J O I N T   C E N T E R
AEI-BROOKINGS JOINT CENTER FOR REGULATORY STUDIES

The Fatality Risks of Sport-Utility Vehicles, Vans, and Pickups Relative to Cars

Ted Gayer

Working Paper 03-13
November 2003

The author is Ted Gayer of Georgetown University. He is grateful for the helpful comments provided by John Cawley, John Graham, Michael Greenstone, Jay Hamilton, Winston Harrington, John Horowitz, Ted Keeler, Arik Levinson, Jens Ludwig, Doug Miller, Edmond Toy, Albert Yoon, Kip Viscusi, Sarah West, and members of the 2000 APPAM Conference, the 2001 RWJ Aspen Conference, the 2002 AEA Conference, and seminar participants at UC Berkeley, Harvard University, Georgetown University, and Resources for the Future. He also gratefully acknowledges support from the Robert Wood Johnson Foundation. The author can be contacted at gayert@georgetown.edu

J O I N T   C E N T E R

In order to promote public understanding of the impact of regulations on consumers, business, and government, the American Enterprise Institute and the Brookings Institution established the AEI-Brookings Joint Center for Regulatory Studies. The Joint Center's primary purpose is to hold lawmakers and regulators more accountable by providing thoughtful, objective analysis of relevant laws and regulations. Over the past three decades, AEI and Brookings have generated an impressive body of research on regulation. The Joint Center builds on this solid foundation, evaluating the economic impact of laws and regulations and offering constructive suggestions for reforms to enhance productivity and welfare. The views expressed in Joint Center publications are those of the authors and do not necessarily reflect the views of the Joint Center.

All AEI-Brookings Joint Center publications can be found at www.aei-brookings.org

Executive Summary

This paper examines the public health consequences of the regulatory subsidy given to light trucks. The empirical challenge is to disentangle the causal effects of light trucks from the selection bias that may occur due to drivers sorting into different vehicle types depending on their unobservable characteristics. I address this by using state variation of snow depth as an instrumental variable for vehicle miles traveled of light trucks and cars. This instrument has strong first-stage explanatory power. Since snow depth is likely a direct determinant of crashes, I meet the exclusion criteria by restricting the dependent variable to those crashes that occurred in the summer. My findings suggest that, given a crash, light trucks are more dangerous to others but less dangerous for those driving them. However, I also find that light trucks are more likely to crash than cars, which neutralizes the safety advantage to those who drive them. My estimates for aggregate fatalities suggest that a world of light trucks leads to substantially more fatalities than a world of cars.

The Fatality Risks of Sport-Utility Vehicles, Vans, and Pickups Relative to Cars

Ted Gayer

1. Introduction

There is currently a regulatory discrepancy between light trucks (i.e., sport-utility vehicles, minivans, and pickups) and passenger cars. Table 1 documents these regulations for new cars and light trucks from 1991 through 1998. The federal government regulates the tailpipe exhaust emissions of nonmethane hydrocarbons, carbon monoxide, nitrogen oxide, and particulates.[1] It regulates fuel economy through corporate average fuel economy (CAFE) standards and through a gas-guzzler tax on consumers.[2] As can be seen from the table, light trucks face a laxer regulatory burden for both emission standards and fuel economy standards.[3] This favorable regulatory treatment effectively represents an implicit relative subsidy for light trucks.

If light trucks pose greater external risks than cars, then efficiency would be achieved through levying a relative tax on light trucks rather than a relative subsidy.[4] However, while reducing or eliminating the favorable regulatory treatment of light trucks may improve efficiency, it is not a priori clear whether this would reduce the number of traffic fatalities. In fact, there is some evidence to suggest that such a change would indeed increase fatalities. The reasoning is straightforward: while light trucks might pose greater risks to others than cars, they might also provide greater safety to those who drive them. Crandall and Graham (1989) use similar reasoning in their study of fuel economy standards. They find that such standards led to smaller cars and that smaller cars led to more fatalities, since a crash between two small cars is expected to lead to more fatalities than a crash between two big cars. This line of reasoning has led many to oppose getting rid of the regulatory subsidy for light trucks, claiming that it would result in more fatalities. For example, the Wall Street Journal cites a study by Douglas Coate and James VanderHoff in which they claim, "The increased safety to occupants of light trucks outweighs the potential increases in fatalities to occupants of other vehicles." During the recent Congressional debate on the regulatory discrepancy, Senator Trent Lott opposed stiffening the regulatory burden for light trucks since, "Many studies have pointed to the enhanced safety of sturdier, full-framed vehicles like trucks and SUVs."

These claims that light trucks result in fewer fatalities are based on fatality estimates given that a crash occurs, and they do not consider whether light trucks are more likely than cars to get into crashes in the first place. The pertinent policy concern is not fatalities given a crash, but rather aggregate fatalities that result from different vehicle types. In this paper I estimate the aggregate fatalities resulting from different vehicle types, and I also separately estimate fatality risk given a crash and crash frequencies. Estimating both of these contributors to fatalities presents a complete picture of whether the light truck subsidy is efficient and whether it improves public health.

The main difficulty with estimating relative crash frequencies (and aggregate fatalities) is that the observational nature of the data makes it difficult to make strong causal claims. For example, if riskier, more reckless drivers are more likely to select to drive light trucks, then the regression coefficients would over-estimate the likelihood of light trucks crashing. I address this selection problem by exploiting cross-sectional state variation of snow depth, which I show is strongly correlated with the vehicle miles traveled (VMT) of light trucks and cars, even after conditioning on regional fixed effects. For state snow depth to be a valid instrument, it must also be orthogonal to unobservable determinants of crashes. Since snow depth is likely to directly contribute to crash frequencies (by affecting road safety conditions during the winter), I restrict the outcome variable to crashes that occur during summer, when snow depth is not a contributing factor in crash outcomes. The thought experiment here is that snow depth variation across states leads to exogenous variation in the amount of miles driven of light trucks relative to cars throughout a year, but it does not directly influence the number of crashes in the summer (non-snow) months of that year. By this innovation, I can examine how relative differences in light truck versus car driving (in which the variation comes about by exogenous snow depth variation) influences non-snow related crashes (in which road conditions and other state-specific unobservables are unchanged). One sign that snow depth may be a valid instrument is that I find that it is uncorrelated with other observable covariates. I also find that snow depth is uncorrelated with alternative outcome measures that may be related to unobserved population characteristics but that should not be directly affected by the state's vehicle-type mix.

The results suggest that, given a crash, light trucks pose significantly higher risk to other drivers than do cars. For example, a car driver is 1.50 to 1.88 times more likely to die given a crash with a light truck relative to a crash with another car. However, given a crash, light trucks provide more safety to the driver than does driving a car. For example, a light truck driver is 0.30 to 0.50 times as likely as a car driver to die given a crash with a car. Using cross-sectional variation in snow depth as an instrument for VMT, the results suggest that light trucks are 2.63 to 4.00 times more likely to crash than cars. This neutralizes the safety advantage that light trucks provide to their drivers. Combining fatality risk given a crash with the crash risk indicates that aggregate fatalities are higher in a world of light trucks relative to a world of cars.

The rest of the paper is organized as follows. In the next section, I describe the identification strategies I use to estimate the relative fatality risks given a crash and the relative crash frequencies, and I also discuss the data sources. In section 3, I present the estimates of fatality risk given a crash, relative crash frequencies, and aggregate fatality risk. Section 3 also discusses the first-stage explanatory power of the IV estimation, as well as the implied aggregate fatalities given different ratios of light trucks to cars on the road. Section 4 offers validity checks on the IV framework, and section 5 concludes.

2. Identification Strategies and Data Description

Estimating Fatality Risk Given a Crash

Recall that Crandall and Graham (1989) studied the effect of CAFE standards on car size and contended that lighter cars would lead to more fatalities since a crash between two small cars results in more deaths than a crash between two large cars. I re-examine this question with respect to light trucks versus cars. That is, I estimate the probability of dying given that a crash has occurred, for drivers of different vehicle types and conditional on crashing into different vehicle types. For simplicity, the analysis focuses on the probability of *driver* death given a two-vehicle crash. For this analysis, I use driver-level data from 1991 through 1998 from the *Fatal Analysis Reporting System* (FARS), which is a census of all the crashes in the United States that involved a fatality.[5] FARS contains detailed information on the characteristics of the crash, the characteristics of each vehicle involved in the crash, and the characteristics of each person involved in the crash.

The analysis considers only the crashes that involve cars, sport-utility vehicles, van-based light trucks, and pickups.[6] This excludes crashes that involve buses, trucks greater than 10,000 pounds, motorcycles, mopeds, all-terrain vehicles, all-terrain cycles, as well as other small vehicles such as snowmobiles and go-carts. That leaves 103,056 drivers involved in two-vehicle crashes from 1991 through 1998 in which at least one of the drivers died. Table 2 reports the number of such crashes by types of vehicles involved, as well as the number of driver fatalities for each vehicle, by vehicle-type pair. The top number in each cell is the number of fatalities of drivers in Type I vehicles who died in a crash with a Type II vehicle. The bottom number in parentheses is the total number of crashes involving Type I and Type II vehicles in which at least one of the drivers died.

One can obtain a sense of the relative risk given a crash of each vehicle type by comparing the symmetrical, off-diagonal cells. Since the goal of this analysis is to analyze the risk differential between cars and light trucks, of particular interest is the comparison between the off-diagonal cells of the first column and their symmetrical cells in the first row. For example, there were 4,749 crashes between a car and a sport-utility vehicle in which at least one driver died. Of these crashes, 945 sport-utility vehicle drivers died, while 3,990 car drivers died. This suggests that, given such a fatal crash, the driver of the car is 4.2 times as likely to be the one who dies than is the sport-utility vehicle driver. Similarly, given a fatal crash between a van and a car, the driver of the car is 5.0 times as likely to be the one who dies; and given a fatal crash between a pickup and a car, the driver of the car is 3.8 times as likely to be the one who dies.

Comparing non-symmetrical cells can be misleading. The naïve method would be to divide the number of fatalities by the number of fatal crashes to obtain a risk measure to compare to other cells. For example, from Table 2 one finds that, given a fatal crash between a car and a sport-utility vehicle, 84% of the drivers of cars died. Similarly, given a fatal crash between a car and a van, 87% of the drivers of the cars died. But one cannot directly compare these numbers since doing so neglects possible sample selection bias from the omission of data on non-fatal crashes. For example, assume that vans and sport-utility vehicles are equally likely to get into a crash and that sport-utility vehicles are less of a threat to car drivers than are vans. In this case, a sport-utility vehicle crashing into a car would be less likely to result in a fatality, and such crashes would be excluded from the data set. In other words, the denominator (4,749) for sport-utility crashes would be under-represented relative to the denominator for vans (5,074). This would result in an upward bias of the risk that sport-utility vehicles pose to cars. However, the direction of the bias is unclear. If the less-threatening sport-utility vehicles are more dangerous to those who drive them (relative to vans), then the sport-utility denominator would be over-represented, and there would be a resulting downward bias of the risk that sport-utility vehicles pose to cars.[7]

The sample selection problem exists because there are no reliable data on non-fatal crashes. As part of the Department of Transportation's Crash Outcome Data Evaluation System (CODES), some states have linked (or are currently linking) crash data (for fatal and non-fatal crashes) with injury and cost data. Unfortunately, this data are not helpful for my study for a number of reasons. First, the data contain a sample of crashes rather than a census. This poses a sample selection problem of its own since the crashes in which injuries occurred are more likely to be reported. Additionally, the reporting rules vary by state: some states only include crashes in which an injury has occurred, some include non-injury crashes in which damages were above $500, some include non-injury crashes in which damages were above $1,000, etc. Another problem with these data is that, relative to the FARS data, there are a limited number of observations. That is, most of the states have only recently started collecting the data, so their availability is limited only to one or two years. The biggest problem with the CODES data with respect to my study is that most states that collect the data do not keep separate classifications for the vehicle types. For the few states that do classify the vehicle types involved in the crashes, the classifications vary greatly across states. Given that these crash data do not clearly distinguish among vehicle types, I instead rely on the FARS data for my analysis.

