# VEHICLE WEIGHT, HIGHWAY SAFETY, AND ENERGY POLICY[1]

Michael Anderson
*University of California, Berkeley*

Maximilian Auffhammer
*University of California, Berkeley & NBER*

**Preliminary Draft**

March 23, 2011

**Abstract**

Heavier vehicles are safer for their own occupants but more hazardous for the occupants of other vehicles. In this paper we estimate the increased probability of fatalities from being hit by a heavier vehicle in a collision. We show that, controlling for own-vehicle weight, being hit by a vehicle that is 1,000 pounds heavier results in a 49 percent increase in the baseline fatality probability. Estimation results further suggest that this risk is even higher if the striking vehicle is a light truck (SUV, pickup truck, or minivan). We calculate that a second-best gasoline tax, which accounts for the external risk generated by the gain in fleet weight since 1989, is approximately 28 cents per gallon. We further calculate that the total fatality externality is roughly equivalent to a gas tax of $1.04 per gallon. We find that the difference between a gas tax and an optimal weight varying mileage tax is modest for most vehicles.

JEL Codes: H23, I18, Q48, Q58, R41

Keywords: Traffic accidents, gasoline tax, arms race

---

[1]Anderson (corresponding author): mlanderson@berkeley.edu; Auffhammer: auffhammer@berkeley.edu. We gratefully acknowledge support from the Robert Wood Johnson Foundation and the University of California Energy Institute. We thank Larry Goulder, Ryan Kellogg, Ian Parry and Ken Small for valuable feedback. Seminar participants at CESIfo München, Duke University, UC Berkeley, the UC Energy Institute, UC Irvine, University of Michigan, and the Occcasional Workshop on Environmental and Resource Economics have provided helpful comments. All errors in the paper are the authors'.

## 1. INTRODUCTION

The average weight of light vehicles sold in the United States has fluctuated substantially over the past 35 years. From 1975 to 1980, average weight dropped almost 1,000 pounds (from 4,060 pounds to 3,228 pounds), likely in response to rising gasoline prices and the passage of the Corporate Average Fuel Efficiency (CAFE) standard. As gasoline prices fell in the late-1980s, however, average vehicle weight began to rise, and by 2005 it had attained 1975 levels (US EPA 2009). A rich body of research examines the effects of CAFE and gasoline prices on consumers' vehicle choices (Goldberg 1998; Portney et al. 2003; Kleit 2004; Austin and Dinan 2005; Klier and Linn 2008; Bento, Goulder, Jacobsen, and von Haefen, 2009; Busse, Knittel, and Zettelmeyer 2009; Li, Timmins, and Von Haefen 2009).

One question that remains unresolved is how the choices consumers make in response to gasoline prices and fuel economy standards affect traffic fatalities. Traffic accidents are the leading cause of death for persons under the age of 40, and they are a major source of life-years lost.[2] Intuitively, heavier cars are safer than lighter cars, and previous research has argued that a heavier vehicle fleet is a safer vehicle fleet (Crandall and Graham 1989). Much of the subsequent transportation safety literature has focused on the effects of average vehicle weight on safety, reaching varying conclusions.[3] Recent work by Jacobsen (2010) explores the traffic safety implications of different fuel economy regulatory schemes across ten vehicle classes. The paper uses micro data on fatal accidents and concludes that tightening fuel economy standards will not increase fatalities as long as the standards are "footprint based" or unified across cars and trucks.[4]

---

[2] Lung cancer, a disease that is generally the result of smoking, kills approximately four times as many Americans each year as traffic accidents. However, the average lung cancer decedent is 71 years old while the average traffic accident decedent is only 39 years old. The number of life-years lost to traffic accidents is thus similar in magnitude to the number of life-years lost to lung cancer.

[3] Much of the transportation safety literature is based on time series correlations between average vehicle weight and aggregate fatality rates (Robertson 1991; Khazzoom 1994; Noland 2004, 2005; Ahmad and Greene 2005). Two exceptions are Kahane (2003) and Van Auken and Zellner (2005), which use micro data containing fatal accidents only. They supplement the fatal accident data with data on police-reported accidents from several states to estimate the rate at which different types of vehicles enter into collisions. These studies come to varying conclusions regarding the sign of the relationship between average vehicle weight and overall fatality rates, but all conclude that the magnitude of the relationship is relatively modest.

[4] If the current separation between cars and trucks is maintained and standards are not footprint based, Jacobsen estimates that raising CAFE standards by one mile per gallon could increase traffic fatalities by 149

From an economic standpoint an unregulated vehicle fleet must be inefficiently heavy. A heavier vehicle is safer for its own occupants but more hazardous for the occupants of other vehicles. The safety benefits of vehicle weight are therefore internal, while the safety costs of vehicle weight are external. Consumers' vehicle choices thus have the important features of an "arms race." To date, however, no detailed attempt has been made to quantify the *external* costs of vehicle weight. This figure is essential for determining the socially optimal weight of the vehicle fleet, and it cannot be inferred from the effects of average vehicle weight or fuel economy regulations on traffic safety.

We quantify the external costs of vehicle weight using a large micro data set on police-reported crashes for a set of 8 heterogeneous states. Unlike the data sets employed in the previous transportation literature or Jacobsen (2010), our data set includes both fatal and nonfatal accidents. Using unique vehicle identifiers (VINs), we determine the curb weight of each vehicle involved in an accident, thereby minimizing concerns about measurement error induced attenuation bias. The rich set of vehicle, person, and accident observables in the data set allow us to minimize concerns about omitted variables bias in our coefficients on weight. Using these data, we estimate the external effects of vehicle weight on fatalities and serious injuries conditional on a collision occurring.

Two key results emerge from our estimates. First, we show that vehicle weight is an important determinant of fatalities in other vehicles in the event of a multivehicle collision; our preferred estimate implies that a 1,000 pound increase in striking vehicle weight raises the probability of a fatality in the struck vehicle by 49%. When we translate this higher probability of a fatality into external costs relative to a small baseline vehicle, the total external costs of vehicle weight from fatalities alone are estimated to be $97 billion per year. Second, by separately controlling for vehicle weight and whether the striking vehicle is a light truck (i.e., a pickup truck or sport utility vehicle), we show that light trucks significantly raise the probability of a fatality in the struck car – in addition to the effect of their already higher vehicle weight.

Our unique data set allows us to condition on a collision occurring and ensures that our results cannot be generated by differences in collision rates between drivers of lighter and heavier vehicles. Nevertheless, driver selection could bias our results if drivers of heavy

---

deaths per year. Jacobsen does not attempt to estimate the causal effect of vehicle weight on fatalities in other vehicles, which is the focus of this paper.

vehicles have a tendency towards severe accidents. We rule out this possibility through three tests. First, we show that vehicle weight does not predict fatalities when two vehicles of equal weight collide. This suggests that drivers of heavy vehicles are not predisposed towards severe accidents. Second, we show that our estimates persist even when controlling for specific vehicle type via make and model fixed effects. Finally, we instrument for striking vehicle weight using the number of occupants in the striking vehicle and find estimates that are close to our least squares estimates. All three tests suggest that we successfully identify the causal effect of vehicle weight on the probability of fatalities in two-car collisions.

We apply our estimates to consider whether a gasoline tax could internalize most of the external costs and conclude that it could. Our calculations suggest that the external costs of vehicle weight eclipse any other component in existing estimates of vehicle externalities per mile driven (Portney, Parry, Gruenspecht and Harrington, 2003). Furthermore, our calculations imply that the level of the optimal gasoline tax is substantially higher than previously estimated (e.g. Parry and Small 2005).

The paper is organized as follows. Section 2 presents the analytic and empirical framework and discusses the previous literature. Section 3 details the data. Section 4 presents the main results, and Section 5 presents falsification tests and alternative sources of identification to check whether selection bias contaminates our results. Section 6 links the results to energy policy implications, focusing in particular on the gasoline tax. Section 7 concludes.

## 2. ANALYTIC AND EMPIRICAL FRAMEWORK

Consumers' vehicle choices represent a classic example of an externality driven "arms race." Purchasing a heavier vehicle enhances safety for each individual, but also makes other roadway users less safe. The net benefit of vehicle weight on traffic fatalities is thus smaller than the private benefit of vehicle weight on traffic fatalities; consumers are incentivized to purchase heavier vehicles than is socially optimal.

Figure 1 presents a stylized plot of the marginal private and social costs per mile of driving a heavier vehicle against the marginal private benefit per mile of driving a heavier vehicle. The marginal private cost of a heavier vehicle is positive due to the higher use of inputs to produce heavier vehicles (e.g. more steel, bigger tires, etc.) and the lower fuel efficiency of heavier vehicles. The marginal private benefit of a heavier vehicle is positive but decreasing in vehicle weight – heavier vehicles provide increased protection in a collision and more cargo capacity, but as size increases the vehicle becomes increasingly difficult to park and handle.[5] The consumer equates marginal private cost and marginal private benefit and buys a vehicle weighing W* pounds. The private operating cost per mile is P*. However, a heavier vehicle may impose a cost on other roadway users in the form of increased risk of fatalities in a collision with this vehicle, and the driver does not bear this external cost. If external costs increase linearly in vehicle weight, as we show is the case, the social marginal cost curve lies above the private marginal cost curve by a fixed amount equal to the external per mile cost. To maximize social welfare, our stylized consumer should purchase a car weighing W** pounds, where W** < W*. The necessary per-mile tax to induce this behavior is the marginal external cost of vehicle weight, t*. If the consumer chooses a vehicle of weight W*, the external cost from this choice over the socially optimal choice of a vehicle weighing W** would be t*·(W**–W*). We calculate this individual cost and aggregate it across all individuals to arrive at the total external costs.

It is important to note that the primary costs of this "arms race" accrue not in the form of traffic fatalities – which on net may change little with a uniform reduction in fleet weight – but rather in the form of purchases of larger vehicles that are more expensive to construct and operate. In this sense it is similar to a conventional arms race, which need not increase the probability of conflict even as both countries spend large amounts on new weapons.

In principle, liability rules and insurance regulations could internalize many of the external costs due to vehicle weight. If drivers of heavy vehicles know that they will be held liable for deaths in other vehicles, then they should take these risks into account when purchasing their own vehicles. If insurance companies understand that heavier vehicles pose more danger to other roadway users, then they should charge higher liability premiums to drivers of heavy vehicles. In practice, however, liability rules and insurance regulations fail to internalize the fatality risks generated by heavy vehicles.

---

[5] At some point the marginal private benefits of weight become negative. For example, few drivers would want a 30 foot stretched limousine as their primary vehicle, even if it were luxuriously appointed and heavily subsidized.

Tort liability rules are inadequate to internalize fatality risks for two reasons. First, liability only applies in cases in which a driver behaves in a negligent manner (White 2004). This implies that the driver of any given vehicle will frequently not be liable in the event of multivehicle accident. Second, even if found liable, few drivers possess assets that are sufficient to cover the cost of a fatality. The value of a statistical life used by the United States Department of Transportation in cost-benefit analyses is $5.8 million (2008 dollars), but only 7 percent of families in the United States had a net worth exceeding $1 million in 2001 (Kennickell 2003).

Though few drivers can cover the cost of a fatality, liability insurance regulations could force most drivers to pay the expected liability costs of operating their vehicles. Again, however, the mandated levels of liability insurance are inadequate to cover the costs of a fatality. Two states (Florida and New Hampshire) require drivers to carry no liability coverage for injuries at all, and 44 states require drivers to carry $25,000 or less in liability coverage for each person injured. Only five states require more than $25,000 of liability coverage for each person injured (Insurance Information Institute 2010).[6] Many drivers remain uninsured despite the regulations, and even drivers that carry more than the mandated minimums rarely have policies that exceed several hundred thousand dollars of coverage.

