Casualty risks are higher in Florida in part because non-injury crashes are under-reported in that state; this issue is addressed in further detail in Section 3, below.

Because the absolute casualty risk per crash is quite different in each state, Figure 3.10 indexes the risk for each vehicle type in each state by the risk for all vehicles in that state. The figure indicates that, for the most part, the relative risks by vehicle type are quite similar in the five states. The exceptions are high risks of sports cars in Pennsylvania and ¾-ton pickups in Pennsylvania and Missouri, and low risks of midsize crossover SUVs in Pennsylvania.

**Figure 3.9. Casualty risks per 10,000 crashes, by state and vehicle type**

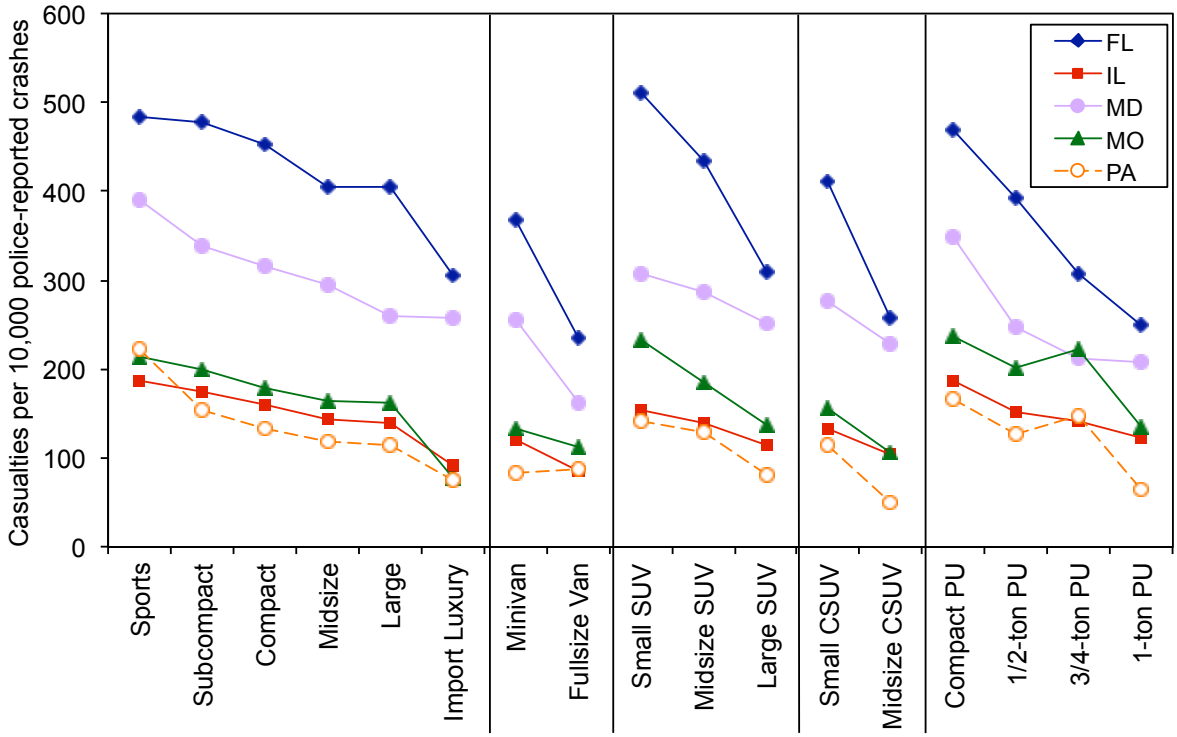



**Figure 3.10. Casualty risks per 10,000 crashes, indexed to all vehicles in each state, by state and vehicle type**

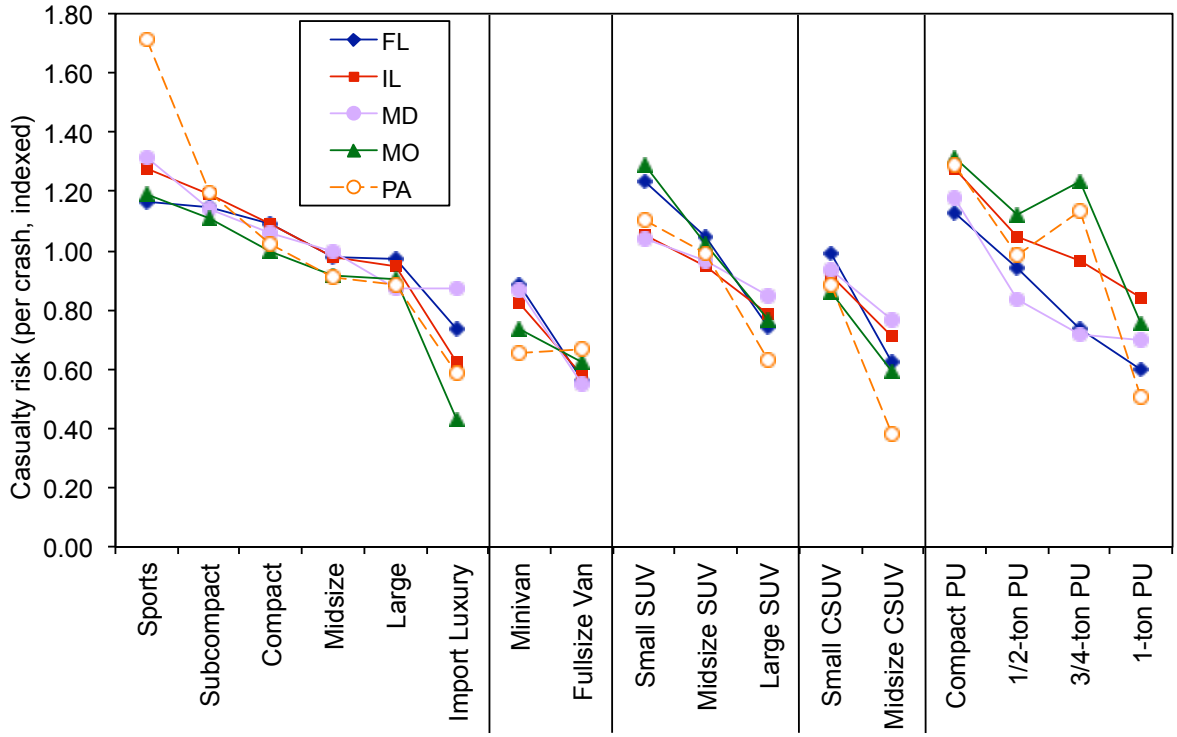


Figure 3.11 compares casualty risk per crash and per vehicle registration-years, by vehicle model, for all five states combined. The figure indicates that there is reasonable agreement between the two measures of risk, with an $R^2$ of 0.62.

**Figure 3.11. Five state casualty risk per vehicle registration-year and per crash, by vehicle make and model**

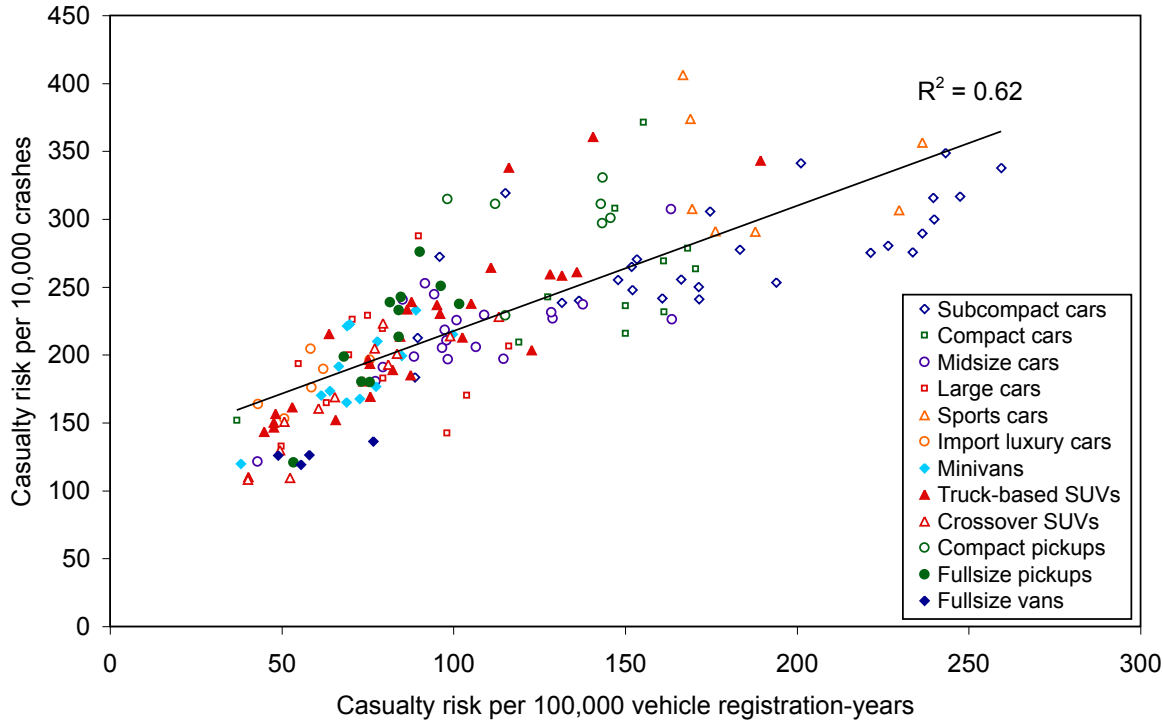



Table 3.1 shows the correlations between casualty risk per crash and casualty risk per vehicle registration-year, in each state. The analysis is done only for models that were involved in at least 500 police-reported crashes. The table indicates that the highest agreement between the two types of risk is in Pennsylvania ($R^2$=0.72), while the lowest agreement is in Maryland ($R^2$=0.51). The table also indicates that calculating casualty risk per mileage-adjusted registration-year slightly lowers the correlation with casualty risk per crash in each state.

**Table 3.1. Correlations ($R^2$) between casualty risk per crash and per vehicle registration-year, by state and vehicle model**

| | Correlation ($R^2$) with casualty risk per crash by vehicle model for models involved in at least 500 crashes | | |
|---|---|---|---|
| State | Number of models | Casualty risk per vehicle reg-yr | Casualty risk per mileage-adjusted vehicle reg-yr |
| Florida | 186 | 0.64 | NA |
| Illinois | 204 | 0.62 | 0.54 |
| Maryland | 116 | 0.51 | 0.46 |
| Missouri | 135 | 0.61 | 0.55 |
| Pennsylvania | 119 | 0.72 | 0.67 |
| Combined | 103 | 0.63 | |

Tables 3.2 and 3.3 compare the correlations in the two measures of casualty risk among the five states. Table 3.2 indicates that the highest correlations in casualty risk per vehicle registration-years are obtained using Illinois as the baseline state, whereas Table 3.3 indicates that the highest correlations in casualty risk per crash are obtained using Florida as the baseline state. The tables suggest that the correlation among states is stronger for casualty risk per vehicle registration-year ($R^2$ between 0.56 and 0.77) than for casualty risk per crash ($R^2$ between 0.28 and 0.45).

**Table 3.2. Correlations ($R^2$) between casualty risk per 100,000 vehicle registration-years in each state, 103 vehicle models involved in at least 500 crashes in each state**

| State | FL | IL | MD | MO | PA |
|---|---|---|---|---|---|
| FL | 1.00 | 0.77 | 0.77 | 0.60 | 0.56 |
| IL | 0.77 | 1.00 | 0.70 | 0.70 | 0.56 |
| MD | 0.77 | 0.70 | 1.00 | 0.58 | 0.52 |
| MO | 0.60 | 0.70 | 0.58 | 1.00 | 0.65 |
| PA | 0.56 | 0.56 | 0.52 | 0.65 | 1.00 |

**Table 3.3. Correlations ($R^2$) between casualty risk per 10,000 crashes in each state, 103 vehicle models involved in at least 500 crashes in each state**

| State | FL | IL | MD | MO | PA |
|---|---|---|---|---|---|
| FL | 1.00 | 0.45 | 0.35 | 0.28 | 0.37 |
| IL | 0.45 | 1.00 | 0.24 | 0.42 | 0.44 |
| MD | 0.35 | 0.24 | 1.00 | 0.08 | 0.21 |
| MO | 0.28 | 0.42 | 0.08 | 1.00 | 0.33 |
| PA | 0.37 | 0.44 | 0.21 | 0.33 | 1.00 |

### 3.3. Accounting for vehicle miles traveled

The differences in casualty risk per vehicle registration-years across states discussed above could be due to differences in the number of miles vehicles are driven in each state. LBNL has obtained vehicle emission inspection and maintenance (I/M) records from four states that have I/M programs (Illinois, Maryland, Missouri, and Pennsylvania), as well as four other states with I/M programs (California, Colorado, Ohio, and Wisconsin). These data can be used to estimate the relative number of miles driven annually by vehicle type and make/model.

Table 3.4 shows the average odometer reading of model year 2002 vehicles in 2008 and 2009, for urban areas in eight states. Average odometer readings are highest in Missouri (87,330) and lowest in Maryland (79,523), a difference of about 10%. (Average odometer readings are even higher in Wisconsin, 88,481, and even lower in Ohio, 77,472). Vehicle types with VMT more than 5% different than the average VMT for all vehicles in that state are shown in red in the table. Note that sports cars have consistently lower, 16% to 33% lower, VMT, and import luxury cars 4% to 14% lower, VMT than the average vehicle in each state. On the other hand, large SUVs, ¾-ton pickups, and fullsize vans have consistently higher VMT than the average vehicle in each state. Note that 1-ton pickups have 3% to 22% higher VMT in four states (Maryland, Missouri, Ohio, and California), but 9% and 13% lower VMT in Illinois and Pennsylvania,

respectively. The data in Table 4 exclude Crown Victoria Police Interceptors, which have dramatically higher (30% to two times higher) VMT than the average vehicle in a given state

**Table 3.4. Average vehicle mileage relative to that of all light-duty vehicles in eight states, by vehicle type**

| Vehicle type | IL | MD | MO | PA | OH | CA | CO | WI |
|---|---|---|---|---|---|---|---|---|
| Subcompact cars | 0.94 | 0.98 | 0.95 | 0.96 | 0.95 | 0.99 | 0.97 | 0.95 |
| Compact cars | 1.04 | 1.03 | 1.04 | 1.00 | 1.02 | 1.05 | 1.04 | 1.01 |
| Midsize cars | 0.98 | 0.99 | 0.98 | 0.97 | 0.97 | 0.99 | 0.97 | 0.98 |
| Large cars | 0.97 | 0.98 | 0.94 | 0.94 | 0.92 | 0.93 | 0.94 | 0.97 |
| Sports cars | 0.70 | 0.84 | 0.73 | 0.83 | 0.67 | 0.80 | 0.69 | 0.67 |
| Import luxury cars | 0.89 | 0.91 | 0.86 | 0.96 | 0.94 | 0.88 | 0.89 | 0.94 |
| Minivans | 1.08 | 1.04 | 1.07 | 1.07 | 1.10 | 1.05 | 1.11 | 1.03 |
| Small SUVs | 0.95 | 0.96 | 0.95 | 0.96 | 0.98 | 0.91 | 0.94 | 0.95 |
| Midsize SUVs | 1.07 | 1.04 | 1.06 | 1.05 | 1.10 | 1.04 | 1.09 | 1.08 |
| Large SUVs | 1.13 | 1.06 | 1.14 | 1.12 | 1.18 | 1.07 | 1.13 | 1.18 |
| Small CSUVs | 0.99 | 0.98 | 1.01 | 1.01 | 1.04 | 0.98 | 0.99 | 1.03 |
| Midsize CSUVs | 1.00 | 0.96 | 1.00 | 1.02 | 1.06 | 0.95 | 0.99 | 1.05 |
| Compact pickups | 1.04 | 1.05 | 0.98 | 0.98 | 0.96 | 1.04 | 1.00 | 0.97 |
| 1/2-ton pickups | 1.11 | 1.07 | 1.08 | 1.05 | 1.04 | 1.11 | 1.14 | 1.06 |
| 3/4-ton pickups | 1.11 | 1.11 | 1.21 | 1.09 | 1.06 | 1.13 | NA | NA |
| 1-ton pickups | 0.91 | 1.08 | 1.31 | 0.87 | 1.22 | 1.03 | NA | NA |
| Fullsize vans | 1.17 | 1.18 | 1.20 | 1.17 | 1.27 | 1.20 | 1.14 | 1.10 |
| Average odometer, all | 80,537 | 79,523 | 87,330 | 83,711 | 77,472 | 85,559 | 84,125 | 88,481 |

Figure 3.12 shows the average odometer by vehicle type in each state. Trends by vehicle type are similar across the states, but there are differences, particularly for full size pickups in Missouri. However, many ¾- and 1-ton pickup trucks have a gross vehicle weight rating over 8,500 lbs, and therefore are exempted from testing in most state I/M programs. As a result, there are very few of these vehicles included in the I/M databases, and therefore their average odometer readings are highly uncertain. Figure 3.13 indexes the average odometer reading by vehicle type in each state to the average odometer reading for all vehicles in that state. The figure indicates that average odometer reading by vehicle type is quite similar across the four states. There is more variability across the states in terms of average miles driven for sports cars, import luxury cars, large SUVs, midsize CSUVs, and compact pickups than for other vehicle types. Figure 3.14 shows the average odometer reading for four other states with vehicle emission I/M programs, California, Colorado, Ohio, and Wisconsin. Again average mileage across vehicle types is similar across states, with some differences for sports cars, import luxury cars, fullsize vans, large SUVs, midsize crossover SUVs. Note that there are no ¾- and 1-ton pickup trucks in the I/M databases for Colorado and Wisconsin.

**Figure 3.12. Average odometer reading, by vehicle type and state**

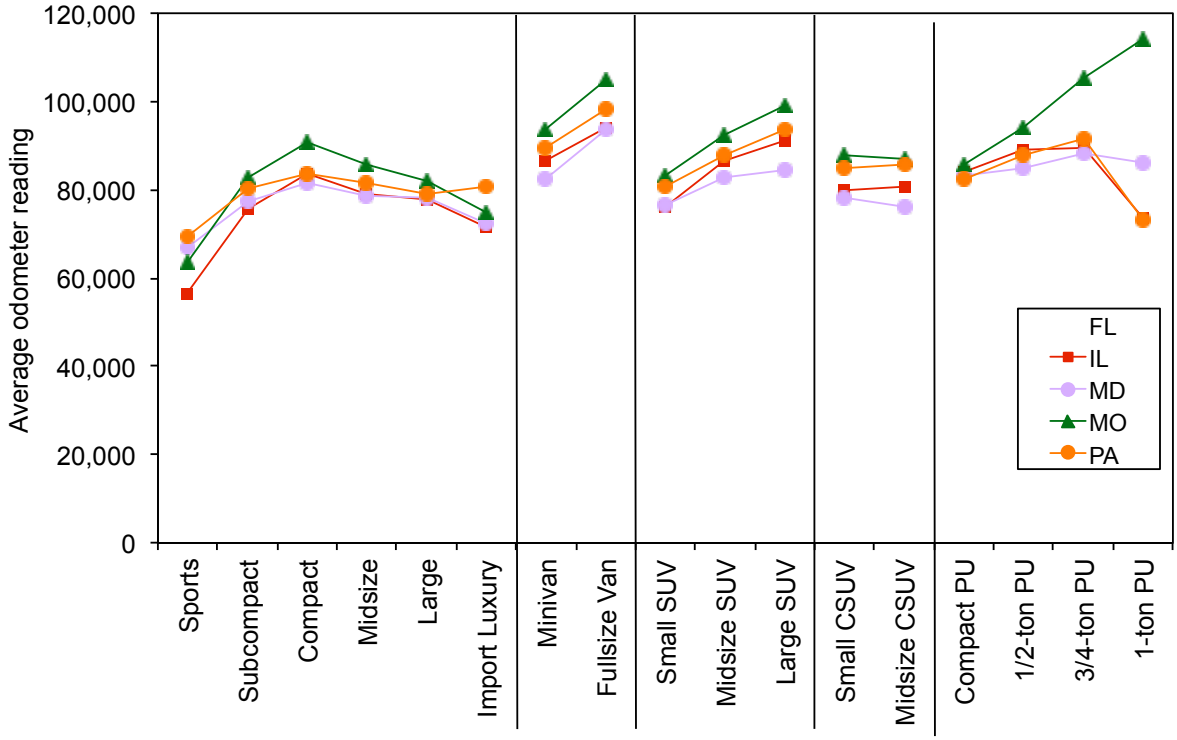



**Figure 3.13. Average odometer reading indexed to average odometer for all vehicles in each state, by vehicle type and state**