Relying on the FARS data introduces a potential sample selection problem both in the estimation of fatality risk given a crash and the estimation of differential crash frequencies. In order to address the possible sample selection problem inherent in FARS, I exploit information on pedestrian fatalities in the data set.[8] For crash risk, a simple logit model using the FARS data gives an estimate of the probability a driver of a vehicle of Type i dies given a *fatal* crash with a vehicle of Type j (with i and j equal to car, sport-utility vehicle, van, or pickup). From this estimate, one could derive an estimate of the predicted number of deaths to drivers of vehicle Type i given a fatal crash with drivers of vehicle Type j, conditional on the other covariates. The problem is that without knowledge of the total number of crashes (fatal and non-fatal) involving Type i and Type j vehicles, one cannot estimate the probability that a driver of a vehicle of Type i dies given a *crash* with a vehicle of Type j. The FARS data set does, however, contain the number of vehicle crashes in which a pedestrian dies. Under a set of reasonable assumptions (discussed below),[9] the product of pedestrian fatalities by vehicle types is proportionate to the number of crashes of the same vehicle types, and the proportionality is constant across different vehicle type combinations. I re-weight the predicted probabilities by these products to obtain unbiased relative risk estimates. Later on in the paper, I estimate aggregate fatalities by vehicle type, which—since it combines fatality risk given a crash with crash frequency—does not suffer from the sample selection bias, and thus does not rely on the assumptions needed to justify the use of pedestrian fatalities. Comparing the estimates of aggregate fatalities to the estimates implied by the component estimates provides a credible consistency test.

Using the product of pedestrian fatalities by vehicle types as a proxy for total crashes by vehicle types is valid given that three assumptions are met. For notational convenience, let y_{ij} equal the number of two-vehicle crashes involving Type i and Type j vehicles, and let y_i equal the total number of crashes involving Type i vehicles. Let p_{ij} equal the probability that a given two-vehicle crash involves vehicles of Type i and j, and let p_i equal the probability that a given crash involves one vehicle of Type i. Let w_i equal the number of crashes of vehicles of Type i into a pedestrian, and let w_i^f equal the number of crashes of vehicles of Type i into a pedestrian in which the pedestrian dies. Finally, let Y equal the total number of two-vehicle crashes.

Assumption 1: $p_{ij} = p_i p_j$.

This assumption is that the probability that one of the vehicles in a given two-vehicle crash is of a certain type is independent of the probability that the other vehicle is of a certain type. This seems reasonable, since there are no clear reasons why different types of vehicles would cluster in crashes disproportionately more than other types.[10] Although not reported in the paper, the empirical results are robust after stratifying the samples by urban vs. rural in order to control for potential vehicle clustering within these road-types.

Assumption 2: $y_i = \alpha_1 w_i$ and $y_j = \alpha_1 w_j$, for all i and j.

This assumption is that the total number of crashes of a certain type of vehicle is proportional to the total number of crashes of that type of vehicle into a pedestrian, *and this* *proportion is the same across vehicle types*. In other words, certain types of vehicles do not hit pedestrians disproportionately.

Assumption 3: $w_i = \alpha_2 w_i^f$ and $w_j = \alpha_2 w_j^f$, for all i and j.

This assumption is that the number of pedestrian crashes involving a certain type of vehicle is proportional to the number of crashes in which the same type of vehicle kills the pedestrian, *and this proportion is the same across vehicle types*. This means, for example, that a sport-utility vehicle hitting a pedestrian is as likely to kill the pedestrian as is a car hitting a pedestrian. This assumes that there exists a threshold vehicle mass above which the probability of a pedestrian fatality given a crash remains constant (and that cars and light trucks are all above this threshold). This seems reasonable because even the lightest car is considerably heavier than a pedestrian, so the forces acting on the pedestrian are relatively independent of vehicle mass.[11] Note that increased pedestrian fatality risk due to the greater dimensions of light trucks (as opposed to the greater mass) would not bias the results.

Assumption 3 is the most restrictive of the three assumptions, since it assumes that a pedestrian hit by a light truck is as likely to die as a pedestrian hit by a car. It is important to note that in what follows, violation of this assumption will lead to an *underestimate* of the external fatality risk of light trucks relative to cars given that a crash has occurred, and it will result in an *overestimate* of the relative crash frequency of light trucks compared to cars. However, combining the component estimates of fatality risk and crash frequency leads to very similar estimates of aggregate fatalities as does the combined analysis (discussed later), suggesting that these assumptions lead to valid results.

Implications of Assumptions:

Given these three assumptions, let us examine what we know about the total number of two-vehicle crashes involving a vehicle of Type i and a vehicle of Type j.

1) $p_{ij} = p_i p_j$ → By Assumption 1

2) $y_{ij} / Y = (y_i / Y)(y_j / Y)$ → By Definition

3) $y_{ij} = y_i y_j / Y$ → Rearranging Terms

4) $y_{ij} = (\alpha_1 w_i)(\alpha_1 w_j) / Y$ → By Assumption 2

5) $y_{ij} = (\alpha_1\alpha_2 w_i^f)(\alpha_1\alpha_2 w_j^f)/Y$ → By Assumption 3

6) $y_{ij} = k w_i^f w_j^f$ → Where $k = \frac{(\alpha_1\alpha_2)^2}{Y}$

Note that the constant term, *k*, is the same for all values of i and j. This is due to the second and third assumptions. The result is that once one uses the FARS data to estimate the predicted number of drivers of vehicle Type i that died in crashes with drivers of vehicle Type j, then this value can be divided by $w_i^f w_j^f$, which is a constant proportion of the number of crashes involving these types of vehicles (and which is obtainable from the FARS data). By doing this for all the vehicle-type crash combinations, one obtains an unbiased estimate of the relative probability of a driver in a Type i vehicle dying given a crash with a Type j vehicle.

Estimating Relative Crash Frequencies

The model of vehicle crashes is given by the following equations:

$$\begin{aligned} y_{jst} &= X'_{jst}\beta_j + \delta_j VMT_{jst} + \varepsilon_{jst} \\ VMT_{jst} &= X'_{jst}\pi_j + \eta_{jst}, \end{aligned} \tag{1}$$

where y_{jst} is the number of crashes of vehicle type *j* (with *j* equal to either light trucks or cars) in state *s* in year *t*, X_{jst} is a vector of observed characteristics, VMT_{jst} is the vehicle miles traveled of vehicle-type *j*, and ε_{jst} and η_{jst} are the unobservable determinants of vehicle crashes and vehicle miles traveled, respectively. The coefficient δ_j captures the impact of VMT on crashes for each vehicle type, and thus the ratio of these coefficients for each vehicle type yields the relative crash frequency of light trucks versus cars.

Again, since there are no reliable data on crashes by vehicle type on a state by year level, I use as a proxy the number of crashes of a given vehicle type in which a pedestrian was killed. Based on assumptions 2 and 3, this is a valid proxy, because $y_{jst} = \alpha_1\alpha_2 w_{jst}^f$, for all *j*. Note that if these assumptions hold, the pedestrian fatality proxy allows for an estimate of the *relative* crash frequencies of light trucks versus cars, not the absolute crash frequencies.

Consistent estimation of equation (1) requires that $E(\varepsilon_{jst}, \eta_{jst}) = 0$. This assumption is likely to be violated since VMT across vehicle types is not randomly assigned. That is, states with more reckless or more risky drivers (characteristics that are unobservable) might have greater selection into certain types of vehicles. (Since I assume that all drivers select either cars or light trucks, if unobservable characteristics of drivers lead them to drive more light trucks, this will also lead to fewer VMT of cars in the state.) Thus, a change in unobservable reckless driving in a state will both influence VMT (of both vehicle types) and will directly influence crashes; this leads to biased estimation of the impact of VMT on crashes.[12]

Instrumental estimation can eliminate this bias if there is a variable, Z_{st}, that is correlated with VMT of each vehicle type but otherwise independent of crashes. Essentially, the instrumental variable offers exogenous variation that approximates random assignment. I use state snow depth as an instrumental variable. There are two conditions that must be met for consistent estimation. The first is that the snow depth must be correlated with the endogenous variable. As I will show later, cross-sectional variation in snow depth is correlated with VMT of both light trucks and cars, and it affects these measures differentially (i.e., an increase in snow depth is associated with a decrease in car and light truck driving, but a greater decrease in the former relative to the latter). The second condition is that the instrument must be independent of the number of crashes (conditional on the other covariates). This is not likely to be met, since more snow depth would likely lead to more crashes. I address this problem by restricting the crash measure to include only those crashes in the state that occurred in the summer (i.e., months June, July, and August). Thus, instead of estimating equation (1), I replace y_{jst} with $\bar{y}_{jst}$, where the latter measures the crashes by vehicle type *j* in state *s* that took place in the summer months of year *t*. Whereas $E(Z, \varepsilon_{jst} \mid X_{jst}) \neq 0$, the identifying assumption of my analysis is that $E(Z, \bar{\varepsilon}_{jst} \mid X_{jst}) = 0$, where $\bar{\varepsilon}_{jst}$ captures the unobservable characteristics of summertime crashes. I examine the validity of this assumption later in the paper.

The data for the VMT of light trucks and cars comes from the Department of Transportation's *Highway Statistics Series* for 1994 through 1998. (Before 1994, the data do not distinguish between light trucks and cars.) Each cell is on the state by year by road-type level, where the road-type categories are designated as rural and urban.[13] Unfortunately, the data combine all information on sport-utility vehicles, vans, and pickups into one "light truck" category, so I am unable to test whether crash frequencies vary by specific types of light trucks.

I obtained the snow depth measure used as the instrumental variable from the *National Climatic Data Center's* "Surface Summary of the Day" file. This file contains weather conditions for each weather station in the country. Using this data set, I computed two different state-by-year measures of snow depth. The first is the average daily snow depth (in inches) across all weather stations for days in January, February, March, October, November, and December. I computed this by summing all the daily snow depth measures across all weather stations (for January, February, March, October, November, and December only) and then dividing by the number of days. For the second measure, I calculate the maximum daily snow depth for each weather station over the entire year. I then averaged these maximum values across weather stations. In other words, this measures the maximum daily snow depth averaged across weather stations.

Tables 3a and 3b present the snow depth of each state from 1994 through 1998, listed by census divisions (there are four census regions and nine census divisions). Within each census division, the states are sorted by average snow depth over this period. Table 3a lists the average daily snow depth (in inches) across January, February, March, October, November, and December. Table 3b lists the maximum daily snow depth (in inches) averaged across all of the states' weather stations. Both these tables indicate that there is substantial variation in snow depth across census divisions. In order for snow depth to serve as a valid instrument for VMT, it must be orthogonal to unobservable determinants of crashes. This assumption may be violated if snow depth is simply a proxy for unobservable regional characteristics that also affect driving and other risky behaviors.