While liability rules and insurance regulations cannot internalize the majority of fatality costs, they may internalize a significant fraction of incapacitating injury costs. Estimates of the value of an incapacitating injury are far lower than the value of a statistical life, and it is plausible that insurance policies carried by many drivers could cover the costs of an incapacitating injury.[7] For this reason, our policy analysis focuses on external fatality costs and ignores external incapacitating injury costs. Accounting for injury costs would increase the magnitude of our results, but we cannot accurately estimate what fraction of injury costs are already internalized.

Previous work on the "arms race" on American roads has focused on the internal and external risks posed by the largest vehicles – pickup trucks and sport utility vehicles

---

[6] Minnesota and North Carolina each require $30,000 of liability coverage for each person injured, and Alaska, Maine, and Wisconsin each require $50,000 of liability coverage for each person injured. None of these states are in our data set.

[7] The National Safety Council, for example, estimates the comprehensive cost of an incapacitating injury at $214,000 (2008 dollars). In comparison, they estimate the comprehensive cost of a fatality at $4.2 million.

(SUVs) – relative to the typical passenger car. White (2004), Gayer (2004), Anderson (2008), and Li (2009) all conclude that light trucks (pickups and SUVs) impose significant risks relative to passenger cars. This study expands upon that literature by considering the fundamental role that vehicle weight plays in determining external risk. We recognize that *any* vehicle that is heavier than the smallest feasible vehicle poses some external risk to other roadway users. We quantify that risk and find that the total external costs of vehicle weight substantially exceed the external costs that accrue only from light trucks. Our comprehensive results span the entire range of the vehicle fleet and allow us to consider the broader implications of vehicle weight for energy policy.

To measure the effect of vehicle weight on external fatalities under ideal conditions, we would randomly assign vehicles of differing weights to drivers and observe the external fatality rates by vehicle type. Such an experiment is infeasible in practice, and even an analogous study using observational data is impractical due to substantial measurement error in vehicle stocks and model-level vehicle miles traveled in most states. Instead, we focus on the risk of a fatality conditional on a collision occurring. A key identifying assumption for our research design is that vehicle weight has no causal effect on the probability of a collision. We discuss this assumption below and conclude that, if it is violated, then the effect of vehicle weight on the probability of a collision is likely positive. Our estimates thus represent a lower bound on the effect of weight on external fatalities.

Consider the expected external fatalities for a vehicle of type *i* during time interval *t*. For simplicity, assume that *t* is short enough that the probability of multiple collisions during *t* is effectively zero.

$$E\left[\text{fatalities}_{it}\right] = E\left[E\left[\text{fatalities}_{it} \mid \text{collision}_{it}\right]\right] = E\left[\text{fatalities}_{it} \mid \text{collision}_{it}\right] \cdot P\left(\text{collision}_{it} = 1\right) \quad (1)$$

Equation (1) must hold via the law of iterated expectations. It implies that if weight has no effect on the probability of a collision, then the total effect of weight on external fatalities is proportional to the effect of weight on external fatalities conditional on a collision occurring. Weight may affect the probability of a collision in two ways, however. First, from an engineering perspective heavier vehicles are less maneuverable and have longer braking distances. Even if driver behavior is unchanged, heavier vehicles may therefore get into more accidents. Second, heavier vehicles may also affect driver behavior. On the margin, drivers may respond to the internal safety benefits of heavy vehicles by increasing their optimal collision rate (Peltzman 1975). Both the physical characteristics of heavier vehicles and the potential driver response to heavier vehicles could therefore generate a *positive* effect of vehicle weight on collision rates.

Empirical evidence also suggests that, if anything, heavier vehicles have higher collision rates than lighter vehicles. Evans (1984) examines the relationship between accident rates and vehicle weight using accident data and vehicle registration data from North Carolina, New York, and Michigan. He finds that, after conditioning on driver age, 4,000 pound vehicles have accident rates that are 39 percent higher than 2,000 pound vehicles. More recently, White (2004) and Anderson (2008) estimate that light trucks are 13 to 45 percent more likely to experience multivehicle collisions than passenger cars.[8] Of course, some of the observed differences in crash rates may be due to driver selection; careless drivers may choose heavier vehicles. Nevertheless, both theory and empirical evidence suggest that weight may directly increase the probability of experiencing a collision. We thus interpret our estimates – which are conditional on a collision occurring – as lower bounds on the causal effect of weight on external fatalities.

## 3. DATA

The data set consists of the population of police-reported accidents for eight states: Florida, Kansas, Kentucky, Maryland, Missouri, Ohio, Washington and Wyoming. These data come from the State Data System, maintained by the National Highway Traffic Safety Administration (NHTSA). We obtained permission from the head of each state's police force to use the data. The SDS data include information on injuries and fatalities, geographic location, weather conditions, use of safety equipment, and driver and occupant characteristics. We selected these eight states out of the 32 states currently participating in the SDS as they report the vehicle identification number (VIN) for the majority of vehicles in the data set. We purchased data tables from DataOne Software to match the first 9 digits of the VIN to curb weight data for each vehicle. We therefore observe curbside vehicle weight for over 70% of the vehicles in our data set. For analytic purposes, we decompose the data set into three sub-samples, two-vehicle crashes, three-vehicle crashes, and single-

[8] Using a different methodology, Gayer (2004) estimates that light truck collision rates may be as much as 200 percent higher than passenger car collision rates.

vehicle crashes. The two-vehicle crash data set is the focus of most of our analyses. It contains 6.5 million vehicles with curbside weight data.[9]

One important feature of the SDS data is that accidents only appear in the data set if the police take an accident report. According to NHTSA documentation, various estimates suggest that only half of all motor vehicle accidents are police reported. While many of the unreported accidents are single vehicle accidents, some no doubt involve two vehicles as well. This sampling frame could affect our estimates if vehicle weight affects the probability of a police report, all other factors held constant. Serious multivehicle accidents will always be reported to the police regardless of vehicle weight, but vehicle weight could affect the probability that a minor accident is reported to the police. Unlike the probability of a collision, there is no a priori reason to believe that vehicle weight must have a positive effect on the probability of a police report. On the one hand, collisions involving heavier vehicles cause more property damage, all other factors held constant, because more kinetic energy must be dissipated.[10] On the other hand, some heavier vehicles – pickup trucks – are more likely to be involved in rugged work. These trucks may have accumulated more dents, reducing the likelihood that they will report property damage from a minor accident.

If vehicle weight positively affects the reporting probability of minor accidents, then our estimates will represent a lower bound on the effect of weight on external fatalities. If vehicle weight negatively affects the reporting probability of minor accidents, however, then our estimates of the effect of weight on external fatalities could be upwardly biased. To test whether the "ruggedness" hypothesis affects our results, we estimate our regressions while limiting the sample to collisions that do not involve any light trucks. This sample restriction does not reduce the coefficient estimates.[11] We also conduct a series of falsification tests in Section 5 that imply that the sampling frame does not bias our results.

Table 1 presents summary statistics from our two-vehicle collision data set. This data set contains all collisions involving two light vehicles built after 1960. We define a light vehicle as any car, pickup truck, SUV, or minivan that weighs between 1,500 and 6,000

---

[9] The dataset contains the population of police reported accidents for Florida (1989-2005), Kansas (1990-2005), Kentucky (1998-2005), Maryland (1989-2005), Missouri (1989-2005), Ohio (1989-2005), Washington (1989-2005), and Wyoming (1998-2005).

[10] Kinetic energy is dissipated in a collision through the deformation of materials (i.e., property damage).

[11] In the sample that excludes all collisions involving light trucks, the estimated effects are 10 to 20 percent larger than the analogous estimates from the main sample, reported in Table 2. This implies that the "ruggedness" hypothesis is not upwardly biasing our main results (see online Appendix Table A1).

pounds. We exclude collisions involving heavy trucks. The first two columns report statistics for the entire two-vehicle collision data set. The mean vehicle weight in this data set is 3,052 pounds, and approximately 17.5 percent of vehicles are light trucks (pickups, SUVs, or minivans). The average model year is 1992, and the average number of occupants per vehicle is 1.44. The probability of a fatality in each vehicle is 0.19 percent (i.e., 0.0019), and the probability of a serious injury in each vehicle is 3.0 percent. Alcohol is involved in approximately 9 percent of collisions.

The last two columns of Table 1 report summary statistics for the estimation sample. The estimation sample is smaller than the overall two-vehicle collision sample because we drop any collisions in which curbside weight is missing for either vehicle. This restriction reduces the sample from 9.7 million observations to 4.8 million observations. Nevertheless, the two samples appear similar along most observable measures. We confirm in Section 4 that the missing weight data do not bias our estimates.

## 4. Specification and Results

Consider a collision involving two vehicles, Vehicle 1 and Vehicle 2. Suppose that we label Vehicle 1 as the "struck vehicle" and Vehicle 2 as the "striking vehicle." These labels are for expositional purposes only – they do not signify which vehicle may be at fault in the collision.[12] The external effects of vehicle weight are given by the effect of striking vehicle weight on the probability of fatalities in the struck vehicle. The internal effects of vehicle weight are given by the effect of struck vehicle weight on the probability of fatalities in the struck vehicle. The former is the quantity of policy interest, but we report results for the latter as well for comparison purposes.

We estimate the conditional expectation of fatalities occurring in the struck vehicle as a function of striking vehicle weight, struck vehicle weight, and a rich set of covariates. We estimate the conditional expectation function (CEF) using either a linear probability model (LPM) or a probit.[13] For robustness, we report estimates for both models.

---

[12] The labels are symmetric in that each vehicle enters our data set twice, once as the striking vehicle and once as the struck vehicle.

[13] The LPM cannot literally be true. Nevertheless, it provides the minimum mean squared error linear approximation to the true CEF, and in our case the LPM coefficients are always close to the corresponding average marginal effects from the probit models.

We specify the linear probability model as follows:

$$E\left[\text{struck veh fatality}_i \mid \text{striking veh weight}_i,\ \text{struck veh weight}_i,\ X_{1i},\ X_{2i},\ W_i\right] \quad (2)$$
$$= \beta_1 \text{striking veh weight}_i + \beta_2 \text{struck veh weight}_i + X_{1i}\delta_1 + X_{2i}\delta_2 + W_i\delta_3$$

In equation (2), $\beta_1$ represents the coefficient of interest, $X_{1i}$ represents a set of characteristics pertaining to the striking vehicle in collision *i*, $X_{2i}$ represents a set of characteristics pertaining to the struck vehicle in collision *i*, and $W_i$ represents a set of characteristics common to both vehicles in collision *i*. The probit model modifies equation (2) as follows:

$$E\left[\text{struck veh fatality}_i \mid \text{striking veh weight}_i,\ \text{struck veh weight}_i,\ X_{1i},\ X_{2i},\ W_i\right] \quad (3)$$
$$= \Phi\left(\alpha_1 \text{striking veh weight}_\text{i} + \alpha_2 \text{struck veh weight}_i + X_{1i}\gamma_1 + X_{2i}\gamma_2 + W_i\gamma_3\right)$$

In equation (3), the link function $\Phi$ is the normal CDF. Therefore, the marginal effect of striking vehicle weight varies with striking vehicle weight. For comparability with the LPM results, for each probit regression we report the average marginal effect across all observations included in that regression.[14]

Table 2 presents results from estimating equations (2) and (3) on the two-vehicle collision data set. The number of observations drops by 50 percent relative to the full data set because our regressions only include collisions in which vehicle weight is coded for both vehicles. Analyses restricted to the subset of states with low rates of missing weight data indicate that this constraint does not bias our results.[15] Also note that each vehicle appears in

---

[14] Some of our probit regressions include fixed effects, raising the possibility of inconsistency due to the incidental parameters problem. However, in most cases we have many observations for each fixed effect, and as shown in Fernandez-Val (2009), the incidental parameters problem generates a trivial degree of bias in the probit model when estimating marginal effects (which are our quantities of interest).