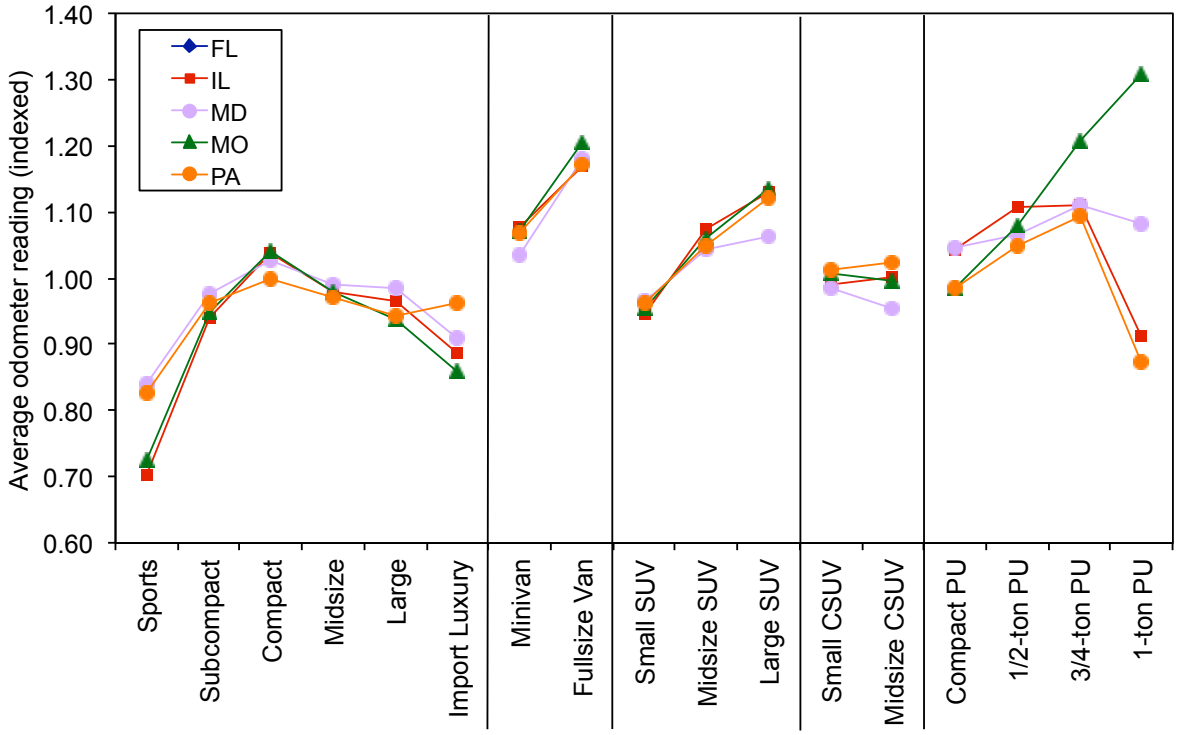
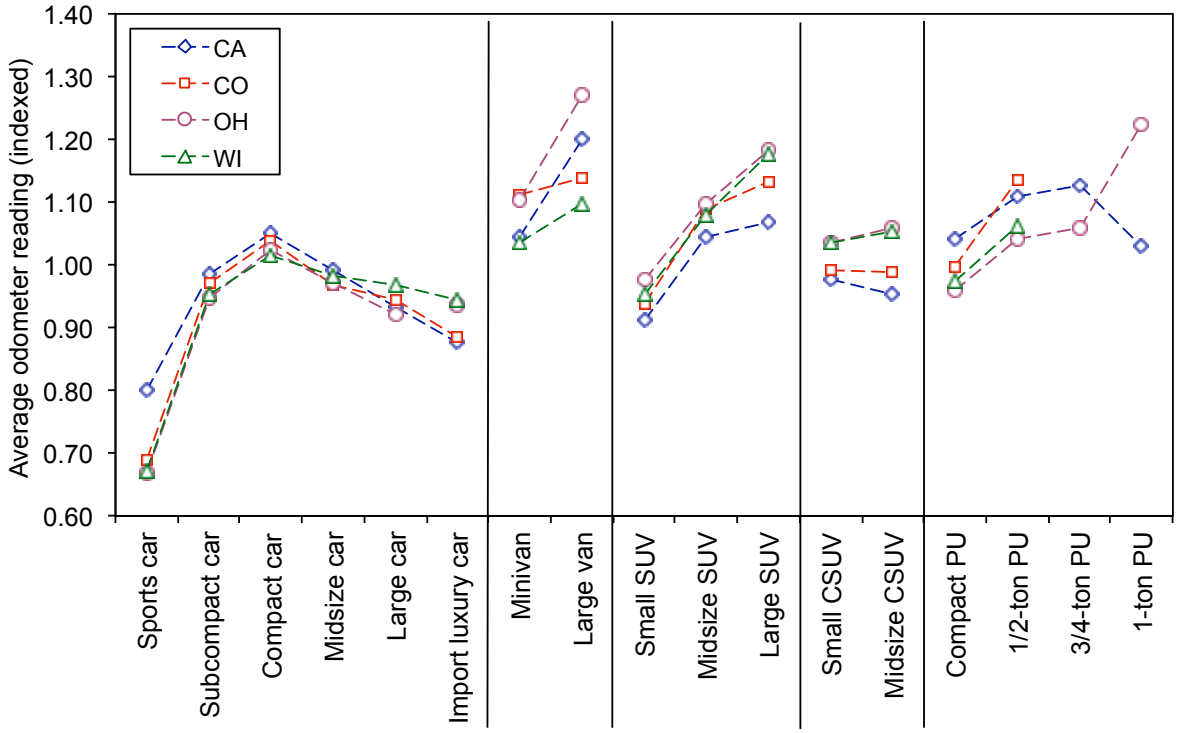


**Figure 3.14. Average odometer reading for four additional states, indexed to average odometer for all vehicles in each state, by vehicle type and state**




Because the I/M programs in the eight states operate predominantly in urban counties, only crashes that occur in the I/M counties in each state are included in the analysis. Table 3.5 compares odometer readings in urban and rural counties from California's I/M program, which covers virtually the entire state, in order to assess how population density influences VMT by model. We consider counties with population density higher than 200 per square mile as "urban", and less dense counties as "rural"; these definitions correspond to the population densities of the urban counties included in the I/M programs in the five states for which we have crash data. California's I/M program applies to all vehicles 6 years old or older in the most urban counties of the state; only vehicles that are changing ownership are subject to I/M testing in the more rural counties. As a result, 72% of the vehicles tested in 2008 and 2009 were registered in urban counties, and only 28% were registered in rural counties.

Table 3.5 indicates that, overall, vehicles registered in urban California counties are driven 5% fewer miles than vehicles registered in rural counties. Almost all vehicle types are driven less in urban counties than rural counties; the exceptions are 1-ton pickups and fullsize vans. There are slight differences in the effect of driving location on miles driven by vehicle type, ranging from nearly identical VMT for compact pickups to 6% more miles in urban counties for fullsize vans, and 6% fewer miles in urban counties for midsize crossover SUVs.

Figure 3.15 compares the average odometer in rural and urban California counties, by vehicle type. One to one correspondence is shown as a long dashed line in the figure, while values for 5 percent higher and lower odometer readings in rural counties are shown as short dashed lines in the figure. The figure indicates that all vehicle types, except fullsize vans and 1-ton pickup trucks, registered in rural counties are driven up to 5 percent more miles than those registered in

**Table 3.5. Average vehicle mileage in California urban and rural counties and percent difference, by vehicle type**

| Vehicle type | Average odometer of MY02 vehicles in 2008 or 2009 | | Percent difference | Percent registered in urban counties |
|---|---|---|---|---|
| | Rural counties | Urban counties | | |
| Subcompact cars | 92,636 | 87,446 | -5.6% * | 73% |
| Compact cars | 99,041 | 94,199 | -4.9% * | 66% |
| Midsize cars | 94,700 | 89,942 | -5.0% * | 73% |
| Large cars | 83,436 | 81,710 | -2.1% * | 69% |
| Sports cars | 73,745 | 71,447 | -3.1% * | 73% |
| Import luxury cars | 82,832 | 78,874 | -4.8% * | 79% |
| Minivans | 97,476 | 94,864 | -2.7% * | 73% |
| Small SUVs | 82,944 | 81,123 | -2.2% | 70% |
| Midsize SUVs | 95,714 | 92,262 | -3.6% * | 74% |
| Large SUVs | 99,433 | 96,014 | -3.4% * | 69% |
| Small CSUVs | 92,381 | 87,205 | -5.6% * | 73% |
| Midsize CSUVs | 91,017 | 85,612 | -5.9% * | 77% |
| Compact pickups | 94,287 | 93,847 | -0.5% | 69% |
| 1/2-ton pickups | 101,656 | 98,887 | -2.7% * | 66% |
| 3/4-ton pickups | 102,245 | 99,637 | -2.6% * | 63% |
| 1-ton pickups | 90,409 | 93,015 | 2.9% | 67% |
| Fullsize vans | 103,376 | 109,542 | 6.0% * | 75% |
| Total | 93,655 | 89,349 | -4.6% * | 72% |

* difference between average odometer in rural and urban counties is statistically significant.

urban counties. Fullsize vans registered in urban counties are driven about 6 percent more miles than those registered in rural counties (1-ton pickup trucks registered in urban counties are also driven more miles than those registered in urban counties, but the difference is not statistically significant because of the small number of 1-ton pickup trucks registered in urban counties).

Figure 3.16 shows a similar plot by vehicle model. The figure indicates that there is good agreement across vehicle models between miles driven in urban and rural counties ($R^2$=0.76). Three vehicle models, Ford Crown Vic, Mercury Grand Marquis, and Lincoln Town Car, have much higher mileage in urban counties than in rural counties; the correlation improves to 0.92 when these three models are excluded (Crown Vic models which are used as police vehicles are not included in the figure). The figure suggests that most models are driven 5 percent more miles when registered in rural counties; a few models are driven as much as 10 percent more miles in rural counties than in urban counties. Therefore using odometer readings from vehicles participating in I/M programs, which tend to be in urban counties, will result in a less than 10 percent under-reporting of miles traveled for vehicles registered in rural counties. Table 3.5 and Figures 3.15 and 3.16 also suggest that the range in vehicle odometer readings across different vehicles is much greater (up to 53% by vehicle type, and up to 70% by vehicle model) than it is between a given vehicle registered in a rural as opposed to an urban county (up to 6% by vehicle type, and up to 23% by vehicle model).

**Figure 3.15. Average odometer of MY02 vehicles tested in California I/M program in 2008 and 2009, by vehicle type and registration county**

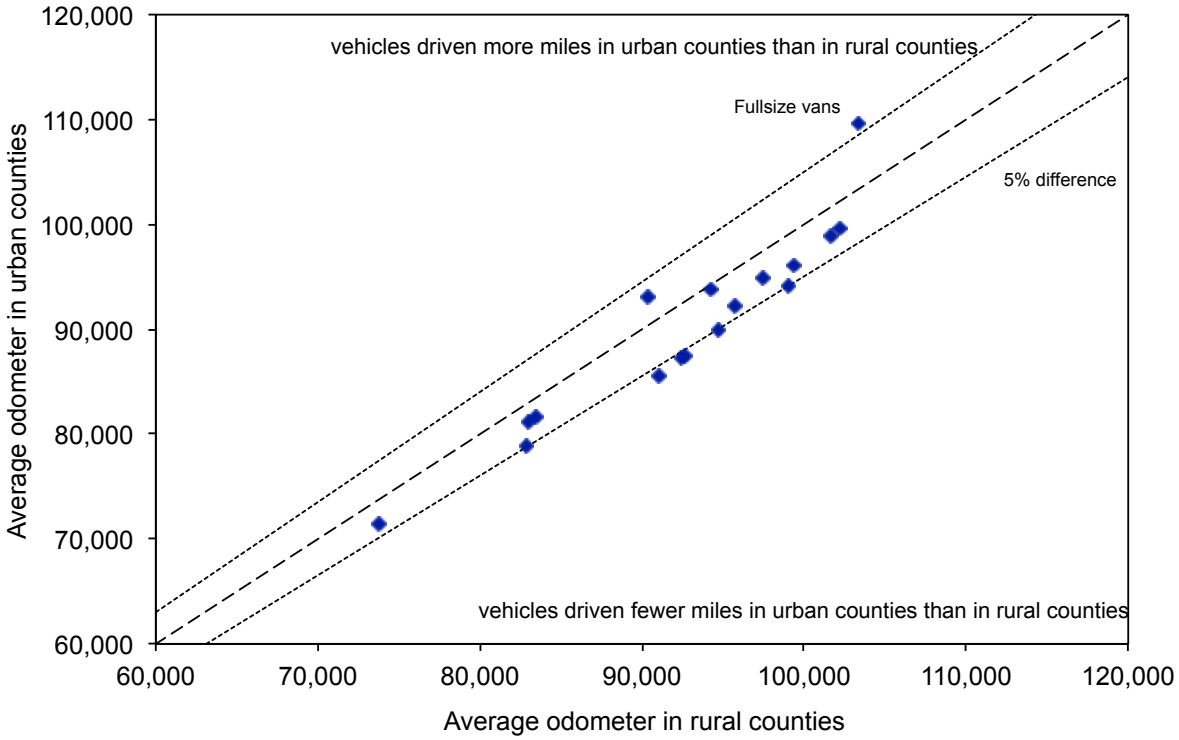



**Figure 3.16. Average odometer of MY02 vehicles tested in California I/M program in 2008 and 2009, by vehicle model and registration county**

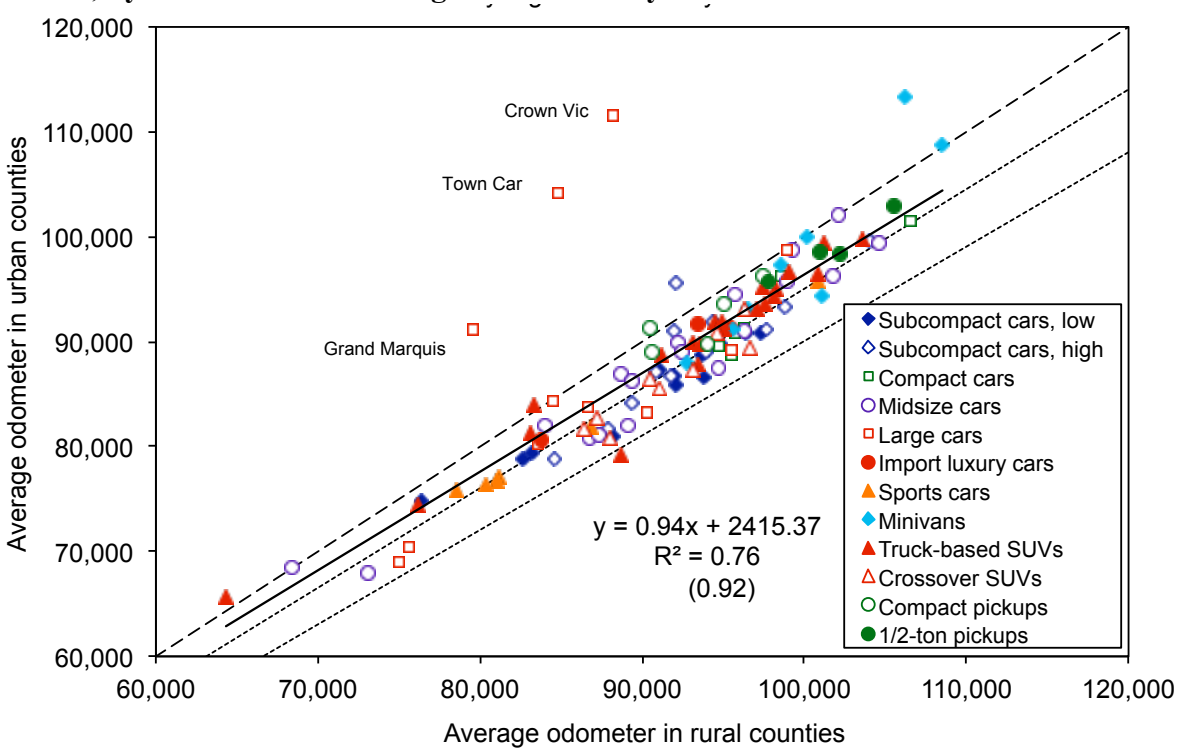


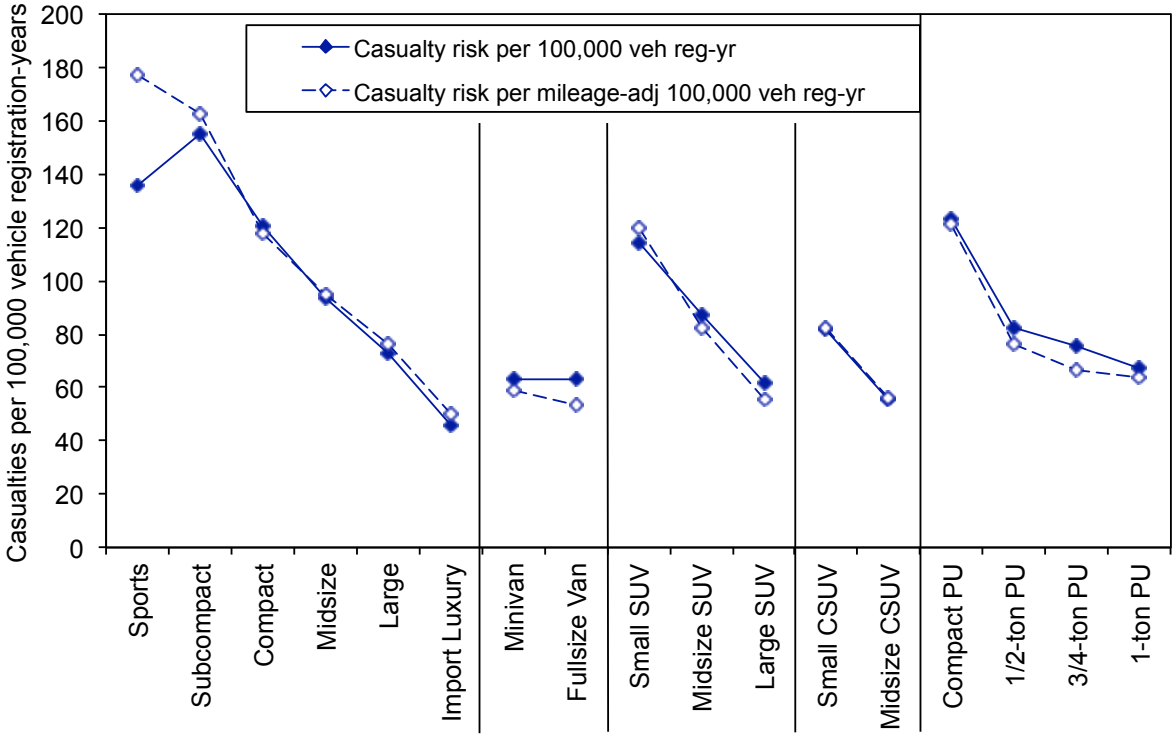
Table 3.6 shows the correlations between average odometer reading among 103 vehicle models in each state. The table indicates that the correlations are highest when Illinois is used as the baseline state, with $R^2$ between 0.68 (for vehicle models in Maryland) and 0.78 (for vehicle models in Pennsylvania).

**Table 3.6. Correlations ($R^2$) between average vehicle odometer reading in each state, 103 vehicle models involved in at least 500 crashes in each state**

| State | FL | IL | MD | MO | PA |
|---|---|---|---|---|---|
| FL | NA | NA | NA | NA | NA |
| IL | NA | 1.00 | 0.68 | 0.74 | 0.78 |
| MD | NA | 0.68 | 1.00 | 0.59 | 0.53 |
| MO | NA | 0.74 | 0.59 | 1.00 | 0.61 |
| PA | NA | 0.78 | 0.53 | 0.61 | 1.00 |

Figure 3.17 shows the effect of accounting for annual miles driven on casualty risks per vehicle registration-years in the four states that operate vehicle emission inspection and maintenance programs, and therefore provide odometer readings (that is, for all the states except Florida). For most vehicle types, adjusting for miles driven has little to no effect on casualty risk pre vehicle registration-year; however, adjusting for miles driven substantially increases casualty risk for sports cars, by 30%, and slightly reduces casualty risk for fullsize vans and ¾-ton pickups.

**Figure 3.17. Four I/M (excluding Florida) state casualty risks per 100,000 vehicle registration-years, with and without mileage-adjustment, by vehicle type**



### 3.4. Crash involvement rates

We next calculated crash involvement rates by vehicle type, model and state. Crash involvement rates are the number of police-reported crashes per 1,000 vehicle registration-years. Table 3.7 shows that crash involvement rates vary consistently by state, with Illinois and Missouri having relatively high crash rates (77 and 66 crashes, respectively, per 1,000 registered vehicle-years) and Florida and Pennsylvania having relatively low crash rates (30 per 1,000 registered vehicle-years).

**Table 3.7. Crash involvement rate per 1,000 vehicle registration-years, by vehicle type and state**

| | Crash involvement rate per 1,000 vehicle registration-years | | | | | |
|---|---|---|---|---|---|---|
| Vehicle type | FL | IL | MD | MO | PA | All |
| Subcompact cars | 44 | 109 | 61 | 94 | 40 | 65 |
| Compact cars | 40 | 97 | 53 | 80 | 32 | 59 |
| Midsize cars | 32 | 83 | 42 | 64 | 27 | 48 |
| Large cars | 23 | 65 | 35 | 51 | 22 | 39 |
| Sports cars | 34 | 77 | 44 | 76 | 33 | 47 |
| Import luxury cars | 23 | 61 | 24 | 47 | 22 | 32 |
| Minivans | 23 | 62 | 33 | 54 | 25 | 39 |
| Small SUVs | 28 | 83 | 47 | 78 | 37 | 50 |
| Midsize SUVs | 27 | 74 | 37 | 62 | 31 | 44 |
| Large SUVs | 24 | 59 | 29 | 52 | 27 | 36 |
| Small CSUVs | 25 | 74 | 35 | 65 | 32 | 43 |
| Midsize CSUVs | 22 | 66 | 29 | 60 | 29 | 38 |
| Compact pickups | 30 | 71 | 44 | 71 | 32 | 45 |
| 1/2-ton pickups | 28 | 58 | 37 | 56 | 25 | 39 |
| 3/4-ton pickups | 27 | 52 | 36 | 49 | 27 | 38 |
| 1-ton pickups | 29 | 64 | 50 | 53 | 29 | 42 |
| Fullsize vans | 29 | 81 | 48 | 70 | 34 | 49 |
| Total | 30 | 77 | 42 | 66 | 30 | 47 |

Figure 3.18 shows the data in Table 3.7 in graphical form, while Figure 3.19 plots the crash frequencies for each vehicle type indexed to the crash frequency for all vehicles in each state. The figures suggest that the trends in crash frequencies among vehicle types are similar across the five states.