However, as can be seen in Tables 3a and 3b, there is substantial variation in snow depth within census regions and even within census divisions. Even within the East South Central division, there is a difference of more than two orders of magnitude between the state with the highest and lowest average daily snow depth and maximum daily snow depth (both averaged across years). In fact, I show below that my findings are robust to including census division fixed effects, suggesting that the instruments are not being driven solely by across-region variation.

Note also from these two tables that there is little variation within states over time in snow depth. For example, whereas there is substantial variation in snow depth across states, the average daily snow depth averaged over all states (shown in the bottom row of each table) varies from only 1.27 to 2.20 inches over time. For the maximum daily snow value the range is from 6.21 to 10.72 inches. As a result, using state fixed effects in the first-stage regressions absorbs all of the snow depth variation, thus leaving very weak first-stage explanatory power. The IV estimates are thus driven by cross-sectional variation in snow, which raises the possibility of confounding influences from unobservable state heterogeneity. I address this potential problem in a number of ways in what follows.

3. Results

Relative Driver Fatality Risk Given a Two-Vehicle Crash

Table 4 reports the estimation results of driver fatality risk for each vehicle type of driver, conditional on opposing vehicle type. I stratify the sample by the type of vehicle that the driver is in, which is represented in the separate columns of Table 4. Each column represents the results of a logit model in which the dependent variable equals one if the driver died, and equals zero otherwise. The variables of interest are a series of dummies denoting the type of vehicle that the driver crashed into (i.e., car, sport-utility, van, pickup, with the car dummy withheld as the comparison group). As discussed previously, I use pedestrian fatalities by vehicle type to adjust for the sample selection bias. Keep in mind that this method yields predicted probabilities within a proportionality factor.

Though not listed in the table, the regression includes covariates for both drivers and for crash conditions. The drivers' covariates are age, age squared, sex, whether an air bag was deployed, whether the driver was wearing a seat belt, whether the driver was drunk, and whether the driver had any major or minor traffic incidents within three years before the crash. Major incidents are accidents, DWI convictions, suspensions and revocations of license. Minor incidents are speeding and other moving violations. The variables describing the crash conditions are the road condition (wet or dry), the type of road (rural interstate, rural non-interstate, urban interstate, or urban non-interstate), the speed limit, the time of day (four six-hour dummy variables), and the year.

Heteroskedastic-consistent standard errors are reported in parentheses beneath the coefficient estimate, and the mean predicted probabilities for each opposing vehicle type are reported in brackets. These bracketed predicted probabilities are the estimates without adjusting for the possible sample selection bias, and thus offer a means of comparison with the adjusted estimates. The sample-selection adjusted predicted probabilities are reported in braces. For the results reported in the tables, I used total counts of pedestrian fatalities by vehicle types to adjust for the sample selection. For robustness, I also used year-specific pedestrian fatality counts to weight each predicted probability. These latter results are not reported in the paper since they are nearly identical to the reported estimates. Since the sample selection adjustment yields *relative* risks across vehicle types, I standardized the predicted probabilities so that the probability of a car driver dying given a crash with another car driver is one.

The results strongly suggest that no matter what type of vehicle one drives, crashing into a sport-utility vehicle, van, or pickup poses a greater risk than does crashing into a car. The results also suggest that no matter what type of vehicle one crashes in to, one is at a significantly greater risk if one is in a car rather than a light truck. The sample-selection adjusted results do suggest a different ordering than the unadjusted results of the risks among the different light trucks. The unadjusted estimates suggest that it is less risky to crash into a pickup than it is to crash into a sport-utility vehicle, and that it is less risky to crash into a sport-utility vehicle than it is to crash into a van. The results after correcting for sample-selection bias indicate that in a crash, vans tend to pose less risk to others than do sport-utility vehicles, and pickups tend to pose the greatest risk to others. These discrepancies suggest that the fatality crash data over-represent the risk posed by vans relative to the risk posed by sport-utility vehicles and pickups.

For ease of exposition, Panel A of Figure 1 reports the predicted relative probabilities of a driver dying given a crash with different types of vehicles. The x-axis groupings are for each possible vehicle choice of the driver, and the vertical bars show the probability of dying conditional on a crash with each type of vehicle. Each x-axis grouping is standardized so that the probability of dying given a crash with a car is equal to one. Thus, only comparisons within each grouping are possible.

The results of Panel A indicate that, given a crash, light trucks pose significantly higher risks to other drivers than do cars. For example, relative to crashing into a car, a car driver is 1.88 times as likely to die if the opposing vehicle is a sport-utility vehicle, 1.50 times as likely to die if the opposing vehicle is a van, and 1.88 times as likely to die if the opposing vehicle is a pickup. The results are similar across the driver’s vehicle designation. No matter what vehicle the driver drives, pickups pose the highest external risks in a crash. And for car, sport-utility, and pickup drivers, crashing into a sport-utility is riskier than crashing into a van.

Panel B of Figure 1 reports the predicted relative probabilities of dying depending on the type of vehicle the driver is driving. The x-axis groupings are for each type of opposing vehicle, and the vertical bars show the probability of dying, given a crash, depending on the driver's vehicle type. Each x-axis grouping is standardized so that the probability of a car driver dying is one. Thus, only comparisons within each grouping are possible.

The results of Panel B indicate that driving a light truck is significantly safer for a driver than is driving a car. For example, given a crash with a car, a sport-utility vehicle driver is 0.44 times as likely to die than is a car driver, a van driver is 0.30 times as likely to die than is a car driver, and a pickup driver is 0.50 times as likely to die than is a car driver. The results are similar across the different types of opposing vehicle. No matter what the opposing vehicle, driving a van poses the lowest risk to the driver in the event of a crash, driving a sport-utility vehicle is the next safest, and driving a pickup is the least safe among the light trucks.

These results are consistent with the implications of Crandall and Graham (1989). That is, while light trucks pose a greater fatality risk to others in a crash, this is dominated by the safety advantage they give to those who drive them (I show the exact net effects later). However, these risk estimates are conditional on crashes having occurred and ignore the possibility that crash frequencies might differ across vehicle types. To the extent that light trucks are more prone to crash than cars, both the internal and external risk estimates for them will increase. We now turn to the crash frequency results.

OLS Results of Relative Crash Frequencies

For purposes of comparison, I first estimate the relative crash frequencies of equation (1) using ordinary least squares. I stratify the sample to estimate the gradient of crashes with respect to VMT for both light trucks and cars. I then divide the estimated light truck gradient by the estimated car gradient to arrive at the crash frequency of light trucks relative to cars.[14]

An examination of the raw aggregate ratio of pedestrian fatalities caused by light trucks (per VMT of light trucks) divided by the ratio of pedestrian fatalities caused by cars (per VMT of cars) indicates that light trucks are 25 percent more likely to crash for a given mile driven. Of course, this raw ratio ignores both confounding factors (such as time trends, speed limit, etc.) and selection bias. As a first step, I perform two sets of OLS regressions. The first includes different combinations of state fixed effects, year indicators, and a state-specific linear time trend. These controls should help alleviate the possible selection bias, since they will absorb any mean shifts in unobservable determinants of crashes across years or states, and will also absorb any unobservable determinants of crashes that track linearly over time within states. The second set of regressions uses more flexible controls of year by road-type indicators and a state by road-type specific linear time trend. Each one of these controls passes an F-test in which the restricted model of comparison does not include road-type specific controls. The variables should further alleviate the selection bias since they control for mean shifts over time within road-types, as well as for trends in crash determinants occurring over time within each state's road-types. Each specification also controls for the road-type, the state's unemployment rate, the legal speed limit, the proportion of the state's population that is young (15-29) and male, young and female, old (65 and up) and male, and old and female. The reported standard errors (in parentheses) adjust for error clustering within states.

Table 5a lists the OLS results for the car and the light truck specifications for the first set of specifications. The first set of rows is for the car specifications, and the next set of rows is for the light truck specifications. The results suggest that more VMT in a state for cars and light trucks increases the number of crashes by these respective vehicle types. Each coefficient estimate is highly statistically significant. Most notably, the gradient with respect to light trucks is greater than the gradient with respect to cars. The final row of Table 5a takes the ratio of the gradients and suggests that light trucks are between 2.57 to 3.03 times more likely to crash than are cars.

Table 5b lists the OLS results for the more flexible OLS specifications (i.e., allowing the year indicators and the state-specific linear time trend to vary by road-type). Once again, the results suggest that an increase in VMT is associated with an increase in crashes, for both cars and light trucks. The coefficient estimates are all statistically significant within the one-percent level. As with the other specifications, the light truck gradient is larger than the car gradient. The last row takes the ratio of the two and suggests that light trucks are 2.14 to 3.02 times more likely to crash than are cars. The results across specifications in Tables 5a and 5b are rather robust. However, these OLS estimates may suffer from the selection bias discussed earlier. We now turn to the IV results as a means of addressing this problem.

First-Stage Explanatory Power of IV Estimation of Crash Frequencies

I examine the first-stage IV regression to check the explanatory power of the instrumental variable. If the instrumental variable has weak explanatory power in the first stage, then the IV coefficient estimates will have large standard errors.[15] Bound, Jaeger, and Baker (1995) recommend using an F-test of the joint significance of the instruments in the first-stage regression to assess the fit.[16] As they also recommend, I report the partial R^2 (which is the R^2 of the first-stage regression with the instruments partialled out), in addition to the F-statistic. Staiger and Stock (1997) suggest that instruments should be declared weak if the first-stage F-statistic is less than ten.

Tables 6a and 6b show the first-stage regression results of the endogenous explanatory variable (VMT) against the instrument (snow depth) for both light trucks and cars. Each regression includes the other exogenous covariates, which are the road-type indicators, the state unemployment rate, the legal speed limit, and the age by sex variables as a proportion of the population. The different columns represent specifications with different combinations of year by road-type indicators, a state by road-type linear time trend variable, and census division indicators. As mentioned earlier, I include the census division indicators in some specifications as a way of addressing the concern that the instrument is only picking up regional unobservable determinants of crashes. Table 6a shows the first-stage results when the average daily snow depth is used as the instrument, and Table 6b shows the first-stage results when the maximum daily snow depth for the year is used as the instrument.

The first-stage results listed in Tables 6a and 6b indicate that a state's snow depth is negatively correlated with both VMT of cars and VMT of light trucks. What's more, since this analysis is interested in exogenous variation of VMT of light trucks *relative* to cars, it is noteworthy to see that the first-stage gradient for light trucks has a lower magnitude than the gradient for cars. That is, while greater snow depth leads to fewer miles driven, the decline is greater for cars than for light trucks. Relatively speaking, worse weather conditions leads to a shift towards light trucks. The last row in each table shows how many millions of miles more of light truck driving occurs relative to car driving given a 1-inch increase in average snow depth (Table 6a) and maximum daily snow depth (Table 6b). Thus, an increase in the average daily snow depth of one inch leads to a relative shift towards light truck driving of between 1,105 and 2,142 million miles. And an increase of one inch in the maximum daily snow depth (averaged across weather stations) leads to a relative shift towards light truck driving of between 285 and 601 million miles.

The bracketed terms under the coefficient estimates report the F-statistics for the instrumental variable, and the terms in braces reports the partial R^2. For all the specifications in Tables 6a and 6b, the fit of the first-stage regressions seems rather strong, with an F-statistic greater than ten. The first-stage results therefore suggest that the instruments are strongly correlated with the endogenous variable.

