Pedestrian Deaths and Large Vehicles

Justin Tyndall

jtyndall@hawaii.edu
University of Hawai'i Economic Research Organization
University of Hawai'i at Mānoa Department of Economics
2424 Maile Way, Saunders Hall 540, Honolulu, HI, USA, 96822

July, 2021

Abstract

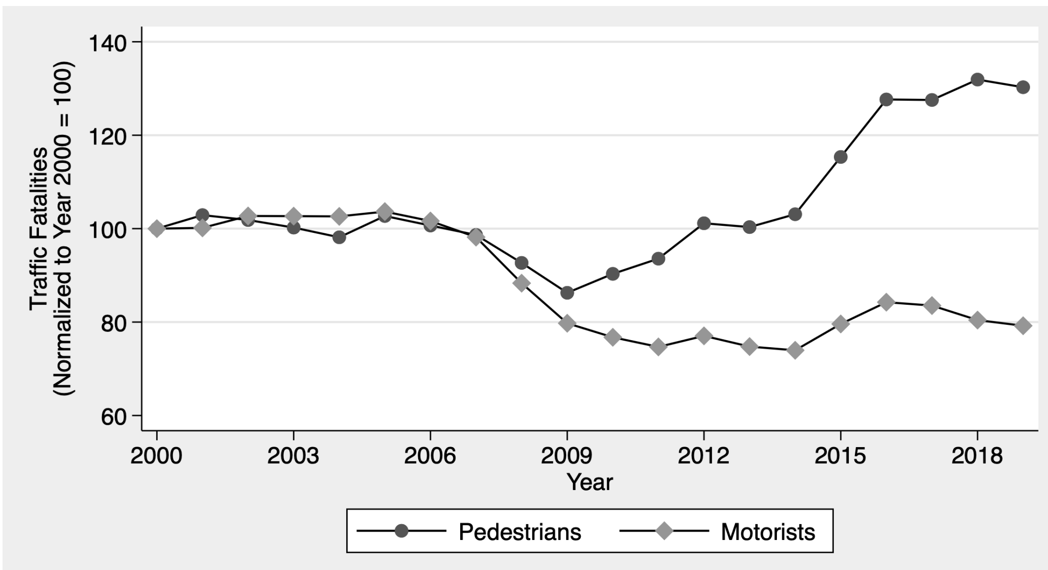
Traffic fatalities in the US have been rising among pedestrians even as they fall among motorists. Contemporaneously, the US has undergone a significant shift in consumer preferences for motor vehicles, with larger Sport Utility Vehicles comprising an increased market share. Larger vehicles may pose a risk to pedestrians, increasing the severity of collisions. I use data covering all fatal vehicle collisions in the US and exploit heterogeneity in changing vehicle fleets across metros for identification. Between 2000 and 2019, I estimate that replacing the growth in Sport Utility Vehicles with cars would have averted 1,100 pedestrian deaths. I find no evidence that the shift towards larger vehicles improved aggregate motorist safety.

Transportation; Safety; Health; Traffic Fatalities; Externalities
I1; R41; R42; R48

1 Introduction

Between 2000 and 2019, motor vehicle crashes killed 741,000 people in the US including 100,000 pedestrians.[1] Figure 1 charts the trends in traffic fatalities for both vehicle occupants and pedestrians over the 2000-2019 period. While deaths among motorists have declined over this period, deaths among pedestrians have risen by 30%. Over the same period the consumer market for private vehicles has shifted towards larger vehicles and particularly towards Sport Utility Vehicles (SUVs). Larger vehicles may impose a negative externality on pedestrians by making crashes involving pedestrians more lethal. I estimate the effect of large vehicle uptake on the pedestrian fatality rate.

Figure 1: Trends in Traffic Fatalities



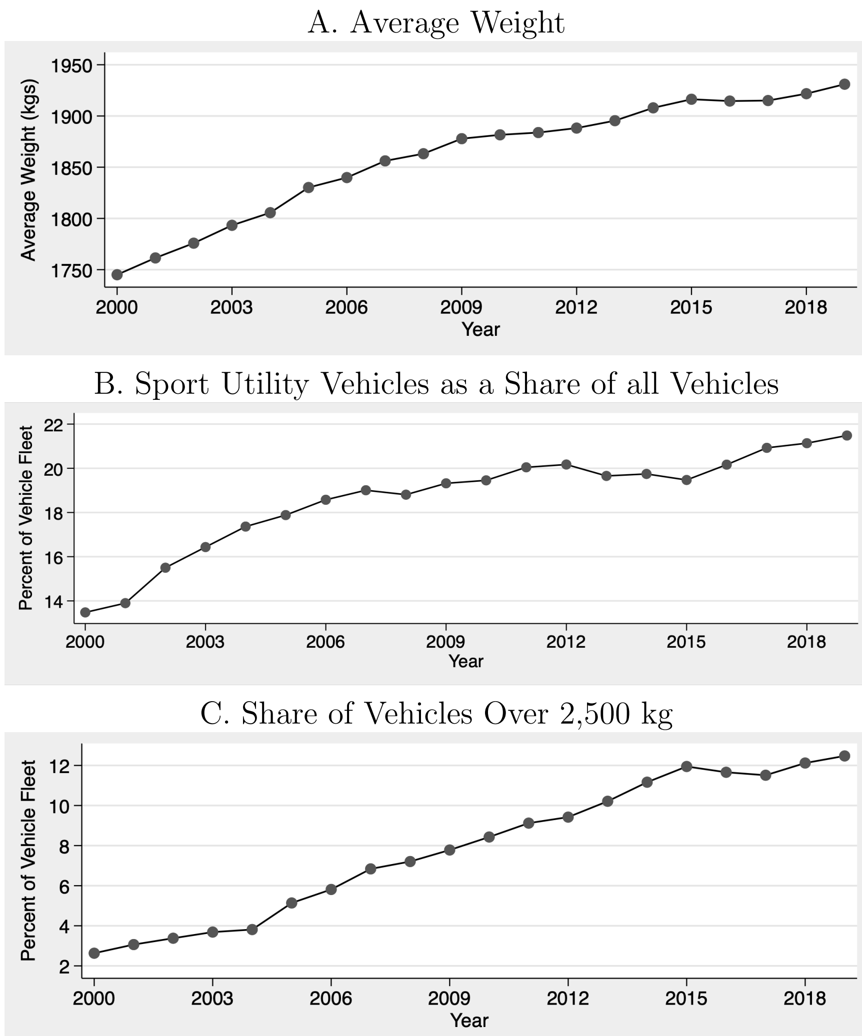
The number of fatalities among drivers and their passengers fell by 21% between 2000 and 2019. Over the same period the number of motor vehicle related fatalities among pedestrians increased by 30%.

Vehicles on US roads became measurably larger between 2000 and 2019. Figure 2 plots changes in vehicle characteristics among all vehicles involved in a fatal crash between 2000 and 2019. While in 2000, the typical vehicle weighed 1,745 kg, by 2019 the average vehicle had increased in weight by 10.7% to 1,931 kg (Figure 2A). Additionally, SUVs increased their prevalence from 13.5% of vehicles to 21.5% (Figure 2B). Over this same period, a new class of very large vehicles began to enter the consumer market.

[1] National Highway Traffic Safety Administration, Fatality Analysis Reporting System.

In 2000 only 2.6% of vehicles involved in fatal crashes weighed more than 2,500 kg, by 2019 the share had increased fivefold to 12.5% (Figure 2C).[2] The increased prevalence of these very large vehicles was mostly attributable to the popularity of a few large SUVs, particularly the Ford Expedition and the Chevrolet Suburban and Tahoe.

Figure 2: Changes in Vehicle Size Among Vehicles Involved in a Fatal Crash



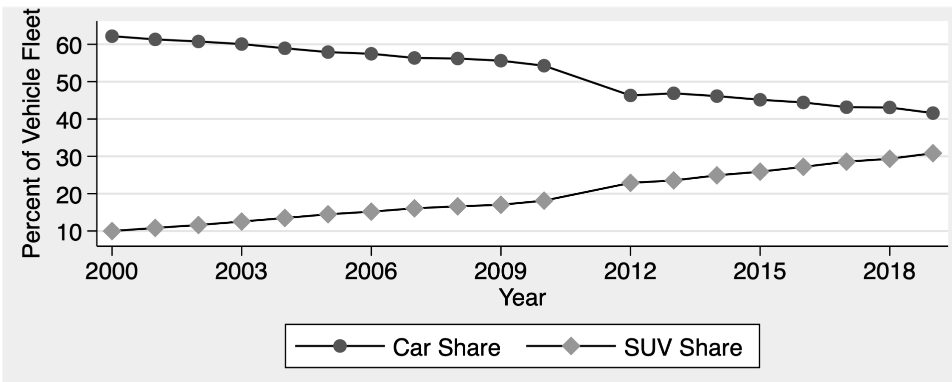
Between 2000 and 2019 the average weight of consumer vehicles involved in a fatal crash increased by 11%, the prevalence of SUVs increased by 59% and the share of vehicles that are more than 2,500 kg increased by 374%.

Data for vehicles involved in fatal incidents are more consistently collected in the

[2]Data comes from the national Fatality Analysis Reporting System and the Environmental Protection Agency's Fuel Economy Test Car List Database. Data details are included in Section 2.

US compared to vehicle registration data, as will be discussed in Section 2. However, vehicles involved in fatal crashes may not be representative of the overall vehicle fleet. Figure 3 uses data reported on registered vehicles by the Federal Highway Administration (FHWA). This source suggests an even more rapid shift towards SUVs and away from cars than is observed in crash data. The share of registered vehicles that are SUVs roughly tripled over this period, while the share of vehicles classified as cars fell by a third.

Figure 3: Registered Cars and SUVs as Share of National Vehicle Fleet



Data is reported by each state and tabulated by the Federal Highway Administration. 2011 data was not recorded consistently across states and is therefore omitted.

While larger vehicles are designed to protect their drivers and passengers in the event of a crash, less concern is given to the effect on pedestrians. Past research in the safety literature has considered the mechanisms that relate vehicle size to motorist and pedestrian safety. There are two primary mechanisms that could lead large vehicles to generate additional harm when hitting a pedestrian. First, the additional weight means the vehicle will take longer to come to a stop and will strike with more force as compared to a lighter vehicle. Second, large vehicles have higher front ends, affecting the point of impact on a pedestrian. A conventional car is likely to strike a pedestrian in the legs, propelling them over the hood of the vehicle. A vehicle with a higher front end is likely to make first contact with the pedestrian's torso or head, harming vital organs and deflecting their body under the vehicle. In transportation safety literature, pedestrians hit by light trucks (a category including SUVs, pickups and minivans) have been found to suffer greater rates of mortality (Simms and Wood, 2006; Tamura et al., 2008) and higher rates of brain injury (Roudsari et al., 2004) than those hit by cars.