**Figure 3.18. Crashes per 1,000 vehicle registration-years, by vehicle type and state**

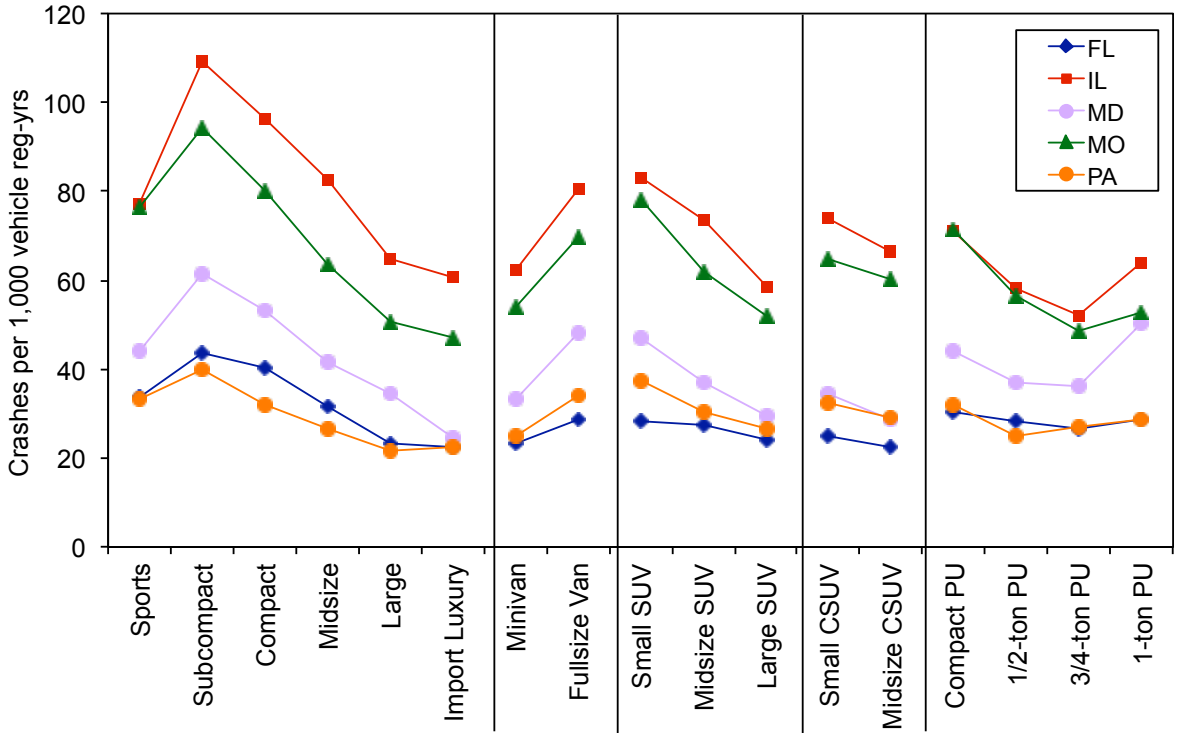



**Figure 3.19. Crashes per 1,000 vehicle registration-years, indexed to crash frequency for all vehicles in each state, by vehicle type and state**

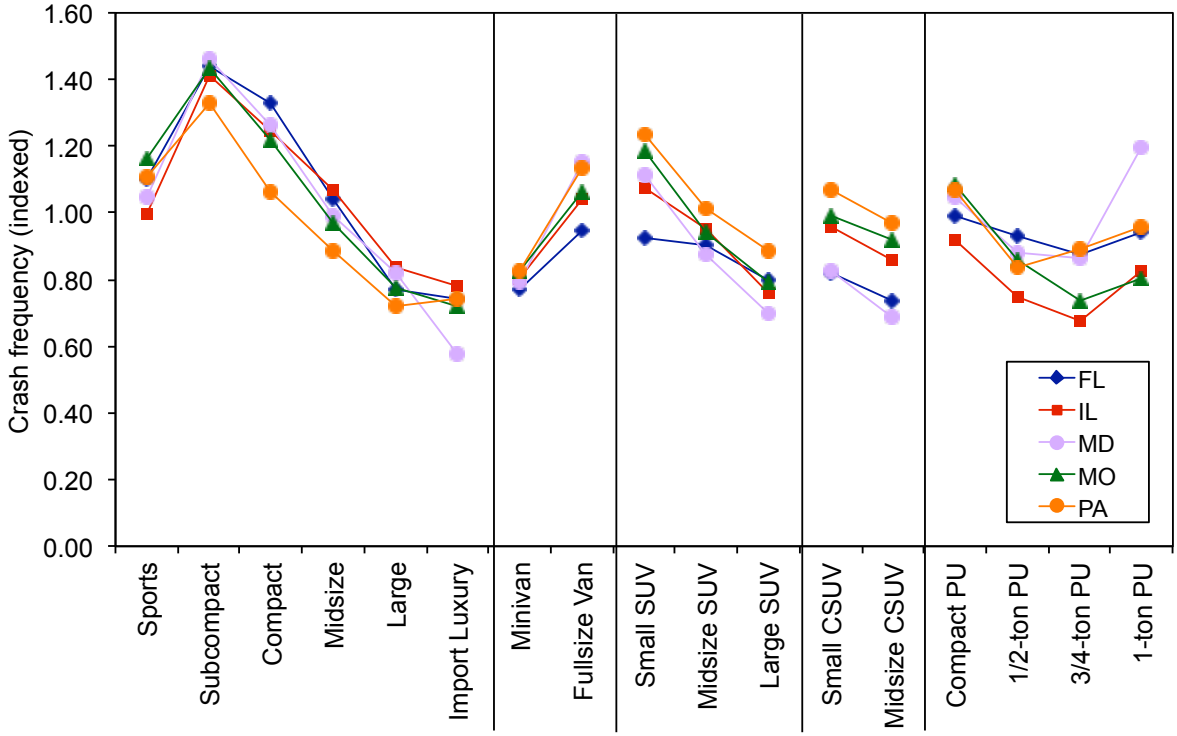


Table 3.8 shows the ratio of the crash involvement rate for each vehicle type relative to the total crash involvement rate in that state, to better allow comparisons across states with different absolute crash rates. Some vehicle types have consistently higher crash involvement rates in each state (subcompact, compact cars), while others have consistently low crash rates (large cars, import luxury cars, minivans, large SUVs, midsize CSUVs, ½- and ¾-ton pickups). Note that sports cars do not have substantially higher crash frequency than other car models, which suggests that their high fatality and casualty risks are due to them being involved in more severe crashes, or to relatively poor crashworthiness compared to other cars. Table 3.8 indicates that some vehicle types have crash rates that differ by more than 10% from the crash rate in all five states; these cases are shown in red in the table. For example, compact cars have a 26% higher crash rate than all vehicles in all five states combined, but in Pennsylvania compact cars have only a 6% higher crash rate than all Pennsylvania cars. On the other hand, small SUVs in Missouri and Pennsylvania have higher crash rate ratios (19% and 24% higher than all vehicles in each state, respectively), while small SUVs in Florida have much lower crash rates (7% lower than all vehicles in Florida), than small SUVs in all five states combined (7% higher than all vehicles).

Table 3.8 suggests that, while there are some similarities across states in crash involvement rates by vehicle type, there also are important differences for particular vehicle types in certain states. The causes of these differences are not clear.

**Table 3.8. Ratio of crash involvement rate by vehicle type to that of all light-duty vehicles in state**

| | Ratio of crash involvement rate by vehicle type to crash involvement rate for all vehicles in state | | | | | |
|---|---|---|---|---|---|---|
| Vehicle type | FL | IL | MD | MO | PA | All |
| Subcompact cars | 1.44 | 1.41 | 1.46 | 1.43 | 1.33 | 1.38 |
| Compact cars | 1.33 | 1.25 | 1.26 | 1.22 | 1.06 | 1.26 |
| Midsize cars | 1.04 | 1.07 | 0.99 | 0.97 | 0.89 | 1.04 |
| Large cars | 0.77 | 0.84 | 0.82 | 0.77 | 0.72 | 0.84 |
| Sports cars | 1.10 | 1.00 | 1.05 | 1.16 | 1.11 | 1.01 |
| Import luxury cars | 0.74 | 0.78 | 0.58 | 0.72 | 0.74 | 0.69 |
| Minivans | 0.77 | 0.80 | 0.79 | 0.82 | 0.83 | 0.83 |
| Small SUVs | 0.93 | 1.07 | 1.11 | 1.19 | 1.24 | 1.07 |
| Midsize SUVs | 0.90 | 0.95 | 0.88 | 0.94 | 1.01 | 0.94 |
| Large SUVs | 0.80 | 0.76 | 0.70 | 0.79 | 0.88 | 0.76 |
| Small CSUVs | 0.82 | 0.96 | 0.82 | 0.99 | 1.07 | 0.92 |
| Midsize CSUVs | 0.74 | 0.86 | 0.69 | 0.92 | 0.97 | 0.82 |
| Compact pickups | 0.99 | 0.92 | 1.05 | 1.09 | 1.07 | 0.96 |
| 1/2-ton pickups | 0.93 | 0.75 | 0.88 | 0.86 | 0.84 | 0.83 |
| 3/4-ton pickups | 0.87 | 0.67 | 0.86 | 0.74 | 0.89 | 0.80 |
| 1-ton pickups | 0.94 | 0.83 | 1.20 | 0.80 | 0.96 | 0.91 |
| Fullsize vans | 0.95 | 1.04 | 1.15 | 1.07 | 1.14 | 1.06 |
| Total | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |

Table 3.9 shows the correlation among crash involvement rates by vehicle model in each state. The highest correlations are obtained when using Missouri as the "baseline" state; the $R^2$ for crash rates between models in Missouri and the other four states exceed 0.80 in each state (shown in red).

**Table 3.9. Correlations ($R^2$) between crashes per vehicle registration-year in each state, 103 vehicle models involved in at least 500 crashes in each state**

| State | FL | IL | MD | MO | PA |
|---|---|---|---|---|---|
| FL | 1.00 | 0.83 | 0.87 | 0.81 | 0.65 |
| IL | 0.83 | 1.00 | 0.79 | 0.85 | 0.69 |
| MD | 0.87 | 0.79 | 1.00 | 0.85 | 0.72 |
| MO | 0.81 | 0.85 | 0.85 | 1.00 | 0.81 |
| PA | 0.65 | 0.69 | 0.72 | 0.81 | 1.00 |

Figure 3.20 plots vehicle registration years and the number of vehicles involved in police-reported crashes, in counties with I/M programs in the four states, by vehicle model. There is a strong correlation between vehicle registration-years and vehicles involved in police-reported crashes, suggesting that crashes can be used as a proxy for registrations as a measure of exposure. The correlation decreases somewhat, to $R^2$=0.89, when one accounts for mileage by vehicle model using average odometer readings from model year 2002 vehicles in 2008 and 2009 using test result records from emission inspection and maintenance programs in the four states.

**Figure 3.20. Vehicle registration-years and vehicles involved in police-reported crashes, vehicles included in four state I/M programs, by vehicle type and model**

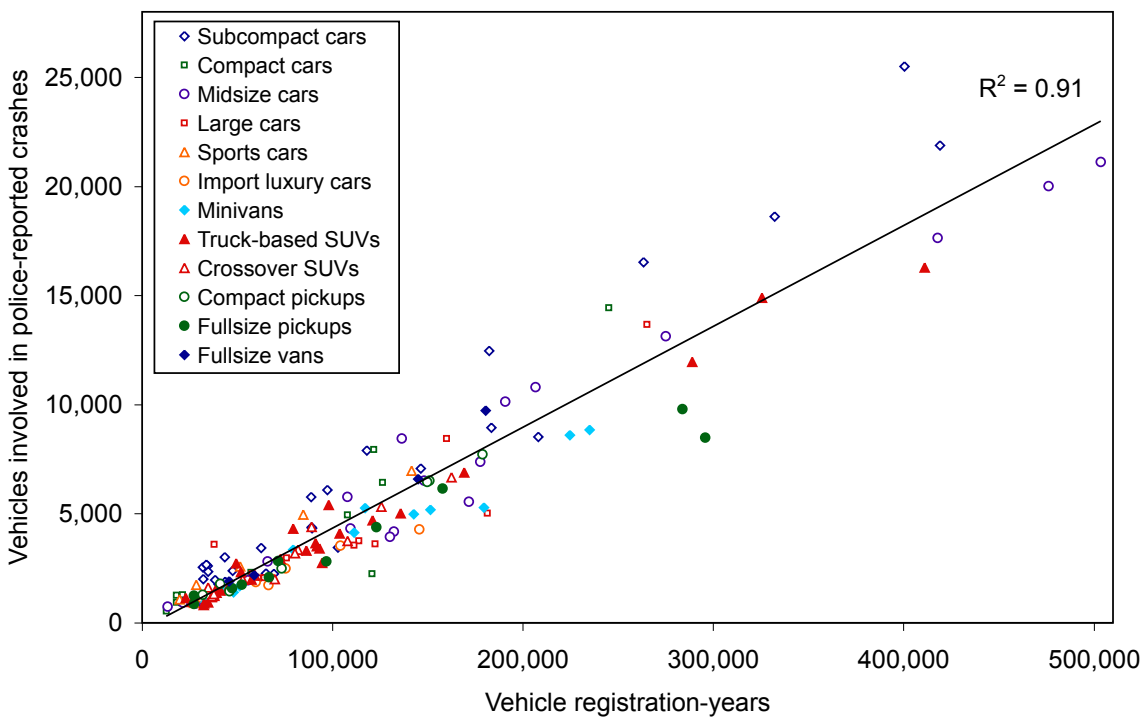



Figure 3.21 compares the average odometer and pooled crash frequency in the four states for which we have odometer data, by vehicle type. The figure uses only vehicles registered in

counties included in each state's I/M program (i.e. urban counties), since these are the vehicles for which we have odometer data. The figure indicates that overall there is no correlation between miles driven and crash frequency across vehicle types. Figure 3.22 indicates the lack of correlation between average odometer and crash frequency across vehicle models.

**Figure 3.21. Average vehicle odometer reading and crash frequency, vehicles included in four state I/M programs, by vehicle type**

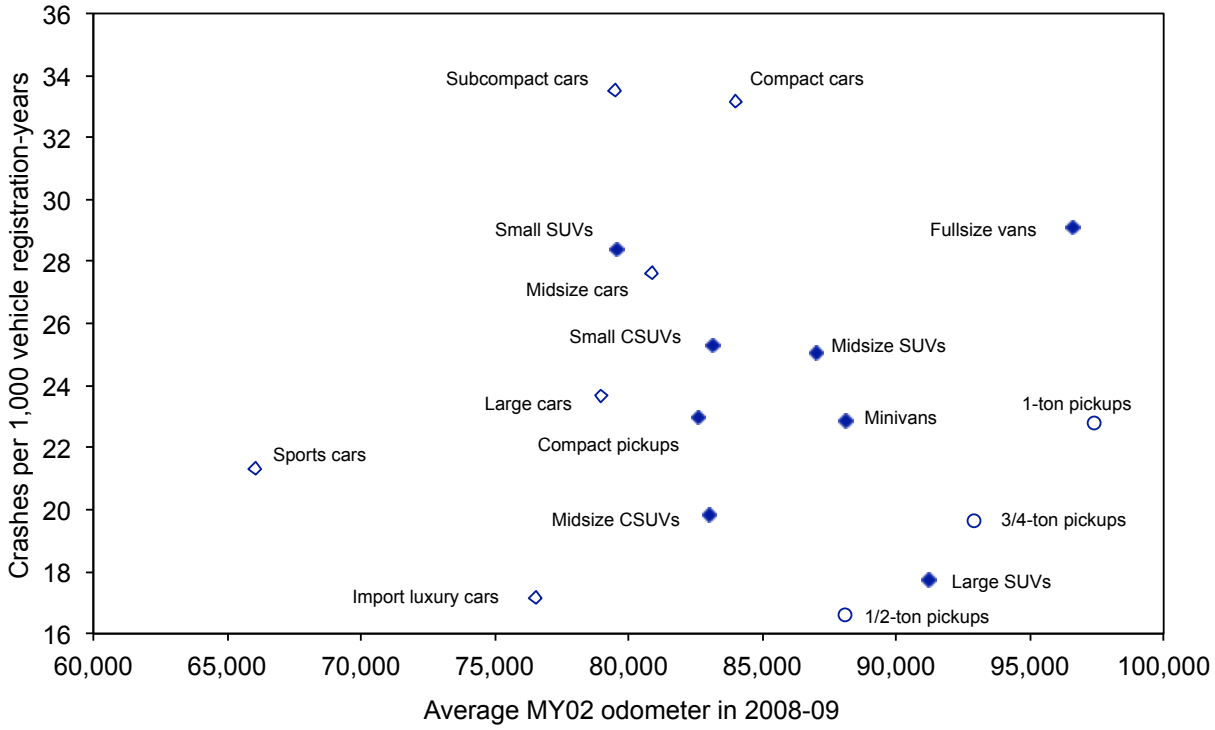


**Figure 3.22. Average vehicle odometer reading and crash frequency, vehicles included in four state I/M programs, by vehicle model**

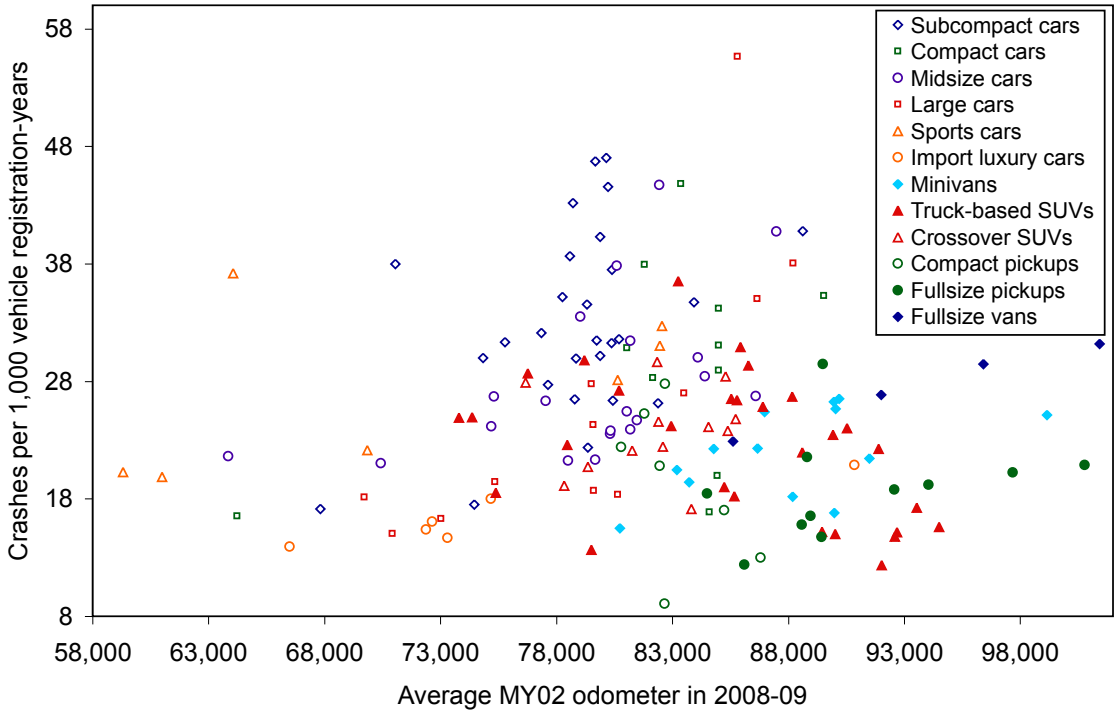



Figure 3.23 compares pooled driver casualty risk by vehicle type in the five states, both per 100,000 vehicle registration-years (in blue) and per 10,000 police-reported crashes (in red); these are the same data from Figure 3.4 above, with risk per vehicle plotted on the left-hand axis and risk per crash plotted on the right-hand axis. Figure 3.23 also shows the number of crashes per 10,000 vehicle registration-years (in green), plotted against the left-hand axis. The figure shows how crash frequency relates to crash risk per vehicle and per crash. For example, the 12,709 casualties in subcompact cars translate into 175 casualties per 100,000 vehicle registration years (see equation 1).

Eq (1) 12,709 casualties / 7,276,294 vehicle reg-yrs = 175 casualties / 100,000 vehicle reg-years

Subcompact cars have a relatively high crash involvement rate, 65 crashes every 1,000 vehicle registration-years. The casualty risk per crash can be calculated by dividing the casualty risk per registration-year by the number of crashes per registration-year:

Eq (2) (175 casualties / 100,000 vehicle reg-years) * (1,000 vehicle reg-years / 65 crashes) * 100 = 270 casualties / 10,000 crashes

Calculating risk using all police-reported crashes, rather than vehicle registration-years, as the measure of exposure changes the risk of certain vehicle types relative to that of other vehicle types. For example, subcompact and compact cars have higher crash involvement rates (green triangles) than midsize, large and import luxury cars, so the difference in casualty risk between smaller and larger cars is dramatically reduced when the measure of exposure is changed from registration-years (blue diamonds) to all police-reported crashes (red circles). On the other hand,

midsize and large cars, and minivans, have relatively low crash involvement rate, so changing the measure of exposure to crashes increases their casualty risk relative to that of other vehicle types. Similarly, the relatively low crash involvement rates of large SUVs, crossover SUVs, and fullsize pickups results in relatively higher casualty risks when the measure of exposure is changed from vehicle registration-years to all police-reported crashes.

**Figure 3.23. Five state casualty risks (per vehicle registration-years and per crash) and crash frequency, by vehicle type**

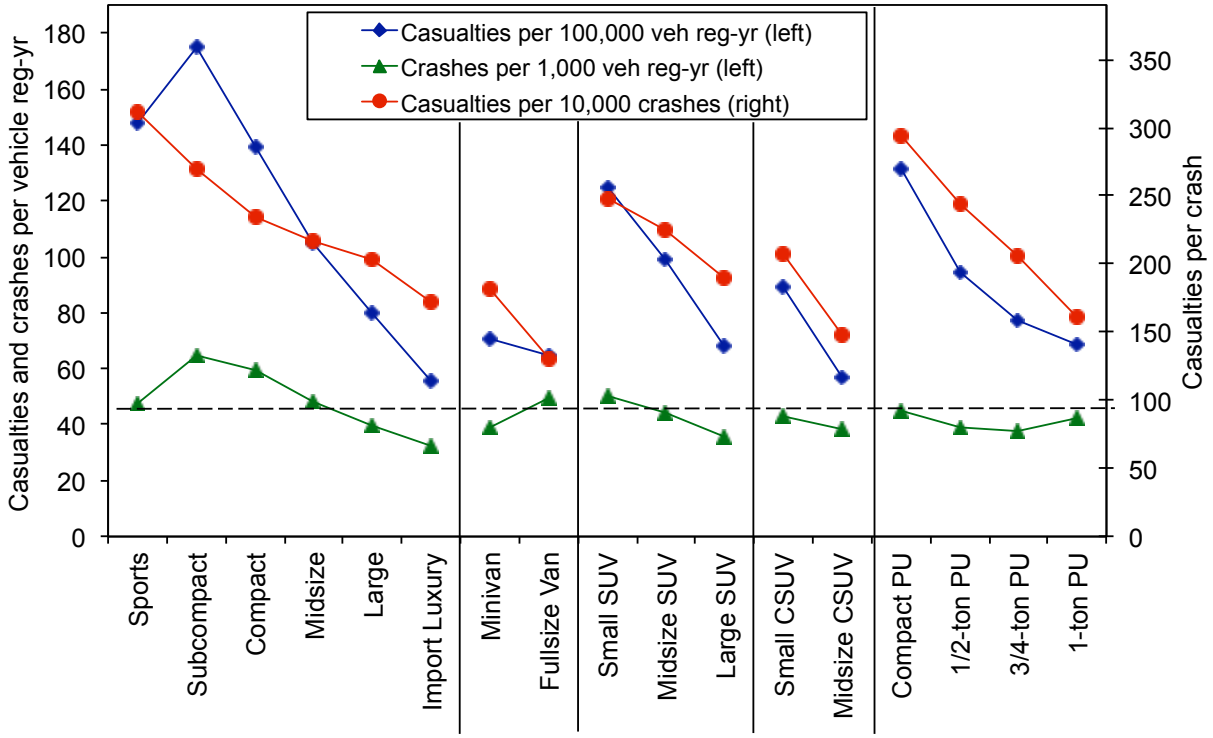


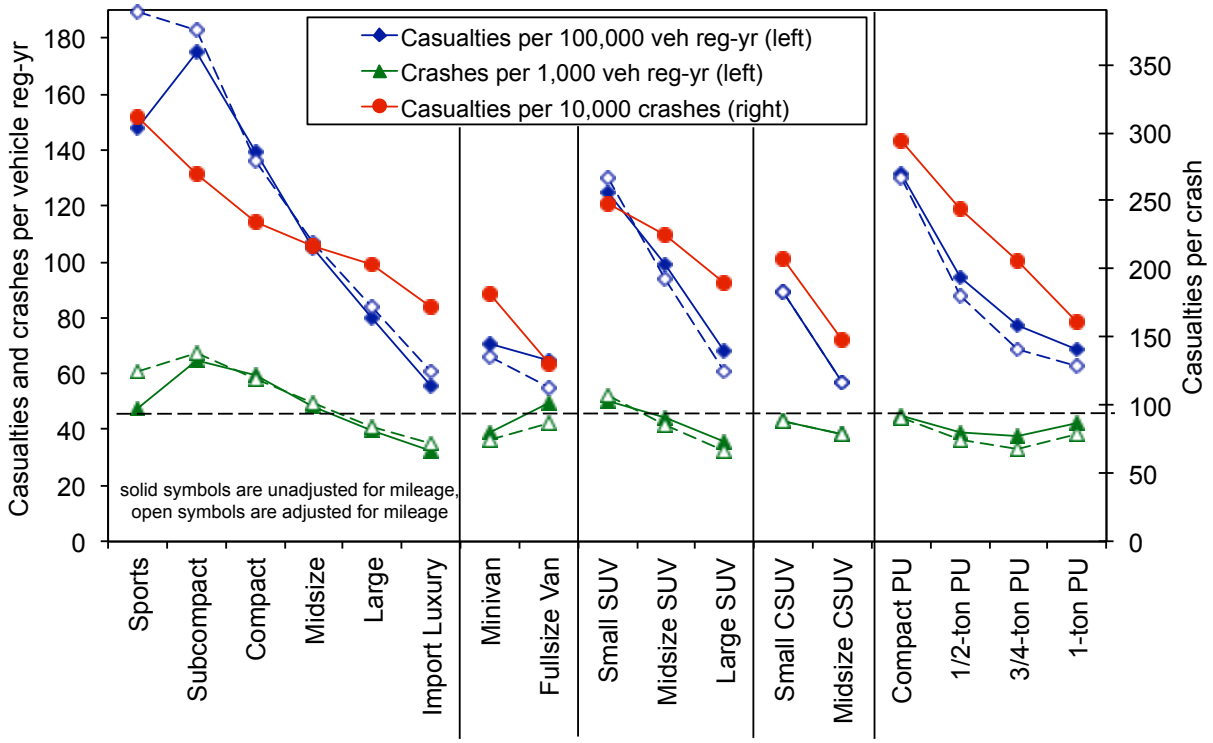

Figure 3.24 shows casualty risk and crash frequency per vehicle registration-year before and after adjusting for vehicle mileage (as open blue diamonds and open green triangles, respectively). Figure 3.24 indicates that adjusting risks by the estimated mileage vehicles are driven in a year has little effect on casualty risks other than for sports cars. However, adjusting crash frequency by mileage has an effect for several vehicle types: sports cars, fullsize vans, and fullsize pickup trucks.