IV Results

Table 7a presents the second-stage IV results when the instrument is the average daily snow depth in the state (in January, February, March, October, November, and December), and Table 7b presents the second-stage IV results when the instrument is the maximum daily snow depth averaged across weather stations. For each specification, the dependent variable is restricted to pedestrian fatalities that occurred in summer months (June, July, and August). Each specification contains road-type indicators, state unemployment rate, speed limit, and age by sex controls. The various specifications use different combinations of year by road-type indicators, a state by road-type specific linear time trend, and census division indicators. The reported standard errors (in parentheses) adjust for error clustering within states.

The top set of rows of Table 7a suggests that there is a positive gradient of car crashes with respect to VMT of cars. All of the estimated coefficients are statistically significant at the one-percent level. The next set of rows indicates that there is also a positive gradient of light truck crashes with respect to VMT of light trucks. These, too, are all significant at the one-percent level. Of particular interest are the relative magnitudes of the light truck versus the car gradient. The bottom row of Table 7a takes the ratios of these coefficient estimates and finds that light trucks are 2.93 to 4.00 times more likely to crash than are cars for a given amount of VMT. These ratios are fairly robust across specifications.

The results in Table 7b also suggest a positive gradient of crashes with respect to VMT, for both cars and for light trucks. The various specifications are all statistically significant at the one-percent level. The bottom set of rows shows the ratios of the light truck gradient relative to the car gradient, and these estimates suggest that light trucks are 2.63 to 3.88 times more likely to crash than are cars. These ratios are fairly robust across specifications.

The OLS and IV results therefore suggest that light trucks are considerably more likely to crash than cars. Given that the IV estimation correctly controls for unobservable behavioral differences among drivers, the main remaining hypothesis on why light trucks crash more frequently than cars has to do with their physical design. In order to receive the regulatory advantage of being a light truck, regulatory officials require that light trucks be "capable of off-highway operation." As reported by Easterbrook (2002), the test for this essentially became whether the vehicle is tall enough to provide ground clearance. This gave automakers an incentive to make their light trucks tall, which, among other things, increases the glare of their headlights to oncoming drivers, diminishes their sight-lines to other drivers, and makes them more likely to roll over. Additionally, light trucks are all built on stiffer, heavier frames than are cars, which also makes them more difficult to handle, and thus more prone to crashing. The greater crash frequency of light trucks could also be due to changes in driver behavior due to the (mis)perception that they offer greater safety to the drivers. Lave and Weber (1970) and Peltzman (1975) hypothesized that the safer drivers feel, the more reckless may be their driving. Thus, the combination of physical construction and behavioral changes could both contribute to greater crash frequencies of light trucks.

Aggregate Fatalities Given Different Vehicle-Type Compositions

As mentioned before, previous studies have indicated that a crash involving two lighter vehicles tends to result in fewer expected deaths than a crash involving two heavier vehicles. Those who oppose increasing the regulatory burdens of light trucks use this evidence to claim that such a change in policy would lead to more fatalities. However, this is potentially misleading, since one must consider whether crash frequencies vary by vehicle type in order to get an accurate estimate of the expected number of traffic fatalities given different ratios of light trucks to cars on the road.

Suppose that there are N vehicles in the world, that the number of cars equals C, and the number of light trucks equals T. Also assume that $C + T = N$. Let γ_{ij} equal the probability that a driver of vehicle type i is killed by a driver of vehicle type j, conditional on a crash between the two vehicles. Let β_{ij} equal the probability of vehicle type i crashing with vehicle type j. Then the probability (P_c) of a car driver being killed in a two-vehicle crash, and the probability (P_T) of a light truck driver being killed in a two-vehicle crash are given as follows:

$$P_C = \gamma_{CC}\beta_{CC}C + \gamma_{CT}\beta_{CT}T$$
$$P_T = \gamma_{TC}\beta_{TC}C + \gamma_{TT}\beta_{TT}T. \qquad (2)$$

The expected number of traffic fatalities (E) is $P_C C + P_T T$. Let us examine three different cases. Case 1 is a world in which all vehicles are cars (i.e., $C=N$, $T=0$). Case 2 is a world in which all vehicles are light trucks (i.e., $C=0$, $T=N$). Case 3 is a world in which half the vehicles are cars and half are light trucks (i.e., $C=N/2$, $T=N/2$). The expected number of traffic fatalities in each state is given as follows:

$$E_1 = \gamma_{CC}\beta_{CC}N^2$$
$$E_2 = \gamma_{TT}\beta_{TT}N^2 \qquad (3)$$
$$E_3 = (\frac{N^2}{4})(\gamma_{CC}\beta_{CC} + \gamma_{TT}\beta_{TT} + \gamma_{CT}\beta_{CT} + \gamma_{TC}\beta_{TC}).$$

This paper has estimated the following: 1) the relative probability of a driver dying conditional on the type of vehicle he or she is driving and the type of vehicle of the opposing driver, and 2) the relative crash frequency of light trucks versus cars. By coupling these results, I have estimates of the ratios of the gammas and betas in equation (3). Thus, I can use the empirical estimates to compute the number of fatalities in one state of the world relative to another state of the world (i.e., E_3/E_1 and E_2/E_1). The top panel of Table 8 presents the estimates of the relative expected number of fatalities given that crash frequencies do not vary across vehicle types (i.e., the betas are assumed constant). According to these results, a world with only sport-utility vehicles or vans (or 50 percent of each) is indeed a world with fewer traffic fatalities than is a world with only cars. This supports the claims of those opposed to increasing the regulatory burden of light trucks. However, given the IV results for crash frequencies, if one chooses a conservative estimate of 2.5 for the greater crash probability of light trucks relative to cars, the results change dramatically. Panel B of Table 8 presents the expected number of fatalities relative to a world in which everyone drives cars, given the crash frequency adjustment. The results suggest that a world in which everyone drives sport-utility vehicles would result in 4.96 times more fatalities than a world in which everyone drives cars. A world of vans would result in 3.63 times more fatalities than a world of cars. A world of pickups would result in 8.15 times more fatalities than a world of cars. The off-diagonal cells give estimates in which half the vehicles are of each respective vehicle type, and also indicate that these states lead to more fatalities than a world with only cars. Minimization of traffic fatalities results when all vehicles are cars.

The primary goal of this paper is to examine both fatality risk given a crash and crash frequencies in order to determine how they both contribute to aggregate fatalities. Of course, the results are contingent on the validity of the assumptions that justify using pedestrian fatalities by vehicle type as a proxy for crashes by vehicle type. Another way to estimate aggregate relative fatalities due to vehicle types is to re-estimate the IV equations using driver fatalities by vehicle types as the outcome measure (instead of the pedestrian fatality measure). This has the advantage of focusing on the outcome of greatest policy interest – fatalities – without involving the assumptions justifying the use of pedestrian fatalities as a proxy variable for crashes.

Tables 9a and 9b present these estimates and suggest that light trucks result in 3.90 to 10.37 times more driver fatalities than cars. These estimates are a near perfect match to the product of the fatality risk given a crash estimates and the crash frequency estimates, suggesting that relying on the pedestrian fatality assumption does not bias the results. For comparison, the third panel of Table 8 uses these IV results to estimate the relative aggregate fatalities (by taking the ratios in equation 3). Again, the results suggest a world of light trucks would result in 3.90 to 10.37 times more fatalities than a world of all cars, and a world of half cars and half light trucks would result in 2.45 to 5.69 times more fatalities than a world of all cars. The results are very similar to the results presented in Panel B. Once again, while light trucks offer a safety advantage to those who drive them in case of a crash, they also offer a substantially greater risk to others and are substantially more likely to crash. The net result is that they lead to more fatalities than do cars.

4. <u>Validity Tests</u>

As mentioned earlier, the validity of the IV results rests on the assumption that the instrumental variable is orthogonal to the unobservable determinants of the dependent variable. Researchers frequently test (as a standard diagnostic) whether the excluded instruments are uncorrelated with the error term by using an over-identification test. However, in this paper there is only one excluded instrument, thus precluding an over-identification test.

A more informal analysis of the validity of an excluded instrument is to examine the relationship between the instrument and other observable covariates. In the ideal case where the instrument is randomly assigned, the instrument would be uncorrelated with both the observable and unobservable determinants of the outcome variable. Examining the relationship between the instrument and the observable determinants serves as a guide to how much selection there is on the unobservables (see Altonji, Elder and Taber 2001a, 2001b). Table 10 shows the mean values of each of the observable covariates partitioned by whether the state was above or below the median level of snow depth. The columns on the left use the average daily snow depth measure, and the columns on the right use the maximum daily snow depth measure. The means are for 1998, but given that there is little variation in snow depth over time within states, the implications are the same for the other years. The t-statistics for the difference in means suggests that the observable covariates are balanced for states with high and low snow depth, which offers some evidence that the instrument is exogenous.

As another test of the exogeneity of the snow depth instrument, I re-estimate the IV models substituting the state unemployment rate for the dependent variable in the model. If the instrument is correlated with VMT yet orthogonal to other determinants, then the coefficient estimate in the second-stage equation should be insignificant when unemployment rate is the outcome variable. Table 11a and 11b confirm this for the most part. Of the twelve specifications, only specification four using the maximum snow depth instrument shows a statistically significant relationship (at the ten-percent level) between VMT and unemployment rate. Again, this serves as an informal check of the validity of the instrument, since the lack of a relationship between the instrument and the observable variables increases our confidence that the instrument is also not correlated to unobservable determinants.

It is conceivable that the instrument is orthogonal to observables but is correlated with an unobservable characteristic of the state that contributes to pedestrian fatalities. For example, Ruhm (2000) finds that a state's economic conditions are correlated with numerous health outcomes. It could be that my instrument is picking up unobservable state economic conditions that affect pedestrian fatalities, thus biasing my coefficient estimate. As a test for this potential bias, I substitute the state level fatality risk of chronic liver disease and cirrhosis for my outcome variable. If the instrument is picking up unobservable state economic effects that influence overall health outcomes, then one would expect this to be captured in the coefficient estimate for the cirrhosis model. Tables 12a and 12b present the coefficient estimates (for VMT) for the different specifications. Of the twelve specifications, only the two that omit the census division indicators and the year by road-type indicators show a statistically significant relationship (at the ten-percent level) between VMT and death from cirrhosis and chronic liver disease. The results in these tables again suggest that the instrument is exogenous for most of the specifications.

Finally, my IV research design rests on the assumption that variation in light truck versus car driving due to snow depth persists throughout the year, including in the summer months. That is, the VMT measures are for the entire year, whereas the outcome crash and aggregate fatality measures are for the summer months. This could lead to biased results if, for example, people own both a light truck and a car, and shift to driving the former in the winter and the latter in the summer. My IV estimation would pick up annual variation in light truck driving due to snow depth, but this variation would not persist in the summer months, which is when I measure the crash outcome variable. If the variation in light truck versus car driving does not exist in the summer months, my exclusion restriction of using summer crashes would not be valid. As a check, I re-estimated the IV equations using only those crashes that occurred in conditions of clear roads and clear skies *throughout the entire year* (i.e., when snow would not directly contribute to crashes). Though not reported in this paper, the results are virtually identical to the estimates using the summer months' restriction.