[15] Weight data are missing for vehicles for which we do not have VINs. The percentage of vehicles with missing weight data ranges from 11.7 percent (Wyoming) to 52.2 percent (Maryland). When estimating our main statistical models on the subset of states with low rates of missing weight data (Ohio, Washington, and Wyoming), we find that an additional 1,000 pounds of striking vehicle weight increases the probability of a fatality in the struck vehicle by 47 to 52 percent. When estimating the same models on the subset of states with high rates of missing weight data (Florida, Kansas, Kentucky, Maryland, and Missouri), we find that an additional 1,000 pounds of striking vehicle weight increases the probability of a fatality in the struck vehicle by 49 percent. The rate of missing weight data thus appears to have no impact on our estimates (see online Appendix Table A2).

the two-vehicle collision data set twice, once as the struck vehicle and once as the striking vehicle. We therefore cluster the standard errors at the collision level to account for correlation between observations that pertain to the same collision.

The first and second columns in Table 2 include the following covariates: vehicle weight, light truck indicators, and year fixed effects. A striking vehicle and struck vehicle version of each of the first two variables is included. The first column implies that an increase in weight of 1,000 pounds in the striking vehicle is associated with a statistically significant 0.10 percentage point increase in the probability of a fatality in the struck vehicle ($t$ = 25.0). This coefficient represents a 53 percent increase over the average probability of a fatality in a struck vehicle in this sample (0.19 percent). In comparison, an increase in weight of 1,000 pounds in the struck vehicle is associated with a smaller 0.04 percentage point decrease in the probability of a fatality in the struck vehicle ($t$ = –14.7). Light trucks increase the probability of a fatality in the struck vehicle by 0.10 percentage points (51 percent of the sample mean), even after controlling for striking vehicle weight ($t$ = 13.7). The results from the probit model in column (2) display $z$-statistics that are similar to the $t$-statistics in column (1), and the average marginal effect generated by the probit model is of similar magnitude to the LPM coefficient (0.10 percentage points versus 0.09 percentage points).

Subsequent columns in Table 2 add additional covariates to the regressions. Columns (3) and (4) add controls for rain, darkness, day of week (weekday versus weekend), Interstate highway, a quadratic in model year for each vehicle, and year, hour, and county fixed effects. The estimated effect of striking vehicle weight changes little in both the LPM and probit models. Columns (5) and (6) add controls for any seat belt usage, a quadratic in driver age, indicators for drivers under 21 or over 60, and indicators for male drivers or young male drivers. A striking vehicle and struck vehicle version of each of these variables is included. The inclusion of these driver characteristics has no impact on the primary coefficient of interest (striking vehicle weight). They do, however, increase the magnitude of the struck vehicle weight coefficient to –0.09 percentage points ($t$ = –18.4).

Column (7) of Table 2 adds city fixed effects and is our preferred specification. City fixed effects should absorb any geographic heterogeneity in fatality rates that could be correlated with average vehicle weight. This issue would arise if, for example, heavy vehicles clustered in rural areas and these areas had deadlier accidents due to a prevalence of undivided highways or a sparseness of hospitals. At this point there are too many regressors to reliably estimate a probit model, and for many cities the city fixed effect perfectly predicts the fatality indicator, forcing the city to be dropped. We thus estimate only linear probability models in columns (7) through (9) of Table 2. The addition of city fixed effects has little impact on the coefficient on striking vehicle weight, changing it from 0.11 percentage points to 0.12 percentage points ($t$ = 19.8). Column (8) estimates the same specification as column (7) but limits the sample to observations for which we have data on the number of occupants per vehicle and the seat belt usage of each occupant (two controls we add in the next column). This restriction shrinks the sample by 50 percent and reduces the coefficient on striking vehicle weight to 0.06 percentage points ($t$ = 9.1). However, the ratio of the coefficient to the average fatality rate in the sample remains stable (45 percent). The change in the coefficient simply reflects the fact that the restricted sample contains states with a lower threshold for reporting accidents, and thus a lower fatality rate per reported accident. Column (9) adds controls for the number of occupants per vehicle and seat belt usage rate of these occupants. The coefficient on striking vehicle weight is unchanged from column (8).

The results in Table 2 suggest that selection bias has little impact on the striking vehicle weight coefficient but is a larger issue for the struck vehicle weight coefficient. In particular, the addition of driver characteristic controls in columns (5) and (6) has a notable impact on the struck vehicle weight coefficient but almost no impact on the striking vehicle weight coefficient. When adding covariates one at a time, we find that virtually all of the change in the struck vehicle weight coefficient between columns (4) and (6) can be attributed to the addition of the controls for driver age. The patterns strongly suggests that older drivers tend to drive heavier vehicles and that older drivers are more susceptible to dying in crashes. Since there is little correlation between the age of the struck vehicle's driver and the weight of the striking vehicle, however, the addition of driver age controls has no impact on the striking vehicle weight coefficient. Stated simply, heavy vehicles do not "seek out" elderly drivers to crash into.

The results in Table 2 also suggest that the external risk posed by light trucks is not due solely to their heavy weight. The coefficient on the indicator for whether the striking vehicle is a light truck is positive and statistically significant in every column. In our preferred specification, column (7), the coefficient implies that being struck by a light truck increases the probability of a fatality by 0.07 percentage points ($t$ = 7.5), even after conditioning on striking vehicle weight. This represents a 30 percent increase over the average fatality rate in the sample. In comparison, if we do not control for vehicle weight, then the light truck coefficient doubles to 0.15 percentage points (i.e., 0.0015).[16] The additional risk posed by light trucks may be due to the stiffness of their frames or their height incompatibility with other vehicles (Hakim 2003). However, the robustness tests that we perform in Section 5 for the vehicle weight coefficient do not carry over to the light truck coefficient. Thus we cannot rule out the possibility that a portion of the light truck coefficient may represent driver selection effects – i.e., consumers that purchase light trucks may drive in an aggressive manner that generates particularly severe collisions. For this reason we do not incorporate the light truck coefficient when calculating the total externality across all vehicles in Section 6.

Table 3 presents results from estimating versions of equations (2) and (3) in which the dependent variable is the presence of serious injuries in the struck vehicle. The regressions are analogous to those in Table 2, but the dependent variable has changed from any fatalities to any serious injuries. The striking vehicle weight coefficients (or marginal effects, in the case of probit regressions) in Table 3 are approximately 6 times larger than the corresponding coefficients in Table 2. This difference arises because the probability of a serious injury in these collisions is approximately 15 times higher than the probability of a fatality. In the preferred specification, column (7), a 1,000 pound increase in striking vehicle weight raises the probability of serious injuries in the struck vehicle by 0.7 percentage points ($t$ = 35.5). This figure represents 20 percent of the average probability of a serious injury in this sample.

Overall, the pattern of coefficients in Table 3 is similar to the pattern of coefficients in Table 2, with one exception. When the dependent variable is the presence of serious injuries (Table 3), the magnitude of the struck vehicle weight coefficient is larger than the magnitude of the striking vehicle weight coefficient. For example, in the preferred specification the striking vehicle weight coefficient is 0.7 percentage points while the struck vehicle weight coefficient is –0.8 percentage points. This contrasts with Table 2, in which the magnitude of the struck vehicle weight coefficient is always smaller than the magnitude of

[16] The 0.15 percentage point coefficient represents 62 percent of the average fatality rate in the sample. This effect is similar in magnitude to the external effects of light trucks in two-vehicle collisions that White (2004) and Anderson (2008) estimate. Anderson (2008), for example, estimates that light trucks increase the probability of a fatality in the struck vehicle by approximately 60 percent of the sample average fatality rate.

the striking vehicle weight coefficient.[17] Since the proportion of serious injuries that represent external costs is ambiguous, we focus on fatalities for the remainder of the paper.

Table 4 presents results testing for heterogeneity in the effect of striking vehicle weight on fatalities. In column (1) we interact the striking vehicle weight variable with indicators for whether the striking vehicle is a light truck and whether the struck vehicle is a light truck. The interaction coefficients are negative, suggesting that the effect of striking vehicle weight is somewhat lower if either vehicle a light truck. However, both coefficients are small in magnitude, and neither attains more than marginal significance. In column (2) we add a quadratic term in striking vehicle weight. The coefficient on the quadratic term is negative, suggesting that the marginal effect of striking vehicle weight may be smaller for heavier vehicles. However, the coefficient is again small and statistically insignificant.

Columns (3) and (4) replicate columns (1) and (2) but are estimated using the probit model instead of the LPM. When using the probit model, the light truck interaction terms and the quadratic weight term are highly significant, suggesting non-linear effects from striking vehicle weight. In fact, the opposite is true. The probit is an inherently non-linear model that forces the marginal effect of vehicle weight to increase in accidents that involve heavier striking vehicles.[18] Including the light truck interactions or the quadratic weight term, however, allows the regression to offset this increase, and the resulting function is much closer to a linear function. Figure 2 demonstrates this fact. It plots the estimated marginal effects of striking vehicle weight for four models: linear OLS, quadratic OLS, linear probit, and quadratic probit.[19] The marginal effects of the linear OLS, quadratic OLS, and quadratic probit are generally similar, particularly between 2,400 to 4,500 pounds of vehicle weight (a range which includes over 80 percent of the vehicles in our sample). In contrast the marginal effects of the linear probit model diverge substantially from the marginal effects of the other three models. Since both the flexible OLS and flexible probit models suggest that the true CEF is approximately linear in striking vehicle weight, and because the probit cannot

[17] For example, in column (7) of Table 2, the striking vehicle weight coefficient is 0.12 percentage points while the struck vehicle weight coefficient is –0.09 percentage points.

[18] The probit marginal effect equals $\phi(X\beta) \cdot \beta$, where $\phi(\cdot)$ represents the standard normal density function. Since probability of a fatality is less than 50 percent, $\phi(X\beta)$ is increasing in $X\beta$. The marginal effect of striking vehicle weight thus increases in striking vehicle weight. The rate of increase is substantial since the effect of striking vehicle weight is large.

[19] The "linear probit" is a model in which there are no higher order terms of striking vehicle weight. It is not literally a linear model. The "quadratic probit" is a model in which both striking vehicle weight and striking vehicle weight squared appear on the right-hand side.

accommodate city level fixed effects, we focus on linear probability models in much of the remaining analysis.[20]

Though 90 percent of multivehicle collisions involve two vehicles, nine percent involve three vehicles, and one percent involve four or more vehicles. Adding 1,000 pounds to a vehicle in a three-vehicle collision should increase the risk of a fatality in the other two vehicles by less than 49 percent (our preferred estimate from the two-vehicle collision data set). This attenuation occurs because the extra mass of the first vehicle is now distributed across two other vehicles rather than one other vehicle. We estimate the relationship between vehicle weight and fatalities in three-vehicle collisions in Table 5. For expositional purposes assume that Vehicle 1 is the struck vehicle and that Vehicles 2 and 3 are the striking vehicles. In Table 5 the striking vehicle weight coefficient represents the average effect of a 1,000 pound increase in the weight of either Vehicle 2 or 3 (but not both) on the probability of a fatality in Vehicle 1. The striking vehicle weight coefficient is positive and statistically significant in all specifications, and the magnitude of the coefficient ranges from 27 to 44 percent of the average probability of a fatality. Our preferred estimate, column (7), implies that a 1,000 pound increase in one vehicle raises the probability of a fatality in either of the other two vehicles by 34 percent.

## 5. FALSIFICATION TESTS AND ALTERNATIVE SOURCES OF IDENTIFICATION

The results in Section 4 demonstrate a strong relationship between striking vehicle weight and struck vehicle fatalities. The robustness of this relationship to the inclusion of a rich set of accident and driver characteristics, as well as very fine geographic fixed effects, suggests that the striking vehicle weight coefficients represent causal effects of weight on fatality risk. However, two potential sources of upward bias seem particularly plausible. First, driver selection might bias the coefficient estimates if heavier vehicles attract aggressive drivers who get into deadlier accidents. Second, the sampling frame might bias the coefficient estimates if minor collisions involving heavier vehicles are less likely to be

[20] For simplicity we assume a linear effect of striking vehicle weight when comparing a gasoline tax to a weight varying mileage tax in Section 6. This assumption is conservative in that the fit between the gasoline tax and the weight varying mileage tax improves if the true marginal effects decrease below 2,400 lbs and above 4,500 lbs, as implied by the quadratic probit.

reported to the police, all other factors held constant.[21] To test whether either of these factors could bias our results, we conduct three exercises. First, we implement a series of falsification tests that we benchmark against engineering safety estimates. Second, we estimate the effect of striking vehicle weight on fatalities using within-model changes in vehicle weight that occur when models are refreshed. Finally, we estimate the effect of striking vehicle weight on fatalities using striking vehicle occupants as an instrument for weight.