Lefler and Gabler (2004) used US data from the 1990s to estimate that a pedestrian struck by a light truck is two to three times more likely to die than a pedestrian struck by a car. In a meta-analysis of papers concerned with pedestrian fatalities, Desapriya et al. (2010) found that the chance of fatal injury among pedestrians was 50% higher when struck by a light truck compared to a car. I will test for the effect of both vehicle weight and body type on pedestrian fatalities.

Significant past research has examined the effect of vehicle size on road safety. The adoption of the Corporate Average Fuel Economy (CAFE) vehicle emission standards in the US encouraged consumers to purchase lower emission vehicles, which were likely to be smaller. Crandall and Graham (1989) argued that this incentive resulted in higher rates of motorist fatalities due to smaller vehicles providing more limited protection to drivers. The authors pointed out that drivers of smaller vehicles are more vulnerable in crashes than they would be in a larger vehicle and extrapolate this effect across the market. However, this method ignores external safety risks that larger vehicles may impart by increasing the severity of injury to other motorists and to pedestrians. Focusing on a subset of crashes from the 1990s, Toy and Hammitt (2003) estimated the effect of vehicle types on injury severity in the US. Results indicated that SUVs fared better in protecting their driver in the event of a crash, but also inflicted more damage onto the drivers of other vehicles compared to cars. Further analysis of the interaction between light trucks in cars is provided in Gayer (2004), who similarly argued that the driver safety improvements provided by large vehicles may come at the expense of externally imposed risks. Estimates suggested that an increase in light trucks would increase overall traffic fatalities. Van Ommeren et al. (2013) focused on the relative weights of opposing vehicles involved in collisions in the Netherlands, estimating that a 500 kg increase in one car's weight increased the risk of a fatality by 70%. Ahmad and Greene (2005) revisited the analysis of Crandall and Graham (1989) specifically, finding little evidence that CAFE led to higher road fatalities in aggregate.

White (2004) attempts to directly estimate the marginal effect of drivers switching from smaller to larger vehicles in the US during 1995-2001. The assessment showed that for every driver whose life was saved on account of being in a larger vehicle, 4.3 fatalities were created among other road users, including motorists, cyclists and pedestrians. The paper also points out the inability of the legal system to provide incentives for drivers to internalize external safety risks, as drivers are typically only held responsible in cases of driver negligence rather than being held responsible for total damages inflicted.

Anderson (2008) examined cross state variation in light truck prevalence and traffic fatalities spanning the 1981-2004 period in the US. The author found that states with higher rates of light truck use had higher rates of traffic fatalities and that the increase in fatalities was primarily due to an increase in deaths among drivers and pedestrians who were struck by a light truck rather than a smaller vehicle. Anderson and Auffhammer (2014) quantified the safety externality of large vehicles, arguing the US vehicle fleet is inefficiently large and the externality could be corrected through gasoline taxes. Li (2012) also attempted to quantify the externality of light trucks, estimating the implied road safety externality of a light truck over its lifetime to be $2,400.

While there are several past studies linking vehicle size and safety, the current study is unique in a number of respects. First, I contribute an analysis covering a much more recent period in the US. The characteristics of the vehicle fleet have changed substantially during the 2010s. Second, I focus on the effect of vehicle size on pedestrian fatalities in particular. The sharp increase in pedestrian fatalities in the US is a recent phenomenon that has not been noted or studied in the economics literature. Third, I provide a new and novel data source by combining vehicle weight data from the Environmental Protection Agency (EPA) with crash level data and vehicle registration data that allows for analysis at the metropolitan level. Fourth, while prior studies have focused on vehicle weight, I focus on differences in vehicle body types, estimating unique effects for different varieties of light trucks.

The relationship between vehicle characteristics and pedestrian fatalities is one element of overall road safety. Significant economic research has been undertaken to investigate other causes of traffic fatalities such as vehicle speed (Ang et al., 2020; Van Benthem, 2015), road congestion (Green et al., 2016) alcohol consumption (Baughman et al., 2001; Green et al., 2014; Hansen, 2015; Jackson and Owens, 2011; Levitt and Porter, 2001; Ruhm, 1996), public policy and regulation (Basili and Belloc, 2020; Borsati et al., 2019; Bourgeon and Picard, 2007; Carpenter and Stehr, 2008; Karaca-Mandic and Ridgeway, 2010; Peltzman, 1975), electronic distractions (Blattenberger et al., 2013; Oviedo-Trespalacios, 2018) and the driver's state of mind (Giulietti et al., 2020). The current study is focused specifically on the effect of vehicle characteristics on road fatalities, with a particular focus on pedestrian effects.

The paper will proceed as follows. Section 2 provides information on data sources. Section 3 discusses the regression methodology. Section 4 provides results and Section 5 will conclude.

2 Data

I combine data from a number of public sources. Traffic fatality data is taken from the National Highway Traffic Safety Administration (NHTSA), Fatality Analysis Reporting System (FARS).[3] The data set is a complete record of all fatal traffic collisions in the US. To be included in the data the collision must have been on a public road, involved any type of motor vehicle and caused the death of one or more individuals.[4] The database contains a large number of variables characterizing the collision, including information on all vehicles and persons that were involved in the incident. The study period will cover 2000 to 2019. During this period FARS recorded 691,000 crashes that resulted in at least one fatality. These crashes included 1,045,000 vehicles and 1,754,000 individuals. 108,000 of the individuals were pedestrians. 760,000 individuals, including 100,000 pedestrians, died due to a crash. The enormous number of fatalities underlines the scope of the public health issue. The national distribution of crashes causing a pedestrian fatality are shown in Figure 4. Incidents cover all populated areas of the US, and are concentrated in city centers as well as extending along the interstate highway system. In the main analysis I analyze only crashes occurring within metropolitan areas, defined according to Core Based Statistical Area (CBSA) boundaries.

The empirical analysis will base estimates on the rate of deaths per 100,000 residents across metropolitan areas. The sample contains all metropolitan areas in the US for which data is available.[5] The final data set is a balanced panel containing 358 US metropolitan areas with annual observations spanning 2000-2019. Every incident recorded in the FARS data is accompanied by precise location information. I use the recorded county of the crash to assign each observation to a metropolitan area. Across all years and metros, the average rate of traffic fatalities was 13.3 deaths per 100,000 residents. The death rate among vehicle drivers and passengers was 11.7 per 100,000 while the rate of pedestrian deaths was 1.5 per 100,000. Summary statistics are provided in Table 1.

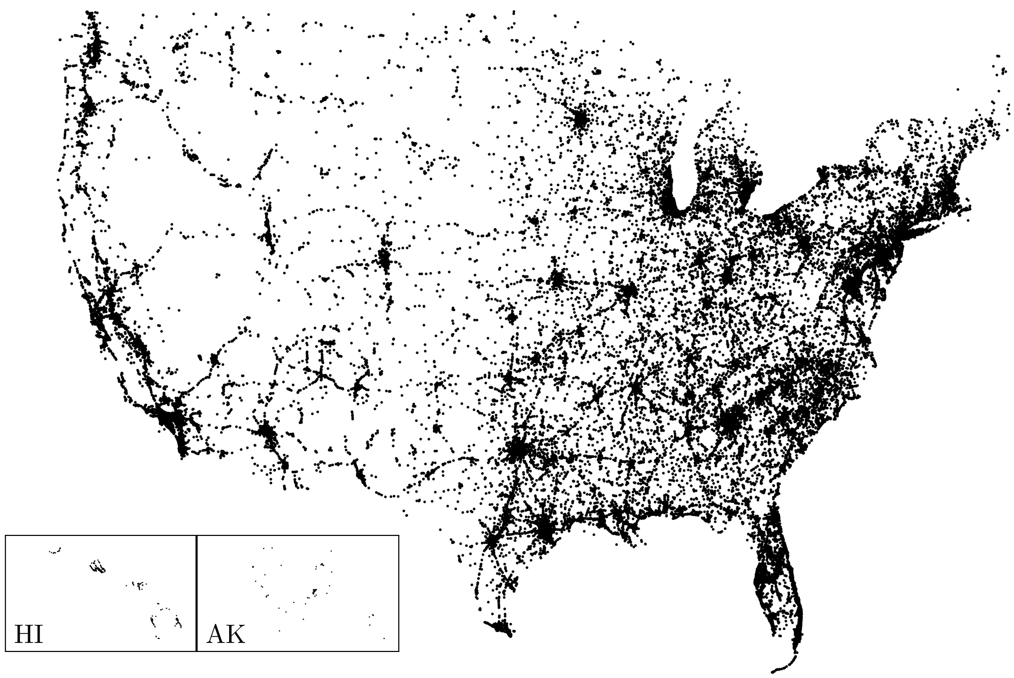
For each metropolitan area in the US I construct annual estimates of vehicle fleets

[3]The FARS data is publicly available at www.nhtsa.gov/research-data/fatality-analysis-reporting-system-fars.

[4]To be considered as caused by the vehicle crash, the death must occur within 30 days of the collision.

[5]Four metros were dropped because they had no fatal crashes that could be merged to EPA vehicle weight data in at least one year of analysis (Aimes, Iowa; Carson City, Nevada; Fairbanks, Alaska; and Sandusky, Ohio). Four metros were dropped because they lacked complete demographic data (Crestview, Fl; Lewiston, ID; Nort Port, Fl; Steubenville OH).