5. Conclusion

The current regulatory framework for motor vehicles was developed in the 1970s and the early 1980s. The Clean Air Act Amendments of 1970 established tailpipe emission standards, the Energy Policy and Conservation Act of 1975 established fuel economy standards for the manufacturers of new vehicles, and the 1980 gas-guzzler tax created a tax on consumers who buy vehicles with poor gas mileage. At the time these regulations were established, there were only 20 million light trucks on the road, and most of them were commercial vehicles. In an attempt to protect industry, the regulations placed on light trucks were considerably more lax than those placed on cars. But in 1984, Chrysler introduced the mini-van, and since it was partly based on a pickup design, it was able to convince regulators to categorize it as a light truck. Since then, the number of light trucks driven for non-commercial purposes has increased dramatically. Today, light trucks make up nearly half of all family vehicles sold, and there are an estimated 63 million light trucks on the road (see U.S. Dept. of Transportation, 1997).

The regulatory differences create an implicit relative subsidy for light trucks. This subsidy is inefficient, given that light trucks cause more externalities than cars. Opponents of doing away with the subsidy argue that light trucks result in fewer fatalities since moving people from light trucks to cars will increase the risk to the drivers by more than it will decrease the risk to other drivers. However, this claim assumes that light trucks and cars crash with equal frequency. In this paper I first confirm that a regulatory-induced shift away from light trucks would lead to more fatalities, given constant crash frequencies. I then use cross-sectional variation in snow depth as an exogenous instrument for VMT of light trucks and cars in order to estimate the relative crash frequencies. The IV results suggest that light trucks crash between 2.63 to 4.00 times more than do cars. While in the event of a crash, light trucks present a safety advantage to their drivers that dominates the extra risk they pose to the opposing drivers, once one adjusts for the greater frequency of crashes by light trucks, the aggregate risk they pose substantially dominates the risk from cars. Indeed, a world of light trucks would lead to three to ten times more fatalities than a world of cars. Thus, eliminating the regulatory subsidy of light trucks would improve efficiency by reducing the relative external risk, and it would also reduce the total number of motor vehicle fatalities.

Notes

[1] For the specific regulations, see Title 40 of the Code of Federal Regulations, Subchapter C, Part 86. For a summary of the regulations, see AAMA (1996). The federal government does not directly regulate the non-stationary emissions of carbon dioxide, a greenhouse gas that contributes to climate change.

[2] CAFE standards, which are standards for harmonic-weighted fleet averages for miles per gallon, were established by the Energy Policy and Conservation Act of 1975 (see Public Law 94-163). The CAFE standards are codified in Title 49 of the U.S. Code, section 32902. The gas-guzzler tax was created by Congress in 1980 and is codified in Title 26 of the U.S. Code, section 4064.

[3] The regulatory discrepancy for nitrogen oxide is scheduled to be phased out by 2007, but the other discrepancies will remain.

[4] Gabler and Hollowell (1998) present evidence that light trucks inflict greater damages in crashes than do cars.

[5] The fatality must occur within 30 days of the crash in order to be included in the data set. FARS was started in 1975 by the Department of Transportation's National Highway Traffic Safety Administration (NHTSA).

[6] For this paper, "pickups" include all light conventional trucks that have a gross vehicle weight range (GVWR) below 10,000 lbs.

[7] Levitt and Porter (2001) discuss the sample selection inherent in FARS in their analysis of seat belt and air bag effectiveness.

[8] I include cyclists among pedestrians, but the results are robust if cyclists are excluded.

[9] See Evans (1985) for a discussion of these assumptions pertaining to passenger car sizes.

[10] This assumption could be violated if certain types of vehicles present conflicting visibility conditions. For example, the greater height of sport-utility vehicles with respect to the height of cars could lead to a disproportionate number of crashes between these two types of vehicles.

[11] There is limited credible evidence of whether or not this assumption is valid. Some studies have examined whether increasing vehicle mass (focusing on cars only) leads to more pedestrian fatalities. Evans (1984) found mixed results of the relationship. A more recent study by NHTSA (1997) using aggregate data found a slightly positive relationship between vehicle mass (of cars) and pedestrian fatalities. However, the dependent variable in this analysis was fatality rates by make, model, and model year. Although the study included a number of controls, it most likely suffers from selection bias since the choice of vehicle model is likely correlated to unobservable characteristics of the drivers. For example, heavier cars may be more likely to be driven by less aggressive middle-aged or older drivers, thus biasing the results.

[12] Note that this identifying assumption is not violated if the type of vehicle changes driving behavior, a phenomenon first discussed by Lave and Weber (1970) and further advanced by Peltzman (1975).

[13] Urban refers to geographic areas with populations over 5,000 people, and rural refers to all other areas. In an unreported part of my analysis, I obtained virtually identical results using four different road-type categories: rural interstate, rural non-interstate, urban interstate, and urban non-interstate.

[14] Although not presented in this paper, I obtained similar results for the OLS and later IV estimation by pooling the data and estimating a single equation that contains an interaction variable that multiplies VMT by a dummy variable that indicates if the observation pertains to light trucks or cars.

[15] Bound, Jaeger, and Baker (1995) show that with weak instruments, even a weak correlation between the instruments and the error in the original equation can lead to a large inconsistency in the IV results, even when the sample is very large (as in Angrist and Krueger 1991).

[16] I use only one instrument in my analysis, so one could equivalently use a t-test; however, for comparability with other studies I compute the F-test.

References

Altonji, Joseph G., Todd E. Elder, and Christopher R. Taber. (2001a). "An Evaluation of Instrumental Variable Strategies for Estimating the Effects of Catholic Schools," *NBER Working Paper Series*, Working Paper 9358.

Altonji, Joseph G., Todd E. Elder, and Christopher R. Taber. (2001b). "Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools," *NBER Working Paper Series*, Working Paper 7831.

American Automobile Manufacturers Association (AAMA). (1996). *Motor Vehicle Facts and Figures*.

Angrist, Joshua D., and Alan B. Krueger. (1991). "Does Compulsory School Attendance Affect Schooling and Earnings?" *Quarterly Journal of Economics* 106, 979-1014.

Automotive News Market Data Book. (1991-1999). Detroit: Crain Communications, Inc.

Bound, John, David A. Jaeger, and Regina M. Baker. (1995). "Problems with Instrumental Variables Estimation When the Correlation between the Instruments and the Endogenous Explanatory Variable is Weak," *Journal of the American Statistical Association* 90(430), 443-450.

Crandall, Robert W., and John D. Graham. (1989). "The Effect of Fuel Economy Standards on Automobile Safety," *Journal of Law and Economics* 32(1), 97-118.

Easterbrook, Gregg. (2002). "Axle of Evil: America's Twisted Love Affair with Sociopathic Cars," *The New Republic*, January 20, 2003.

Evans, Leonard. (1984). "Driver Fatalities versus Car mass Using a New Exposure Approach," *Accident Analysis and Prevention* 16, 19-36.

Evans, Leonard. (1985). "Car Size and Safety: Results from Analyzing U.S. Accident Data," *The International Technical Conference on Experimental Safety Vehicles*, 548-560.

Gabler, Hampton C., and William T. Hollowell. (1998). "NHTSA's Vehicle Aggressivity and Compatibility Research Program," *U.S. National Highway Traffic Safety Administration*, Paper No. 98-S3-O-01.

Lave, Lester B., and Warren E. Weber. (1970). "A Benefit-Cost Analysis of Auto Safety Features," Applied Economics 2(4), 265-275.

Levitt, Steven D., and Jack Porter. (2001). "Sample Selection in the Estimation of Air Bags and Seat Belt Effectiveness," *Review of Economics and Statistics* 83(4), 603-615.

National Highway Traffic Safety Administration (NHTSA). (1997). "Relationship of Vehicle Weight to Fatality and Injury Risk in Model Year 1985-93 Passenger Cars and Light Trucks," Summary Report Number DOT HS 808 569.

Peltzman, Sam. (1975). "The Effects of Automobile Safety Regulation," *Journal of Political Economy* 83, 677-725.

Ruhm, Christopher J. (2000). "Are Recessions Good for Your Health?" *Quarterly Journal of Economics* 115(2), 617-650.

Staiger, Douglas, and James H. Stock. (1997). "Instrumental Variables Regression with Weak Instruments," *Econometrica* 65(3), 557-586.

U.S. Department of Transportation. (1997). "Federal Highway Cost Allocation Study."

Table 1: Federal Emission and Fuel Economy Regulations for Cars and Trucks

Vehicle Type	Year	NMHC	CO	NOx	Particulates	CAFE Standard	Gas Guzzler Tax (0-12.5 mpg)	Gas Guzzler Tax (22.0-22.5 mpg)
Passenger Cars	1991-1993	0.41	3.4	1.0	0.20	27.5	$7,700	$1,000
	1994-1998	0.25	3.4	0.4	0.08	27.5	$7,700	$1,000
Light-Duty Trucks	1991	0.80	10.0	1.7	0.13	20.2	$0	$0
(under 5,750 lbs.)	1992	0.80	10.0	1.7	0.13	20.2	$0	$0
	1993	0.80	10.0	1.7	0.13	20.4	$0	$0
	1994	0.80	10.0	1.7	0.13	20.5	$0	$0
	1995	0.80	10.0	1.7	0.13	20.6	$0	$0
	1996	0.46	6.4	0.98	0.10	20.7	$0	$0
	1997	0.32	4.4	0.7	0.10	20.7	$0	$0
	1998	0.32	4.4	0.7	0.10	20.7	$0	$0
Light-Duty Trucks	1991	0.80	10.0	1.7	0.13	20.2	$0	$0
(over 5,750 lbs.)	1992	0.80	10.0	1.7	0.13	20.2	$0	$0
	1993	0.80	10.0	1.7	0.13	20.4	$0	$0
	1994	0.80	10.0	1.7	0.13	20.5	$0	$0
	1995	0.80	10.0	1.7	0.13	20.6	$0	$0
	1996	0.56	7.3	1.53	0.12	20.7	$0	$0
	1997	0.39	5.0	1.1	0.12	20.7	$0	$0
	1998	0.39	5.0	1.1	0.12	20.7	$0	$0

Notes: NMHC stands for Nonmethane hydrocarbons, CO stands for carbon monoxide, and NO_x stands for nitrogen oxide. The emission standards are measured in grams per mile, the corporate average fuel economy (CAFE) standards are measured as a harmonic-weighted fleet averages in miles per gallon.

Sources: For the gas guzzler tax, see 26 U.S.C.S. 4064. For the emission regulations, see 40 C.F.R. 80. For CAFE standards, see 49 U.S.C.S. 32902.

Table 2: The Number of Fatalities of Drivers in Type I Vehicles in Crashes with Type II Vehicles (1991-1998)

		Type II				
		Car	Utility	Van	Pickup	
	Car	21728	3990	4414	13225	43357
		(40152)	(4749)	(5074)	(15895)	(65870)
Type I	Utility	945	164	190	598	1897
		(4749)	(296)	(326)	(1091)	(6462)
	Van	882	158	234	586	1860
		(5074)	(326)	(432)	(1213)	(7045)
	Pickup	3496	598	721	2979	7794
		(15895)	(1091)	(1213)	(5480)	(23679)
		27051	4910	5559	17388	54908
		(65870)	(6462)	(7045)	(23679)	(103056)

Notes: The top number in each cell reports the number of fatalities of drivers in Type I vehicles in crashes with Type II vehicles. The bottom number in parentheses reports the total number of two-vehicle fatal crashes that occurred between vehicles of Type I and Type II. The data are from the Federal Analysis Reporting System (FARS) from 1991 through 1998.