### 5.1 FALSIFICATION TESTS

Suppose that heavier vehicles pose no additional risk to other vehicles than lighter vehicles do, and that the estimates reported in Section 4 simply reflect the possibility that drivers of heavier vehicles are more aggressive (regardless of vehicle weight) or that heavier vehicles are less likely to generate police reports. In that case, there should be a strong positive correlation between vehicle weight and fatalities or injuries when analyzing two-vehicle collisions between vehicles of the same weight. These accidents therefore provide an opportunity to test whether driver selection bias or sampling frame bias are generating our results.

It is possible, however, that heavier vehicles are safer than lighter vehicles. In that case, a positive driver selection effect might be mitigated by a negative weight effect. Put simply, even if drivers of heavier vehicles drive aggressively, our falsification test might generate a small coefficient because the heavier vehicles are fundamentally safer. We therefore benchmark the results of our falsification tests against the results of NHTSA crash tests. NHTSA crash tests entail colliding a vehicle with a concrete barrier; they are meant to simulate the results of a collision with a stationary object or a head-on collision with another vehicle of similar weight. The primary outcome in the NHTSA crash test is the Head Injury Criterion (HIC). This variable is derived from an accelerometer mounted on the crash test

[21] Note that, unlike the struck vehicle weight coefficients, striking vehicle weight coefficients are unlikely to be biased by any correlation between vehicle weight and vehicle safety features. It is plausible that heavier vehicles may be more or less likely to have safety features such as airbags, side impact protection beams, and unibody construction. However, these safety features will generally be much more helpful to the striking vehicle's own occupants than they are to the occupants of other vehicles that the striking vehicle hits.

dummy's head and measures the forces that the head is exposed to. A higher HIC value corresponds to a higher probability of severe or fatal head injury.

Table 6 presents results from regressions of HIC scores on vehicle weight using the NHTSA crash test data. All regressions include as controls a light truck indicator, a quadratic in vehicle model year, and a quadratic in collision speed. The estimation sample in the first two columns contains all NHTSA vehicle-to-barrier frontal crash tests conducted from 1973 to 2009 (the average year is 1996). Column (1) reports regression results when the dependent variable is HIC. The results indicate that an additional 1,000 pounds of vehicle weight is associated with a statistically insignificant 2.6 percent increase in HIC (15.8 points). Column (2) reports regression results when the dependent variable is is an indicator for whether HIC exceeds 700. This threshold is of interest because it represents the point at which there is a significant (5 percent) chance of severe brain injury (Mertz, Prasad, and Irwin 1997). The results indicate that an additional 1,000 pounds of vehicle weight is associated with a statistically insignificant 5.6 percent increase in the probability that HIC exceeds 700 (1.6 percentage points). The composition of vehicles that NHTSA tests is not identical to the composition of vehicles on the roadways, however. To account for this fact, we estimate regressions in which each test result is weighted by the sales share of the tested vehicle.[22] Columns (3) and (4) report the results from these regressions. The sample size falls because we do not have sales share data for every tested vehicle, but the results are qualitatively unchanged. An additional 1,000 pounds of vehicle weight is associated with a small, statistically insignificant increase in HIC or the probability that HIC exceeds 700. Overall there is a weak positive relationship between vehicle weight and HIC values. The point estimates suggest that an additional 1,000 pounds of vehicle weight could raise the fatality rate by 3 to 7 percent. We thus expect a weak relationship between vehicle weight and fatalities in collisions between two equal weight vehicles if our research design is sound.

Table 7 presents results from regressions in which the estimation sample consists of collisions involving two vehicles of similar weight – the difference in vehicle weight cannot exceed 200 pounds. In each regression, an indicator for fatalities in the struck vehicle is regressed on the average weight of the two vehicles and the set of controls from the preferred specification. Column (1) indicates that an increase of 1,000 pounds in average vehicle weight predicts a statistically insignificant 7 percent increase in the probability of a

[22] Vehicle sales share data come from *Ward's Automotive Yearbook*.

fatality (0.02 percentage points). Column (2) restricts the sample to head-on collisions between two vehicles of the same weight, the type of collision simulated by NHTSA. In this sample, an increase of 1,000 pounds in average vehicle weight predicts a statistically insignificant 14 percent decrease in the probability of a fatality (0.09 percentage points).[23] Columns (3) and (4) replicate columns (1) and (2) but restrict the sample so that the difference in vehicle weight cannot exceed 100 pounds. The estimates remain small and statistically insignificant but are less precisely estimated.

Overall the estimates in Table 7 indicate that there is a weak relationship between vehicle weight and fatalities in collisions between two vehicles of equal weight, and we cannot reject the hypothesis that this relationship is zero. This finding is consistent with NHTSA crash test results (Table 6) and inconsistent with the hypothesis that driver selection bias or sampling frame bias is generating the results in Section 4. The most precise estimate in Table 7 – column (1) – suggests that increasing average vehicle weight by 1,000 pounds raises the fatality rate by 7 percent. This figure falls within the 3–7 percent range implied by the NHTSA crash test data.

In contrast, if the relationship between striking vehicle weight and struck vehicle fatalities were generated by driver selection bias or sampling frame bias, then we would expect a large positive coefficient on average vehicle weight when two vehicles of equal weight collide. The preferred estimate from Section 4 indicates that a 1,000 pound increase in striking vehicle weight raises the probability of a fatality in the struck vehicle by 49 percent. If this coefficient represented driver selection bias, and if two aggressive drivers were twice as dangerous as one aggressive driver, then we might expect a 1,000 pound increase in both vehicles to raise the probability of a fatality by 98 percent (2*49 = 98). However, no coefficient in Table 7 even reaches 10 percent.

As an additional set of falsification tests, we examine the relationship between vehicle weight and fatalities in collisions involving a single vehicle. If drivers of heavier vehicles are more aggressive, then we expect to observe a strong positive relationship between vehicle weight and fatalities in these collisions. Table 8 presents results for single-vehicle collisions. In these collisions, we regress a fatality indicator on vehicle weight and other controls. The results in column (1) pertain to all single-vehicle collisions; a 1,000

[23] The average probability of a fatality is much higher in column (2) than in column (1) because head-on collisions are more dangerous than the average collision.

pound increase in vehicle weight is associated with a 6 percent increase in the probability of a fatality (0.09 percentage points).[24] Column (2) pertains to single-vehicle frontal collisions, the type of collision simulated by NHTSA. A 1,000 pound increase in vehicle weight is associated with a 4 percent increase in the probability of a fatality (0.6 percentage points). Columns (3) and (4) present results that are analogous to columns (1) and (2) but are estimated using a probit specification instead of a linear probability model. In both columns a 1,000 pound increase in vehicle weight is associated with a 3 percent increase in the probability of a fatality. In all columns the percentage effects fall within the 3–7 percent range implied by the NHTSA crash test data, suggesting no substantial bias due to driver selection.

### 5.2 VEHICLE MODEL FIXED EFFECTS RESULTS

To further establish the robustness of our results, we explore two alternative sources of identification. Our first alternative leverages within-model changes in vehicle weight to estimate the effect of striking vehicle weight on fatalities. To implement this design we include vehicle model fixed effects for the striking vehicle in our preferred specification. The effect of striking vehicle weight on fatalities is thus identified off of changes in vehicle weight that occur when a vehicle model is refreshed. This design minimizes the impact of driver selection as long as the composition of customers for a particular vehicle model remains relatively stable when the model is refreshed.

Table 9 reports results from estimation models that include vehicle model fixed effects. Column (1) presents results from our preferred specification estimated on the sample for which we have complete vehicle model data. The sample size is substantially smaller than our main analytic sample because only five states – Kansas, Kentucky, Maryland, Ohio, and Wyoming – report detailed vehicle model data. In this subsample a 1,000 pound increase in striking vehicle weight is associated with a 50 percent increase in the probability of a fatality in the struck vehicle (0.08 percentage points, $t$ = 8.7). This effect is consistent with the estimates from Section 4. Column (2) presents results from the same specification with

[24] The raw magnitude of the coefficients is much larger in Table 8 than in Table 7 because the fatality rate in single-vehicle collisions is approximately 7 times higher than the fatality rate in two-vehicle collisions. This occurs because observed single-vehicle collisions tend to be more severe; drivers have no incentive to report minor single-vehicle collisions to their insurers or the police.

vehicle model fixed effects added.[25] A 1,000 pound increase in striking vehicle weight is now associated with a 52 increase in the probability of a fatality in the struck vehicle (0.08 percentage points, $t$ = 4.7). The close correspondence between the two coefficient estimates suggests that driver selection does not seriously bias our results.

### 5.3 INSTRUMENTAL VARIABLES RESULTS

Our second alternate source of identification leverages the number of occupants in the striking vehicle as an instrument for striking vehicle weight. The number of occupants in the striking vehicle directly affects total striking vehicle weight, so the first condition for an instrumental variable – that it be correlated with the regressor of interest – is satisfied. The instrumental variables (IV) regression we estimate is:

$$\begin{aligned}\text{struck veh fatality}_i = \alpha_1 \widehat{\text{striking veh added weight}}_i + \alpha_2 \text{striking veh curb weight}_i + \\ \mathrm{X}_{1i}\gamma_1 + \mathrm{X}_{2i}\gamma_2 + W_i\gamma_3 + \varepsilon_i\end{aligned} \qquad (4)$$

In this regression, $\widehat{\text{striking veh added weight}}_i$ equals the number of occupants in the striking vehicle multiplied by 164 pounds, which is the average weight of an additional occupant circa 2000.[26] The regression controls for the curb weight of each vehicle (i.e., vehicle weight absent any passengers or cargo) as well as all the covariates from our preferred specification. The identification thus comes from variation in the number of occupants in the striking vehicle after controlling for the curb weight of the striking vehicle. This means that the identifying variation in the IV regression is orthogonal to the variation in curb weight that we use in Section 4.

---

[25] Across the five states with detailed vehicle model data there are 19,105 make-model combinations. Our specification thus includes 19,104 vehicle model fixed effects.

[26] We calculate this figure as follows. First, for the nonrandom subset of accidents for which we have detailed occupant characteristics, we tabulate the share of additional occupants that are male adults, female adults, male children, and female children. We find that 22.7 percent of additional occupants are male adults, 37.6 percent are female adults, 19.5 percent are male children, and 20.2 percent are female children. Using national statistics on body weight by gender and age we then compute the average weight of an additional occupant as 0.227*190 lbs + 0.376*163 lbs + 0.195*110 lbs + 0.202*114 lbs = 149 lbs (Ogden et al. 2004). Finally, we add 15 lbs per occupant to account for clothing, outerwear, and personal belongings (149 lbs + 15 lbs = 164 lbs).

Nevertheless, it is not obvious that the number of occupants in the striking vehicle satisfies the second condition for an instrumental variable – that it be uncorrelated with any other factors that affect fatalities in the struck vehicle. It is possible that, even after controlling for vehicle curb weight and other characteristics, drivers who carry additional occupants in their vehicles drive more aggressively than drivers who do not carry additional occupants. If this were true, then our IV estimates would be biased upward. We thus do not interpret our IV estimates as being more robust than our OLS estimates. Instead, we recognize that the identifying variation in the IV regression is fundamentally different than the identifying variation in the OLS regression. If the two regressions produce qualitatively similar estimates, this suggests that both are estimating causal effects. If the two regressions produce very different estimates, this suggests that one (or both) may be biased.