Figure 4: All US Pedestrian Fatalities 2001-2019



Each dot corresponds to a vehicle crash that resulted in at least one pedestrian fatality. The year 2000 is included in analysis but not in this figure as observations from 2000 lack latitude and longitude information.

by vehicle body type, relying on both FARS data and vehicle registration data. FARS includes variables on vehicle body type. In the average metro, 46.2% of consumer vehicles involved in a fatal crash were cars while 42.6% were light trucks and 11.2% were motorcycles. I examine the subgroups within the light truck category, including SUVs (16.3% of vehicles), pickup trucks (20.1%) and minivans (6.0%). I further brake down the SUV category into Small SUVs (12.3%) and Large SUVs (4.0%). Large SUVs are defined by FARS to be "full-size multi-purpose vehicles primarily designed around a shortened pickup truck chassis." I also include the FARS category of "Utility Station Wagon" in my definition of Large SUVs.[6] Utility Station Wagons have a similar body

[6]The Large SUV and Utility Station Wagon categories in FARS data includes the following vehicle models: Acura MDX; AMC Hummer; Avanti Studebaker XUV; Buick Enclave (2013 on); Cadillac Escalade/Escalade ESV; Chevrolet Full-size Blazer/Suburban/Tahoe/Travellall/Traverse (2013 on)/Yukon XL (2000 on); Chrysler Aspen; Dodge Durango (2004 on); Ford Full-size Bronco (1978 on)/Expedition/Excursion; GMC Acadia (2013 on)/Jimmy (1991-1994)/Yukon (Denali/XL); Honda Pilot; Hyundai Veracruz (2008 on); Infiniti QX56/QX80; Isuzu Ascender; Jeep Grand Cherokee/Grand Wagoneer; Kia Mesa/Borrego; Land Rover LR2/LR3/Freelander (2004 on)/Range Rover; Lexus LX450/470; Lincoln Navigator; Mazda CX-9; Mercedes Benz GL; Nissan Armada; Porsche Cayenne;

type to Large SUVs but are typically even larger, including an extended passenger area. Identifying Large SUVs will be important as these vehicles are likely to have body designs that include very high front ends, which safety tests have suggested could lead to increased pedestrian mortality. The most common Large SUVs in the data are the Chevrolet Tahoe, Ford Expedition and Chevrolet Suburban, which make up 24%, 15% and 14% of Large SUVs respectively.

When computing metropolitan vehicle fleet shares I focus on consumer vehicle shares, omitting the FARS vehicle categories for commercial buses and heavy trucks. Heavy trucks are classified as those exceeding 4,536 kg (10,000 pounds) and account for 9.2% of vehicles involved in fatal crashes over the study period. I also omit crashes involving vehicles that fall outside of the typical categories, including construction and farm equipment, golf carts, and snowmobiles. These unclassified vehicles comprised 1.7% of all vehicle observations.

To construct vehicle fleet shares at the metropolitan level, I augment FARS data with data on all vehicles registered in the US from the FHWA Highway Statistics data set.[7] The data contains the number of vehicles registered in each state, broken out by vehicle type. The FHWA state registration data includes categories for light trucks, SUVs, pickup trucks, minivans and motorcycles which correspond to FARS categories. The distinction between "Small" and "Large" SUVs that is made in the FARS data is not available in the FHWA data. Annual reports from FHWA cover all years of analysis. However, because vehicle registration data is reported to FHWA by individual states, the categorization of specific vehicles into categories is somewhat variable across states and across time. The registration data reported in 2011 uses a different methodology and will not be used. There are also instances where individual states introduce revised vehicle classification systems across the study period. The variability in reporting practices in the FHWA data limit the ability to make use of panel variation in the FHWA registration data. I therefore pool FHFA data cross all years to avoid relying on spurious time-series variation.

While prior studies have focused on state level variation, I choose to focus on the metropolitan area as the unit of analysis. The use of a smaller unit of geography allows estimates to be based on a larger set of observations that preserves more spatial variation in the data. There is significant heterogeneity in variables across metros within

Toyota Land Cruiser/Sequoia; and Volkswagen Touareg.

[7]The FHWA Highway Statistics data is publicly available at www.fhwa.dot.gov/policyinformation/statistics.cfm.

Table 1: Metropolitan Summary Statistics

Variable	Mean	Standard Deviation	Min	Max
Pedestrian deaths per 100,000	1.527	1.122	0.000	10.949
Motorist deaths per 100,000	11.727	5.801	0.616	60.023
Car share	0.462	0.106	0.000	1.000
Light truck share	0.426	0.101	0.000	1.000
SUV share	0.163	0.070	0.000	0.667
Small SUV share	0.123	0.062	0.000	0.600
Large SUV share	0.040	0.037	0.000	0.333
Pickup share	0.201	0.090	0.000	0.800
Minivan share	0.060	0.045	0.000	0.400
Motorcycle share	0.112	0.069	0.000	0.714
Average vehicle weight (kg)	1,855	113	1,304	2,517
Average model year	2000.1	4.8	1985.8	2014.0
Drunk driver related share	0.287	0.122	0.000	1.000
Average driver age	41.4	4.1	23.5	70.0
Population	718,334	1,614,325	49,832	20,320,876
College education share	0.245	0.090	0.061	0.805
High School education share	0.811	0.140	0.270	1.818
Median household income	48,428	10,579	24,863	130,865
Male share	0.492	0.011	0.456	0.583
GDP (millions $)	36,878,702	102,733,220	219,082	1,861,147,412
Health care quality score†	-0.024	0.479	-1.222	1.227

N = 7,160. Data is at the CBSA-year level. Each of the 358 CBSAs in the data set have 20 observations, one for each year in 2000-2019. Vehicle shares are derived from unadjusted FARS data. † Health care quality score is a measure of local health care quality given by the United Health Foundation America's Health Rankings system.

the same state, suggesting state level analysis may be masking important variation. Relevant transportation system differences are more likely to be homogenous within metros than states as metro residents share the same transportation infrastructure for commuting and daily travel.

FARS data includes the make, model and model-year of every vehicle involved in an incident. Using this information I merge on vehicle weight data from the EPA fuel economy testing data.[8] I am able to match EPA vehicle weights to 67.1% of FARS vehicle observations. The EPA data includes information for every vehicle that

[8]The EPA testing data is publicly available at www.epa.gov/compliance-and-fuel-economy-data/data-cars-used-testing-fuel-economy.

underwent EPA testing across model years. Data is available from 1984 to present. The FARS observations that cannot be matched to EPA data include vehicles manufactured prior to 1984 (2.3% of FARS observations) as well as cases where the vehicle make and model have no corresponding entry in the EPA data set (32.2% of FARS observations). The latter case is due to the inconsistent classifications of vehicles between the NHTSA and the EPA. For example, different assumptions are made regarding when two slightly different versions of a vehicle should be considered as different models. There are some cases were the EPA data has information on a particular make and model but not for all years. In such cases I assume the vehicle weight is the same as the most recent model year for which data is available. A significant amount of manual coding was required so that these data sets could be merged reliably.

I use demographic variables from the US Census (2000) and from the American Community Survey (2006-2019 1-year estimates). The data is available at the county level which I collapse to the CBSA level. I linearly interpolate the data to impute missing observations from 2001-2005. I also use the Bureau of Economic Analysis (BEA) county level GDP data. The data set includes GDP estimates for the US between 2002 and 2019. I linearly extrapolate the GDP estimates to impute values for 2000 and 2001.

Finally, I use the America's Health Rankings data set from the United Health Foundation. The data provides annual measures of overall health care quality at the state level. The metrics are constructed from measures of different components of health care quality to construct a state level score. I assign each metro an annual health care quality score for the state in which it is located. For metros that span multiple states, I use a population weighted average of the health care scores of the states containing the metro.

3 Methodology

US metros differ in the average characteristics of their vehicle fleets and have experienced heterogeneous adoption of light trucks and large vehicles over the study period. I estimate the impact of vehicle fleet characteristics on road deaths by regressing the metropolitan pedestrian fatality rate against several measures of vehicle fleet characteristics.

Equation 1 captures the regression equation for estimating the effect of average vehicle weight. D_{mt} is the number of deaths per hundred thousand people, where m indexes a particular metro and t indexes a particular year. W_{mt} is the average weight of a vehicle for a particular metro-year observation. Ψ_{mt} is a vector of metro-year control variables. Φ_m is a vector of metro fixed effects and Λ_t is a vector of year fixed effects. I cluster errors at the metro level in all specifications.

$$D_{mt} = \beta_0 + \beta_1 W_{mt} + \Psi_{mt} + \Phi_m + \Lambda_t + \varepsilon_{mt} \tag{1}$$

Time-invariant differences across metros that may affect fatalities are fully controlled for by metro fixed effects. National trends across the study period are fully controlled by time fixed effects. The source of any omitted variable bias must therefore arise from changes in metro characteristics that are occurring differentially across metros and across time. I include an array of control variables in Ψ_{mt} to control for this potential omitted variable bias. I include the average model year of the vehicle fleet to provide a proxy for omitted vehicle characteristics and safety features that change through time. I control for the share of crashes where alcohol was a factor to remove the effect of potentially changing rates of drunk driving across metros. I control for overall population to absorb differential population growth which may affect congestion and also proxy for urban growth. I control for educational attainment using shares of the adult population with high school diplomas and college degrees. I also control for the changing share of the population who are male. Education level and gender have been considered as potentially correlated with driving safety, though empirical evidence of this relationship is sparse (Lourens et al., 1999). I control for median household income and overall GDP of the metro. Controlling for GDP, and income, is potentially important to capture the changing purchasing power of local residents. If residents are becoming richer they will be more likely to purchase a new vehicle, which is likely to be larger. Finally, Ψ_{mt} includes a measure of health care quality for each metro's state, as described in Section 2. The likelihood that a crash involves a fatality is endogenous to the quality of medical care provided to crash victims. The control variable is meant to capture any changes in health care quality that may be occurring differentially across metros during the study period. Transportation policy and infrastructure changes that are made by individual metros during the study period are not easily available in data and are therefore not controlled for. I assume such changes are negligible in terms of their ability to jointly affect vehicle fleets and road safety. Average values for all regression control variables are shown in Table 1.

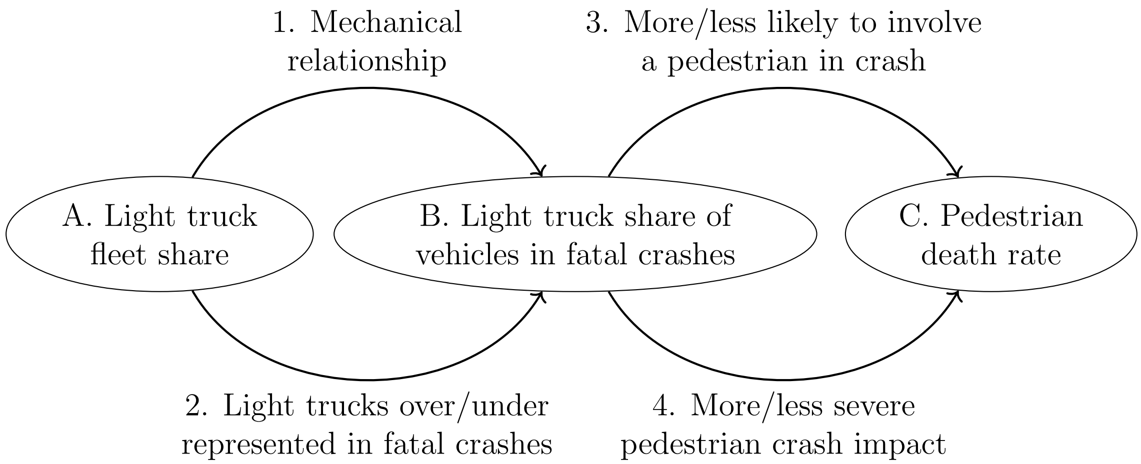
In addition to testing for the effect of average vehicle weight I test for the effect of changing shares of different vehicle body types. One limitation of conducting analysis at the metro level is that vehicle registration data is not available at levels below the state. If estimates of metro vehicle fleets are derived from crash data, the estimated fleet shares may be biased if particular vehicle types are more likely to be involved in fatal vehicle crashes. I summarize the chain of causation in Figure 5, using the light truck category as an example. I have precise information on the share of light trucks that are involved in fatal crashes, and can therefore precisely estimate the impact of this share on the rate of pedestrian deaths. This coresponds to the effect of node B on node C in Figure 5. If light trucks are no more or less likely to be involved in fatal crashes than other vehicle types, their prevalence in crash data should be equal to their actual share of the vehicle fleet. In terms of Figure 5, if edge 2 is assumed to be null, the effect of node A on node C is identical to the effect of node B on node C. If this assumption holds then the estimated causal effect of vehicle shares derived from crash data will be equal to that derived from registered fleet data. However, if light trucks are more or less likely to be involved in a fatal crash, the estimates would differ. I consider this potential bias in a number of ways, which I describe below.