Table 3a: State by Year Snow Depth by Census Division
(Average in Inches for Jan, Feb, March, October, November, December)

Census Divisions	1994	1995	1996	1997	1998	State Mean Snow Depth
Northeast Region:						
New England:						
Rhode Island	2.42037	0.73257	2.20333	0.31625	0.12400	1.15930
Connecticut	3.10252	1.03191	2.07259	0.34622	0.17589	1.34583
Massachusetts	4.09800	1.47267	3.17287	0.75872	0.70289	2.04103
New Hampshire	8.34453	4.19066	5.52323	5.47541	4.62255	5.63128
Vermont	10.31725	5.28225	5.61839	7.21530	5.87592	6.86182
Maine	8.42267	6.76766	5.21848	8.24113	7.36037	7.20206
Middle Atlantic:						
New Jersey	1.55854	0.69400	1.50106	0.11025	0.04518	0.78181
Pennsylvania	4.11805	1.39552	1.91165	0.69688	0.34877	1.69418
New York	6.97515	2.77392	2.81225	2.46738	1.43300	3.29234
Midwest Region:						
East North Central:						
Indiana	0.40159	0.39372	0.63023	0.35689	0.11644	0.37977
Illinois	0.63873	0.43222	0.33600	0.57008	0.25379	0.44616
Ohio	0.83931	0.79767	0.81125	0.17463	0.06695	0.53796
Wisconsin	3.74445	2.48532	5.41483	5.15338	1.99596	3.75879
Michigan	4.27910	4.00023	4.75049	5.13977	2.12403	4.05873
West North Central:						
Missouri	0.09762	0.19407	0.13493	0.25886	0.04866	0.14683
Kansas	0.13008	0.20021	0.07273	0.27146	0.11377	0.15765
Nebraska	0.87944	0.47114	0.52007	0.57193	0.48973	0.58646
Iowa	2.24079	0.94352	1.28061	1.43061	0.67589	1.31429
South Dakota	3.01149	0.94785	3.10867	4.30679	1.04857	2.48468
North Dakota	5.37961	2.82660	5.29528	5.77231	2.16025	4.28681
Minnesota	4.92799	3.64537	7.68513	8.62977	2.16964	5.41158
South Region:						
South Atlantic:						
Florida	0.00037	0.00000	0.00000	0.00000	0.00000	0.00007
Georgia	0.00019	0.00089	0.00347	0.00132	0.00113	0.00140
South Carolina	0.00054	0.00012	0.01022	0.00173	0.00066	0.00266
North Carolina	0.03192	0.02703	0.14763	0.03306	0.11553	0.07103
Virginia	0.17541	0.16565	0.81851	0.13741	0.14335	0.28806
Delaware	0.18408	0.12460	0.81441	0.03906	0.00559	0.37436
Maryland	0.92410	0.33810	1.31479	0.18212	0.10838	0.57350
West Virginia	1.18979	0.95095	1.30661	0.36172	0.45597	0.85301
District of Columbia	NA	NA	NA	NA	NA	NA
East South Central:						
Alabama	0.00000	0.00104	0.00383	0.00107	0.00013	0.00122
Mississippi	0.00116	0.00252	0.00933	0.00249	0.00473	0.00405
Tennessee	0.06393	0.04505	0.19355	0.05476	0.09338	0.09013
Kentucky	0.19115	0.04846	0.22174	0.03890	0.11582	0.12322

West South Central:						
Louisiana	0.00000	0.00000	0.00028	0.00038	0.00023	0.00018
Texas	0.00598	0.00453	0.00674	0.01817	0.00323	0.00773
Arkansas	0.02507	0.01198	0.02995	0.03199	0.00100	0.02000
Oklahoma	0.02141	0.05150	0.02937	0.05594	0.01072	0.03379
West Region:						
Mountain:						
Arizona	0.23276	0.32386	0.11853	0.41692	0.33632	0.28568
New Mexico	0.30303	0.25962	0.17047	0.57420	0.22151	0.30577
Nevada	0.52940	0.34671	0.85274	0.45663	0.62624	0.56234
Montana	2.06889	1.26069	3.20945	2.55755	1.41586	2.10249
Utah	2.36829	1.82518	2.61960	3.03577	2.71013	2.51179
Colorado	2.37428	2.18225	2.99326	3.79142	2.44865	2.75797
Wyoming	2.79609	2.38886	3.79142	3.76865	3.43137	3.23528
Idaho	3.80625	2.17144	4.20037	4.52905	3.37963	3.61735
Pacific:						
Hawaii	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
California	0.67097	0.87680	0.53090	0.50896	0.87740	0.69301
Oregon	1.01954	0.68555	1.14015	0.87705	1.21510	0.98748
Washington	2.04710	1.71375	2.96140	3.25648	1.80336	2.35642
Alaska	13.01603	11.82305	10.44126	12.02777	10.78633	11.61889
Annual Mean	2.19950	1.38618	1.96028	1.93848	1.27058	

Table 3b: State by Year Snow Depth by Census Division
(Maximum Daily in Inches Averaged Across State's Weather Stations)

Census Divisions	1994	1995	1996	1997	1998	State Mean Snow Depth
Northeast Region:						
New England:						
Connecticut	13.55882	9.03125	16.03030	5.58621	2.52000	9.34532
Rhode Island	16.42857	7.33333	20.80000	13.00000	3.00000	12.11238
Massachusetts	17.84848	11.53846	24.95238	14.82143	6.16364	15.06488
New Hampshire	25.02128	18.75000	25.58333	17.95745	16.45833	20.75408
Vermont	29.70833	20.91304	26.37778	19.93333	20.57778	23.50205
Maine	25.46667	23.77632	24.80000	22.64103	22.75949	23.88870
Middle Atlantic:						
New Jersey	10.56250	6.82979	16.59575	4.43478	1.83333	8.05123
Pennsylvania	18.68132	11.71038	21.29474	6.90270	5.43889	12.80561
New York	21.53478	13.14027	19.06635	11.66990	9.65686	15.01363
Midwest Region:						
East North Central:						
Illinois	7.18571	4.59009	5.01333	6.85281	3.75431	5.47925
Indiana	5.98077	4.46584	9.11392	5.14024	4.81595	5.90335
Ohio	8.20000	6.68794	10.52857	2.56738	2.17808	6.03239
Michigan	16.29054	13.68027	14.12418	15.91250	10.05917	14.01333
Wisconsin	14.80000	10.17778	17.51955	16.40556	11.24731	14.03004
West North Central:						
Missouri	3.44898	4.22959	3.40306	5.00529	2.01075	3.61954
Kansas	3.08856	3.49265	2.45455	5.70968	4.09929	3.76894
Nebraska	7.60494	6.49796	5.05714	8.50800	7.48810	7.03123
Iowa	11.53333	7.73810	11.05952	9.72189	7.92941	9.59645
South Dakota	11.01899	9.36076	13.34591	15.11321	8.50000	11.46777
North Dakota	14.43046	7.90667	13.21088	15.93793	9.60959	12.21911
Minnesota	16.43386	12.52941	20.24176	21.78889	10.33702	16.26619
South Region:						
South Atlantic:						
Florida	0.05714	0.00000	0.00000	0.00000	0.00000	0.01143
Georgia	0.01961	0.09396	0.27891	0.12925	0.10884	0.12611
South Carolina	0.04082	0.02083	0.48000	0.13131	0.05155	0.14490
North Carolina	0.88679	0.83544	3.59873	1.65385	2.25949	1.84686
Virginia	4.21094	3.67669	14.21805	4.26357	3.64567	6.00298
Delaware	3.85714	3.28571	11.14286	2.00000	0.60000	6.09524
Maryland	5.906977	5.95455	19.45238	4.25000	2.41861	8.01888
West Virginia	13.07692	8.18269	18.97143	7.95283	8.27193	11.29116
District of Columbia	NA	NA	NA	NA	NA	NA
East South Central:						
Alabama	0.00000	0.11628	0.26154	0.12698	0.01527	0.10401
Mississippi	0.10294	0.28467	0.49286	0.32857	0.51049	0.34391
Tennessee	1.31579	1.96117	4.57692	2.65455	2.28319	2.55832
Kentucky	5.02439	1.54601	6.36145	1.64780	5.29375	3.97468

West South Central:						
Louisiana	0.00000	0.00000	0.02454	0.06433	0.02326	0.02242
Texas	0.42080	0.21361	0.50154	0.67077	0.33489	0.42832
Arkansas	1.23288	0.97333	1.36184	1.88742	0.09032	1.10916
Oklahoma	1.99539	2.43396	1.66038	2.30516	0.79412	1.83780
West Region:						
Mountain:						
Arizona	2.02273	2.23256	1.64706	4.37423	2.23781	2.50288
New Mexico	3.33793	3.71329	4.07303	7.82955	3.41437	4.47363
Nevada	4.70192	3.16822	6.41667	3.88350	5.20755	4.67557
Montana	8.40329	6.78571	13.84100	10.44770	7.68465	9.43247
Utah	8.82353	8.96970	12.10119	12.85714	11.69461	10.88923
Colorado	10.15714	11.04673	11.46729	17.10314	10.97273	12.14941
Idaho	11.36364	8.30303	17.51111	13.29688	11.79845	12.45462
Wyoming	11.59231	10.90909	14.27692	14.74803	13.87597	13.08046
Pacific:						
Hawaii	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
California	3.22922	4.36935	3.44529	2.52632	4.25064	3.56416
Oregon	5.16114	3.93365	6.94231	4.07882	6.08654	5.24049
Washington	7.77564	6.83117	16.95333	11.74830	8.71429	10.40455
Alaska	29.67832	30.51449	23.43750	22.36420	21.75333	25.54957
Annual Mean	8.86444	6.89472	10.72138	8.14152	6.20877	

Table 4: Estimated Logit Models for Driver Fatality Risk, by Vehicle Type
(With Sample Selection Adjustment, Two-Vehicle Crashes, 1991-1998)

Independent Variables	*Driver in Car*	*Driver in Utility Vehicle*	*Driver in Van*	*Driver in Pickup*
Intercept	-0.0432	-0.9023 [b]	-2.0467 [a]	-1.2447 [a]
	(0.1227)	(0.3941)	(0.4314)	(0.2106)
	[0.5418]	[0.1981]	[0.1731]	[0.2201]
	{1.0000}	{0.4417}	{0.2972}	{0.4964}
Opposing Vehicle	1.2780 [a]	1.4896 [a]	1.3019 [a]	1.2982 [a]
is a Utility Vehicle	(0.0461)	(0.1667)	(0.1428)	(0.0785)
	[0.8419]	[0.5588]	[0.4839]	[0.5490]
	{1.8771}	{0.7932}	{0.5452}	{0.8680}
Opposing Vehicle	1.6490 [a]	1.7330 [a]	1.6811 [a]	1.5841 [a]
is a Van	(0.0491)	(0.1490)	(0.1297)	(0.0767)
	[0.8717]	[0.5806]	[0.5394]	[0.5938]
	{1.4966}	{0.6541}	{0.5804}	{0.7522}
Opposing Vehicle	1.4191 [a]	1.6290 [a]	1.5653 [a]	1.4442 [a]
is a Pickup	(0.0286)	(0.0930)	(0.0928)	(0.0429)
	[0.8323]	[0.5471]	[0.4833]	[0.5431]
	{1.8770}	{0.8650}	{0.6123}	{1.3034}
Pseudo-R^2	0.3097	0.3198	0.3126	0.3020
Number of Observations	62,395	6,167	6,661	22,703
Number of Missing Obs.	3,475	295	384	976

Notes: The sample consists of all two-vehicle crashes from 1991 through 1998 in which at least one driver died. Each column pulls from this sample those observations involving drivers of the type of vehicle listed in the column heading. The dependent variable of the logit model equals one if the driver of the vehicle died. The driver covariates are for both drivers. They include age, age squared, sex, air bag deployment, seat belt use, drunk driver, and previous major and minor traffic incidents. The crash covariates are the road condition, type of road, speed limit, time of day, and year. Heteroskedastic consistent standard errors are reported in parentheses. The predicted probabilities given a crash with each opposing vehicle are reported in brackets. These probabilities are adjusted for sample selection bias and standardized, and the values are reported in braces.