The bottom line from Figures 3.23 and 3.24 is that changing the measure of exposure from registration years to all police-reported crashes reduces the sharp decline in casualty risk as car size increases. Casualty risk per crash does decrease as size increases, within each major vehicle type, but for cars and truck-based SUVs the decrease is not as large as when registration-years is used as the measure of exposure.

The figures also indicate that adjusting for estimated miles driven does not have a large impact on casualty risk by vehicle registration-years. And Figure 3.20 shows that total police-reported crashes are a good proxy for total vehicle registration-years and/or total miles driven, even though Figures 3.21 and 3.22 show that there is no correlation between crashes per vehicle or

miles driven per vehicle. Using the number of police-reported crashes, rather than vehicle registration-years or miles, as the measure of exposure estimates the casualty risk based on the crashworthiness of vehicle designs; that is, the risk of casualty given that a crash occurred. However, because the majority of crashes reported to police are non-injury, property-damage-only crashes, casualty risk per crash measures vehicle crashworthiness in all types of crashes.

**Figure 3.24. Five state casualty risks (per vehicle registration-years and per crash) before and after mileage adjustment, and crash frequency, by vehicle type**



## 4. Potential bias using state crash data

There are two potential biases in using police-reported crash data from several states. The first involves whether there are differences among state requirements regarding the types of crashes that require a police report be filed, as well as what criteria each state uses include crashes in the datafiles they provide to NHTSA's State Data System. Differences in reporting requirements and practices among the states can bias casualty rates across the states. The second type of bias involves whether drivers are less likely to report a crash for certain vehicle types than for others; for example, some have reported that pickup truck drivers are less likely to report a crash to police than drivers of other vehicle types. Different reporting practices among vehicle drivers can result in biased risk estimates across vehicle types.

### 4.1. Reporting bias by states

Not all states provide police-reported crashes to NHTSA's State Data System (SDS); of those that do, only about half include the vehicle identification number, which is essential to determine the year, make and model of vehicles involved in crashes. We used the NASS GES data to analyze whether casualty risks per crash from the states providing data to the SDS are representative of national casualty risks per crash. The GES is a national sample of police-reported crashes in abour 60 diverse geographical areas, called primary sampling units. We assigned the PSUs in the GES to their states, and grouped each state based on whether it provides crash and VIN data to the SDS. We calculated casualty risks per crash for model year 2000 to 2004 vehicles through GES 2007.

Figure 4.1 compares the GES casualty risk for model year 2000 to 2004 vehicles by vehicle type, for PSUs in the five states for which we have police-reported crash data, PSUs in the other twelve states that report VIN, and PSUs in states that do not report VIN or are not included in NHTSA's SDS. The risks are calculated using the unweighted number of casualties and crashes in the GES. The figure suggests that, for most vehicle types, casualty risks are highest in the five VIN states, followed by those in the twelve other VIN states, and lowest in the remaining states. For minivans and fullsize vans, however, casualty risks are nearly identical in the twelve other VIN states and all other states, while for pickups casualty risks are nearly identical for the five VIN and twelve other VIN states. The blue line in the figure shows the casualty risk for all PSUs in the GES; it suggests that the twelve other VIN states are most representative of national casualty risk for all vehicle types except pickups; national casualty risks for pickups fall between those in the VIN states and those in the non-VIN states, according to the GES. Based on Figure 4.1 we expect that calculating casualty risk using data from all of the SDS VIN states will overstate national casualty risk, for all vehicle types.

Figure 4.2 compares the GES casualty risks from the VIN states in Figure 4.1 (using the left-hand scale) with the casualty risks calculated using all police-reported crashes in the five states (using the right-hand scale). Here we see that GES casualty risks match well with casualty risks calculated using state data, for cars, minivans, and compact crossovers. However, casualty risks for fullsize vans, midsize and large truck-based SUVs, and midsize crossovers are substantially lower compared to other vehicle types in the state data than in the GES, while casualty risks for pickups are relatively higher in the state data than in the GES.

**Figure 4.1. GES driver casualty risk by vehicle type and State Data System VIN status of state**

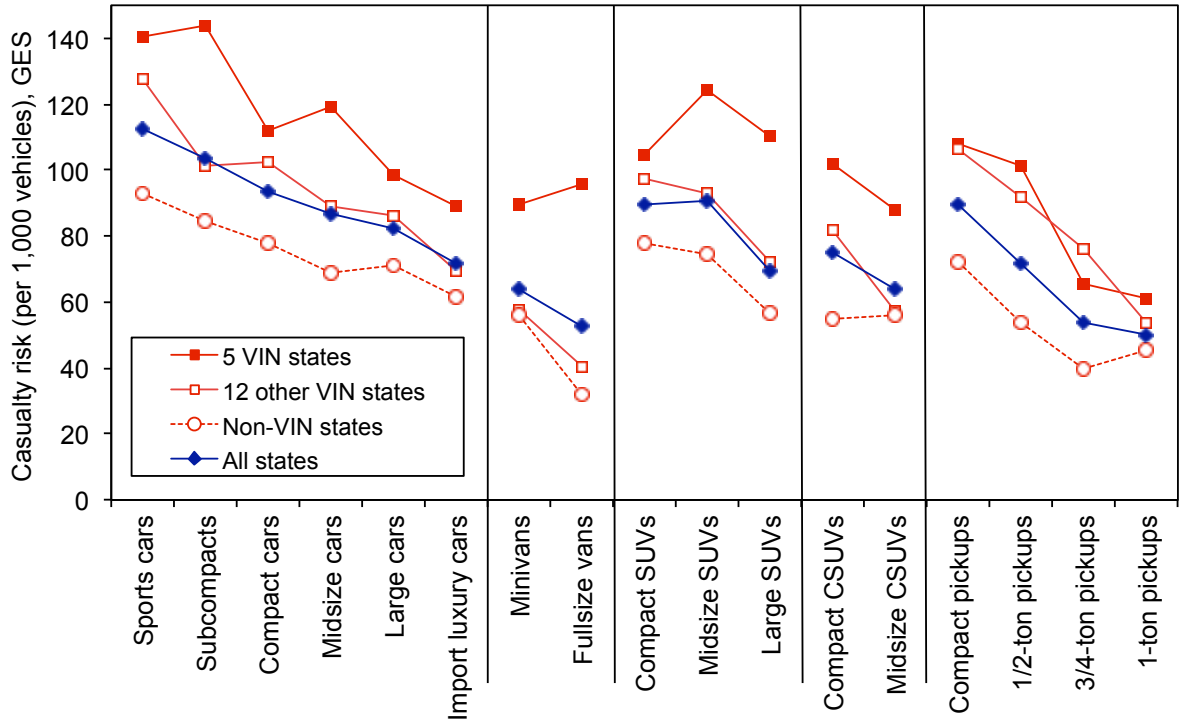



**Figure 4.2. GES and state data driver casualty risk, by vehicle type**

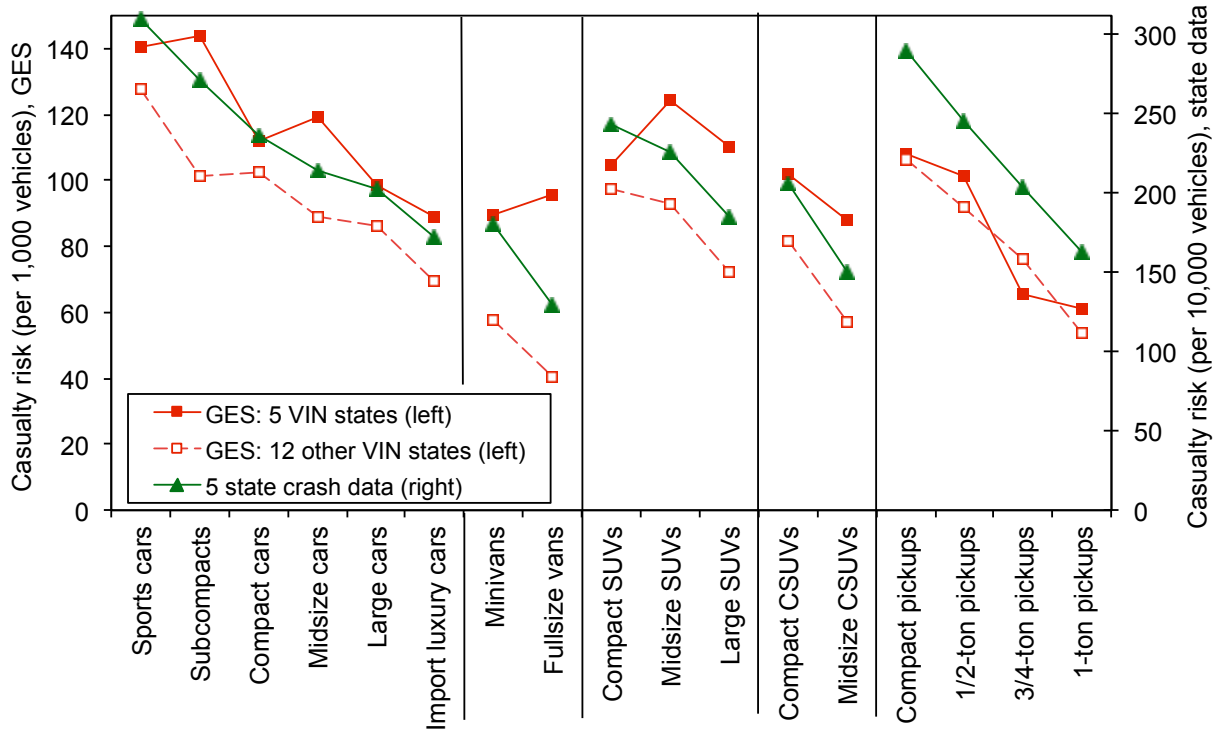
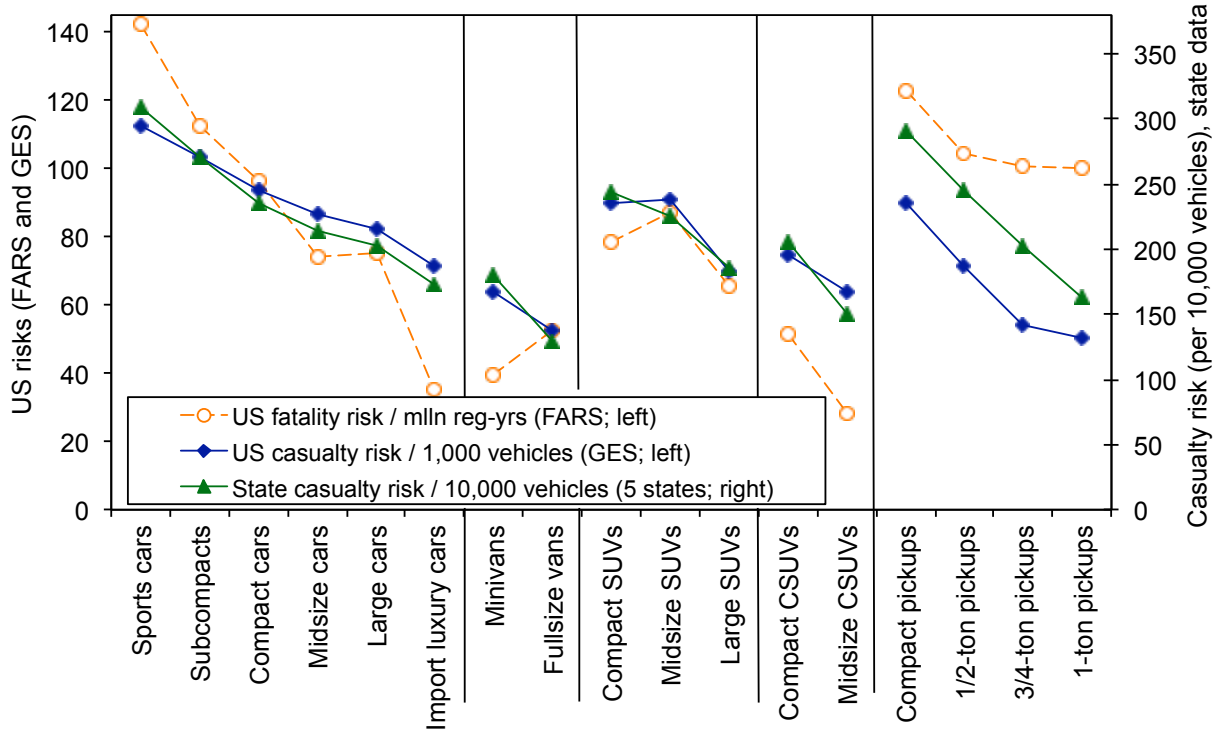


Figure 4.3 compares national casualty risk per police-reported crash using GES (blue line in Figure 4.1), casualty risk per police-reported crash using the five states (green line in Figure 4.2), and national fatality risk per million vehicle registration-years using FARS. The national GES casualty risks are remarkably similar to the casualty risks from all police-reported crashes in the five states, for most vehicle types; the exception is pickup trucks, which have substantially lower casualty risks in the national GES data than in the police-reported crash data from the five states.

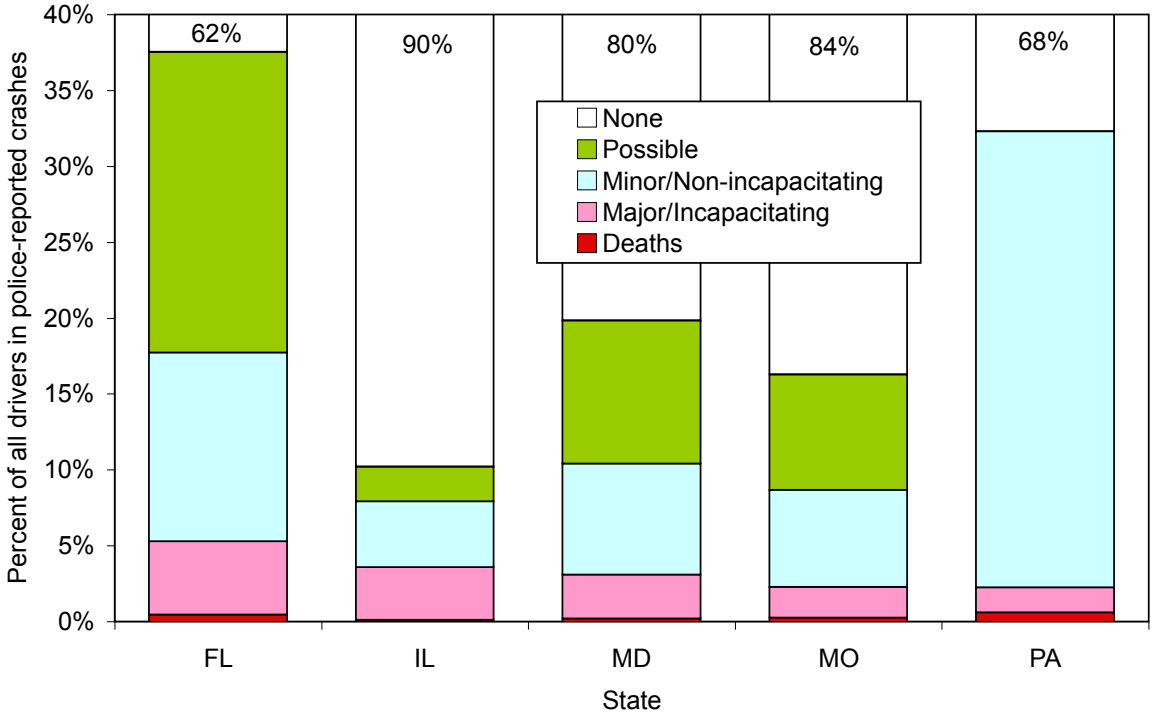
**Figure 4.3. US GES and state data driver casualty risk, and US FARS fatality risk, by vehicle type**




Different states have different requirements for whether a crash should be reported to police or included in a statewide database. These requirements often take the form of a threshold damage amount; crashes are only required to be reported, or included in a database, if the total damage exceeds a certain dollar amount. Figure 4.4 shows the distribution of all police-reported crashes by the most serious injury outcome of the crash, for the five states. The figure indicates that, combined, serious injuries and fatalities account for less than 5% of all drivers in police-reported crashes, from about 5% in Florida to about 2% in Missouri and Pennsylvania.

Note that there is a wide range in non-injury crashes reported in each state, from 62% of all reported crashes in Florida, to about 90% of all reported crashes in Illinois. Since non-injury crashes are the least likely to be reported to police, this gives an indication of which states' data may be biased by unreported non-injury crashes: Florida and to a lesser extent Pennsylvania. The relatively high fraction of serious injury and fatality crashes in Illinois suggests that Illinois has a higher threshold for inclusion of a crash in their database than the other states.

**Figure 4.4. Distribution of drivers in 2000-05 police-reported crashes, by injury severity and state**




It is possible that a substantial number of non-injury crashes in Florida are not reported; substantial under-reporting of non-injury crashes would bias our measure of exposure, all police-reported crashes, and inflate our risk estimates.

The apparent under-reporting of non-injury crashes in Florida biases our estimate of casualty risk per crash in that state; as shown in Figure 3.9 above, Florida has the highest casualty risk per crash of the five states. Indexing the risk by vehicle type to the risk for all vehicles in each state is one method to reduce the bias, as shown in Figure 3.10 above.

In its 2003 report, NHTSA speculated that the high fatality rate in Florida (from FARS) was in part caused by a large number of high-risk young drivers in that state. However, Figure 4.5 indicates that Florida has an average, if not slightly low, fraction of young male drivers (and a slightly high fraction of elderly drivers) involved in all police-reported crashes, relative to the other states. Florida has a slightly higher percentage of 18 to 20 year-old drivers, but a lower percentage of 21 to 25 year-old drivers, than the other states, as shown in Figure 4.6.