Edges 3 and 4 in Figure 5 represent the mechanisms that link more light trucks being involved in crashes with the potential for more pedestrian deaths. These include more frequent collisions with pedestrians, conditional on overall crash likelihood (edge 3). For example, light trucks may provide worse visibility to the driver, reducing their ability to notice pedestrians. The second mechanism is, conditional on striking a pedestrian, a light truck may cause more harm due to the dimensions of its body type (edge 4). My main specification will capture the combined effect of these two mechanisms (edges 3 and 4), providing the partial effect of higher light truck shares on the pedestrian death rate.

Table 2 compares the vehicle shares from the FARS crash data, with the national FHWA vehicle registration data. The largest discrepancy is the stark overrepresentation of motorcycles in fatal crashes. While only 3.0% of vehicles registered nationally are motorcycles, 10.7% of vehicles involved in fatal crashes are motorcycles. For light trucks, I find that 45% of vehicles registered nationally are light trucks, while 43% of vehicles involved in fatal crashes are light trucks. Table 2 suggests that vehicle shares derived from crash data are relatively representative of the vehicle fleet at the national level.

To correct metro vehicle share estimates I calculate the over and under representation of vehicle shares in crash data relative to registration data. National data suggests that cars are underrepresented in fatal crash data by 10.3%, SUVs are underrepresented

Figure 5: Causal Effect Diagram



The figure diagrams the causal paths linking a metropolitan area's overall light truck fleet share, to light trucks appearing in fatal crash data, to the effect on pedestrian deaths. Given data constraints, I am able to precisely estimate the effect of B on C, but impose additional assumptions in order to estimate the effect of A on C.

Table 2: Vehicle Shares by Data Source

	FARS Metros Only	**FARS National**	**FHWA Registrations National**
Car share	0.478	0.468	0.522
Light truck share	0.413	0.425	0.448
SUV share	0.162	0.161	0.197
Pickup share	0.189	0.201	0.177
Minivan share	0.061	0.061	0.074
Motorcycle share	0.109	0.107	0.030

FARs data is a selective sample of vehicles that have been involved in a crash that resulted in a fatality. FHWA data covers all registered vehicles nationally. The data from both sources span 2000-2019, with 2011 data omitted for FHWA data. Fleet share estimates from the two sources are relatively consistent, with the exception of motorcycles.

in crash data by 18.3%, pickups are overrepresented by 13.6%, minivans are underrepresented by 17.6% and motorcycles are overrepreseted by 256.7%. These gaps could be the result of particular vehicle types being intrinsically less safe, or the gaps could be due to the characteristics of the drivers who choose particular vehicle types.

I propose a correction method to adjust fleet estimates derived from crash data so they more closely reflect overall fleet shares. I derive adjusted estimates of metro vehicle fleets by assuming that the different propensity to be involved in a fatal crash across vehicle types is uniform within a metro's home state.

The adjustment for light truck share is described by equation 2. $L_{mt}^{\text{unadjusted}}$ is light truck share in metro m during year t derived from crash data. $\mathcal{L}_{s(m)}$ is the share of registered vehicles in the metro's home state that are light trucks, pooled across all years of data. $L_{s(m)}$ is the share of vehicles involved in a fatal crash that are light trucks in the metro's home state, pooled across all years in the data. Adjusted shares for the other vehicle types are calculated analogously.

$$L_{mt}^{\text{adjusted}} = L_{mt}^{\text{unadjusted}} \times (\mathcal{L}_{s(m)}/L_{s(m)}) \tag{2}$$

For example, if a metropolitan area reported that 20.0% of vehicles involved in fatal crashes for a particular year were SUVs, and SUVs were found to be underrepresented in crash data by 10% in that state, I assume 22% of the vehicle fleet were SUVs in that year. The methodology allows me to generate metro level vehicle fleet estimates that account for the differing crash propensity across vehicle types. The necessary assumption is that differences in fatal crash likelihood across vehicle types are identical across metros in the same state. For metros that span multiple states I use population weighted data from the relevant states to estimate vehicle registration shares. Because FHWA does not distinguish between Large and Small SUVs I assume that, within the SUV category, Large and Small SUV shares in crash data are representative of fleet shares.

Equation 3 is used to estimate the impact of vehicle fleet shares on the pedestrian fatality rate. When constructing vehicle fleet shares, all vehicles fall into categories of either cars, light trucks or motorcycles. Within the light truck category there are SUVs, pickups and minivans. Within the SUV category I further distinguish between Large and Small SUVs. Equation 3 is composed similarly to Equation 1 but rather than using vehicle weight I use the share of a metro's vehicle fleet in each category for a particular year. In the basic form, I include variables for light truck share (L_{mt}) and motorcycle share (C_{mt}). The omitted category is cars. This model setup allows β_1 to be interpreted as the effect of converting a share of cars to light trucks on pedestrian fatalities. For example $0.1 \times \beta_1$ is the effect of converting 10% of the local vehicle fleet from cars to light trucks on the pedestrian fatality rate. I also perform regressions that are analogous to Equation 3 but where I break out the light truck category into the more disaggregated categories.

$$D_{mt} = \beta_0 + \beta_1 L_{mt} + \beta_2 C_{mt} + \Psi_{mt} + \Phi_m + \Lambda_t + \varepsilon_{mt} \tag{3}$$

A central concern with arriving at causal estimates will be the possible presence of omitted variable bias. There may exist unobserved metropolitan characteristics that are correlated with both road fatalities and vehicle ownership choices. For example, metros that are constructing more highways or wider roads may provide an incentive for owning a larger vehicle, but this type of road infrastructure may directly contribute to road deaths by accommodating higher vehicle speeds (Lewis-Evans and Charlton, 2006; Manuel et al., 2014). Supportive of the identification strategy is the fact that, while metropolitan characteristics such as urban form and road characteristics evolve slowly, the shift towards larger vehicles has happened relatively quickly over the study period. If vehicle fleet characteristics contribute to pedestrian fatalities I expect to find that metros that had different shifts in vehicle fleets experienced different shifts in pedestrian fatalities.

An additional barrier to identification is the possibility of reverse causation. While larger vehicles may contribute to pedestrian fatalities for the reasons given above, rising traffic fatalities may cause local residents to look for ways to improve their road safety, potentially influencing their vehicle purchase decisions. This concern is less relevant to the study of pedestrian fatalities than it is for motorist fatalities. Motorists have a clear incentive to purchase a larger vehicle when confronted with deteriorating road safety among motorists. A changing pedestrian fatality rate is not likely to directly influence the decision of drivers regarding what vehicle to purchase, as the driver does not bear the risks imposed on pedestrians. However, pedestrian fatalities and motorist fatalities may be correlated. In a robustness check I will estimate Equations 1 and 3 while directly controlling for the rate of motorist fatalities, which can proxy for road safety.

4 Results

4.1 Main Results

In this section I provide results from the pedestrian fatality panel regression models as well as results from robustness checks, and alternative specifications. I also provide estimates of the effect of vehicle characteristics on motorist safety. Overall, I find strong evidence of larger vehicles causing a deterioration of pedestrian safety but no evidence of improved motorist safety in aggregate.

Table 3, Columns 1-5 show estimates of the effects of average vehicle weight and vehicle fleet shares on the annual number of pedestrian deaths per 100,000 population. In addition to metro and year fixed effects, regressions include the array of control variables listed in the previous section. I supply coefficient estimates for control variables in Appendix A. I adjust estimates of vehicle fleets using state level registration data, according to the method described in Section 3. Overall, I find that larger vehicle fleets are related to more pedestrian fatalities.

Column 1 regresses the pedestrian fatality rate against the average weight of vehicles involved in fatal crashes, following Equation 1. Every 100 kg increase in average vehicle weight is associated with an additional .03 fatalities per 100,000 residents. The median observation has an annual pedestrian fatality rate of 1.34 fatalities per 100,000 residents, meaning that a 100 kg increase in average vehicle weight is related to a 2.4% increase in pedestrian fatalities for a metro with the median fatality rate.

Table 3, Column 2 estimates the effect of light trucks. In columns 2-5 the omitted category is cars, so that the partial effects on vehicle types can be interpenetrated as the effect of substituting cars with the various vehicle categories. Converting 10% of vehicles from cars to light trucks is associated with an increase in the pedestrian fatality rate of .05, or a 3.6% increase for the median metro. Column 3 breaks out light trucks into the constituent categories. I find pickup trucks, minivans and SUVs all significantly increase pedestrian fatalities relative to cars. Converting 10% of the vehicle fleet from cars to pickups is estimated to increase the pedestrian fatality rate by .04 deaths per 100,000 residents (3.4% in the median metro). I find that converting 10% of cars to minivans would increase pedestrain deaths by .05 deaths per 100,000 residents (3.9% in the median metro). Converting 10% of cars to SUVs would increase pedestrian deaths by .03 deaths per 100,000, or 2.6% in the median metro. In column 4 I further break out SUVs into Large and Small SUVs. I am unable to recover statistically significant effects for SUVs at this level of disaggregation.

Across specifications I find that the share of motorcycles has a highly significant, negative effect on pedestrian deaths. Motorcycles are commonly involved in fatal crashes, but in most cases the fatality is only the driver of the motorcycle and pedestrians are rarely victims of fatal crashes involving motorcycles.