[a] Significant at 1% level, two-sided test; [b] Significant at 5% level, two-sided test; [c] Significant at the 10% level, two-sided test.

Table 5a: Estimated OLS Models for Pedestrian Fatalities, by Vehicle Type

	(1)	(2)	(3)	(4)	(5)	(6)
Cars						
Vehicle Miles Traveled (millions)	0.00198 [a]	0.00223 [a]	0.00198 [a]	0.00223 [a]	0.00217 [a]	0.00225 [a]
	(0.00016)	(0.00020)	(0.00017)	(0.00020)	(0.00019)	(0.00021)
R-squared	0.8871	0.9371	0.8877	0.9372	0.9235	0.9436
Light Trucks						
Vehicle Miles Traveled (millions)	0.00517 [a]	0.00672 [a]	0.00517 [a]	0.00674 [a]	0.00557 [a]	0.00681 [a]
	(0.00112)	(0.00173)	(0.00111)	(0.00174)	(0.00142)	(0.00185)
R-squared	0.7411	0.8878	0.7428	0.8888	0.8437	0.8943
State Fixed Effects	No	Yes	No	Yes	No	Yes
Year Indicators	No	No	Yes	Yes	Yes	Yes
State-Specific Linear Time Trend	No	No	No	No	Yes	Yes
Light Truck Crash Frequency Relative to Car Crash Frequency	2.61	3.01	2.61	3.02	2.57	3.03

Table 5b: Estimated OLS Models for Pedestrian Fatalities, by Vehicle Type (Less Restrictive Controls)

	(1)	(2)	(3)	(4)	(5)	(6)
Cars						
Vehicle Miles Traveled (millions)	0.00224 [a]	0.00260 [a]	0.00198 [a]	0.00223 [a]	0.00223 [a]	0.00260 [a]
	(0.00024)	(0.00030)	(0.00017)	(0.00020)	(0.00025)	(0.00030)
R-squared	0.9529	0.9759	0.8889	0.9385	0.9534	0.9762
Light Trucks						
Vehicle Miles Traveled (millions)	0.00479 [a]	0.00592 [a]	0.00517 [a]	0.00673 [a]	0.00479 [a]	0.00592 [a]
	(0.00117)	(0.00195)	(0.00111)	(0.00173)	(0.00116)	(0.00196)
R-squared	0.9199	0.9601	0.7437	0.8893	0.9225	0.9608
State Fixed Effects	No	Yes	No	Yes	No	Yes
Year by Roadtype Indicators	No	No	Yes	Yes	Yes	Yes
State by Roadtype-Specific Linear Time Trend	Yes	Yes	No	No	Yes	Yes
Light Truck Crash Frequency Relative to Car Crash Frequency	2.14	2.27	2.61	3.02	2.15	2.28

Notes: For the first equation in each panel the independent variable of interest is the VMT by cars, and the dependent variable is the number of pedestrian fatalities caused by cars. For the second equation the independent variable of interest is the VMT by light trucks, and the dependent variable is the number of pedestrian fatalities caused by light trucks. Each model contains roadtype indicators, and each controls for the state unemployment rate, the legal speed limit, the proportion of the population that is male and 65 years old and older, the proportion that is female and 65 years old and older, the proportion that is male and 15-29 years old, and the proportion that is female and 15-29 years old. Consistent standard errors allowing for error clustering within states are reported in parentheses. The data set is on a state by year by roadtype level. There are fifty-one states (including DC), and the years of available data for VMT distribution by vehicle type are from 1994 through 1998. The two roadtypes are rural and urban. This yields 510 observations; however, 116 observations either have not availabe data or did not have any vehicle travel within the roadtype (e.g., DC rural). Thus, the car and truck regressions contain 394 observations, and the single regression equation contains 788 (394 x 2) observations.

[a] Significant at 1% level, two-sided test; [b] Significant at 5% level, two-sided test; [c] Significant at the 10% level, two-sided test.

Table 6a: First-Stage Coefficient Estimates of IV Models
(Instrument = Average Daily Snow Depth in Jan, Feb, March, Oct, Nov, Dec)

	(1)	(2)	(3)	(4)	(5)	(6)
Cars						
First-Stage Snow Coefficient Estimate	-1417.98	-2065.79	-1546.04	-1650.34	-2619.47	-1877.26
	[15.95]	[29.27]	[16.38]	[13.98]	[28.66]	[15.25]
	{0.0510}	{0.0724}	{0.0536}	{0.0461}	{0.0724}	{0.0515}
Light Trucks						
First-Stage Snow Coefficient Estimate	-312.39	-509.14	-368.83	-266.27	-477.63	-324.81
	[23.36]	[53.79]	[28.57]	[11.76]	[30.51]	[14.97]
	{0.0729}	{0.1254}	{0.0900}	{0.0391}	{0.0768}	{0.0499}
Census Division Indicators	No	No	No	Yes	Yes	Yes
Year by Roadtype Indicators	No	Yes	Yes	No	Yes	Yes
State by Roadtype-Specific Linear Time Trend	Yes	No	Yes	Yes	No	Yes
Million miles more of truck VMT relative to car VMT given additional Inch of Avg. Snow Depth	1105.59	1556.65	1177.21	1384.07	2141.84	1552.45

Table 6b: First-Stage Coefficient Estimates of IV Models
(Instrument =Maximum Daily Snow Depth in a Year Averaged Across Weather Stations)

	(1)	(2)	(3)	(4)	(5)	(6)
Cars						
First-Stage Snow Coefficient Estimate	-372.07	-547.18	-399.95	-465.73	-714.44	-525.91
	[12.86]	[20.86]	[12.30]	[11.59]	[17.46]	[11.81]
	{0.0415}	{0.0527}	{0.0408}	{0.0386}	{0.0454}	{0.0403}
Light Trucks						
First-Stage Snow Coefficient Estimate	-86.52	-141.65	-101.16	-74.31	-113.75	-89.08
	[17.37]	[30.28]	[17.91]	[17.03]	[23.09]	[19.08]
	{0.0657}	{0.1000}	{0.0764}	{0.0320}	{0.0366}	{0.0374}
Census Division Indicators	No	No	No	Yes	Yes	Yes
Year by Roadtype Indicators	No	Yes	Yes	No	Yes	Yes
State by Roadtype-Specific Linear Time Trend	Yes	No	Yes	Yes	No	Yes
Million miles more of truck VMT relative to car VMT given additional Inch of Avg. Snow Depth	285.55	405.53	298.79	391.42	600.69	436.83

Notes: For the top panel, the first-stage regressions use as an instrument the average snow depth in a given day (in January, February, March, October, November, December) over all reporting weather stations in a state for a given year. For the bottom panel, the first-stage regressions use as an instrument the maximum snow depth in a year (averaged across all weather stations in the state). The IV variables instrument for VMT of light trucks and cars. Each model contains roadtype indicators, and each controls for the state unemployment rate, the legal speed limit, the proportion of the population that is male and 65 years old and older, the proportion that is female and 65 years old and older, the proportion that is male and 15-29 years old, and the proportion that is female and 15-29 years old. Consistent F-statistics for the excluded instrument (allowing for error clustering within states) are reported in brackets, and the partial-R2 are reported in braces.

Table 7a: Estimated IV Models for Crash Frequencies
(Instrument = Average Daily Snow Depth in Jan, Feb, March, Oct, Nov, Dec)

Dependent Variable:	Pedestrian Deaths in Summer Months					
	(1)	(2)	(3)	(4)	(5)	(6)
Cars						
Vehicle Miles Traveled (millions)	0.00035 [a]	0.00034 [a]	0.00034 [a]	0.00040 [a]	0.00037 [a]	0.00042 [a]
	(0.00008)	(0.00005)	(0.00008)	(0.00009)	(0.00006)	(0.00008)
Light Trucks						
Vehicle Miles Traveled (millions)	0.00121 [a]	0.00101 [a]	0.00112 [a]	0.00160 [a]	0.00139 [a]	0.00151 [a]
	(0.00035)	(0.00015)	(0.00027)	(0.00044)	(0.00029)	(0.00037)
Census Division Indicators	No	No	No	Yes	Yes	Yes
Year by Roadtype Indicators	No	Yes	Yes	No	Yes	Yes
State by Roadtype-Specific Linear Time Trend	Yes	No	Yes	Yes	No	Yes
Light Truck Crash Frequency Relative to Car Crash Frequency	3.45	2.93	3.29	4.00	3.76	3.60

Table 7b: Estimated IV Models for Crash Frequencies
(Instrument =Maximum Daily Snow Depth in a Year Averaged Across Weather Stations)

Dependent Variable:	Pedestrian Deaths in Summer Months					
	(1)	(2)	(3)	(4)	(5)	(6)
Cars						
Vehicle Miles Traveled (millions)	0.00040 [a]	0.00040 [a]	0.00037 [a]	0.00048 [a]	0.00043 [a]	0.00048 [a]
	(0.00008)	(0.00006)	(0.00009)	(0.00010)	(0.00006)	(0.00009)
Light Trucks						
Vehicle Miles Traveled (millions)	0.00116 [a]	0.00105 [a]	0.00108 [a]	0.00146 [a]	0.00167 [a]	0.00142 [a]
	(0.00032)	(0.00018)	(0.00026)	(0.00036)	(0.00043)	(0.00032)
Census Division Indicators	No	No	No	Yes	Yes	Yes
Year by Roadtype Indicators	No	Yes	Yes	No	Yes	Yes
State by Roadtype-Specific Linear Time Trend	Yes	No	Yes	Yes	No	Yes
Light Truck Crash Frequency Relative to Car Crash Frequency	2.90	2.63	2.92	3.04	3.88	2.96

Notes: For the top panel, I use as an instrument the average snow depth in a given day (in January, February, March, October, November, December) over all reporting weather stations in a state for a given year. For the bottom panel, I use as an instrument the maximum snow depth in a year (averaged across all weather stations in the state). The IV variables instrument for VMT of light trucks and cars. In order to meet the exclusion restriction, I restrict the dependent variable to the number of pedestrian fatalities that occurred only during June, July, and August. Each model contains roadtype indicators, and each controls for the state unemployment rate, the legal speed limit, the proportion of the population that is male and 65 years old and older, the proportion that is female and 65 years old and older, the proportion that is male and 15-29 years old, and the proportion that is female and 15-29 years old. Consistent standard errors allowing for error clustering within states are reported in parentheses. [a] Significant at 1% level, two-sided test.