**Figure 4.5. Distribution of drivers in all police-reported crashes, by age group and state**

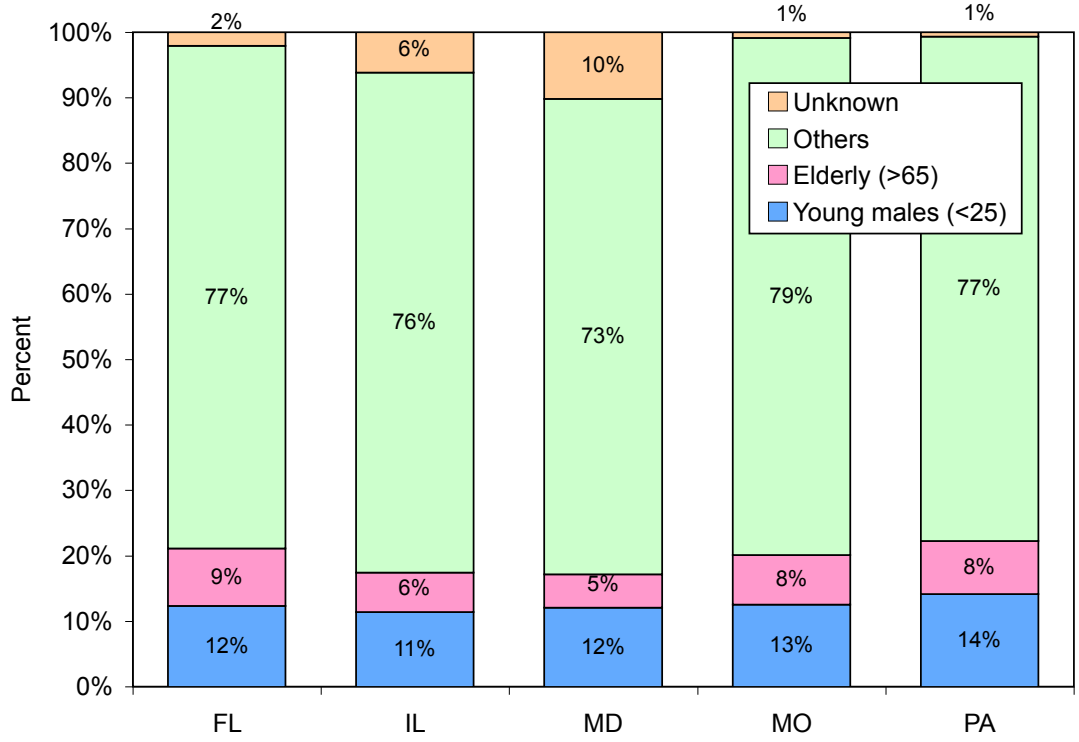
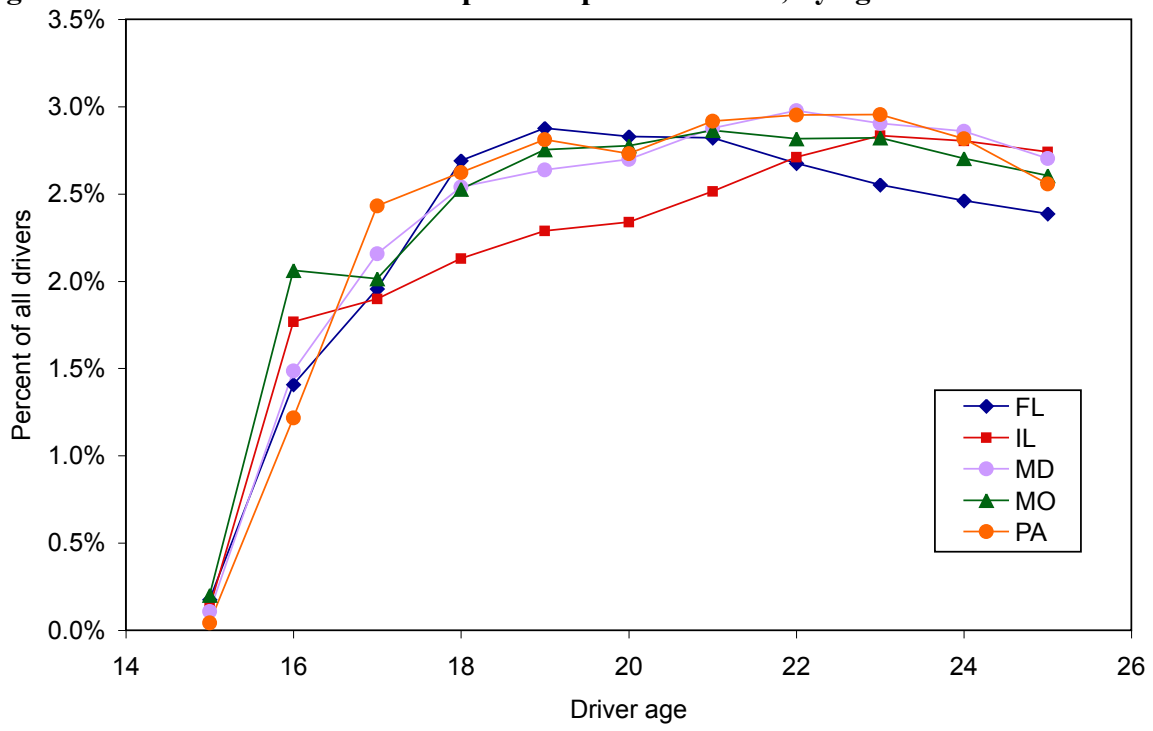



**Figure 4.6. Percent of drivers in all police-reported crashes, by age and state**

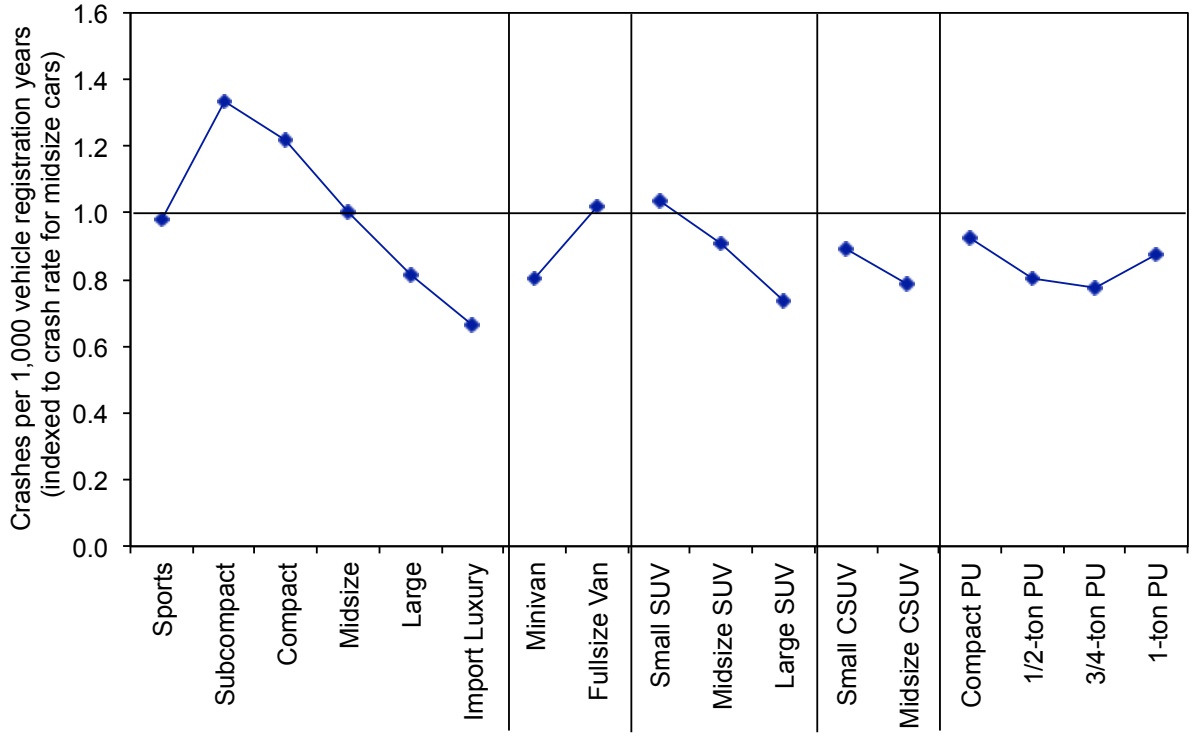


### 4.2. Reporting bias by drivers

If crash rates are under-reported, for whatever reason, actual risk per crash would be lower than the risks per crash we estimate here. For example, if crashes involving pickup trucks were less likely to be reported than crashes involving other vehicle types, the risk per crash in pickups would be over-stated relative to the risk for other vehicle types.

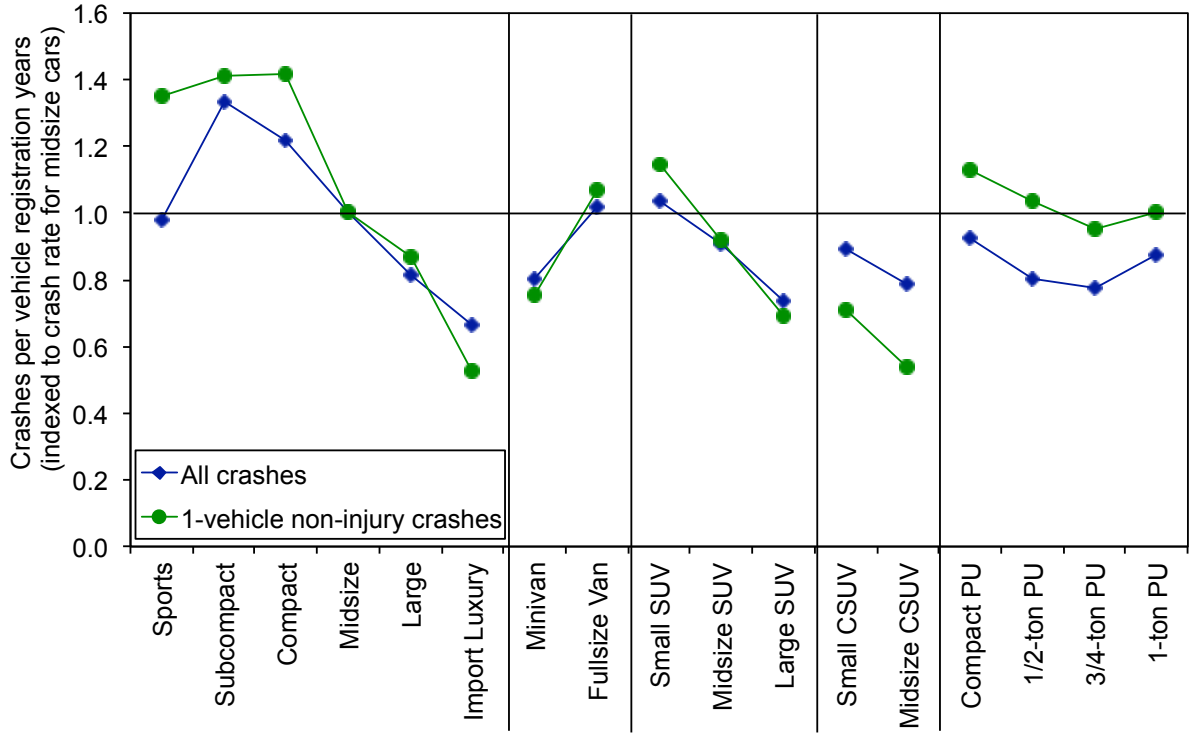
Figure 4.7 compares crash rates per vehicle registration-years by vehicle type, for all five states combined. The crash rate for each vehicle type is indexed to the crash rate for midsize cars. Although pickups have lower crash rates than midsize cars, they are about the same as most other vehicle types. Only subcompact and compact cars have higher crash rates than midsize cars.

**Figure 4.7. Crashes per 1,000 vehicle registration-years, by vehicle type and state**




We suspect that one-vehicle, non-rollover, non-injury crashes by pickup trucks are under-reported. Two-vehicle crashes are more likely to be reported, because two parties are involved, while rollover and injury crashes are more likely to be reported because they tend to be more severe. If pickup truck owners were not reporting one-vehicle non-injury crashes, we would expect the crash rate for pickup trucks in one-vehicle non-injury crashes to be lower relative to that of other vehicles than in all crashes. Figure 4.8 compares the number of one-vehicle, non-injury crashes per 10,000 vehicle registration years, and the number of all crashes per 1,000 registration years, by vehicle type; again, crash rates for each type of crash are indexed to that for midsize cars. Pickup trucks have relatively higher crash rates in one-vehicle non-injury crashes (in green) than in all crashes combined (in blue), relative to the crash rates of other types of vehicles. This suggests that pickup truck owners are not under-reporting the type of crash least likely to be reported, one-vehicle, non-injury crashes, in the state crash databases.

**Figure 4.8. One-vehicle, non-injury crashes and all crashes per 10,000 vehicle registration-years, by vehicle type and state**



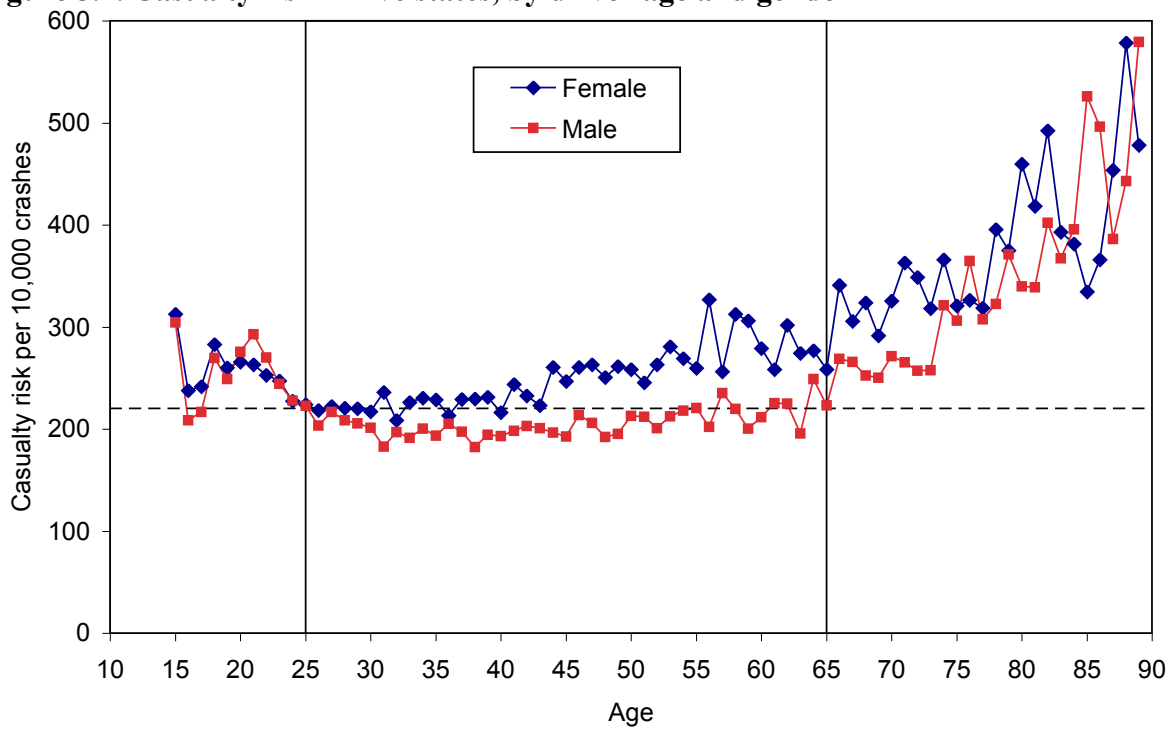
## 5. Accounting for driver behavior

Driver behavior is perhaps the most important contributor to whether a vehicle is involved in a crash. Therefore, risks by vehicle type and make/model should account for differences in driver behavior across vehicle types and makes/models. However, using police-reported crashes as the measure of exposure, rather than registered vehicles or miles driven, in a sense already accounts for driver behavior, in that vehicle models that have high crash involvement rates (presumably because of the "bad" or antisocial behavior of their drivers) have fewer casualties per crash than casualties per registered vehicle or annual miles driven.

### 5.1. Accounting for driver age and gender

Perhaps the simplest method to account for driver behavior is to account for driver characteristics, using age and gender. Figure 5.1 shows the casualty risk for model year 2000 to 2004 cars in the police-reported crash data for the five states; the dashed line denotes a casualty risk of 220. The figure indicates that casualty risks are high (above 220) for all drivers in their teens and early twenties; however, while risks remain consistently low (below 220) for males through age 65, risks for women increase consistently starting in their 40s, so that risks for women at age 50 are about the same as risks teenagers. For all drivers, casualty risks increase dramatically after age 65. Higher casualty risks for women, and older drivers, are not necessarily the result of risky driving by these population groups; rather, they are likely an indication of their relative frailty and capability of surviving a serious crash uninjured.

**Figure 5.1. Casualty risk in five states, by driver age and gender**




Accounting for driver age and gender can be done easily using state crash data by simply excluding both crashes and casualties from the analysis that involve drivers with risky

characteristics, i.e. young drivers and old drivers (Wenzel, 2010). Figure 5.2 compares casualty risk per police-reported crash, by driver and vehicle type. The figure indicates that elderly drivers have higher casualty risk than other drivers, in virtually all vehicle types. However, the figure also indicates that, for most vehicle types, young drivers have the same casualty risk as middle age drivers, and all drivers. Young females have relatively low casualty risk in sports cars, while young males have relatively low risk in small CSUVs, relative to middle aged drivers. Conversely, young females have relatively high risks in fullsize vans, midsize CSUVs, and all sizes of pickups, while young males and females have relatively high casualty risks in import luxury cars, midsize and large SUVs, and 1/2-ton pickups.

**Figure 5.2. Effect of driver age and gender on driver casualty risks by vehicle type**

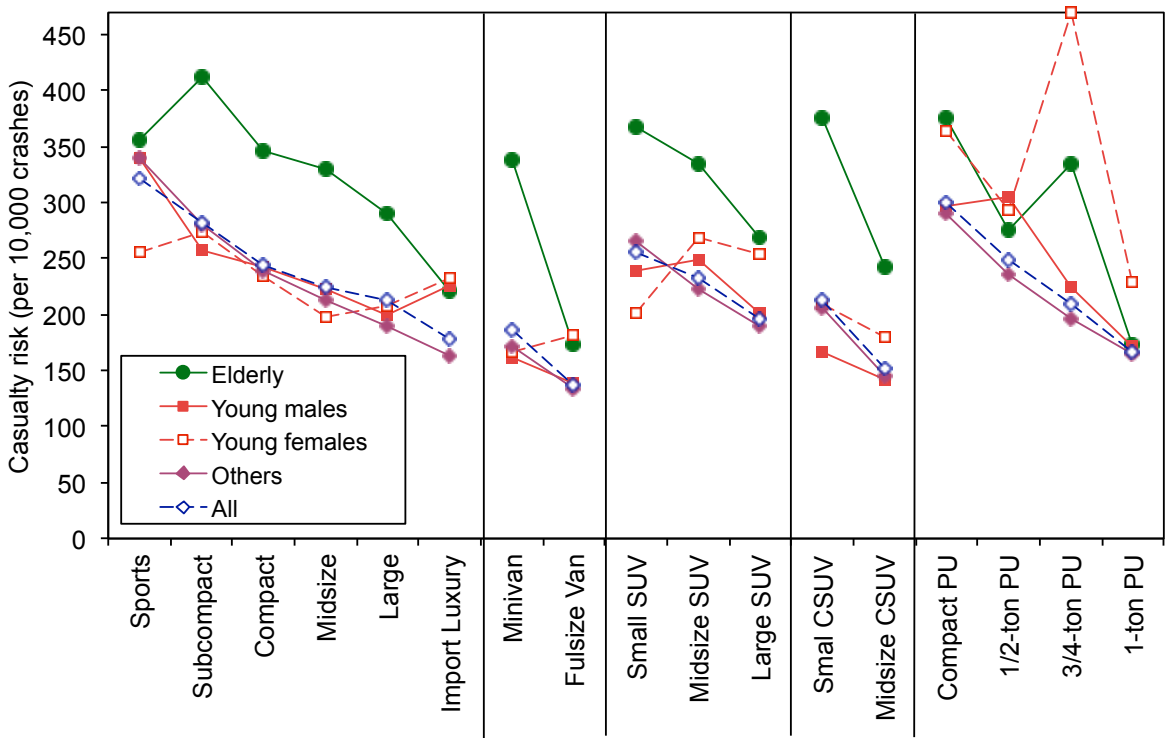



Figure 5.2 suggests that calculating risk per police-reported crash accounts for much of the effect of driving behavior observed in calculating risk per vehicle or vehicle-mile.

## 5.2. Accounting for bad driver behavior

Driver age and gender is a crude measure of driving behavior; a more accurate measure would be the circumstances of the current crash, and past driving record. In its analysis of FARS data, NHTSA developed a variable, "bad driver rating," based on whether alcohol, drugs, or reckless driving was involved in the current crash, as well as the driver's driving record over the last three years. Figure 5.3 shows that young (under 25) males have the highest, and elderly (over 65) drivers the lowest, bad driver rating, across all vehicle types, using FARS data for model years 2003 through 2007.

Note that young females have much lower bad driver ratings than young males; this suggests that the high casualty risks for women compared to men, shown in Figure 5.1 above, are the result of

their relative frailty rather than risky driving behaviors on their part. Similarly, elderly drivers are very safe drivers, but they face high casualty risks because of their frailty.

The blue line in Figure 5.3 shows the bad driver rating for all drivers, while the purple line shows the rating after removing young and elderly drivers. Removing these driver classes has little impact on bad driver ratings for most vehicle types; the exception is sports cars and subcompact cars, whose bad driver ratings are lowered substantially after removing these driver classes.

Although some states report whether alcohol or drugs were involved in the crash, no state reports information on the driver's record. NHTSA's National Analysis Sampling System (NASS) General Estimates System (GES) is a national sample of police-reported crashes. Alcohol or drug involvement, speeding, and reckless driving are consistently reported in the NASS GES. We looked at the fraction of drivers of 2000 to 2004 vehicles receiving a citation for one of these offenses, by vehicle type and driver, as shown in Figure 5.4. Figure 5.4 indicates a similar trend as Figure 5.3.

Figure 5.5 shows casualty risk by vehicle type using GES data for model year 2000 through 2004 vehicles, through 2007, for both "bad" drivers as defined above (i.e. drivers subsequently cited for alcohol or drug use, speeding, reckless driving, or driving on a suspended or revoked license) and "not bad" drivers who were not cited in the current crash. The casualty risks for bad drivers are substantially higher than those for not bad drivers, for all vehicle types; this suggests that bad drivers are involved in a greater number of serious crashes, perhaps involving higher speeds, than not bad drivers.

Figure 5.6 shows the same data as Figure 5.5, but adjusts the scale for the risks for not bad drivers, so that the risks for both bad and not bad drivers of midsize cars overlap. Figure 5.6 indicates that, for certain vehicle types (import luxury cars, SUVs, crossover SUVs, and pickups), bad drivers account for a greater increase in casualty risk than they do for most types of cars (including sports cars), minivans and fullsize vans.

Each of the five states records whether the driver was, or likely was, under the influence of alcohol or drugs at the time of the crash; however, not all states consistently report posted speed limit and likely travel speed, or whether the driver was speeding or driving recklessly, or driving with a suspended or revoked license. We hope that enough of the other states report this information so we can test what influence "bad" driving behavior has on casualty risks using the police-reported crash databases from the states.