Column 5, includes the four light truck categories and average vehicle weight in a single regression. The multicollinearity between weight and body types cause the

Table 3: Effect of Vehicle Characteristics on Pedestrian Fatality Rate

	(1)	(2)	(3)	(4)	(5)
Vehicle weight (100 kg)	0.032*				0.002
	(0.015)				(0.020)
Light truck share		0.477**			
		(0.127)			
SUV share			0.342*		
			(0.151)		
Small SUV share				0.288	0.283
				(0.176)	(0.175)
Large SUV share				0.521	0.504
				(0.271)	(0.290)
Pickup share			0.448*	0.447*	0.431
			(0.176)	(0.176)	(0.222)
Minivan share			0.521*	0.524*	0.514
			(0.238)	(0.239)	(0.261)
Motorcycle share		-2.327**	-2.423**	-2.417**	-2.434**
		(0.578)	(0.582)	(0.581)	(0.608)
CBSA fixed effects?	Y	Y	Y	Y	Y
Year fixed effects?	Y	Y	Y	Y	Y
Control variables?	Y	Y	Y	Y	Y
R^2	0.043	0.050	0.049	0.050	0.050
N	7160	7160	7160	7160	7160

Significance levels: $*$: 5% $**$: 1%. Standard errors are shown in parenthesis and are clustered at the CBSA level. The dependent variable is the number of pedestrian fatalities per 100,000 population. Control variables include: population, share of population with a high school diploma, share of population with a college degree, median household income, share of population who are male, CBSA GDP, average model year of vehicles involved in fatal crashes, share of fatal crashes that involved alcohol, average age of drivers involved in fatal crashes, and a measure of state health care quality.

resulting estimates to be statistically insignificant.

As discussed in Section 3, metropolitan fleet shares derived from crash data will not accurately reflect the effect of changing overall fleet share if particular vehicle types are involved in fatal crashes at different rates. Table 4 tests the sensitivity of results to alternative vehicle share corrections. In columns 1-3 I estimate the effect of substituting cars for light trucks, repeating the Equation 2 specification. In column 1 I use unadjusted metro vehicle shares taken directly from FARS crash data. In column 2 I adjust vehicle shares according to their relative likelihood of being involved in a fatal crash by using a uniform adjustment based on the difference between national FARS statistics and national FHWA vehicle registration data. In column 3 I use adjustments based on state level differences in vehicle prevalence between the two data sources, matching the main specification. In columns 4-6 I estimate the effect of substituting cars for different light truck types, testing the sensitivity of results to the different vehicle fleet adjustments. I find that the estimates of vehicle shares on pedestrian fatalities are robust to differing approaches to correcting the bias associated with a vehicle type's propensity to be involved in a fatal crash. I do find that the adjustment has a significant effect on my estimate of motorcycles' effect of pedestrian fatalities, consistent with the large overrepresentation of motorcycles in crash data.

The need to adjust vehicle shares with registration data could be avoided if the analysis were conducted at the state level where registration data is available. Due to inconsistency in the collection of registration data across states, the state level analysis may give unreliable results. However, I provide full state level results in Appendix B. State level analysis returns statistically insignificant results for all light truck categories and the measure of vehicle weight.

A concern with identification may be that if there existed time varying omitted variables that affected both road safety and vehicle choice this could lead to biased estimates. Also, issues with reverse causality could arise if drivers choose to purchase larger vehicles at times when road safety is deteriorating. Controlling for the motorist fatality rate should eliminate much of this bias by introducing a strong proxy for road safety conditions. On the other hand, motorist fatalities may be an inappropriate control variable because it is not exogenous to pedestrian fatalities. Table 5 compares regression results to an alternative specification where I add a control for the rate of motorist fatalities. Columns 1, 3 and 5 repeat the main regressions using the different levels of vehicle type aggregations. Columns 2, 4 and 6 add the additional motorist fatality rate control variable. The estimated effects of vehicle characteristics are almost identical regardless of whether motorist fatalities are controlled for. If omitted variable bias or reverse causation issues existed that were related to the general state of road safety I would expect main coefficient estimates to change substantially. This result provides additional evidence that the specification is able to isolate exogenous variation in the vehicle fleet that has a causal effect on the rate of pedestrian fatalities.

I introduce an array of control variables, some of which require interpolation for some years. I test the sensitivity of results to the inclusion of the control variables. Due to the use of metropolitan fixed effects and the relatively gradual change in metropolitan

Table 4: Effect of Fleet Shares on Pedestrian Fatality Rate, Effect of Vehicle Share Adjustments

	(1)	(2)	(3)	(4)	(5)	(6)
Light truck share	0.435**	0.412**	0.477**			
	(0.140)	(0.132)	(0.127)			
SUV share				0.370*	0.302*	0.342*
				(0.186)	(0.152)	(0.151)
Pickup share				0.409*	0.464*	0.448*
				(0.169)	(0.192)	(0.176)
Minivan share				0.607*	0.499*	0.521*
				(0.273)	(0.224)	(0.238)
Motorcycle share	-1.069**	-3.784**	-2.327**	-1.078**	-3.818**	-2.423**
	(0.201)	(0.712)	(0.578)	(0.201)	(0.714)	(0.582)
Vehicle share adjustment:	None	National	State	None	National	State
CBSA fixed effects?	Y	Y	Y	Y	Y	Y
Year fixed effects?	Y	Y	Y	Y	Y	Y
Control variables?	Y	Y	Y	Y	Y	Y
R^2	0.051	0.051	0.049	0.051	0.051	0.049
N	7160	7160	7160	7160	7160	7160

Significance levels: $*$: 5% $**$: 1%. Standard errors are shown in parenthesis and are clustered at the CBSA level. The dependent variable is the number of pedestrian fatalities per 100,000 population. Control variables include: population, share of population with a high school diploma, share of population with a college degree, median household income, share of population who are male, CBSA GDP, average model year of vehicles involved in fatal crashes, share of fatal crashes that involved alcohol, average age of drivers involved in fatal crashes, and a measure of state health care quality.

demographic and economic conditions, results prove to be insensitive regarding the inclusion of control variables, suggesting that omitted variable bias is not a significant concern for the specification (Oster, 2019). I show this empirically in Appendix A.

While all regressions include CBSA fixed effects, this would not control for the possibility that particular CBSAs have long run temporal trends that are correlated with both changing vehicle shares and changing road safety. Such trends would potentially bias estimates if they are not perfectly correlated with the included control variables. In Appendix C I provide an alternative specification where I add CBSA specific linear time trends. I find results are almost identical regardless of whether CBSA time trends are included.

A commonly noted motivation for purchasing a large vehicle is the presumed increase in driver and passenger safety. Potentially, the above estimated increases in

Table 5: Effect of Vehicle Characteristics on Pedestrian Fatality Rate, Controlling for Motorist Fatalities

	(1)	(2)	(3)	(4)	(5)	(6)
Vehicle weight (100 kg)	0.032*	0.033*				
	(0.015)	(0.015)				
Light truck share			0.477**	0.482**		
			(0.127)	(0.126)		
SUV share					0.342*	0.337*
					(0.151)	(0.149)
Pickup share					0.448*	0.459**
					(0.176)	(0.176)
Minivan share					0.521*	0.527*
					(0.238)	(0.237)
Motorcycle share			-2.327**	-2.235**	-2.423**	-2.334**
			(0.578)	(0.573)	(0.582)	(0.577)
Motorist deaths per 100,000		0.018**		0.017**		0.017**
		(0.004)		(0.004)		(0.004)
Vehicle share adjustment:	State	State	State	State	State	State
CBSA fixed effects?	Y	Y	Y	Y	Y	Y
Year fixed effects?	Y	Y	Y	Y	Y	Y
Control variables?	Y	Y	Y	Y	Y	Y
R^2	0.043	0.048	0.050	0.054	0.049	0.053
N	7160	7160	7160	7160	7160	7160

Significance levels: $*$: 5% $**$: 1%. Standard errors are shown in parenthesis and are clustered at the CBSA level. The dependent variable is number of pedestrian fatalities per 100,000 population. Control variables include: population, share of population with a high school diploma, share of population with a college degree, median household income, share of population who are male, CBSA GDP, average model year of vehicles involved in fatal crashes, share of fatal crashes that involved alcohol, average age of drivers involved in fatal crashes, and a measure of state health care quality.

pedestrian fatalities have been offset by an improvement in motorist safety. I repeat the panel regression specifications (Equations 1 and 3), but rather than estimate the effect on pedestrians I estimate the effect of vehicle characteristics on the traffic fatality rate among drivers and their passengers (Table 6).[9] Overall, I find no evidence of a relationship between the change in vehicle characteristics and changes in motorist fatalities across metros. None of the light truck categories appear statistically significant,

[9] I estimated the same regression on the cyclist fatality rate but found no statistically significant coefficients. Cyclist fatalities are rare relative to motorist or pedestrian fatalities, causing imprecision in estimates.

which suggests that shifting fleet shares have not contributed to improved motorist safety. The statistically insignificant effect of large vehicles on motorist fatalities can be attributed to the mechanism proposed in prior literature wherein the safety benefits imparted to the occupants of large vehicles are counteracted by the negative safety impacts on other motorists.

Table 6: Effect of Vehicle Characteristics on Vehicle Occupant Fatality Rate

	(1)	(2)	(3)	(4)	(5)
Vehicle weight (100 kg)	-0.077				-0.102
	(0.063)				(0.094)
Light truck share		-0.275			
		(0.490)			
SUV share			0.319		
			(0.622)		
Small SUV share				0.209	0.483
				(0.672)	(0.743)
Large SUV share				0.686	1.492
				(1.069)	(1.325)
Pickup share			-0.696	-0.698	0.056
			(0.742)	(0.742)	(1.034)
Minivan share			-0.357	-0.352	0.135
			(0.948)	(0.950)	(1.019)
Motorcycle share		-5.391**	-5.299*	-5.287*	-4.478*
		(2.066)	(2.052)	(2.050)	(2.052)
CBSA fixed effects?	Y	Y	Y	Y	Y
Year fixed effects?	Y	Y	Y	Y	Y
Control variables?	Y	Y	Y	Y	Y
R^2	0.325	0.325	0.325	0.325	0.326
N	7160	7160	7160	7160	7160

Significance levels: $*$: 5% $**$: 1%. Standard errors are shown in parenthesis and are clustered at the CBSA level. The dependent variable is the number of driver and passenger fatalities per 100,000 population. Control variables include: population, share of population with a high school diploma, share of population with a college degree, median household income, share of population who are male, CBSA GDP, average model year of vehicles involved in fatal crashes, share of fatal crashes that involved alcohol, average age of drivers involved in fatal crashes, and a measure of state health care quality.