The last two columns of Table 9 report coefficients from the instrumental variables sample. The IV sample is approximately half the size of our main analytic sample because data on the number of occupants is not available in every state. Column (3) presents results from estimating the preferred OLS specification (column (7) of Table 2) on the IV sample. A 1,000 pound increase in striking vehicle weight is associated with a statistically significant 0.064 percentage point increase in the probability of a fatality in the struck vehicle ($t$ = 9.1). This coefficient represents a 45 percent increase over the average probability of a fatality in a struck vehicle, which is consistent with the results in Section 4. Column (4) presents results from the IV regression in equation (4). The reported coefficient is $\propto_1$, the coefficient on predicted additional weight in the striking vehicle. An additional 1,000 pounds of occupant weight in the striking vehicle is associated with a statistically significant 0.078 percentage point increase in the probability of a fatality in the struck vehicle ($t$ = 2.9). This coefficient represents a 54 percent increase over the average probability of a fatality in the struck vehicle. It is not far from the coefficient in column (3), and we cannot reject the hypothesis that the two coefficients are identical. The correspondence between the OLS and IV results thus increases our confidence in both estimators.

## 6. Policy Implications

The econometric evidence demonstrates that the impact of heavier striking vehicles on fatalities in struck vehicles is statistically significant and robust to the inclusion of an extensive set of vehicle, driver and accident confounders, estimation methods and identification strategies. This section explores whether the estimated causal effect of vehicle weight on fatalities is economically significant and compares two possible price based policies to distribute the external costs across drivers.

To quantify the magnitude of the external costs of increased vehicle weight, we design the following counterfactual experiment. We consider the externality at the individual level, whereby purchasing and driving a heavy vehicle increases the probability of a fatality in a collision with other vehicles on the road. We conduct a thought experiment in which an individual chooses a vehicle of a certain weight and calculate the external costs from this individual's vehicle choice. We carry out this calculation for each driver on the road while holding the remainder of the fleet constant. We sum across individuals to get the total externality from all individuals' vehicle choices. For the purposes of this calculation we assume that the individual chooses a vehicle weighing as much as the average car sold in 2005 (3,449 pounds). We calculate the total external costs against two baseline vehicles that the individual could buy – a slightly lighter vehicle and the lightest possible vehicle. The first counterfactual vehicle represents the average model sold in 1989, which weighed 2,824 pounds. The second counterfactual vehicle is the smallest drivable car currently in mass production, which weighs approximately 1,900 pounds. In addition we run a scenario that incorporates the external cost from pedestrian and motorcycle fatalities. All of our scenarios represent partial equilibrium approaches to arriving at total external costs – they assume that our regression estimates would not change if the vehicle fleet changed in response to the policies considered. Constructing a general equilibrium model is beyond the scope of this paper, and we restrict ourselves to providing estimates of the total external costs in the context of the current vehicle fleet.

Table 10 presents the results from our counterfactual experiments. We assume the individual purchases a vehicle weighing 3,449 lbs, or the mean weight of 2005 model year vehicles in our estimation sample. The first counterfactual vehicle the individual could have purchased weighs 2,824 pounds, or the mean weight of 1989 model year vehicles in our estimation sample. From 1989 to 2005, the average model sold gained 625 pounds, with heterogeneity in weight gain by model. During this period the Honda Civic gained 457 pounds, the Toyota Camry gained 515 pounds, and the Ford Explorer gained 490 pounds. The Honda Odyssey, a premier minivan, gained 1,060 pounds. Honda's smallest compact car, the 2010 Honda Fit, weighs 360 pounds more than the 1981 Accord, which is now a full size sedan. Similar patterns emerge for most other manufacturers.

When summed across all individuals, our counterfactual scenario computes the total external cost of a 2005 model year vehicle fleet over the representative 1989 model year vehicle. Our experiment is robust to the specific distribution of vehicle weight within the fleet as the probability of a fatality is linear in striking vehicle weight – the linearity ensures that the mean vehicle weight is a sufficient statistic for our policy analysis. We choose 2005 as our cutoff year as many of the parameters necessary for our full simulation are only available as recently as 2005.

The change in the probability of an external fatality for an individual buying vehicle model *i* weighing $w_i$ over a lighter vehicle weighing $w_{cf}$ is given by:

$$\text{External Cost}_i = \beta \cdot (w_i - w_{cf}) \cdot P(accident) \cdot VSL \tag{5}$$

For $\beta$ we employ our preferred estimate of the causal effect of weight on the probability of a fatality in an accident, or 0.119 percentage points for each additional 1,000 lbs in striking vehicle weight.[27] In all experiments we set $w_i$ at 3,449 pounds. We calculate the probability of a vehicle being involved in a multivehicle collision at 3.65% per year (NHTSA, 2007).[28] We use the DOT value of a statistical life of US$ 5.8 million.

If our simulated individual chooses a vehicle weighing 3,449 lbs ($w_i$) instead of one weighing 2,824 lbs ($w_{cf}$), she causes an additional 0.000027 external fatalities per year in expectation, valued at $157.45. Column (1) of Table 10 shows that the total external cost of vehicle weight gain relative to the 1989 baseline vehicle is therefore $35.6 billion per year. This figure represents the "weight gain since 1989" scenario but does not encompass the total external costs from vehicle weight.

Our second counterfactual scenario assumes the individual purchases the 3,449 lb vehicle ($w_i$) over a vehicle weighing 1,896 lbs ($w_{cf}$), which represents the lightest vehicle in mass production that can transport at least two adult passengers and is classified as an automobile. This is the weight of Toyota's iQ or roughly the weight of the Mercedes Benz Smart Car or the first generation Honda Insight. The intuition behind calculating the total external cost using this baseline vehicle is that individuals privately choose the size of the

---

[27] For comparison, the probability in our sample of a fatality in a two-vehicle collision (conditional on the collision occurring) is 0.23 percent.

[28] We estimate the probability of being involved in an accident by dividing the total number of vehicles involved in multivehicle collisions by the total number of registered vehicles in 2005 (BTS, 2010 Table 1-11).

externality by choosing a heavier vehicle than required to provide baseline transportation services. This calculation recognizes that a driver of a Smart Car poses little risk to other roadway users except bicyclists or motorcyclists.[29]

If our simulated individual chooses a vehicle weighing 3,449 lbs instead of one weighing 1,896 lbs, she causes an additional 0.000067 external fatalities per year in expectation, valued at $391.41. Column (2) in Table 10 shows that summed across all vehicles this translates into a total external cost of US$88.5 billion per year.

This scenario, however, ignores the external fatality risks that vehicles pose to pedestrians and motorcyclists. Column 3 in Table 10 adds this baseline risk to the simulation. In 2005, there were 2,659 motorcycle crash fatalities and 5,864 non-motorist fatalities due to fatal crashes (NHTSA 2010). This is equivalent to an external "baseline" fatality cost of $49.4 billion. The total external cost of "excess" vehicle weight and baseline fatality risk is $137.9 billion.

The above calculations ignore the impact of higher striking vehicle weight in multivehicle collisions with more than two vehicles. The majority of these accidents involve three vehicles. Columns 4–6 in Table 10 repeat the simulation above but add the external costs in three-vehicle collisions. We assume that striking vehicle weight has half the causal effect in three-vehicle accidents as compared to its effect in two-vehicle collisions. This assumption is conservative in comparison to our three-vehicle collision estimates in Table 5. These calculations add 1,459 fatalities to the iQ scenario and raise the total external costs to $146.4 billion.

While the magnitude of the total external costs is a straightforward calculation, translating it into an optimal policy is not. The externality consists of fatalities in collisions with pedestrians, motorcyclists, and other vehicles. These costs, as discussed above, are not currently reflected in liability insurance because most coverage levels are far below the VSL of US$ 5.8 million. One way to incorporate these external risks is to include them in a per mile insurance charge. But in contrast to existing proposals for pay as you drive (PAYD) insurance (e.g. Bordoff and Noel 2008), our results demonstrate that the per mile insurance charge should vary sharply by weight – a heavier car generates greater expected external

[29] We do not consider a "zero weight" baseline vehicle because that weight lies far outside the support of our data. Furthermore, it is unclear what the counterfactual is if the vehicle does not exist at all. Would the collision not occur, or would the struck vehicle instead hit a different vehicle or a roadside object?

costs than a lighter car with the same usage. In order to assess a tax that varies per pound and per mile, one needs accurate information on vehicle miles travelled (VMT) for each individual vehicle, which creates substantial monitoring challenges.

A practical policy alternative is to distribute the total external costs by raising the gasoline tax assessed per gallon. Taxing gasoline is appealing because it is simple and because gasoline usage is positively related to both miles driven and vehicle weight. The United States consumed 140 billion gallons of gasoline in 2005 (EIA 2010). If we spread the total external costs calculated above across 140 billion gallons of gasoline, this translates into 28 cents per gallon in the "weight gain since 1989" scenario. The total externality due to vehicle fatalities when the baseline vehicle is 1,896 pounds translates into a tax of 69 cents per gallon. Including pedestrian and motorcycle fatalities translates into a tax of US$1.04 per gallon.

While the gasoline tax does not differ by the type of vehicle fueled, it is correlated with vehicle weight since heavier vehicles have lower fuel economy. Figure 3 plots a lowess smoother of miles per gallon (mpg) against vehicle weight, estimated for cars in model year 2005 using the data from Knittel (forthcoming).[30] There is a strong negative, slightly nonlinear, relationship between the two variables. A linear regression indicates that an additional 1,000 pounds in vehicle weight decreases fuel economy by 4.5 mpg. A gas tax thus results in heavier vehicles indirectly paying a higher per mile tax through its correlation with vehicle weight. In this sense the gas tax approximates a weight varying mileage tax.

A natural question is how close the gasoline tax comes to achieving the desired weight varying mileage tax. We perform a back of the envelope calculation using a large set of vehicles for which we have vehicle weight and mpg ratings from Knittel (forthcoming). The weight based external cost for vehicle type *i* per VMT is given by

$$c_i^e = \frac{\beta\left(w_i - w_{cf}\right) \cdot VSL \cdot P(accident)}{VMT} + c_{pedmot}^e \qquad (6)$$

where $\beta$ is again the estimated causal effect of vehicle weight (0.00119), $w_i$ is the chosen vehicle's weight, $w_{cf}$ is the baseline vehicle's weight, VSL is the value of a statistical life and *P(accident)* is the probability of being involved in a multivehicle collision. We calculate $c_i^e$ for each model in our database. VMT are held constant for each model at 11,000 miles per year

[30] For this comparison we require vehicle weight and EPA fuel economy ratings. The latter are not contained in our VIN decoder database, but Chris Knittel has graciously shared his model level data on weight and fuel economy ratings.

and the per mile motorcycle and pedestrian cost is set at 1.65 cents per mile. The average value of $c_i^e$ across all models is 5.6 cents per mile.

The gas tax per mile for model *i* is given by

$$c_i^g = \frac{\overline{c}_e}{mpg_i} \tag{7}$$

where $\overline{c}_e$ is the average external cost per gallon, calculated as \$1.04 in column (6) of Table 10. This figure includes the motorcycle and pedestrian components as well as three-vehicle collisions. For $mpg_i$ we use the standard 45/55 weighting of the EPA city and highway fuel economy ratings.[31] The gas tax per mile therefore only varies across models through differences in fuel economy.

In the following analysis we compare the gas tax per mile to the weight based mileage tax. For the analysis we remove boutique vehicles, which essentially have zero market share (e.g. Lamborghini, Ferrari, Bentley), flex fuel vehicles, which have inflated mpg ratings for accounting reasons, and a few miscoded observations. We examine vehicles built from 1997 to 2006 to approximate the vehicle fleet in the last year of our sample. This sample contains 8,201 model-year combinations and includes most cars and light trucks sold in the United States during this period. The weight-based tax displays higher variability, with a standard deviation of 1.71 cents per mile compared to the gas tax's standard deviation of 1.25 cents per mile.