**Figure 5.3. Bad driver rating by driver age and gender and vehicle type, US fatal crash data (FARS)**

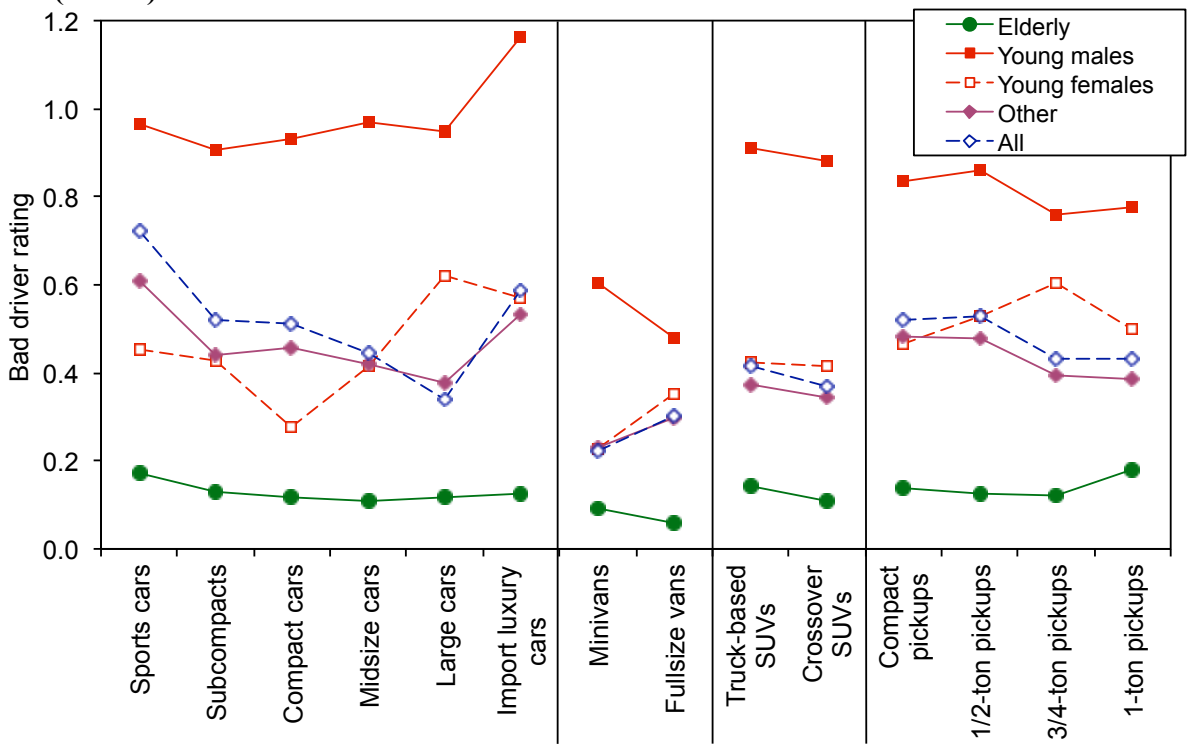



**Figure 5.4. Bad driver rating by driver age and gender and vehicle type, US police-reported crash data (NASS GES)**

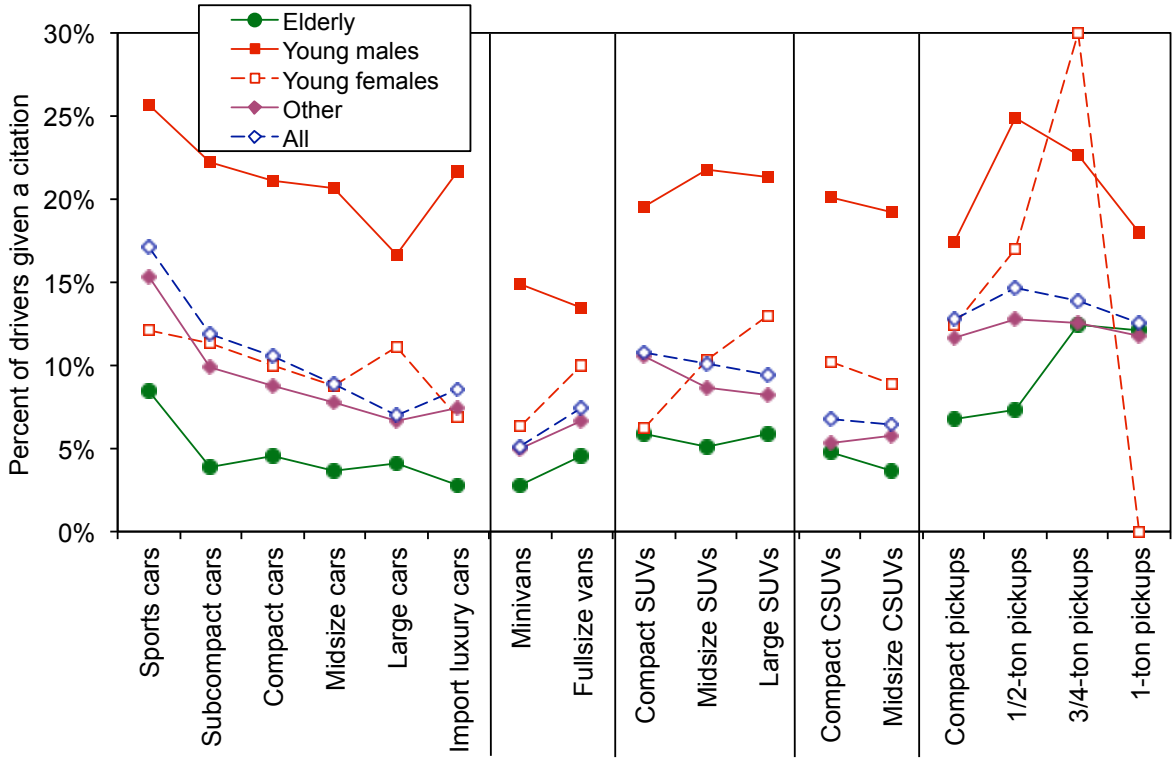


**Figure 5.5. Casualty risk by vehicle type and driver behavior, NASS GES data**

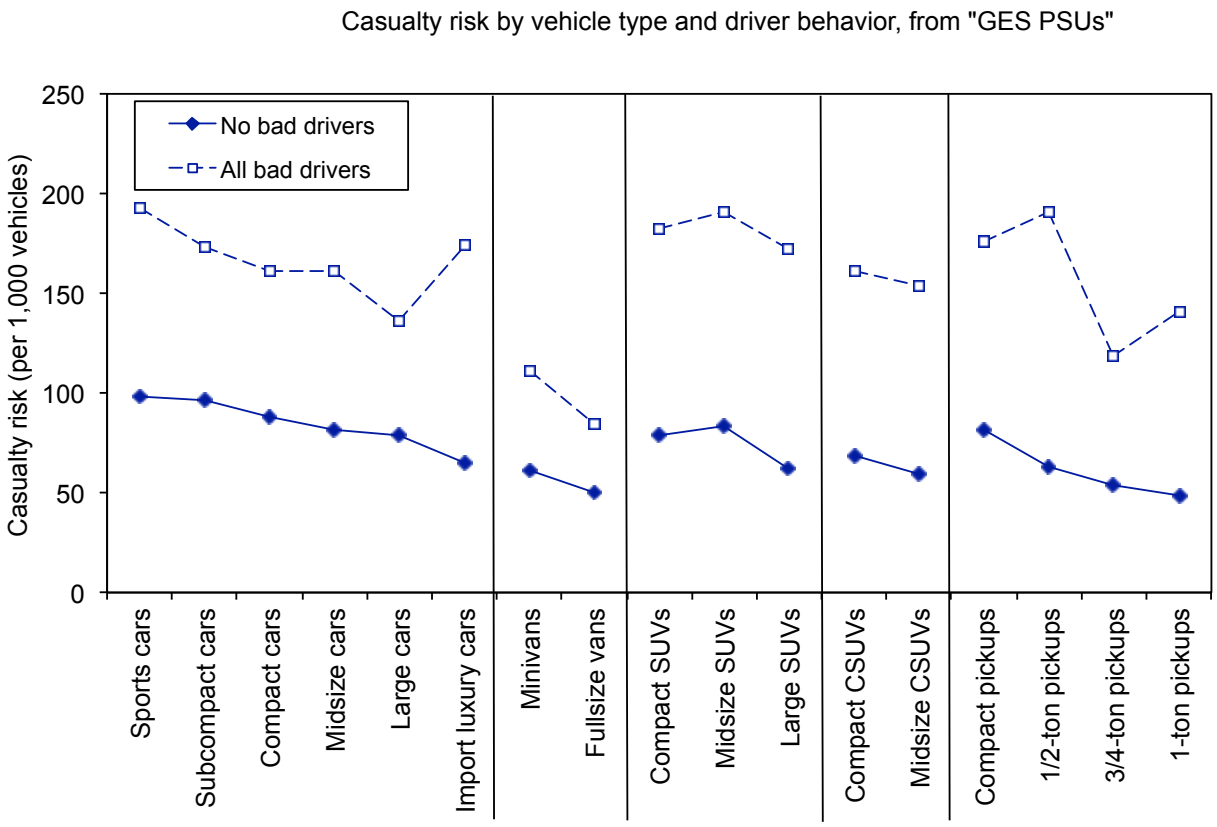
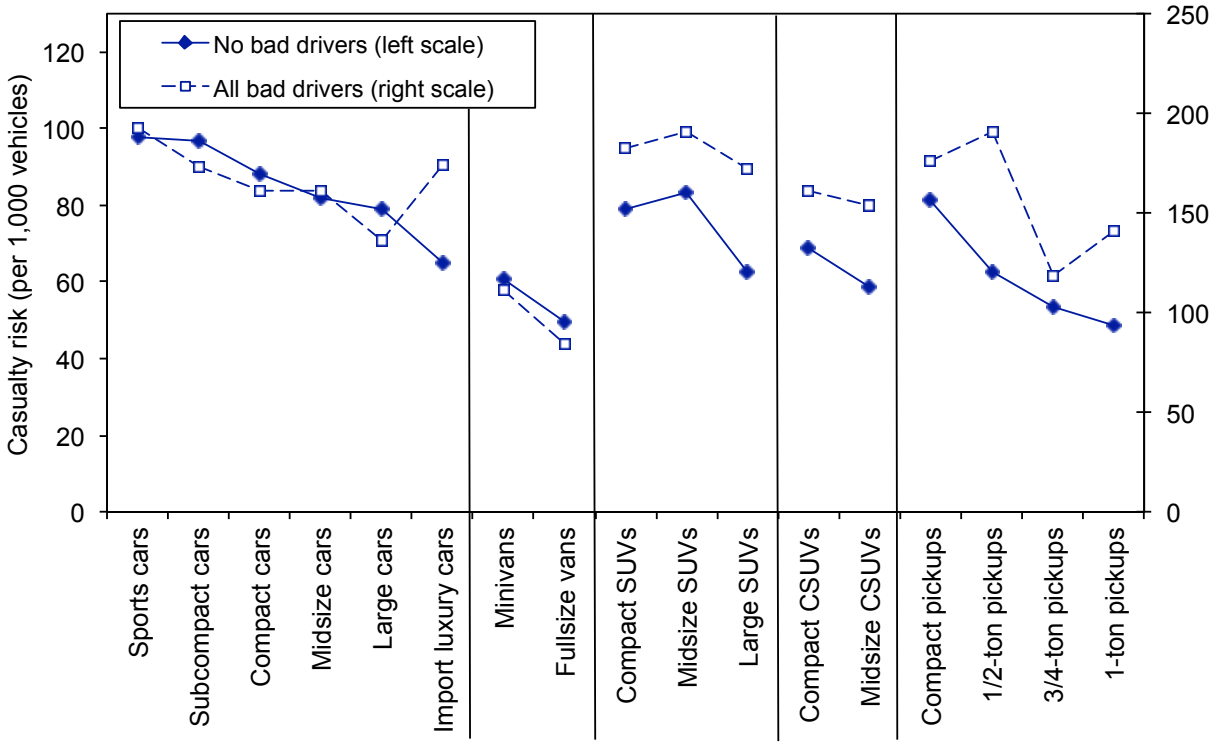


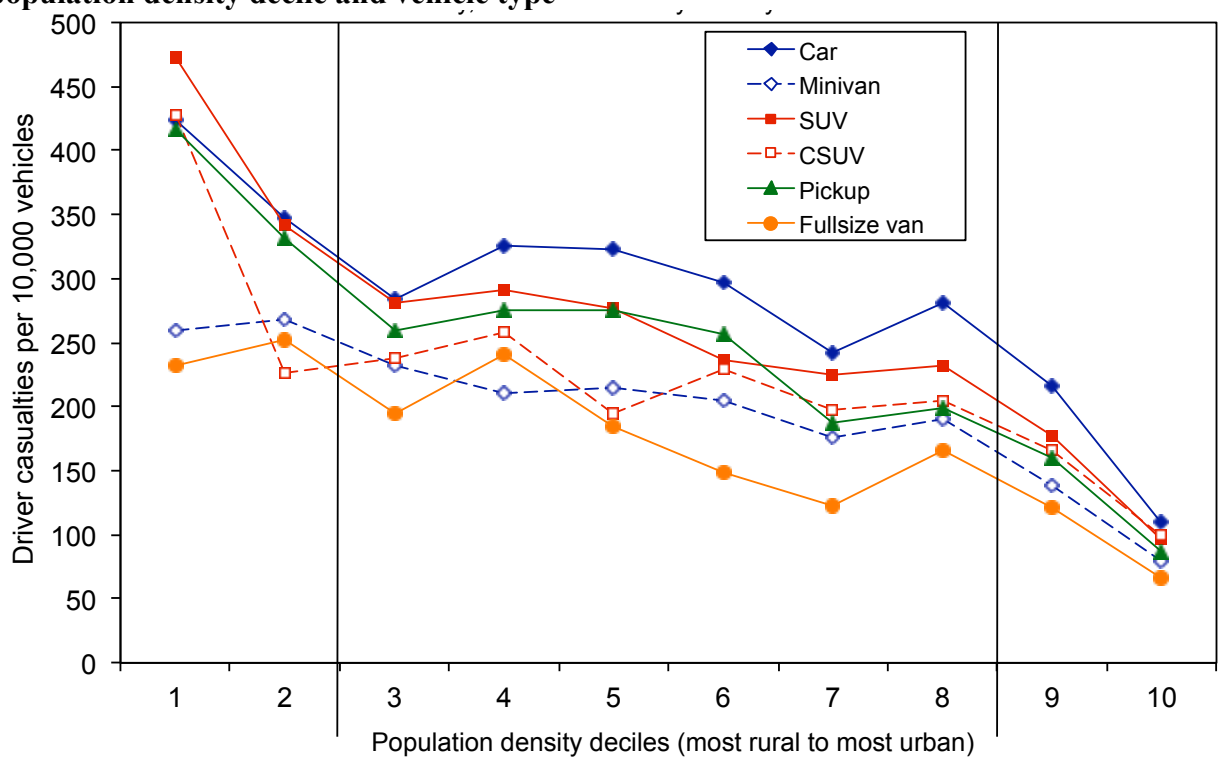

**Figure 5.6. Casualty risk by vehicle type and driver behavior, NASS GES data (adjusted scale)**



## 6. Accounting for crash location (urban vs. rural counties)

As shown in Figure 5.2 above, excluding young male and elderly drivers from the analysis has little effect on casualty risks; however, Figure 6.1 indicates that casualty risk decreases as county population density increases, with the highest casualty risks in the most rural counties, and the lowest risks in the most urban counties, for all vehicles types. However, because a large fraction of pickup trucks are driven in rural areas (Figure 6.2), accounting for the population density of the county in which a crash occurs reduces the casualty risk in pickup trucks by about 15%, relative to that of other vehicle types (Figure 6.3).

**Figure 6.1. Casualty risk for drivers (excluding young males and elderly drivers), by population density decile and vehicle type**




We used the counties with emission inspection and maintenance programs to assess how casualty risk and crash rates vary in urban (I/M) and rural (non-I/M) counties. Figure 6.4 indicates that casualty risk per crash in rural areas (right-hand scale) is roughly twice that in urban areas (left-hand scale); however, casualty risk per vehicle registration-year in rural areas is about ten times that in urban areas. Figure 6.5 indicates that crash rates are about four times higher in rural counties (right-hand scale) than in urban counties (left-hand scale).

Because crash frequency in rural counties is four times higher than in urban counties, the effect of driving in rural counties on casualty risk per crash is smaller (twice as high as in urban counties) than the effect on casualty risks per vehicle registration-year (ten times as high in rural counties as in urban counties).

**Figure 6.2. Distribution of police-reported crashes, by population density decile and vehicle type**

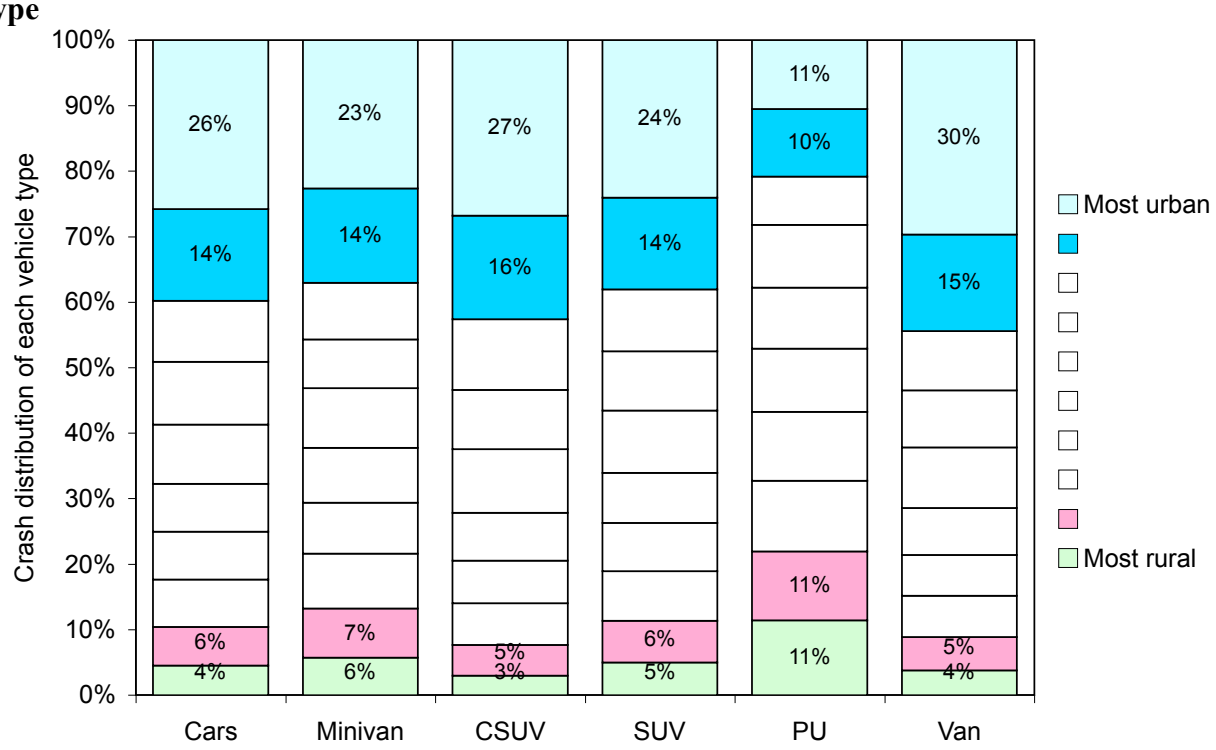



**Figure 6.3. Casualty risk per 10,000 crashes, by crash location and vehicle type**

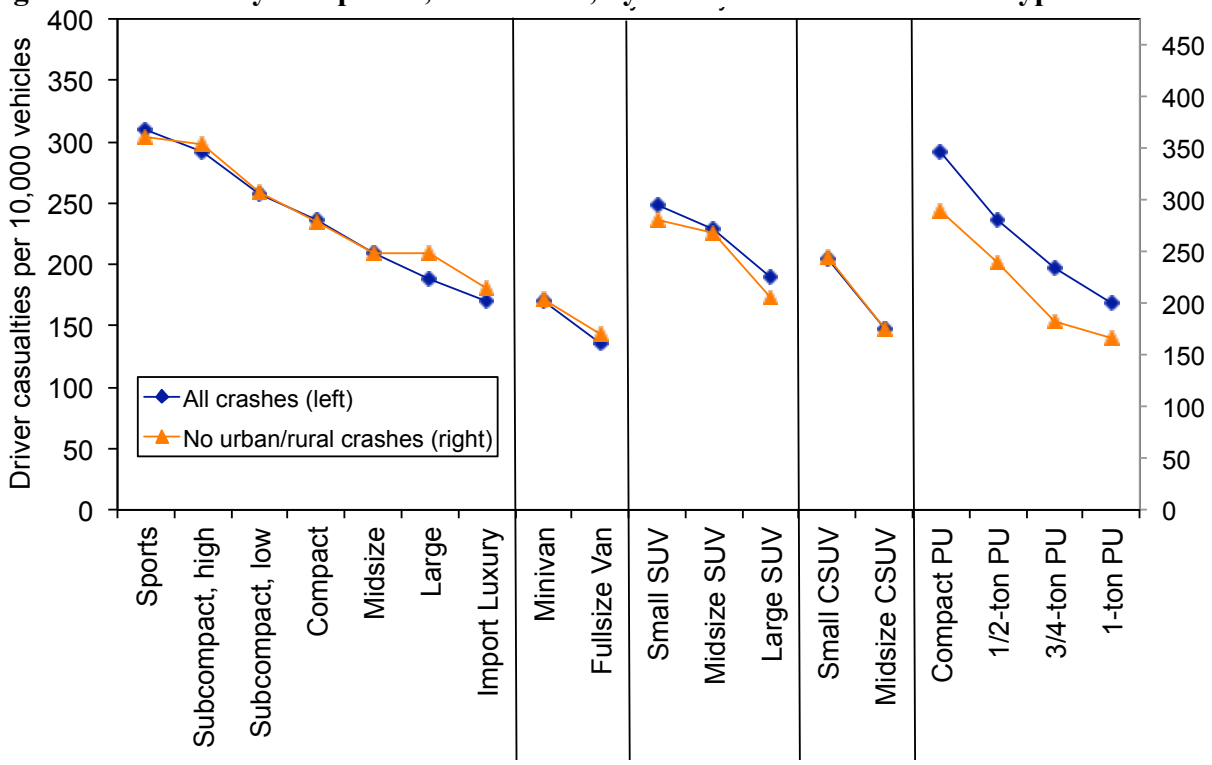


**Figure 6.4. Five state casualty risks per vehicle registration-year and per crash, in urban I/M areas and rural non-I/M areas, by vehicle type**