The share of motorcycles appears to have a large negative effect on the rate of motorist fatalities. The result is mainly an artifact of motorcycles rarely having passengers. The average motorcycle involved in a fatal crash carried 1.13 people, while the average car carried 1.60, the average light truck carried 1.69 and the average SUV carried 1.79. In the counterfactual where cars are replaced with motorcycles, the implication is there would be fewer people on the road generally, which would lower fatalities mechanically. Motorcycles also have a limited ability to harm the occupants of other vehicles due to their small size. For interpreting coefficients in the motorist regression results (Table 6), it is possible that larger vehicles may lead to more fatalities simply because they can accommodate more passengers. Therefore, coefficient estimates for larger vehicles may be biased upwards slightly. Notably, this source of bias is not relevant to the pedestrian findings.

The insignificant effect of light trucks on motorist fatalities contrasted with the highly significant results of the pedestrian fatality regressions provide additional support to the validity of the main estimation strategy. If the meaningful variation was related to omitted variables regarding general changes in road safety I would expect the motorist regressions to also indicate significant effects.

While I find null effects of vehicle shares on aggregate motorist safety, drivers may be incentivized to drive larger vehicles to improve their own safety and that of their passengers. I find some evidence that larger vehicle types improve the safety of their occupants. For each metro-year, I calculate the share of deaths among vehicle occupants across each vehicle type. In Table 7 regressions I include the same fixed effects and control variables as in Equations 1 and 3 but use the share of motorist fatalities that occurred in each vehicle category as a dependent variable and regress this against that vehicle type's share of the fleet, using the registration adjusted fleet shares. If all vehicles were equally safe for their occupants, I would expect each coefficent in Table 7 to be equal to one. If, for example, the share of light trucks on the road increased by 10 percentage points and light trucks provided no more or less safety to their occupants, I would expect to see a proportional increase in the share of motorist fatalities that occurred in light trucks. Coefficient estimates below one in Table 7 indicate the vehicle is safer for its occupants than the average vehicle on the road. Results suggest that cars provide average safety to their occupants, while light trucks provide improved occupant safety and motorcycles provide significantly lower occupant safety. Within the light truck category, I estimate that Large SUVs provide the largest improvement in the safety of the vehicle's own occupants. I find that a 10 percentage point increase in a metro vehicle fleet's share of large SUVs relates to only a 7.8 percentage point increase in the share of fatalities among the occupants of Large SUVs. Jointly considering that light trucks do not appear to improve aggregate road safety, but do improve driver and occupant safety suggests that driving a larger vehicle offloads fatality risk from the occupants to other road users.

Table 7: Effect of Increase in Vehicle Fleet Share on Vehicle Specific Motorist Fatality Share

	Coefficient Estimate	$\mathbf{R^2}$
Car share	0.983**	0.710
	(0.012)	
Light truck share	0.931**	0.666
	(0.013)	
SUV share	0.875**	0.577
	(0.019)	
Small SUV share	0.894**	0.576
	(0.020)	
Large SUV share	0.781**	0.522
	(0.019)	
Pickup share	0.795**	0.659
	(0.013)	
Minivan share	0.822**	0.511
	(0.019)	
Motorcycle share	1.284**	0.900
	(0.017)	

Each estimate corresponds to a separate regression. Significance levels: $*$: 5% $**$: 1%. Standard errors are shown in parenthesis and are clustered at the CBSA level. The dependent variable is the share of annual motorist deaths that occurred among occupants of the particular vehicle type. Control variables include: population, share of population with a high school diploma, share of population with a college degree, median household income, share of population who are male, CBSA GDP, average model year of vehicles involved in fatal crashes, share of fatal crashes that involved alcohol, average age of drivers involved in fatal crashes, and a measure of state health care quality.

4.2 Counterfactual Analysis

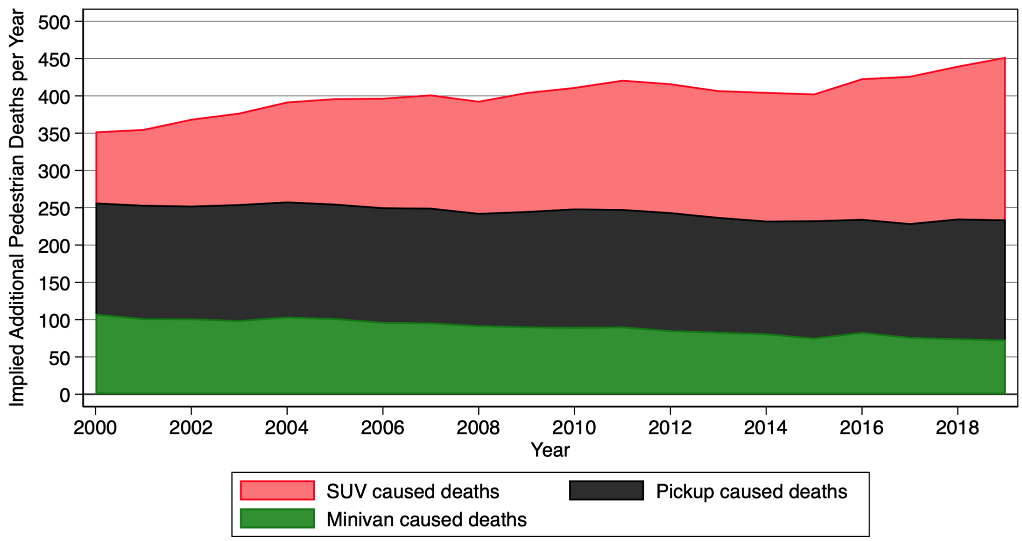
I use coefficient estimates to estimate counterfactual scenarios of alternative vehicle adoption. I use the main vehicle share regression coefficients (Table 3, column 3) to estimate the number of pedestrian fatalities that were caused by the presence of particular vehicle types, compared to the counterfactual scenario where those vehicles were replaced with cars. I multiply the estimated partial effects by the fleet share held by that vehicle type in each year and then scale the figure up by the overall population across all metros in the data set. This empirical exercise can be characterised as a "back-of-the-envelope" approach to roughly quantifying the magnitude of effects, given that the approach ignores the role of standard errors in estimates and given the low explanatory power of the overall model.

Figure 6 graphs the implied number of pedestrian fatalities caused by the light truck categories, compared to the counterfactual where all of these vehicles were substituted with cars. Across 2000-2019 I estimate that 8,131 pedestrian lives would have been saved if all light trucks had been cars. The reduction would be equal to avoiding 9.5% of all pedestrian deaths. In 2000, converting all light trucks to cars would have spared 353 pedestrians, while by 2019 the figure had grown by 30% to 459 pedestrians. Accounting for the overall population increase of the metros, the number of pedestrian deaths attributable to light trucks increased by 7.6% on a per capita basis.

I find an increasing impact of SUVs on pedestrian fatalities. In 2000, if all SUVs were substituted with cars, there would have been 99 fewer pedestrian fatalities across all metros. By 2019, the substitution of SUVs for cars would have averted 227 pedestrian fatalities. The change represents a 129% increase or 90% on a per capita basis. Across all years in the sample period I estimate that replacing all SUVs with cars would have averted 3,283 pedestrian deaths. Maintaining the share of SUVs across the study period at 2000 levels and replacing that growth with cars would have averted 1,084 pedestrian fatalities.

Figure 6 shows how the categories of light trucks changed in their contribution to pedestrian deaths over the study period. The sharp increase in SUVs as a share of metropolitan consumer vehicles (13.6% to 25.7%) caused a significant increase in pedestrian deaths. However, over this period the share of pickups and minivans both fell. Pickups as a share of vehicles fell from 15.5% to 13.8% and minivans fell from 9.5% to 5.3%. I estimated in Table 3 that both pickups and minivans have a significantly harmful effect on pedestrian safety relative to cars. The decline in pickups and minivans

Figure 6: Annual Pedestrian Deaths Averted if all Light Trucks had been Cars



Relying on estimated partial effects, the figure plots the number of pedestrian fatalities that would have been averted if all light trucks were replaced by cars. Over the entire study period, converting all light trucks to cars would have prevented 8,131 pedestrian deaths.

worked to counteract the negative pedestrian safety effects of increased SUVs.

As noted in the introduction, there was a substantial increase in pedestrian fatalities during the 2000-2019 period. Between 2000 and 2019 the pedestrian fatality rate across all metro areas increased by 11%, from 1.72 to 1.91 deaths per 100,000 residents. The overall rise in light trucks over this period was modest, as the rise in SUVs was buffered by the decline in pickups and minivans. Table 3, column 3 estimates imply that if the prevalence of all light truck categories had remained at the 2000 level across the study period there would have been 314 fewer pedestrian deaths between 2000 and 2019, including 31 fewer in 2019. Rather than being 1.91 deaths per 100,000 in 2019, the pedestrian fatality rate would have been 1.90 if light truck shares had remained at 2000 levels. The result suggests that a shift in the vehicle fleet is not responsible for the overall increase in pedestrian deaths. However, converting all light trucks to cars would have reduced the rate to 1.75.

The above estimates ignore incidents occurring outside of metropolitan areas. Metropolitan areas contained 77% of the US population across the study period. To the extent that non metropolitan areas are experiencing negative pedestrian safety effects from larger vehicles, the estimates understate the national effect. However, the effects in less urbanized areas may be markedly different.

A single consumer's decision to substitute a car for a light truck raises the predicted number of pedestrian deaths marginally. I calculate the marginal external cost of a consumer switching from a car to a light truck. To calculate marginal external costs I use the US Department of Transportation's value of a statistical life ($10.45 million in 2020 USD), main regression estimates (Table 3, column 3) and the fact that there were 47.6 registered vehicles in the US for every 100 residents according to the FHWA 2019 data. The marginal external cost of switching from a car to a SUV, pickup truck, or minivan in terms of increased pedestrian fatalities is $75 , $98 , and $114 respectively. Optimal Pigouvian taxes that internalize the external costs of pedestrian fatalities attributable to driving a light truck over a car could be implemented with annual taxes by vehicle types that are equal to these marginal external costs. These taxes would be in addition to other taxes that may address other externalizes of vehicles. Assuming a 10 year vehicle lifespan suggests that if the tax were applied at the time of sale, the one time tax would need to be roughly 10 times the rates calculated above. For example, SUVs would be assessed a point of sale tax equal to $750 in order to internalize the pedestrian fatality risk.

Using the value of a statistical life, the implied economic cost of the 8,131 pedestrian deaths attributable to the presence of light trucks between 2000 and 2019 is $85 billion. The possibility of reducing the pedestrian safety externalities imposed by large vehicles through regulation could provide significant societal welfare improvements.