The difference between the two taxes for model *i* can be expressed as:

$$\Delta_i = c_i^g - c_i^e = \alpha + \frac{\overline{c}_e}{mpg_i} - \gamma w_i \tag{8}$$

where $\alpha = \dfrac{\beta \cdot w_{cf} \cdot VSL \cdot P(accident)}{VMT} - c^e_{pedmot}$ and $\gamma = \dfrac{\beta \cdot VSL \cdot P(accident)}{VMT}$. We keep both α and γ fixed at the values stated previously in this simulation. From equations (6)

[31] Pre-2008 EPA fuel economy ratings are widely recognized to overstate the actual mileage achieved by the average driver. This affects our subsequent analysis because the \$1.04 gas tax was derived from actual fuel economy rather than the EPA's forecast fuel economy. We thus rescale the EPA ratings so that the average fuel economy in this sample matches the average fuel economy observed nationwide (17.8 mpg), after adjusting for weight differences between the two samples. The rescaling factor that achieves this equivalence is 0.73. Our conclusions in the subsequent analysis are unchanged if we instead leave the EPA ratings untouched and recalculate the gas tax using EPA mileage ratings – in both cases the per mile gas tax closely tracks the weight based mileage tax.

and (7) we see that the gas tax paid per mile decreases in fuel efficiency and the per mile weight tax increases in weight, as expected. However, an interesting relationship emerges for the difference between the two pricing tools, given in equation (8). $\Delta_i$ is decreasing in better fuel efficiency and higher weight. A negative $\Delta_i$ means that for vehicle model *i*, the weight tax is higher than the gas tax. Cars most heavily advantaged by the gas tax are therefore heavy fuel efficient vehicles (e.g. Mercedes E320 CDI [35.5 mpg; 3835 lbs]). Cars most heavily advantaged by the weight tax are gas guzzling lighter vehicles (e.g. Ford GT [18.7 mpg; 3351 lbs]).

Figure 4 presents a scatter plot of the gas tax versus the weight tax for all models from 1997–2006 in the cleaned Knittel (forthcoming) database. The correspondence between the two taxes is quite close. Some models lie on the 45 degree line (e.g. 2002 VW Jetta [31.6 mpg; 2893 lbs] and the 2003 Ford Windstar Wagon [23.0 mpg; 3762 lbs]). For most models the difference between the two taxes is small, but it can be significant at the extremes, ranging from –4.4 cents to 5.1 cents per mile. A one cent difference per mile equates to $110 dollars on an annual basis. For 71% of the models in our database the absolute value of the difference between the two taxes is less than one cent per mile, and for 97% of the models the absolute value of the difference is less than 2 cents per mile. The average difference between the two taxes is 0.76 cents per mile, which represents 13.5% of the average value of the per mile weight tax.

A related question is what level of gas tax would best mimic the weight tax across all models. For any gas tax set at $\overline{c}$ per gallon, we can write the following equation:

$$c_i^e = \overline{c} \cdot {}^{1}\!/_{mpg_i} + \varepsilon_i \qquad (9)$$

In this equation, $\bar{c} \cdot {}^{1}\!/_{mpg_i}$ equals the gas tax per mile for vehicle $i$ (dollars/gallon divided by miles/gallon = dollars/mile), and $\varepsilon_i$ equals the difference between the gas tax per mile and the weight based mileage tax. If we specify a quadratic loss function we can estimate the gas tax that best mimics the weight tax via a least squares regression of the weight based mileage tax on the inverse of miles per gallon.[32] Each observation in this regression is a different vehicle model. The resulting estimate from this OLS regression is $\overline{c} = \$0.99$ per gallon, which is close to our previously computed tax of $1.04 in magnitude

[32] Equation (9) reveals that we must constrain the intercept to be zero in this regression – the only degree of freedom in setting the gas tax comes from choosing $\overline{c}$.

(though the two quantities are significantly different at the 5% level).[33] The total revenue raised by either of the gas taxes is close to the total revenue raised by the mileage based tax. In all cases the revenues could be redistributed to make the taxes revenue neutral.

While many countries charge high gasoline taxes in part to encourage fuel efficiency, the United States encourages fuel efficiency through CAFE standards. Though CAFE standards represent a de facto tax on weight, they are insufficient to internalize the externality presented in this paper. Goldberg (1998) estimates that CAFE increases the price of pickup trucks by 0.6 percent and reduces the price of subcompacts by 0.5 percent. This equates to a tax on pickup trucks (relative to subcompacts) of approximately $200. The gasoline tax discussed above, however, equates to tax on pickup trucks (relative to subcompacts) of over $4,000 over the life of the vehicle.

## 7. CONCLUSION

The US vehicle fleet has become significantly heavier over the past two decades. The average car on the road in 2008 was roughly 530 pounds heavier than the average car on the road in 1988, representing a 20 percent increase. This trend has been widely discussed by policymakers when contemplating more stringent fuel economy standards or greenhouse gas emissions standards. However, it is less widely recognized that an unregulated vehicle fleet is inefficiently heavy due to the "arms race" nature of vehicle choice. In this paper we estimate the external effects of choosing a heavier vehicle on fatalities in two-vehicle collisions. We present robust evidence that increasing striking vehicle weight by 1,000 pounds increases the probability of a fatality in the struck vehicle by 40 to 50 percent. This finding is unchanged across different specifications, estimation methods, and for different subsets of the sample. We show that there are also significant impacts on serious injuries.

The external costs of fatalities are currently not internalized in the form of a first- or second-best policy. We calculate that the second-best gasoline tax that internalizes the fleet weight gain since 1989 is 28 cents per gallon. We further calculate that internalizing the total cost of external fatalities due to vehicle weight and operation, including crashes with

[33] The least squares regression minimizes the sum of squared deviations. We can alternatively estimate a median regression to minimize the sum of absolute deviations. The median regression coefficient is $0.98, which is virtually identical to the OLS coefficient of $0.99.

motorcycles and pedestrians, requires a tax on the order of $1.04 per gallon. Parry and Small (2005), applying a lower VSL to monetize other external costs and not accounting for the vehicle weight externality, calculate an optimal value of $1.01 per gallon for the US gas tax (approximately 60 cents above its current level). Internalizing the vehicle weight externality, however, would increase this optimal value substantially.

**References**

Ahmad, S., and D. L. Greene. 2005. "Effect of Fuel Economy on Automobile Safety: A Reexamination." *Journal of the Transportation Research Board* (1941): 1–7.

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

Austin, D., and T. Dinan. 2005. "Clearing the air: The costs and consequences of higher CAFE standards and increased gasoline taxes." *Journal of Environmental Economics and Management* 50(3): 562–582.

Bento, A. M et al. 2009. "Distributional and efficiency impacts of increased US gasoline taxes." *The American Economic Review* 99(3): 667–699.

Bordoff, J., and P. Noel. 2008. "Pay-As-You-Drive auto insurance: A simple way to reduce driving-related harms and increase equity." *Hamilton Project Discussion Paper*.

Busse, M. R, C. R Knittel, and F. Zettelmeyer. 2009. "Pain at the Pump: How Gasoline Prices Affect Automobile Purchasing in New and Used Markets."

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. (Accessed March 19, 2010).

Evans, L. 1984. "Accident involvement rate and car size." *Accident Analysis & Prevention* 16(5): 387–406.

Fernández-Val, I. 2009. "Fixed effects estimation of structural parameters and marginal effects in panel probit models." *Journal of Econometrics* 150(1): 71–85.

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.

Goldberg, Pinelopi Koujianou. 1998. "The Effects of the Corporate Average Fuel Efficiency Standards in the US." *The Journal of Industrial Economics* 46(1): 1-33. (Accessed March 19, 2010).

Hakim, Danny. 2003. "S.U.V.'s to Be Redesigned to Reduce Risk to Cars." *The New York Times*. http://www.nytimes.com/2003/12/04/business/suv-s-to-be-redesigned-to-reduce-risk-to-cars.html?pagewanted=all (Accessed March 19, 2010).

"Insurance Information Institute - Compulsory Auto/Uninsured Motorists." http://www.iii.org/IU/Compulsory-Auto-Uninsured-Motorists/ (Accessed March 19, 2010).

Jacobsen, M. 2010. "Fuel Economy, Car Class Mix, and Safety." *University of California San Diego Working Paper*.

Kahane, C. 2003. "Vehicle weight, fatality risk and crash compatibility of model year 1991-99 passenger cars and light trucks." *NHTSA Technical Report*.

Kennickell, A. B. 2003. "A Rolling Tide: Changes in the Distribution of Wealth in the US." *Federal Reserve Occasional Staff Study*.

Khazzoom, J. D. 1994. "Fuel Efficiency and Automobile Safety: Single-Vehicle Highway Fatalities for Passenger Cars." *The Energy Journal* 15(4): 49–102.

Kleit, A. N. 2004. "Impacts of long-range increases in the corporate average fuel economy (CAFE) standard." *Economic Inquiry* 42(2): 279–94.

Klier, T., and J. Linn. 2008. "New Vehicle Characteristics and the Cost of the Corporate Average Fuel Economy Standard."

Knittel, C. R. Forthcoming. "Automobiles on Steroids: Product Attribute Trade-Offs and Technological Progress in the Automobile Sector." *American Economic Review*.

Li, S. 2009. "Traffic Safety and Vehicle Choice." *RFF Discussion Paper*.

Li, S., C. Timmins, and R. H von Haefen. 2009. "How do gasoline prices affect fleet fuel economy?" *American Economic Journal: Economic Policy* 1(2): 113–137.

Mertz, H. J., P. Prasad, and A. L. Irwin. 1997. "Injury risk curves for children and adults in frontal and rear collisions (SAE 973318)." In *Proceeding of the 41st Stapp Car Crash Conference*, , p. 13–30.

National Highway Traffic Safety Administration. "FARS Encyclopedia." http://www-fars.nhtsa.dot.gov/Main/index.aspx (Accessed March 20, 2010).

———. 2007. "Traffic Safety Facts 2005."

Noland, R. B. 2005. "Fuel economy and traffic fatalities: multivariate analysis of international data." *Energy Policy* 33(17): 2183–2190.

———. 2004. "Motor Vehicle Fuel Efficiency and Traffic Fatalities." *The Energy Journal* 25(4): 1–22.

Ogden, Cynthia L et al. 2004. "Mean body weight, height, and body mass index, United States 1960-2002." *Advance Data* (347): 1-17. (Accessed February 10, 2011).

Parry, Ian W. H., and Kenneth A. Small. 2005. "Does Britain or the United States Have the Right Gasoline Tax?" *The American Economic Review* 95(4): 1276-1289. (Accessed March 19, 2010).

Peltzman, Sam. 1975. "The Effects of Automobile Safety Regulation." *The Journal of Political Economy* 83(4): 677-725. (Accessed March 19, 2010).

Portney, Paul R., Ian W. H. Parry, and Howard K. Gruenspecht, Winston Harrington.

Autumn. "Policy Watch: The Economics of Fuel Economy Standards." *The Journal of Economic Perspectives* 17(4): 203-217. (Accessed March 19, 2010).

Robertson, L. S. 1991. "How to save fuel and reduce injuries in automobiles." *The Journal of Trauma* 31(1): 107.

United States Environmental Protection Agency. 2009. "Light-Duty Automotive Technology, Carbon Dioxide Emissions, and Fuel Economy Trends: 1975 Through 2009."

US Energy Information Administration. "Use of Gasoline - Energy Explained, Your Guide To Understanding Energy." http://tonto.eia.doe.gov/energyexplained/index.cfm?page=gasoline_use (Accessed March 20, 2010).

Van Auken, M., and J. W Zellner. 2005. "An Assessment of the Effects of Vehicle Weight and Size on Fatality Risk in 1985 to 1998 Model Year Passenger Cars and 1985 to 1997 Model Year Light Trucks and Vans." *SAE International*.

Ward's Reports Incorporated. 2005. *Ward's Automotive Yearbook*.