Table 8: The Expected Relative Number of Fatalities

Panel A: Assuming Constant Crash Frequencies
(Cell [i,j] represents vehicle composition of 50% i and 50% j vehicles)

	Car	Utility	Van	Pickup
Car	1.00	1.03	0.84	1.17
Utility		0.79	0.64	0.96
Van			0.58	0.81
Pickup				1.30

Panel B: Combining Fatality Risk with Estimated Variable Crash Frequencies
(Cell [i,j] represents vehicle composition of 50% i and 50% j vehicles)

	Car	Utility	Van	Pickup
Car	1.00	2.94	2.14	3.77
Utility		4.96	4.02	5.98
Van			3.63	5.08
Pickup				8.15

Panel C: Using IV Estimates with Driver Fatalities by Vehicle Type as Outcome Measure
(Cell [i,j] represents vehicle composition of 50% i and 50% j vehicles)

	Car	Light Truck
Car	1.00	2.45-5.69
Light Truck		3.90-10.37

Table 9a: Estimated IV Models for Total Driver Fatalities in Two-Vehicle Crashes
(Instrument = Average Daily Snow Depth in Jan, Feb, March, Oct, Nov, Dec)

Dependent Variable:	Total Driver Deaths in Two-Vehicle Crashes in Summer Months					
	(1)	(2)	(3)	(4)	(5)	(6)
Cars						
Vehicle Miles Traveled (millions)	0.00059 [b]	0.00071 [a]	0.00067 [b]	0.00035 [b]	0.00048 [a]	0.00041 [a]
	(0.00029)	(0.00020)	(0.00029)	(0.00016)	(0.00010)	(0.00014)
Light Trucks						
Vehicle Miles Traveled (millions)	0.00431 [a]	0.00338 [a]	0.00403 [a]	0.00364 [a]	0.00272 [a]	0.00321 [a]
	(0.00126)	(0.00050)	(0.00097)	(0.00105)	(0.00036)	(0.00068)
Census Division Indicators	No	No	No	Yes	Yes	Yes
Year by Roadtype Indicators	No	Yes	Yes	No	Yes	Yes
State by Roadtype-Specific Linear Time Trend	Yes	No	Yes	Yes	No	Yes
Light Truck Crash Frequency Relative to Car Crash Frequency	7.31	4.75	6.04	10.37	5.73	7.77

Table 9b: Estimated IV Models for Total Driver Fatalities in Two-Vehicle Crashes
(Instrument =Maximum Daily Snow Depth in a Year Averaged Across Weather Stations)

Dependent Variable:	Total Driver Deaths in Two-Vehicle Crashes in Summer Months					
	(1)	(2)	(3)	(4)	(5)	(6)
Cars						
Vehicle Miles Traveled (millions)	0.00094 [b]	0.000989 [a]	0.001091 [b]	0.00056 [a]	0.00062 [a]	0.00064 [a]
	(0.00043)	(0.00099)	(0.00052)	(0.00020)	(0.00014)	(0.00019)
Light Trucks						
Vehicle Miles Traveled (millions)	0.00483 [a]	0.00386 [a]	0.00458 [a]	0.00422 [a]	0.00320 [a]	0.00373 [a]
	(0.00165)	(0.00075)	(0.00145)	(0.00149)	(0.00075)	(0.00107)
Census Division Indicators	No	No	No	Yes	Yes	Yes
Year by Roadtype Indicators	No	Yes	Yes	No	Yes	Yes
State by Roadtype-Specific Linear Time Trend	Yes	No	Yes	Yes	No	Yes
Light Truck Crash Frequency Relative to Car Crash Frequency	5.14	3.90	4.20	7.60	5.18	5.84

Notes: For the top panel, I use as an instrument the average snow depth in a given day (in January, February, March, October, November, December) over all reporting weather stations in a state for a given year. For the bottom panel, I use as an instrument the maximum snow depth in a year (averaged across all weather stations in the state). The IV variables instrument for VMT of light trucks and cars. In order to meet the exclusion restriction, I restrict the dependent variable to the number of pedestrian fatalities that occurred only during June, July, and August. Each model contains roadtype indicators, and each controls for the state unemployment rate, the legal speed limit, the proportion of the population that is male and 65 years old and older, the proportion that is female and 65 years old and older, the proportion that is male and 15-29 years old, and the proportion that is female and 15-29 years old. Consistent standard errors allowing for error clustering within states are reported in parentheses. [a] Significant at 1% level, two-sided test; [b] Significant at 5% level, two-sided test.

Table 10: Sample Means by High vs. Low Snow Depth (1998 Cross Section of States)

	Average Daily Snow Depth			Maximum Daily Snow Depth		
State Variable	Below Median	Above Median	T-Statistic	Below Median	Above Median	T-Statistic
State Unemployment Rate	4.4480	4.2280	0.7582	4.5400	4.1360	1.4126
	(0.1726)	(0.2332)		(0.1735)	(0.2274)	
Rural Speed Limit 70 or Over	0.4800	0.6000	0.8402	0.5600	0.5200	0.2782
	(0.1020)	(0.1000)		(0.1013)	(0.1020)	
Proportion Male and 65 and Over	0.0528	0.0529	0.0319	0.0530	0.0527	0.1194
	(0.0015)	(0.0017)		(0.0016)	(0.0016)	
Proportion Female and 65 and Over	0.0761	0.0725	1.0287	0.0756	0.0730	0.7380
	(0.0020)	(0.0028)		(0.0020)	(0.0028)	
Proportion Male and 15-29 Years Old	0.1046	0.1062	0.7813	0.1047	0.1060	0.6256
	(0.0012)	(0.0017)		(0.0013)	(0.0017)	
Proportion Male and 15-29 Years Old	0.1030	0.1033	0.1282	0.1026	0.1038	0.5231
	(0.0014)	(0.0017)		(0.0014)	(0.0017)	

Table 11a: Testing IV Models using Unemployment Rate as Outcome Variable
(Instrument = Average Daily Snow Depth in Jan, Feb, March, Oct, Nov, Dec)

	Dependent Variable: State by Year Unemployment Rate					
	(1)	(2)	(3)	(4)	(5)	(6)
Cars						
Vehicle Miles Traveled (millions)	0.00002	0.00004	0.00006	-0.00007	-0.00004	-0.00002
	(0.00004)	(0.00003)	(0.00004)	(0.00005)	(0.00003)	(0.00004)
Light Trucks						
Vehicle Miles Traveled (millions)	0.00011	0.00018	0.00026	-0.00042	-0.00021	0.00010
	(0.00021)	(0.00015)	(0.00019)	(0.00033)	(0.00016)	(0.00022)
Census Division Indicators	No	No	No	Yes	Yes	Yes
Year by Roadtype Indicators	No	Yes	Yes	No	Yes	Yes
State by Roadtype-Specific Linear Time Trend	Yes	No	Yes	Yes	No	Yes

Table11b: Testing IV Models using Unemployment Rate as Outcome Variable
(Instrument =Maximum Daily Snow Depth in a Year Averaged Across Weather Stations)

	Dependent Variable: State by Year Unemployment Rate					
	(1)	(2)	(3)	(4)	(5)	(6)
Cars						
Vehicle Miles Traveled (millions)	2.92E-06	0.00005	0.00005	-0.00010 [c]	-0.00006	-0.00004
	(0.00005)	(0.00004)	(0.00005)	(0.00006)	(0.00004)	(0.00005)
Light Trucks						
Vehicle Miles Traveled (millions)	0.00001	0.00019	0.00021	-0.00060 [c]	-0.00034	-0.00023
	(0.00022)	(0.00016)	(0.00021)	(0.00036)	(0.00024)	(0.00028)
Census Division Indicators	No	No	No	Yes	Yes	Yes
Year by Roadtype Indicators	No	Yes	Yes	No	Yes	Yes
State by Roadtype-Specific Linear Time Trend	Yes	No	Yes	Yes	No	Yes

Notes: For the top panel, I use as an instrument the average snow depth in a given day (in January, February, March, October, November, December) over all reporting weather stations in a state for a given year. In the bottom panel, I use the maximum snow depth in year (averaged across all weather stations in the state). These variables instrument for VMT of light trucks and cars. I use state unemployment rate as the outcome variable. Each model contains roadtype indicators, the legal speed limit, the proportion of the population that is male and 65 and over, the proportion that is female and 65 and over, the proportion that is male and 19-25 years old, and the proportion that is female and 19-25 years old. Consistent standard errors allowing for error clustering within states are reported in parentheses. [c] Significant at the 10% level, two-sided test.

Table 12a: Testing IV Models using Chronic Liver and Cirrhosis Death Rate (per 100,000) as Outcome Variable
(Instrument = Average Daily Snow Depth in Jan, Feb, March, Oct, Nov, Dec)

	Dependent Variable: State by Year Chronic Liver and Cirrhosis Death Rate					
	(1)	(2)	(3)	(4)	(5)	(6)
Cars						
Vehicle Miles Traveled (millions)	0.00013 [c]	0.00008	0.00011	0.00007	0.00005	0.00007
	(0.00007)	(0.00005)	(0.00007)	(0.00006)	(0.00005)	(0.00006)
Light Trucks						
Vehicle Miles Traveled (millions)	0.00057 [c]	0.00032	0.00045	0.00043	0.00025	0.00038
	(0.00032)	(0.00020)	(0.00028)	(0.00036)	(0.00026)	(0.00034)
Census Division Indicators	No	No	No	Yes	Yes	Yes
Year by Roadtype Indicators	No	Yes	Yes	No	Yes	Yes
State by Roadtype-Specific Linear Time Trend	Yes	No	Yes	Yes	No	Yes

Table 12b: Testing IV Models using Chronic Liver and Cirrhosis Death Rate (per 100,000) as Outcome Variable
(Instrument =Maximum Daily Snow Depth in a Year Averaged Across Weather Stations)

	Dependent Variable: State by Year Chronic Liver and Cirrhosis Death Rate					
	(1)	(2)	(3)	(4)	(5)	(6)
Cars						
Vehicle Miles Traveled (millions)	0.00012 [c]	0.00007	0.00010	0.00007	0.00004	0.00005
	(0.00007)	(0.00005)	(0.00007)	(0.00005)	(0.00005)	(0.00006)
Light Trucks						
Vehicle Miles Traveled (millions)	0.00054 [c]	0.00029	0.00040	0.00043	0.00027	0.00030
	(0.00032)	(0.00019)	(0.00028)	(0.00035)	(0.00031)	(0.00034)
Census Division Indicators	No	No	No	Yes	Yes	Yes
Year by Roadtype Indicators	No	Yes	Yes	No	Yes	Yes
State by Roadtype-Specific Linear Time Trend	Yes	No	Yes	Yes	No	Yes

Notes: For the top panel, I use as an instrument the average snow depth in a given day (in January, February, March, October, November, December) over all reporting weather stations in a state for a given year. In the bottom panel, I use the maximum snow depth in the year (averaged across all weather stations in the state). These variables instrument for VMT of light trucks and cars. I use state by year chronic liver and cirrhosis death rate as the outcome variable. Each model contains roadtype indicators, the legal speed limit, the proportion of the population that is male and 65 and over, the proportion that is female and 65 and over, the proportion that is male and 19-25 years old, and the proportion that is female and 19-25 years old. Consistent standard errors allowing for error clustering within states are reported in parentheses. [c] Significant at the 10% level, two-sided test.

Figure 1: Relative External and Internal Driver Risk in a Two-Vehicle Crash