Reducing the prevalence of large vehicles could also affect the safety of motorists. However, I find no evidence of a relationship between larger vehicles and improved motorist safety in the aggregate. The finding is supported by prior literature (Anderson, 2008; White, 2004).

Switching from a car to a light truck also carries additional externalizes. Excess air pollution is a particularly large externality that arises from the comparatively low fuel economy of larger vehicles. In EPA data on new, 2019 vehicles, the average reported fuel economy is 29.9 miles per gallon (mpg) for a car, 27.5 for a SUV, 19.0 for a pickup, and 22.4 for a van (EPA 2020 Automotive Trends Report). I assume vehicle miles traveled does not depend on vehicle choice, and average annual vehicle miles traveled for a private vehicle in the US is 11,500.[10] The numbers imply that switching from

[10]Federal Highway Administration. Highway Statistics 2018, Table VM-1. https://www.fhwa.dot.gov/policyinformation/statistics/2018/pdf/vm1.pdf

a car to a SUV, pickup, or van results in an additional annual consumption of 34, 221 or 129 gallons of fuel, respectively. A recent National Academy of Science report found that the marginal external cost in terms of air pollution of a gallon of burned vehicle fuel is \$0.81 (National Academies of Sciences, Engineering, and Medicine, 2020). Switching from a car to a SUV, pickup or van therefore implies annual excess external air quality costs of \$28, \$179, and \$104 respectively. The external costs of light trucks on pedestrian fatalities appear roughly the same magnitude as the external cost of excess air pollution.

5 Conclusion

I estimate that the popularity of light trucks on US roads is responsible for a large number of pedestrian deaths. If all light trucks were replaced with cars, over 8,000 pedestrian deaths would have been averted between 2000 and 2019. Vehicle body types appear to be an important determinant of pedestrian deaths in the aggregate, strengthening arguments made in the transportation safety literature regarding the link between larger light trucks and more severe pedestrian injuries.

Average vehicle size has undergone a sustained increase over the past 20 years, with no signs of abating. If the popularity of large vehicles continues to rise there is likely to be a corresponding increase in pedestrian fatalities. Given strict federal regulation of vehicle safety standards, it is perhaps surprising that there is limited legislation that restricts the overall size and body type of vehicles with the intent of improving pedestrian safety. It is unlikely that the purchase decision of vehicle owners will take account of the safety externalities that large vehicle body types impose on pedestrians (Lindberg, 2005). These facts suggest there could be societal benefits from restricting sales of large vehicles, or implementing a Pigouvian tax on particular vehicles, as was suggested in Anderson (2008), Anderson and Auffhammer (2014) and Li (2012).

The shift in vehicle types over the study period is unable to account for the dramatic rise in overall pedestrian deaths. While the increased popularity of SUVs caused a significant number of deaths, the declining popularity of pickup trucks and minivans offset the majority of this trend. Other changes to vehicles and road conditions over this period are deserving of future study and may be able to account for the rise in aggregate pedestrian deaths. In particular, the rapid shift in personal consumer technologies may have impacted road safety during the same period. The proliferation of smartphones among both drivers and pedestrians presented a new distraction for road users (Lin and Huang, 2017; Ortiz et al., 2018; Vollrath et al., 2016). Additionally, the decision of automakers to include complex navigation and entertainment consoles in vehicles may have served to reduce drivers' ability to monitor for pedestrians.

References

Ahmad, S. and Greene, D. L. (2005). Effect of fuel economy on automobile safety: A reexamination. *Transportation Research Record*, 1941(1):1–7.

Anderson, M. (2008). Safety for whom? The effects of light trucks on traffic fatalities. *Journal of Health Economics*, 27(4):973–989.

Anderson, M. L. and Auffhammer, M. (2014). Pounds that kill: The external costs of vehicle weight. *Review of Economic Studies*, 81(2):535–571.

Ang, A., Christensen, P., and Vieira, R. (2020). Should congested cities reduce their speed limits? evidence from são paulo, brazil. *Journal of Public Economics*, 184:104155.

Basili, M. and Belloc, F. (2020). The deterrent effect of "Vehicular Homicide Laws": Microeconometric evidence from Italy. *Research in Transportation Economics*, Forthcoming.

Baughman, R., Conlin, M., Dickert-Conlin, S., and Pepper, J. (2001). Slippery when wet: The effects of local alcohol access laws on highway safety. *Journal of Health Economics*, 20(6):1089–1096.

Blattenberger, G., Fowles, R., and Loeb, P. D. (2013). Determinants of motor vehicle crash fatalities using bayesian model selection methods. *Research in Transportation Economics*, 43(1):112–122.

Borsati, M., Cascarano, M., and Bazzana, F. (2019). On the impact of average speed enforcement systems in reducing highway accidents: Evidence from the italian safety tutor. *Economics of Transportation*, 20:100123.

Bourgeon, J.-M. and Picard, P. (2007). Point-record driving licence and road safety: an economic approach. *Journal of Public Economics*, 91(1-2):235–258.

Carpenter, C. S. and Stehr, M. (2008). The effects of mandatory seatbelt laws on seatbelt use, motor vehicle fatalities, and crash-related injuries among youths. *Journal of Health Economics*, 27(3):642–662.

Crandall, R. W. and Graham, J. D. (1989). The effect of fuel economy standards on automobile safety. *The Journal of Law and Economics*, 32(1):97–118.

Deery, H. A. (1999). Hazard and risk perception among young novice drivers. *Journal of Safety Research*, 30(4):225–236.

Desapriya, E., Subzwari, S., Sasges, D., Basic, A., Alidina, A., Turcotte, K., and Pike, I. (2010). Do light truck vehicles (LTV) impose greater risk of pedestrian injury than passenger cars? A meta-analysis and systematic review. *Traffic Injury Prevention*, 11(1):48–56.

Gayer, T. (2004). The fatality risks of sport-utility vehicles, vans, and pickups relative to cars. *Journal of Risk and Uncertainty*, 28(2):103–133.

Giulietti, C., Tonin, M., and Vlassopoulos, M. (2020). When the market drives you crazy: Stock market returns and fatal car accidents. *Journal of Health Economics*, 70:102245.

Green, C. P., Heywood, J. S., and Navarro, M. (2014). Did liberalising bar hours decrease traffic accidents? *Journal of Health Economics*, 35:189–198.

Green, C. P., Heywood, J. S., and Navarro, M. (2016). Traffic accidents and the London congestion charge. *Journal of Public Economics*, 133:11–22.

Hansen, B. (2015). Punishment and deterrence: Evidence from drunk driving. *American Economic Review*, 105(4):1581–1617.

Jackson, C. K. and Owens, E. G. (2011). One for the road: public transportation, alcohol consumption, and intoxicated driving. *Journal of Public Economics*, 95(1-2):106–121.

Karaca-Mandic, P. and Ridgeway, G. (2010). Behavioral impact of graduated driver licensing on teenage driving risk and exposure. *Journal of Health Economics*, 29(1):48–61.

Lefler, D. E. and Gabler, H. C. (2004). The fatality and injury risk of light truck impacts with pedestrians in the United States. *Accident Analysis and Prevention*, 36(2):295–304.

Levitt, S. D. and Porter, J. (2001). How dangerous are drinking drivers? *Journal of Political Economy*, 109(6):1198–1237.

Lewis-Evans, B. and Charlton, S. G. (2006). Explicit and implicit processes in behavioural adaptation to road width. *Accident Analysis and Prevention*, 38(3):610–617.

Li, S. (2012). Traffic safety and vehicle choice: Quantifying the effects of the "arms race"on American roads. *Journal of Applied Econometrics*, 27(1):34–62.

Lin, M.-I. B. and Huang, Y.-P. (2017). The impact of walking while using a smartphone on pedestrians' awareness of roadside events. *Accident Analysis & Prevention*, 101:87–96.

Lindberg, G. (2005). Accidents. *Research in Transportation Economics*, 14:155–183.

Lourens, P. F., Vissers, J. A., and Jessurun, M. (1999). Annual mileage, driving violations, and accident involvement in relation to drivers' sex, age, and level of education. *Accident Analysis & Prevention*, 31(5):593–597.

Manuel, A., El-Basyouny, K., and Islam, M. T. (2014). Investigating the safety effects of road width on urban collector roadways. *Safety Science*, 62:305–311.

National Academies of Sciences, Engineering, and Medicine (2020). Reducing the fuel consumption and greenhouse gas emissions of medium- and heavy-duty vehicles, phase two.

Ortiz, C., Ortiz-Peregrina, S., Castro, J., Casares-López, M., and Salas, C. (2018). Driver distraction by smartphone use (whatsapp) in different age groups. *Accident Analysis & Prevention*, 117:239–249.

Oster, E. (2019). Unobservable selection and coefficient stability: Theory and evidence. *Journal of Business & Economic Statistics*, 37(2):187–204.

Oviedo-Trespalacios, O. (2018). Getting away with texting: Behavioural adaptation of drivers engaging in visual-manual tasks while driving. *Transportation Research Part A: Policy and Practice*, 116:112–121.

Peltzman, S. (1975). The effects of automobile safety regulation. *Journal of Political Economy*, 83(4):677–725.

Roudsari, B. S., Mock, C. N., Kaufman, R., Grossman, D., Henary, B. Y., and Crandall, J. (2004). Pedestrian crashes: Higher injury severity and mortality rate for light truck vehicles compared with passenger vehicles. *Injury Prevention*, 10(3):154–158.

Ruhm, C. J. (1996). Alcohol policies and highway vehicle fatalities. *Journal of Health Economics*, 15(4):435–454.

Simms, C. K. and Wood, D. P. (2006). Pedestrian risk from cars and sport utility vehicles-a comparative analytical study. *Proceedings of the Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering*, 220(8):1085–1100.

Tamura, A., Nakahira, Y., Iwamoto, M., Watanabe, I., Miki, K., Hayashi, S., Kitagawa, Y., and Yasuki, T. (2008). Analysis of traumatic brain injury due to primary head contact during vehicle-to-pedestrian impact. *International Journal of Crashworthiness*, 13(4):375–385.

Toy, E. L. and Hammitt, J. K. (2003). Safety impacts of SUVs, vans, and pickup trucks in two-vehicle crashes. *Risk Analysis: An International Journal*, 23(4):641–650.

Van Benthem, A. (2015). What is the optimal speed limit on freeways? *Journal of Public Economics*, 124:44–62.

Van Ommeren, J., Rietveld, P., Hop, J. Z., and Sabir, M. (2013). Killing kilos in car accidents: Are external costs of car weight internalised? *Economics of Transportation*, 2(2-3):86–93.