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: 333-355.

Figure 1: External Costs of Vehicle Weight

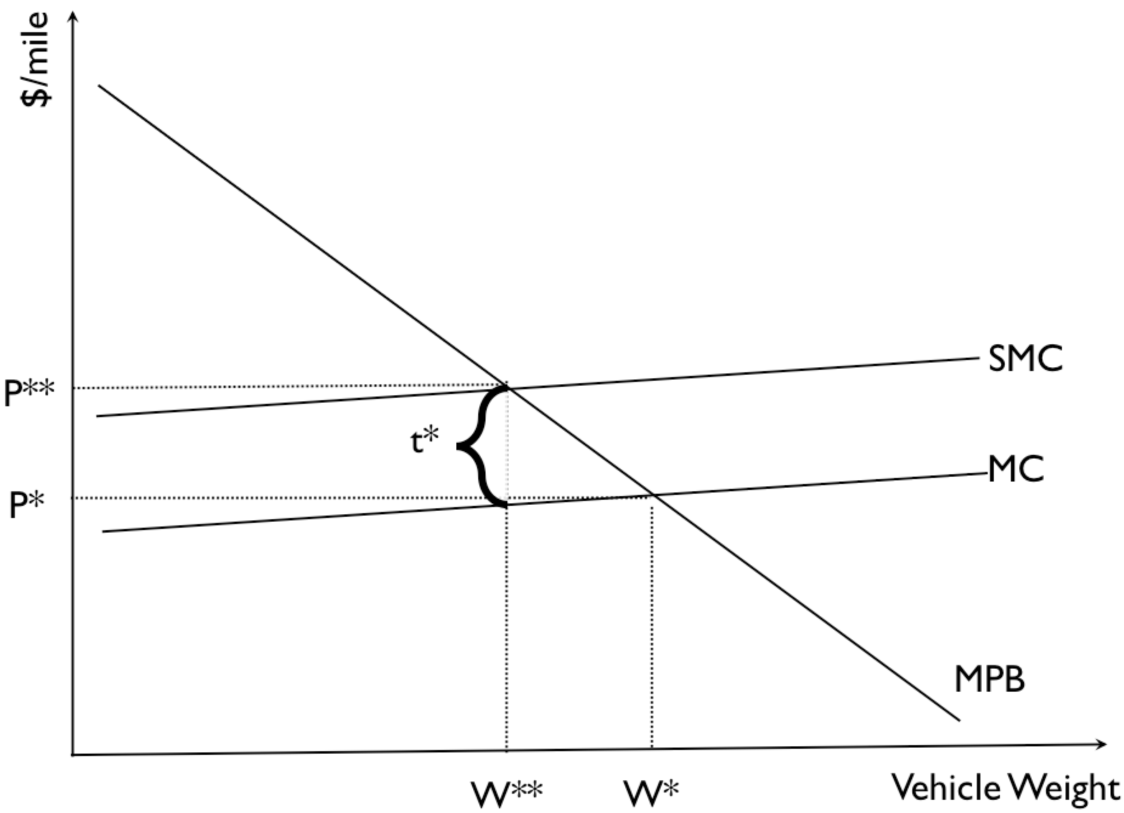


Figure 2: Marginal Effects of Striking Vehicle Weight

Figure 3: Fuel Economy vs. Weight for 2005 Model Year Light Vehicles

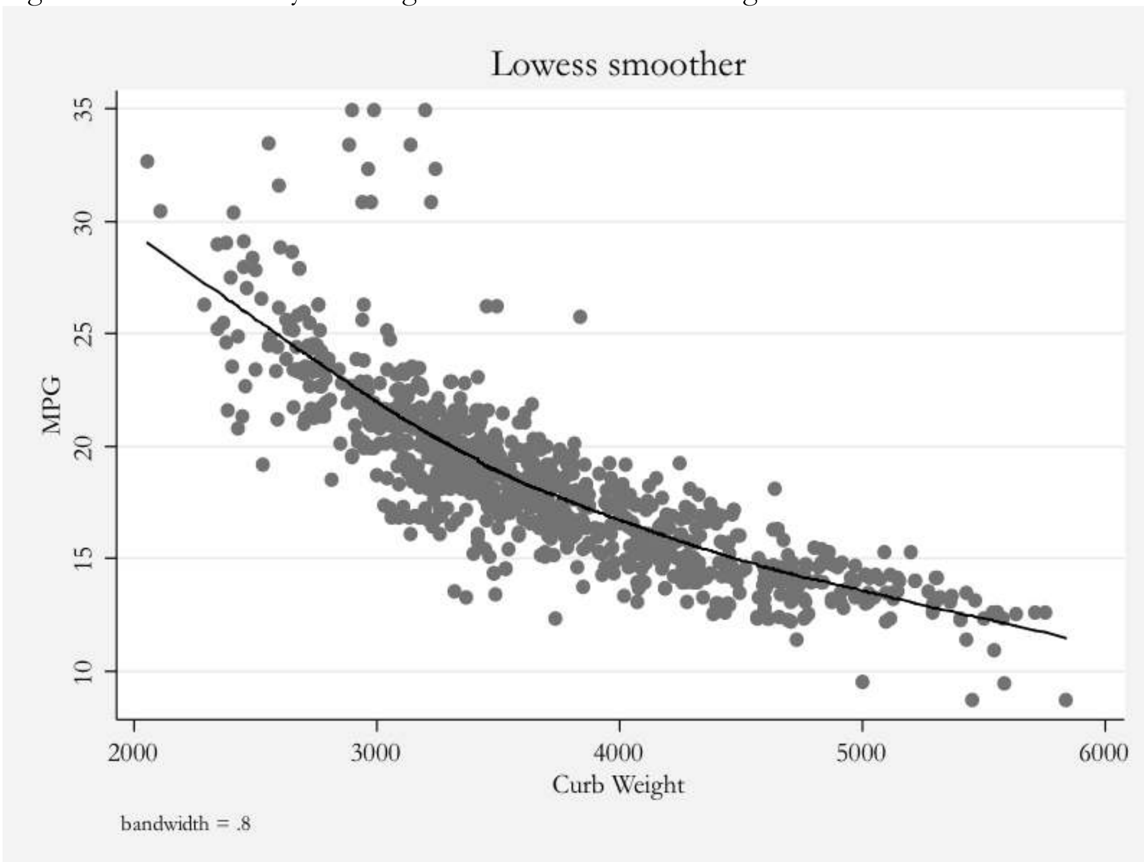
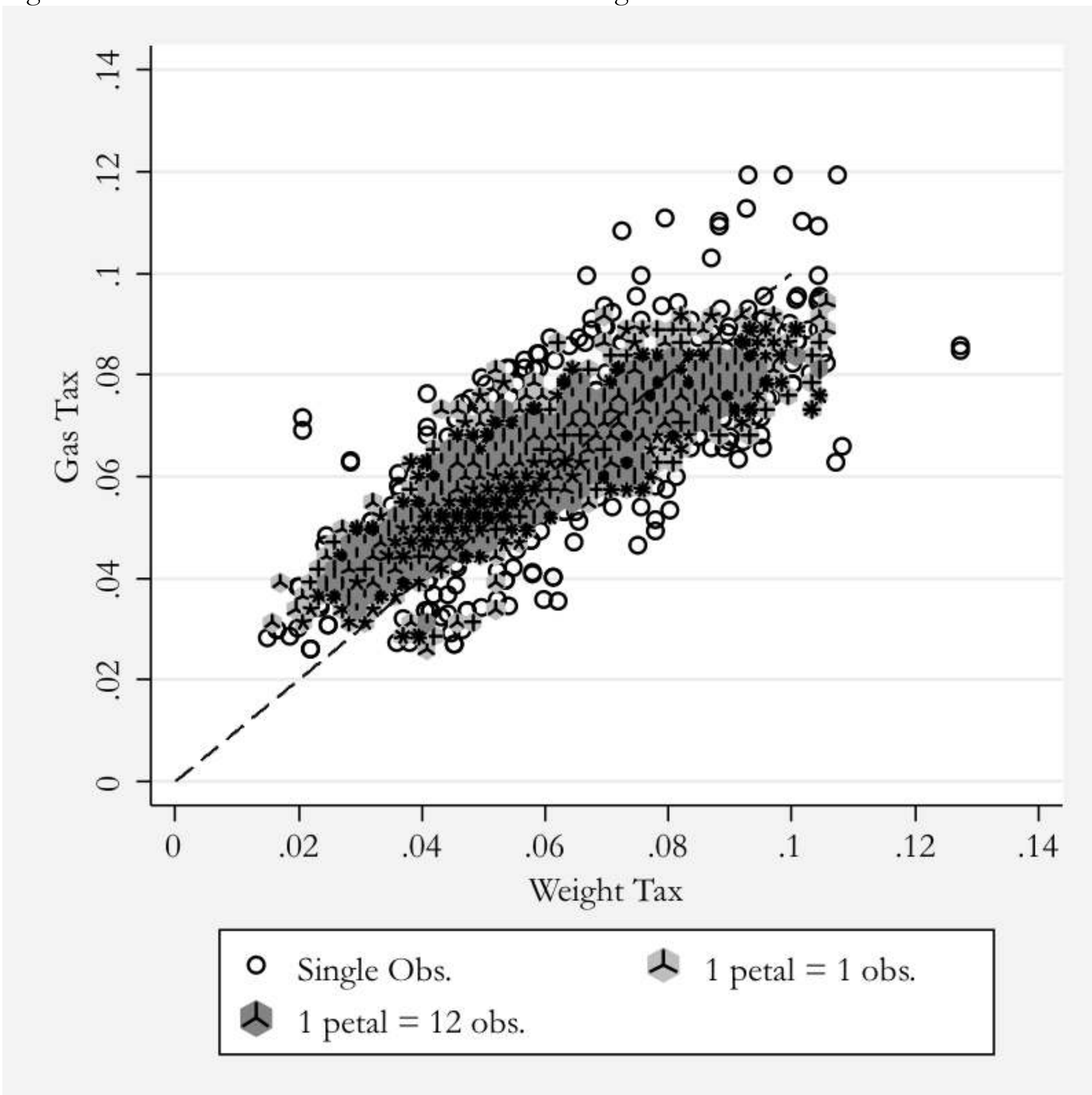


Figure 4: Sunflower Scatter Plot of Gas Tax vs. Weight Tax for Cars and Trucks




**Notes:** The graph above displays the joint distribution of the weight tax and gas tax per mile for the sample of cars and trucks with model years 1997-2006 from the database provided by Knittel (forthcoming). We remove boutique cars, flex fuel vehicles, and a few outliers with incorrectly recorded fuel ratings. The sunflower plot bunches multiple observations into single flowers, where the number of petals indicates the total number of observations represented by the flower. The petals of light flowers represent one observation each and the petals of darker flowers represent 12 observations each.

Table 1: Summary Statistics for Two-Vehicle Collision Data Set

| | Overall Sample | | Estimation Sample | |
|---|---|---|---|---|
| | Mean (Std Dev) | Sample Size | Mean (Std Dev) | Sample Size |
| Weight | 3,051 lbs (683) | 6,540,582 | 3,059 lbs (684) | 4,766,645 |
| Light Truck | 17.5% (38.0) | 9,684,978 | 16.5% (37.1) | 4,766,645 |
| Model Year | 1992 (6.5) | 8,295,450 | 1993 (5.6) | 3,822,427 |
| Accident Year | 1998 (4.7) | 9,684,978 | 1998 (4.4) | 4,766,645 |
| Occupants | 1.44 (1.21) | 4,662,294 | 1.41 (0.84) | 2,572,454 |
| Fatality | 0.19% (4.40) | 9,684,978 | 0.19% (4.35) | 4,766,645 |
| Serious Injury | 3.0% (17.1) | 9,684,978 | 2.7% (16.1) | 4,766,645 |
| Alcohol Involved | 8.9% (28.5) | 4,785,153 | 8.3% (27.6) | 2,710,906 |

*Notes:* Both samples are limited to collisions involving two light vehicles built post-1960. The estimation sample is further limited to collisions in which vehicle weight is non-missing for both vehicles.