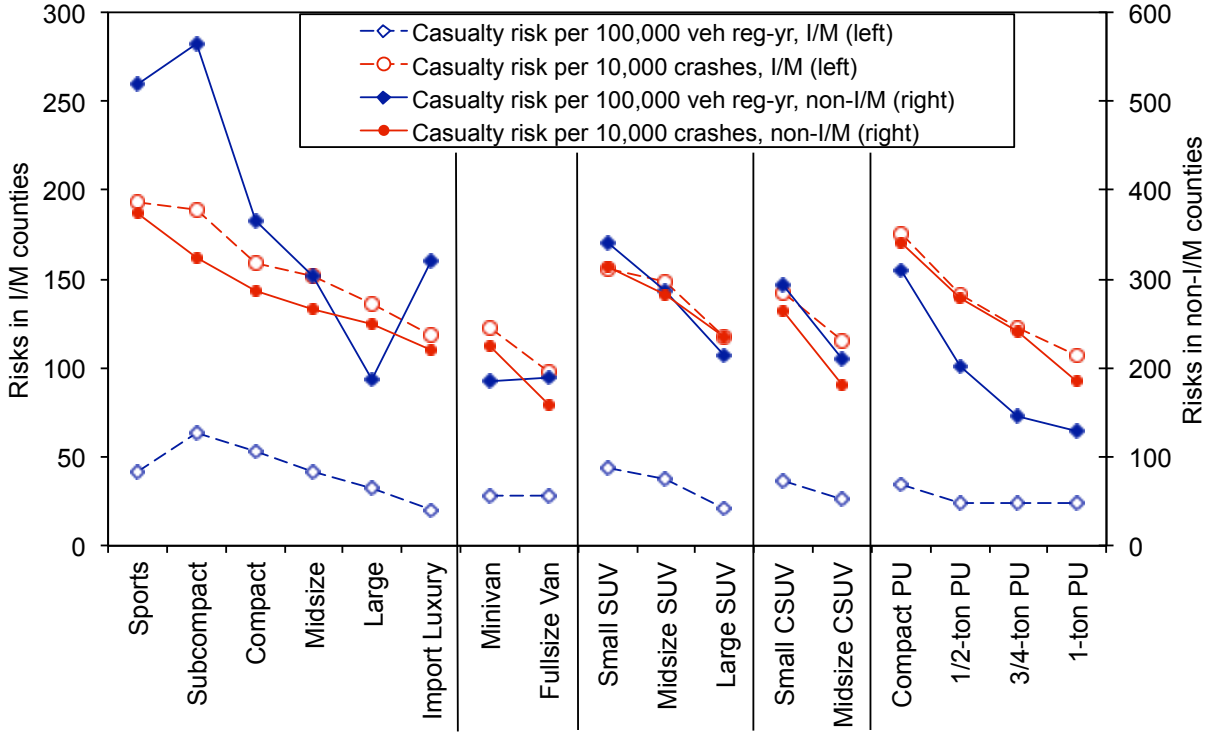
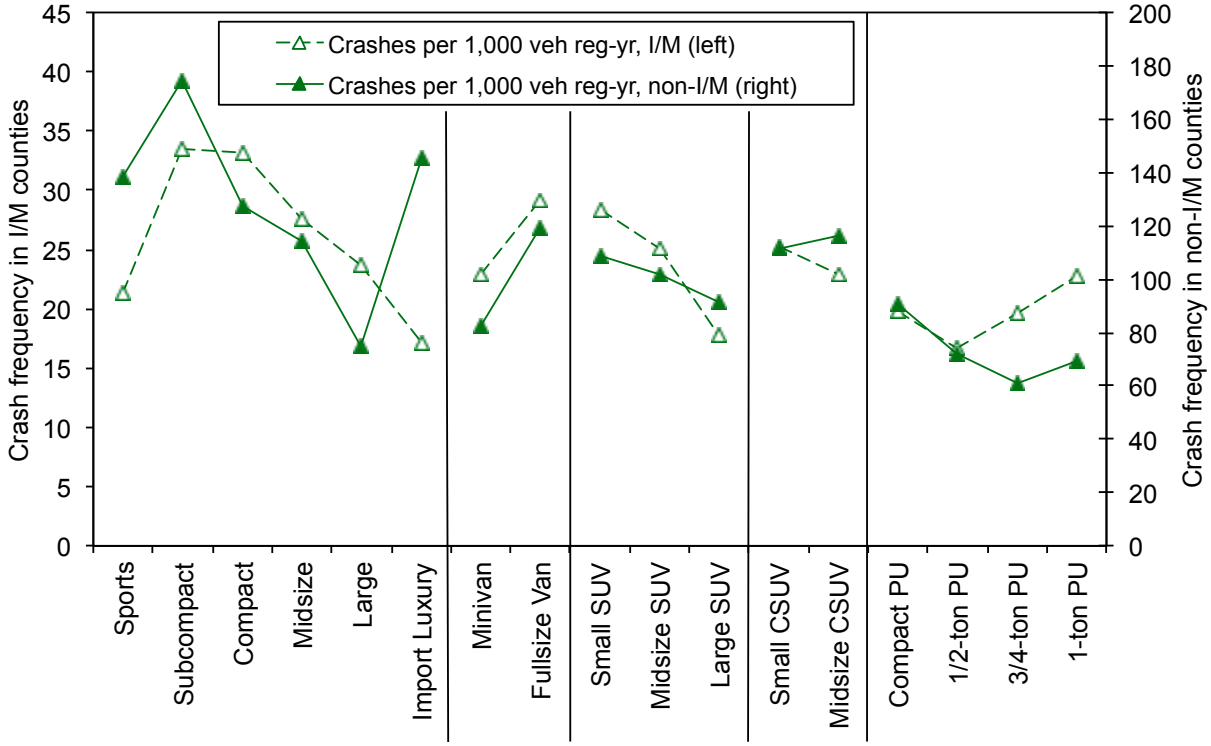


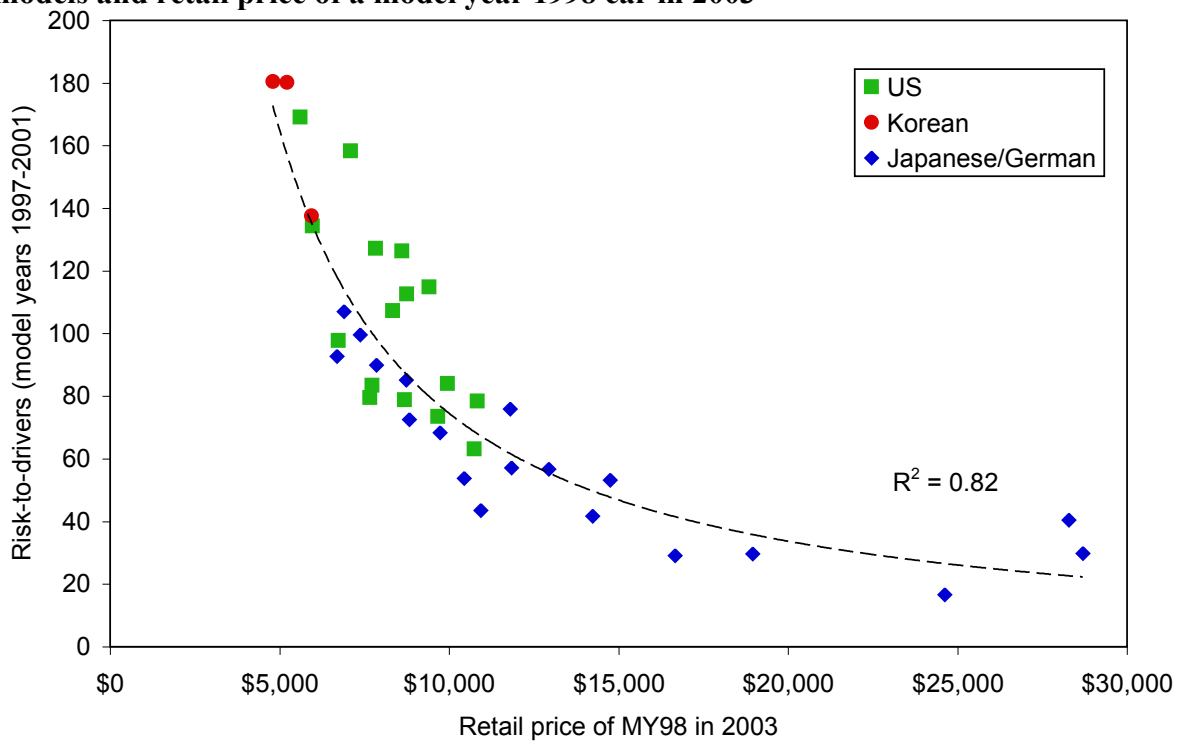

**Figure 6.5. Five state crashes per 1,000 vehicle registration-year, in urban I/M areas and rural non-I/M areas, by vehicle type**



## 7. Assessing general quality of vehicle design

Wenzel and Ross found that there was a strong correlation between driver fatality risk (uncorrected for driver behavior or crash location) and resale value (as defined by Kelley Blue Book national retail price of a 1998 model in 2003) by car model, as shown in Figure 7.1 (Wenzel and Ross, 2008). We hope to obtain data on vehicle resale value from Kelley Blue Book in order to assess whether the general quality of vehicle design, as measured by resale value or resale value per vehicle pound, has an influence on casualty risk.

**Figure 7.1. Relationship between fatality risk to drivers of model year 1997 to 2001 car models and retail price of a model year 1998 car in 2003**

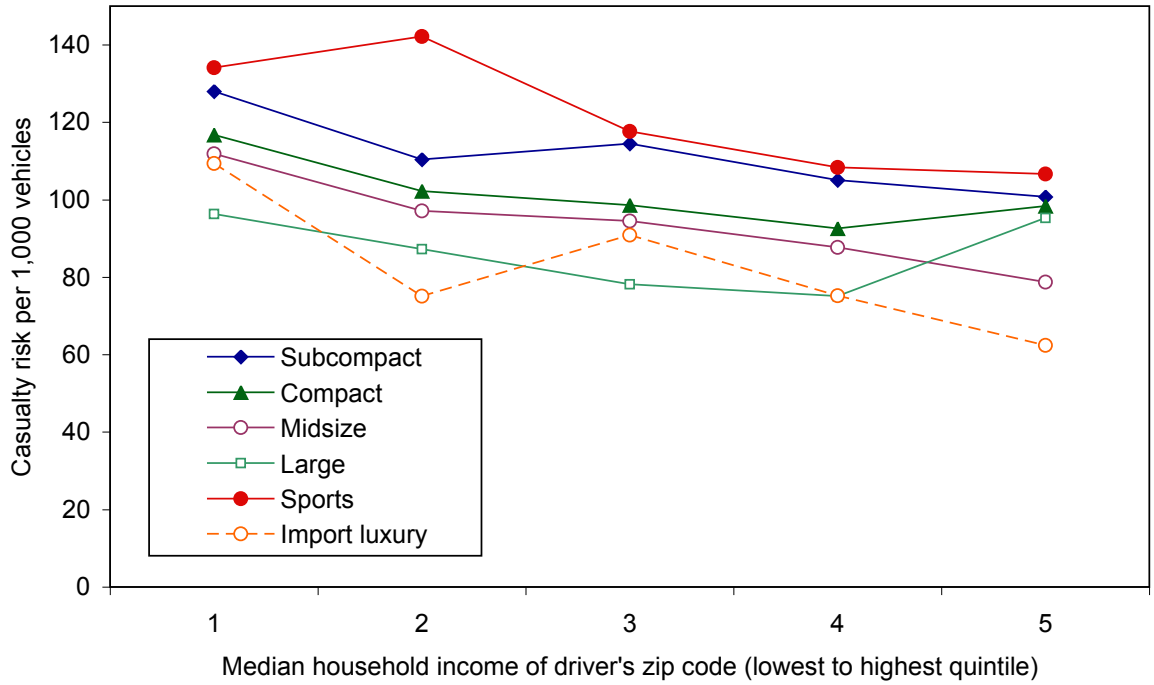



In its response to comments on the 2003 report, NHTSA addressed this issue, by adding car price (in addition to driver age and gender, and crash location and state) to its regression models, and found that controlling for car price had little effect on the relationship between vehicle weight and safety, for all crash types (Kahane, 2004). NHTSA speculated that driver behavior, and not any general aspect of vehicle design, causes the reduction in risk as car resale value or price increases. To test this theory, NHTSA ran a regression model for head-on collisions between two cars, under the assumption that such crashes are not influenced by driver behavior. The model suggested that more expensive cars do not have higher crashworthiness in head-on collisions with other cars.

One factor that may confound an analysis of the relationship between vehicle resale value and casualty risk is driver income. For example, is the relatively low risk in import luxury cars a function of their design or the behavior of their drivers? Or is there a driver income effect independent of vehicle design or driver behavior? Figure 7.2 plots the relationship between casualty risk per crash, using NASS GES data through 2007, and the median household income

of the driver's zip code (based on the driver license zip code), by car type. There clearly is a relationship for cars, with casualty risk per crash decreasing as driver income increases.

**Figure 7.2. Casualty risk by car type and median household income, model year 2000 to 2004 vehicles in NASS GES through 2008**




Figures 7.3 and 7.4 plot the same data for the ten most prevalent 2000 to 2004 vehicle models in the national GES crash data between 2000 and 2007. Plotting the relationship between casualty risk and household income for vehicle models accounts for differences in the design of vehicle models. Figure 6.3 shows a fairly consistent relationship between casualty risk and household income for five vehicle models, with risk decreasing as income increases. However, there is no similar relationship for the five models shown in Figure 7.4.

The relationship between casualty risk and driver income, with risk decreasing as income increases, could be explained by better driving behavior by drivers with higher incomes. However, as shown in Figure 7.5, the percent of bad drivers does not increase with increasing income, with the possible exception of drivers with the highest incomes.

Unfortunately, none of the five states for which we have police-reported crash data report the zip code for the drivers involved in crashes. We hope that this information is reported in the crash databases of some of the additional states we will analyze as part of the expanded study. If so, we may be able to more definitively analyze the relationship between casualty risk and driver income, independent of differences in vehicle design.

**Figure 7.3. Casualty risk by vehicle model and median household income, selected model year 2000 to 2004 models in NASS GES through 2008**

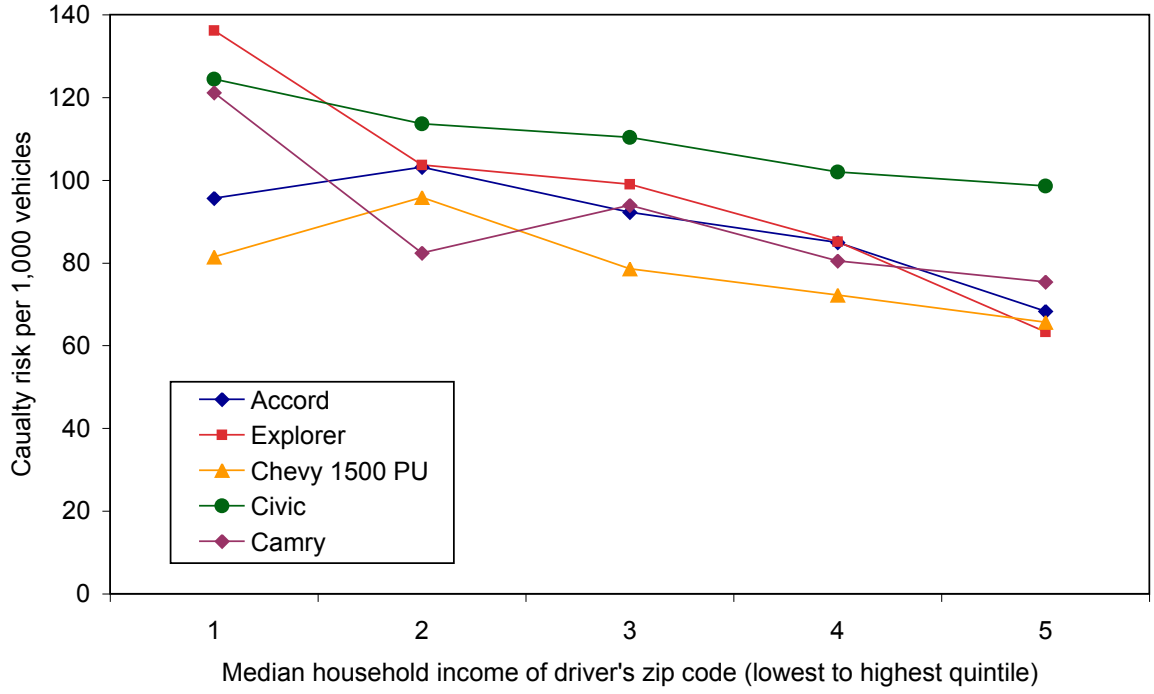



**Figure 7.4. Casualty risk by vehicle model and median household income, selected model year 2000 to 2004 models in NASS GES through 2008**

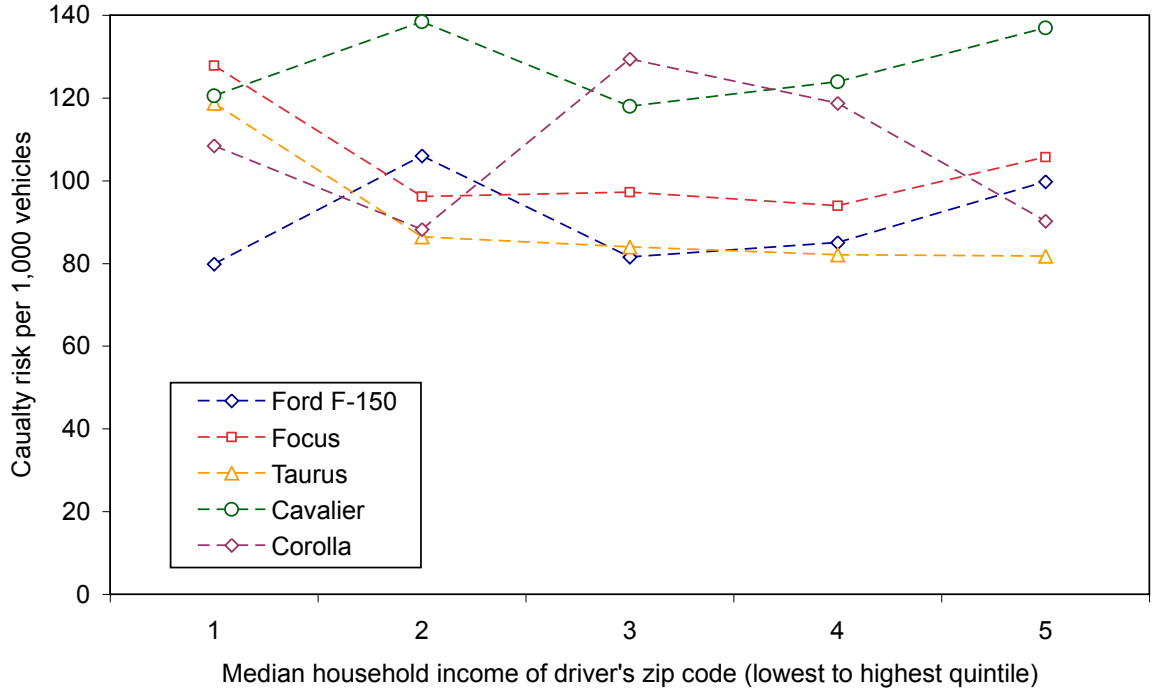
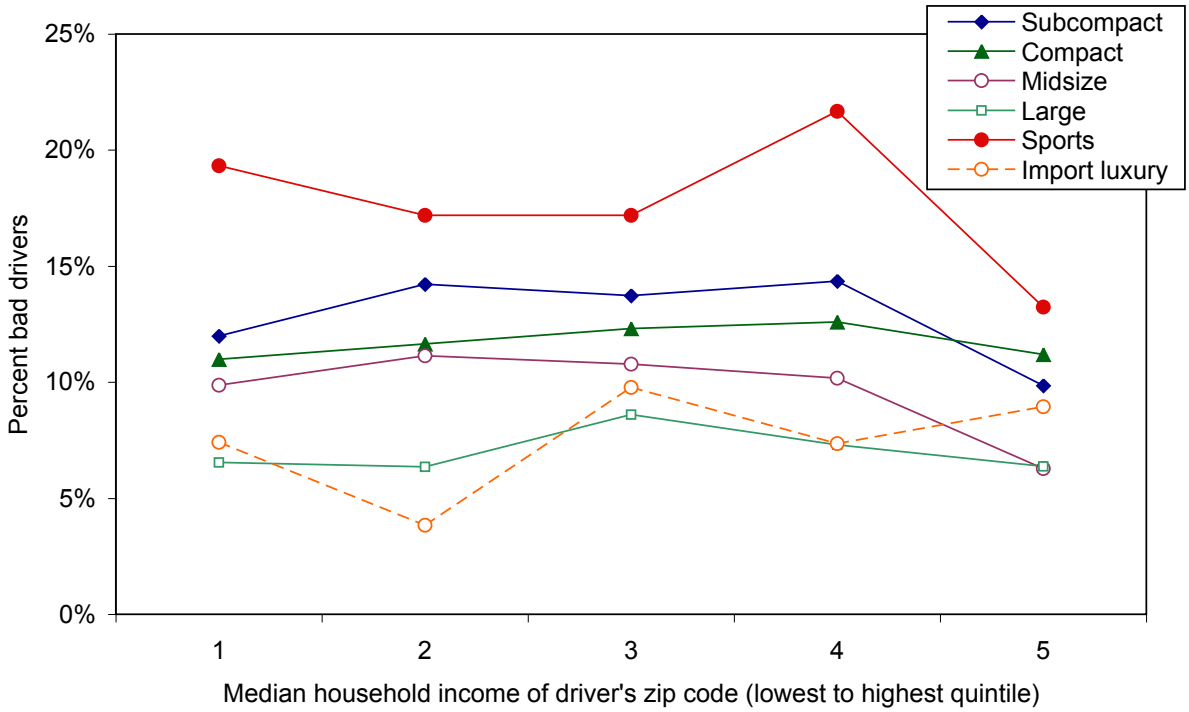


**Figure 7.5. Percent bad drivers by vehicle type and median household income, model year 2000 to 2004 models in NASS GES through 2008**



## 8. Conclusions

A combination of corrections made to the truncated VINs provided in the state crash databases increased the number of model year 2000 to 2004 vehicles available for analysis by 18%.

Our analysis of the two types of risk, fatality risk per registered-vehicle and casualty risk per police-reported crash, indicates that there is reasonable agreement between the two types of risk, especially by vehicle type. Based on our analysis of data from five states, there is good agreement between fatality and casualty risk per registered vehicle across most vehicle types, although sports cars and pickup trucks have substantially lower casualty risk than fatality risk. Using police-reported crashes rather than vehicle registration-years as the measure of exposure results in relatively lower casualty risks for subcompact and compact cars, and relatively higher risk for large and import luxury cars, minivans, large SUVs, and ½-ton and ¾-ton pickups.

Adjusting casualty risks per registration year by the average odometer reading of individual vehicle models increases casualty risk of sports cars, by about 30%, as sports cars are driven many fewer miles than other vehicle types. However, adjusting for mileage has no effect on the casualty risks of other vehicle types.

There is a strong correlation between vehicle registration years and the number of vehicles involved in police-reported crashes, by vehicle model. This indicates that crashes, rather than registration-years, can be safely used as the measure of exposure.

Calculating risk using all police-reported crashes, rather than vehicle registration-years, as the measure of exposure changes the risk of certain vehicle types relative to that of other vehicle types. For example, subcompact and compact cars have higher crash involvement rates than midsize, large and import luxury cars, so the difference in casualty risk between smaller and larger cars is dramatically reduced when the measure of exposure is changed from registration-years to all police-reported crashes. On the other hand, midsize and large cars, and minivans, have relatively low crash involvement rates, so changing the measure of exposure to crashes increases their casualty risk relative to that of other vehicle types. Similarly, the relatively low crash involvement rates of large SUVs, crossover SUVs, and fullsize pickups results in relatively higher casualty risks when the measure of exposure is changed from vehicle registration-years to all police-reported crashes. In general, changing the measure of exposure from registration years to all police-reported crashes reduces the sharp decline in casualty risk as car size increases. Casualty risk per crash does decrease as size increases, within each major vehicle type, but for cars and truck-based SUVs the decrease is not as large as when registration-years is used as the measure of exposure.

National GES casualty risks are remarkably similar to the casualty risks from all police-reported crashes in the five states, for all vehicle types except pickup trucks, which suggests that the five-state casualty risks are representative of national casualty risks. However, pickup trucks do have substantially lower casualty risks in the national GES data than in the police-reported crash data from the five states.