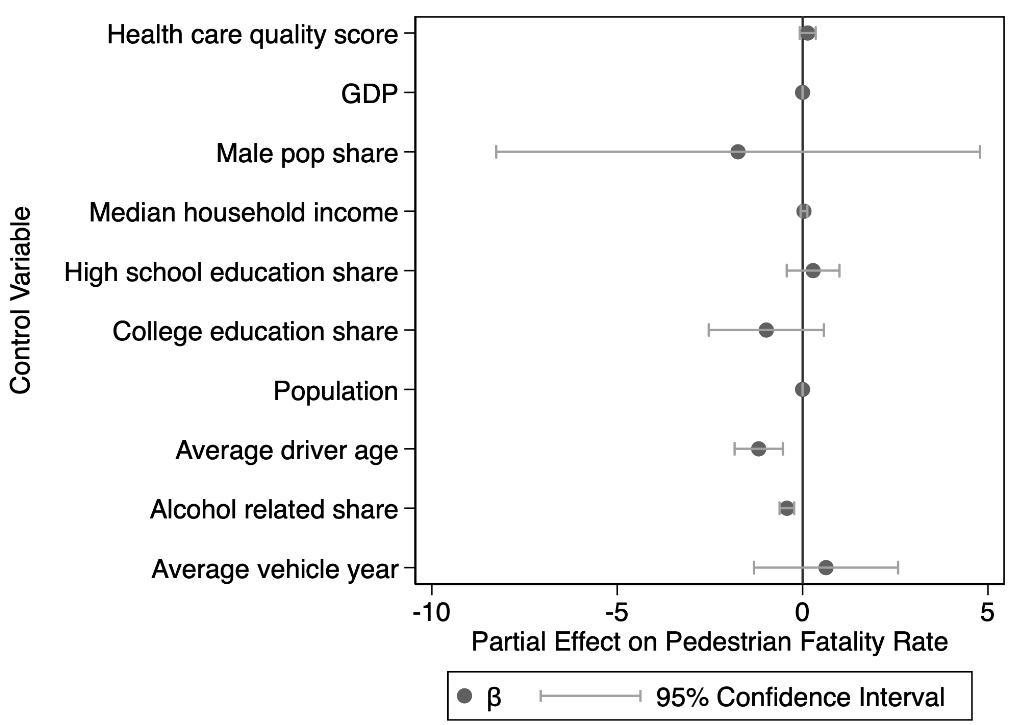
Vollrath, M., Huemer, A. K., Teller, C., Likhacheva, A., and Fricke, J. (2016). Do German drivers use their smartphones safely?-not really! *Accident Analysis & Prevention*, 96:29–38.

White, M. J. (2004). The “Arms Race” on American roads: The effect of sport utility vehicles and pickup trucks on traffic safety. *The Journal of Law and Economics*, 47(2):333–355.

Appendix A

In the main tables I omit the partial effects of control variables to focus on the effects of vehicle fleet characteristics. Figure A1 provides information on the estimated coefficients of control variables. I show results that correspond to the estimates from Table 3, column 3. The estimated partial effects of control variables are very similar across all specifications in Table 3.

A1. Control Variable Coefficient Estimates



Coefficient estimates for control variables from the main model specification are shown. I convert the units of some variables to improve the readability of the chart. All "shares" range from 0-1. Median household income is in $10,000s, population is in millions, GDP is in billions and average vehicle year and age are in 100s of years.

I find two control variables have a statistically significant effect on the pedestrian fatality rate; average driver age and the rate of alcohol related incidents. An increase in the average age of drivers involved in fatal crashes within a metro is correlated with fewer pedestrian fatalities. The negative correlation potentially suggests that driver experience improves the safety of pedestrians (Deery, 1999). The share of fatal crashes that involved alcohol is correlated with *fewer* pedestrian fatalities. This at first seems counterintuitive; however, a rise in the share of incidents involving alcohol could be caused by increased drunk driving, or equally by a decline in incidents among sober drivers. Among motorist fatalities, 35% were the result of incidents involving alcohol, while among pedestrian fatalities, only 27% of incidents involved alcohol. Therefore, a rise in the share of incidents involving alcohol is correlated with fewer pedestrian deaths but more motorist deaths. I find the drunk driving control variable is significant and positively related to motorist deaths in the Table 6 regressions.

Despite some control variables being statistically significant, their inclusion does not significantly affect the main results of the paper. Table A1 repeats the main regressions of the paper but omits all metro-year control variables. Comparing the results to those of Table 3 demonstrates that estimates are insensitive to the inclusion of metro-year level control variables.

Table A1. Effect of Vehicle Characteristics on Pedestrian Fatality Rate, No Control Variables

	(1)	(2)	(3)	(4)	(5)
Vehicle weight (100 kg)	0.032*				0.005
	(0.015)				(0.019)
Light truck share		0.461**			
		(0.127)			
SUV share			0.342*		
			(0.152)		
Small SUV share				0.276	0.264
				(0.176)	(0.176)
Large SUV share				0.564*	0.528
				(0.274)	(0.295)
Pickup share			0.410*	0.410*	0.376
			(0.177)	(0.177)	(0.218)
Minivan share			0.502*	0.505*	0.484
			(0.237)	(0.237)	(0.260)
Motorcycle share		-2.424**	-2.524**	-2.514**	-2.551**
		(0.579)	(0.583)	(0.583)	(0.606)
CBSA fixed effects?	Y	Y	Y	Y	Y
Year fixed effects?	Y	Y	Y	Y	Y
Control variables?	N	N	N	N	N
R^2	0.037	0.044	0.044	0.044	0.044
N	7160	7160	7160	7160	7160

Significance levels: $*$: 5% $**$: 1%. Standard errors are shown in parenthesis and are clustered at the CBSA level. The dependent variable is the number of pedestrian fatalities per 100,000 population.

Appendix B

Table B1 provides estimates of the partial effect of vehicle fleet characteristics on the pedestrian fatality rate with analysis completed at the state level. Table B1 follows the same model specification as the metropolitan level analysis but uses states as the unit of observation. Column 1 estimates the effect of vehicle weight derived from EPA data, columns 2-4 use fleet estimates from FHWA vehicle registration data and columns 5-7 use FARS data. In estimating state fleet shares from FARS data I perform the vehicle share adjustment according to Equation 2.

Table B1. Effect of Vehicle Characteristics on Pedestrian Fatality Rate, State Level Analysis

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Vehicle weight (100 kg)	-0.067			-0.068			-0.086
	(0.070)			(0.071)			(0.073)
Light truck share		-0.782			-0.324		
		(0.564)			(0.422)		
SUV share			-0.949	-0.974		-0.790	-0.483
			(1.033)	(1.007)		(0.501)	(0.438)
Pickup share			-1.184	-1.162		-0.175	0.381
			(1.006)	(1.022)		(0.522)	(0.460)
Minivan share			0.250	0.116		0.453	0.850
			(1.328)	(1.280)		(0.512)	(0.560)
Motorcycle share		-2.675**	-2.768**	-2.797**	-2.030	-1.964	-1.461
		(0.998)	(0.999)	(1.024)	(1.424)	(1.461)	(1.474)
State fixed effects?	Y	Y	Y	Y	Y	Y	Y
Year fixed effects?	Y	Y	Y	Y	Y	Y	Y
Control variables?	Y	Y	Y	Y	Y	Y	Y
Fleet data source	.	FHWA	FHWA	FHWA	FARS	FARS	FARS
R^2	0.366	0.368	0.368	0.372	0.367	0.371	0.374
N	1000	1000	1000	1000	1000	1000	1000

Significance levels: $*$: 5% $**$: 1%. Standard errors are shown in parenthesis and are clustered at the state level. The dependent variable is the number of pedestrian fatalities per 100,000 population. Control variables include: population, share of population with a high school diploma, share of population with a college degree, median household income, share of population who are male, CBSA GDP, average model year of vehicles involved in fatal crashes, share of fatal crashes that involved alcohol, average age of drivers involved in fatal crashes, and a measure of state health care quality.

Unlike the metro level analysis, I find no statistically significant effects of light trucks or vehicle weight on pedestrian fatalities. A number of reasons could contribute to this. First, for columns 2-4, FHFA data is reported by individual states and different states changed reporting practices at different times throughout the study period, which undermines the identification strategy by introducing time varying changes in the independent variable of interest across the units of analysis. Second, using metropolitan areas may be more sensible because they capture units that share common transportation systems whereas aggregating to states may erase important statistical variation. Finally, the sample size is reduced by 86% when switching from metropolitan to state analysis, drastically reducing the statistical power of the estimates.

Appendix C

I apply a conventional difference in difference regression design as the main model in the above analysis. In Table C1 I repeat the main analysis but add controls for metro specific linear time trends. The models estimated are identical to Equations 1 and 3, with the addition of a vector of metro level linear time trends.

Table C1. Effect of Vehicle Characteristics on Pedestrian Fatality Rate, Linear Metro Time Trends

	(1)	(2)	(3)	(4)	(5)
Vehicle weight (100 kg)	0.023				-0.014
	(0.015)				(0.021)
Light truck share		0.460**			
		(0.129)			
SUV share			0.301		
			(0.165)		
Small SUV share				0.266	0.306
				(0.192)	(0.192)
Large SUV share				0.416	0.530
				(0.278)	(0.300)
Pickup share			0.468*	0.467*	0.574*
			(0.182)	(0.182)	(0.237)
Minivan share			0.606*	0.607*	0.676*
			(0.244)	(0.244)	(0.270)
Motorcycle share		-1.905**	-1.976**	-1.973**	-1.855**
		(0.611)	(0.617)	(0.616)	(0.649)
CBSA fixed effects?	Y	Y	Y	Y	Y
Year fixed effects?	Y	Y	Y	Y	Y
CBSA time trends?	Y	Y	Y	Y	Y
Control variables?	Y	Y	Y	Y	Y
R^2	0.121	0.126	0.126	0.126	0.126
N	7160	7160	7160	7160	7160

Significance levels: $*$: 5% $**$: 1%. Standard errors are shown in parenthesis and are clustered at the CBSA level. The dependent variable is the number of pedestrian fatalities per 100,000 population. Control variables include: population, share of population with a high school diploma, share of population with a college degree, median household income, share of population who are male, CBSA GDP, average model year of vehicles involved in fatal crashes, share of fatal crashes that involved alcohol, average age of drivers involved in fatal crashes, and a measure of state health care quality.

I find results are robust to the inclusion of metro time trends, with point estimates and standard errors changing very little between Tables 3 and C1. The estimate of vehicle weight's impact on pedestrian fatalities falls short of statistical significance when time trends are added (column 1), providing evidence that weight is relatively less important than vehicle body type.