Bias introduced by different states' definitions of police reported crashes, or serious injury crashes can be addressed by indexing risk by vehicle type or model to the risk for all vehicles in that state. This bias can be addressed in the logistic regression model by introducing a dummy variable for each state. Analysis of crash frequencies of one-vehicle, non-rollover, non-injury crashes suggests that crashes involving pickup trucks are not under-reported.

Elderly drivers (over 65 years old) have higher casualty risk than other drivers, in virtually all vehicle types. For most vehicle types, young drivers (under 25 years old) have the same casualty risk as middle age drivers, and all drivers. Apparently calculating risk per police-reported crash accounts for much of the effect of poor driving behavior of young drivers that is observed in calculating risk per vehicle or vehicle-mile.

Driver age and gender is a crude measure of driving behavior; a more accurate measure would be the circumstances of the current crash, and past driving record. Using NHTSA's "bad driver rating" variable and FARS data for model years 2003 through 2007, young males have the highest, and elderly drivers the lowest, bad driver rating, across all vehicle types. Young females have much lower bad driver ratings than young males, suggesting that the higher casualty risks for women compared to men are the result of their relative frailty rather than risky driving behaviors on their part. Similarly, elderly drivers are very safe drivers, but they face high casualty risks because of their frailty.

Casualty risk decreases as county population density increases, with the highest casualty risks in the most rural counties, and the lowest risks in the most urban counties, for all vehicles types. However, because a large fraction of pickup trucks are driven in rural areas (22%, as opposed to 10% for other vehicle types), accounting for the population density of the county in which a crash occurs reduces the casualty risk in pickup trucks by about 15% relative to that of other vehicle types. Because crash frequency in rural counties is four times higher than in urban counties, the effect of driving in rural counties on casualty risk per <u>crash</u> is smaller (twice as high as in urban counties) than the effect on casualty risks per <u>vehicle registration-year</u> (ten times as high in rural counties as in urban counties).

NASS GES data indicate that casualty risk per crash decreases as driver income increases, for each car type. Although risk consistently decreases as income increases for five of the ten most prevalent vehicle models, there is no similar relationship for the other five prevalent vehicle models; therefore it is not clear how strong the relationship between driver income and risk is when one accounts for vehicle model. The decrease in casualty risk as income increases is not the result of better driving behavior by drivers with higher incomes, as the percent of bad drivers does not increase with increasing income, with the possible exception of drivers with the highest incomes.

## 9. References

Kahane, C.J., 2003. *Vehicle weight, fatality risk and crash compatibility of model year 1991-99 passenger cars and light trucks*. NHTSA DOT HS 809 662. U.S. Department of Transportation, National Highway Traffic Safety Administration, Washington, D.C.

Kahane, C.J., 2004. *Response to docket comments on NHTSA technical report* "*Vehicle weight, fatality risk and crash compatibility of model year 1991-99 passenger cars and light trucks"*. Submission to docket no. NHTSA-2003-16318. U.S. Department of Transportation, National Highway Traffic Safety Administration, Washington, D.C.

Wenzel, T.P., Ross, M., 2005. "The effects of vehicle model and driver behavior on risk." *Accident Analysis and Prevention* 37: 479-494.

Wenzel, Tom and Marc Ross, 2008. "The Relationship between Vehicle Weight/Size and Safety," in *The Physics of Sustainable Energy: Using Energy Efficiently and Producing It Renewably.* Hafemeister, David, Barbara Levi, Mark Levine, Peter Schwartz, editors. The American Institute of Physics: Melville, New York.

Wenzel, Tom, 2010. *Analysis of the Relationship Between Vehicle Weight/Size and Safety, and Implications for Federal Fuel Economy Regulation*. LBNL-3143E

## Appendix

All of the figures in this report consider only the risk to drivers (not passengers) in 2000 to 2004 vehicles, and not the risks to drivers or passengers of other vehicles involved in the crash. Casualties to passengers are not included, as this might introduce bias to risks for vehicle types that typically have relatively high occupancy rates, i.e. minivans and fullsize vans.

Below are companion versions of Figures 3.2, 3.4, and 3.5 from the report; Figures A.1 through A.3 show the risk to drivers of other vehicles, and Figures A.4 through A.6 the combined, or societal, risk including drivers in both vehicles, in terms of fatality and casualty risk per vehicle registration-year and crash.

Analysis of the risks to drivers of other vehicles (Figures A.1 through A.3) suggests that risks to others increase as truck-based SUV and pickup size increases, and for fatality risk by fullsize vans. Comparison of the risks to drivers (Figures 3.2, 3.4, and 3.5) with the combined risks to all drivers (Figures A.4 through A.6) indicates that the trend in decreasing risk with increasing size is muted, and in the case of pickups reversed, when societal risk is considered. The relationship between combined casualty risk per vehicle and per crash (Figure A.5), and combined fatality risk per vehicle and per crash (Figure A.6), is quite similar to that when only the risk to drivers is considered (Figures 3.4 and 3.5), for cars and crossovers. However, societal casualty risk does not decrease as SUV or pickup size increases (Figure A.5 vs. Figure 3.4), and societal fatality risk actually increases as van and pickup size increases (Figure A.6 vs. Figure 3.5).

**Figure A.1. Five state fatality and casualty risks to others per vehicle registration-year, by vehicle type**

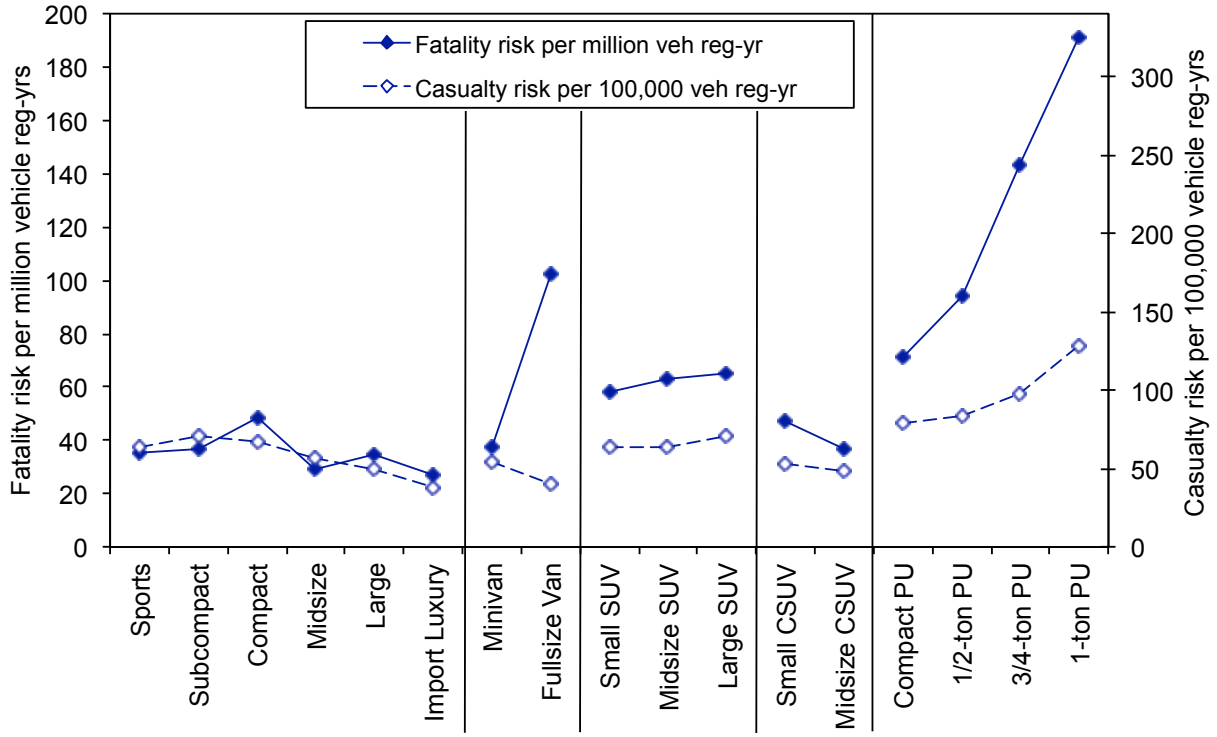


**Figure A.2. Five state casualty risks <u>to others</u> per vehicle registration-year and per-crash, by vehicle type**

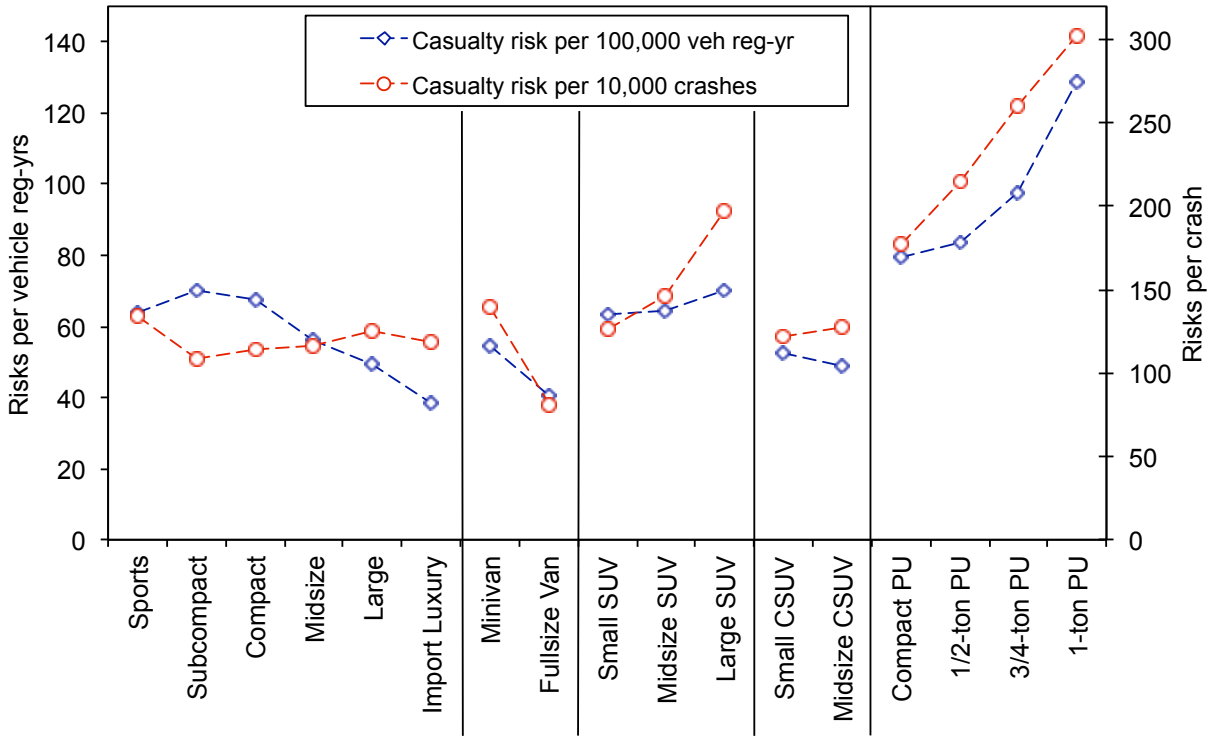



**Figure A.3. Five state fatality risks <u>to others</u> per vehicle registration-year and per crash, by vehicle type**

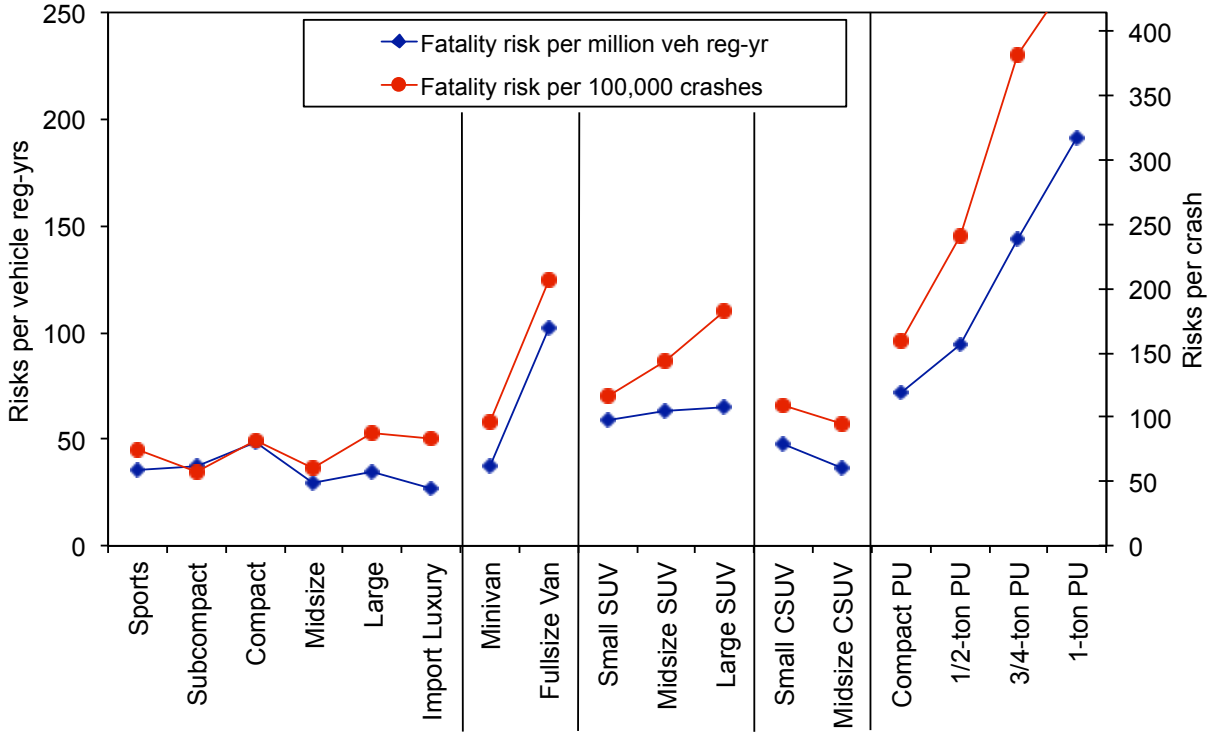


**Figure A.4. Five state fatality and casualty combined risks per vehicle registration-year, by vehicle type**

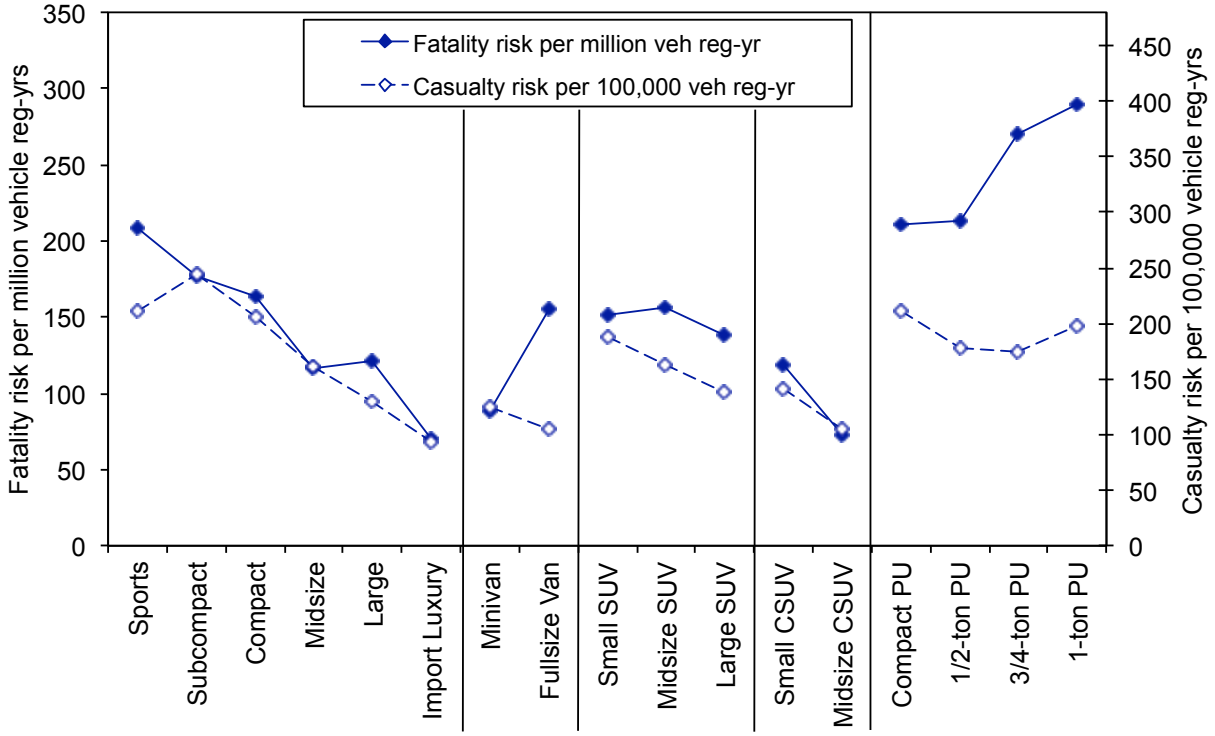



**Figure A.5. Five state casualty combined risks per vehicle registration-year and per-crash, by vehicle type**

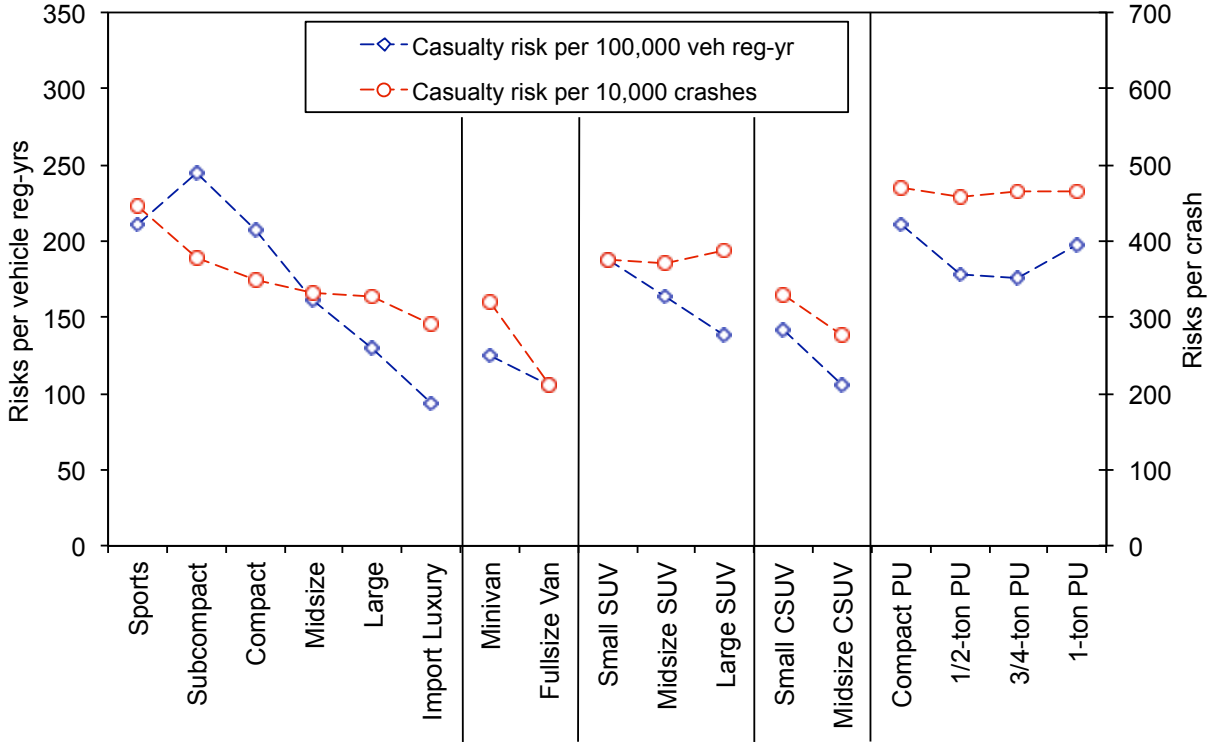
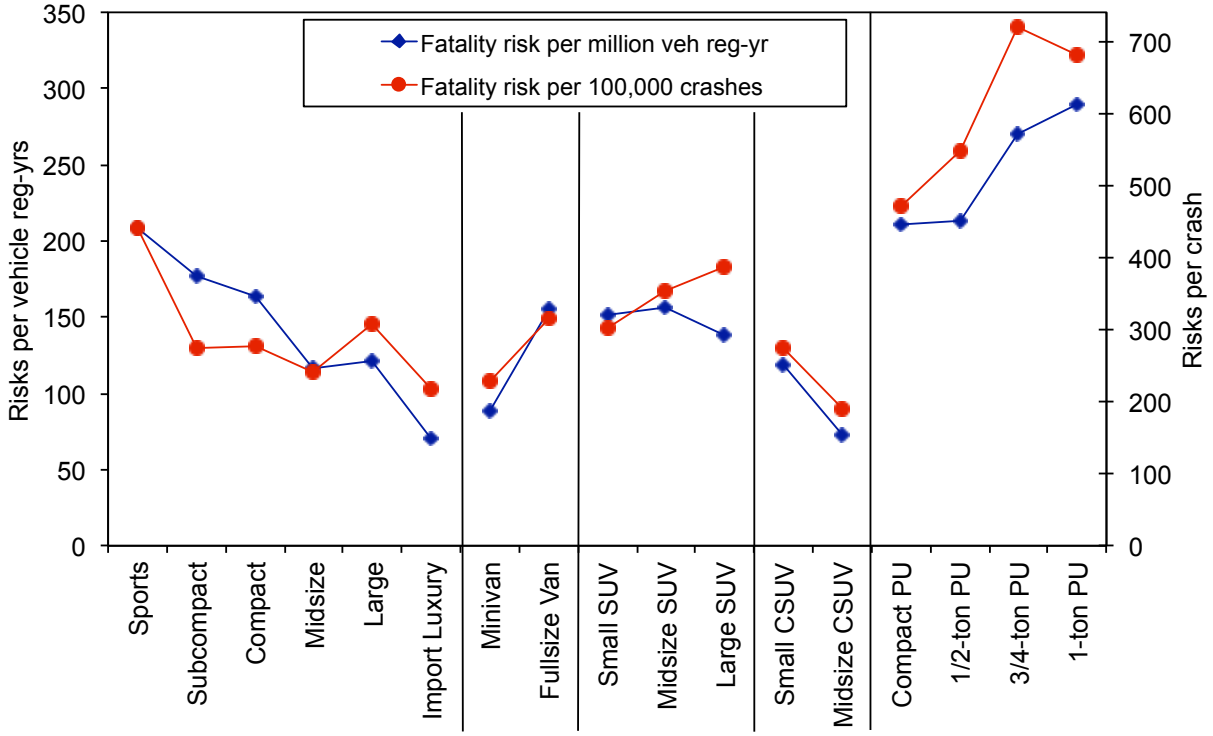



**Figure A.6. Five state fatality <u>combined</u> risks per vehicle registration-year and per crash, by vehicle type**