Table 2: Effect of Vehicle Weight on Fatalities

Dependent Variable: Presence of Fatalities in Struck Vehicle

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) |
|---|---|---|---|---|---|---|---|---|---|
| Weight of Striking Vehicle (1000s of lbs) | 0.00100 | 0.14634 | 0.00101 | 0.14041 | 0.00113 | 0.14375 | 0.00119 | 0.00064 | 0.00063 |
| | (0.00004) | (0.00500) | (0.00005) | (0.00566) | (0.00005) | (0.00639) | (0.00006) | (0.00007) | (0.00007) |
| Effect of 1000 lb Increase in Striking Vehicle Weight/ | 0.00100 | 0.00089 | 0.00101 | 0.00094 | 0.00113 | 0.00098 | 0.00119 | 0.00064 | 0.00063 |
| Percent Increase Over Sample Mean | 53% | 47% | 46% | 42% | 46% | 39% | 49% | 45% | 44% |
| Weight of Struck Vehicle (1000s of lbs) | -0.00044 | -0.07781 | -0.00052 | -0.08474 | -0.00092 | -0.14676 | -0.00085 | -0.00056 | -0.00062 |
| | (0.00003) | (0.00588) | (0.00004) | (0.00650) | (0.00005) | (0.00756) | (0.00005) | (0.00006) | (0.00006) |
| Striking Vehicle is Light Truck | 0.00096 | 0.12034 | 0.00093 | 0.12181 | 0.00075 | 0.07915 | 0.00075 | 0.00065 | 0.00065 |
| | (0.00007) | (0.00904) | (0.00008) | (0.00973) | (0.00010) | (0.01099) | (0.00010) | (0.00011) | (0.00011) |
| Struck Vehicle is Light Truck | -0.00030 | -0.05527 | -0.00039 | -0.07862 | -0.00032 | -0.08661 | -0.00004 | -0.00020 | -0.00014 |
| | (0.00006) | (0.01065) | (0.00007) | (0.01139) | (0.00008) | (0.01279) | (0.00008) | (0.00009) | (0.00009) |
| Specification | OLS | Probit | OLS | Probit | OLS | Probit | OLS | OLS | OLS |
| Weather, Time, and County Fixed Effects | | | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Driver Characteristics | | | | | Yes | Yes | Yes | Yes | Yes |
| City Fixed Effects | | | | | | | Yes | Yes | Yes |
| Occupants and Seat Belt Usage | | | | | | | | | Yes |
| Sample Size | 4,766,645 | 4,766,645 | 3,615,381 | 3,573,406 | 3,038,122 | 3,000,738 | 2,639,086 | 1,313,542 | 1,313,542 |

*Notes:* Each column represents a separate regression. The estimation sample is limited to collisions involving two vehicles. Parentheses contain standard errors clustered at the collision level. Effects of a 1,000 lb increase in striking vehicle weight are computed as the average effect of a 1,000 lb increase in weight across all observations included in the regression. All regressions include as right-hand-side variables the weight of each vehicle, indicators for whether each vehicle is a light truck, and year fixed effects. Weather, time, and county fixed effects controls include rain, darkness, day of week (weekday versus weekend), Interstate highway, a quadratic in model year for each vehicle, and year, hour, and county fixed effects. Driver characteristic controls include quadratics in driver age, indicators for drivers under 21 or over 60, indicators for male drivers and young male drivers, and indicators for any seat belt usage in the vehicle. The results indicate that there is a strong positive relationship between striking vehicle weight and struck vehicle fatalities and that this relationship is not sensitive to the inclusion of a large set of controls.

Table 3: Effect of Vehicle Weight on Serious Injuries

Dependent Variable: Presence of Serious Injuries in Struck Vehicle

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) |
|---|---|---|---|---|---|---|---|---|---|
| Weight of Striking Vehicle (1000s of lbs) | 0.00588 | 0.09248 | 0.00566 | 0.08648 | 0.00637 | 0.08734 | 0.00709 | 0.00311 | 0.00303 |
| | (0.00013) | (0.00194) | (0.00015) | (0.00221) | (0.00018) | (0.00241) | (0.00020) | (0.00019) | (0.00019) |
| Effect of 1000 lb Increase in Striking Vehicle Weight/ | 0.00588 | 0.00568 | 0.00566 | 0.00550 | 0.00637 | 0.00595 | 0.00709 | 0.00311 | 0.00303 |
| Percent Increase Over Sample Mean | 22% | 21% | 19% | 18% | 19% | 18% | 20% | 19% | 19% |
| Weight of Struck Vehicle (1000s of lbs) | -0.00613 | -0.10688 | -0.00698 | -0.11768 | -0.00844 | -0.13029 | -0.00810 | -0.00419 | -0.00459 |
| | (0.00012) | (0.00210) | (0.00014) | (0.00237) | (0.00017) | (0.00260) | (0.00018) | (0.00018) | (0.00018) |
| Striking Vehicle is Light Truck | 0.00144 | 0.02390 | 0.00398 | 0.07006 | 0.00367 | 0.05397 | 0.00415 | 0.00273 | 0.00274 |
| | (0.00023) | (0.00354) | (0.00027) | (0.00388) | (0.00032) | (0.00420) | (0.00034) | (0.00034) | (0.00034) |
| Struck Vehicle is Light Truck | -0.00501 | -0.09981 | -0.00305 | -0.07783 | -0.00254 | -0.06572 | -0.00134 | -0.00149 | -0.00105 |
| | (0.00020) | (0.00389) | (0.00023) | (0.00426) | (0.00027) | (0.00462) | (0.00030) | (0.00030) | (0.00030) |
| Specification | OLS | Probit | OLS | Probit | OLS | Probit | OLS | OLS | OLS |
| Weather, Time, and County Fixed Effects | | | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Driver Characteristics | | | | | Yes | Yes | Yes | Yes | Yes |
| City Fixed Effects | | | | | | | Yes | Yes | Yes |
| Occupants and Seat Belt Usage | | | | | | | | | Yes |
| Sample Size | 4,766,645 | 4,766,645 | 3,615,381 | 3,613,483 | 3,038,122 | 3,036,352 | 2,639,086 | 1,313,542 | 1,313,542 |

*Notes:* Each column represents a separate regression. The estimation sample is limited to collisions involving two vehicles. Parentheses contain standard errors clustered at the collision level. Effects of a 1,000 lb increase in striking vehicle weight are computed as the average effect of a 1,000 lb increase in weight across all observations included in the regression. All regressions include as right-hand-side variables the weight of each vehicle, indicators for whether each vehicle is a light truck, and year fixed effects. Weather, time, and county fixed effects controls include rain, darkness, day of week (weekday versus weekend), Interstate highway, a quadratic in model year for each vehicle, and year, hour, and county fixed effects. Driver characteristic controls include quadratics in driver age, indicators for drivers under 21 or over 60, indicators for male drivers and young male drivers, and indicators for any seat belt usage in the vehicle. The results indicate that there is a strong positive relationship between striking vehicle weight and struck vehicle injuries and that this relationship is not sensitive to the inclusion of a large set of controls.

Table 4: Heterogeneous Effects of Vehicle Weight on Fatalities

Dependent Variable: Presence of Fatalities in Struck Vehicle

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Weight of Striking Vehicle (1000s of lbs) | 0.00120 | 0.00128 | 0.17045 | 0.38894 |
| | (0.00006) | (0.00031) | (0.00794) | (0.03962) |
| Effect of 1000 lb Increase in Striking Vehicle Weight/ | 0.00114 | 0.00114 | 0.00103 | 0.00106 |
| Percent Increase Over Sample Mean | 47% | 46% | 41% | 43% |
| Weight of Striking Vehicle | -0.00010 | | -0.07322 | |
| *Striking Vehicle is Light Truck | (0.00012) | | (0.01215) | |
| Weight of Striking Vehicle | -0.00024 | | 0.00861 | |
| *Struck Vehicle is Light Truck | (0.00012) | | (0.01522) | |
| Weight of Striking Vehicle Squared | | -0.00002 | -0.00002 | -0.03435 |
| | | (0.00004) | (0.00004) | (0.00643) |
| Specification | OLS | OLS | Probit | Probit |
| Sample Size | 3,038,122 | 3,038,122 | 3,000,738 | 3,000,738 |

*Notes:* Each column represents a separate regression. The estimation sample is limited to collisions involving two vehicles. Parentheses contain standard errors clustered at the collision level. Effects of a 1,000 lb increase in striking vehicle weight are computed as the average effect of a 1,000 lb increase in weight across all observations included in the regression. All regressions include the following right-hand-side variables: weight of each vehicle, a quadratic in model year for each vehicle, indicators for whether each vehicle is a light truck, rain, darkness, day of week (weekday versus weekend), Interstate highway, quadratics in driver age, indicators for drivers under 21 or over 60, indicators for male drivers and young male drivers, indicators for any seat belt usage in the vehicle, and year, hour, and city fixed effects. The results indicate that the effects of vehicle weight on fatalities do not vary strongly with striking vehicle weight or body type.

Table 5: Effect of Vehicle Weight on Fatalities in Three-Vehicle Accidents

Dependent Variable: Presence of Fatalities in Struck Vehicle

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) |
|---|---|---|---|---|---|---|---|---|---|
| Weight of Striking Vehicle (1000s of lbs) | 0.00092 | 0.12408 | 0.00081 | 0.10791 | 0.00083 | 0.10401 | 0.00087 | 0.00057 | 0.00055 |
| | (0.00009) | (0.01097) | (0.00011) | (0.01228) | (0.00012) | (0.01440) | (0.00013) | (0.00018) | (0.00018) |
| Effect of 1000 lb Increase in Striking Vehicle Weight/ | 0.00092 | 0.00082 | 0.00081 | 0.00083 | 0.00083 | 0.00080 | 0.00087 | 0.00057 | 0.00055 |
| Percent Increase Over Sample Mean | 44% | 39% | 34% | 31% | 32% | 27% | 34% | 33% | 32% |
| Weight of Struck Vehicle (1000s of lbs) | -0.00056 | -0.08773 | -0.00069 | -0.10660 | -0.00108 | -0.17143 | -0.00087 | -0.00052 | -0.00056 |
| | (0.00010) | (0.01648) | (0.00012) | (0.01844) | (0.00015) | (0.02203) | (0.00015) | (0.00021) | (0.00021) |
| Striking Vehicle is Light Truck | 0.00060 | 0.07090 | 0.00059 | 0.08169 | 0.00064 | 0.07823 | 0.00062 | 0.00051 | 0.00052 |
| | (0.00016) | (0.01926) | (0.00018) | (0.02075) | (0.00022) | (0.02425) | (0.00023) | (0.00031) | (0.00031) |
| Struck Vehicle is Light Truck | -0.00044 | -0.07320 | -0.00045 | -0.08538 | -0.00012 | -0.04690 | 0.00020 | -0.00048 | -0.00046 |
| | (0.00019) | (0.03270) | (0.00022) | (0.03466) | (0.00027) | (0.03975) | (0.00028) | (0.00035) | (0.00035) |
| Specification | OLS | Probit | OLS | Probit | OLS | Probit | OLS | OLS | OLS |
| Weather, Time, and County Fixed Effects | | | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Driver Characteristics | | | | | Yes | Yes | Yes | Yes | Yes |
| City Fixed Effects | | | | | | | Yes | Yes | Yes |
| Occupants and Seat Belt Usage | | | | | | | | | Yes |
| Sample Size | 510,777 | 510,777 | 404,797 | 363,923 | 332,046 | 298,872 | 301,380 | 99,366 | 99,366 |

*Notes:* Each column represents a separate regression. The estimation sample is limited to collisions involving three vehicles. Striking vehicle weight coefficients represent the average effect of increasing the weight of one striking vehicle by 1,000 pounds; they are the average of the coefficients on the first and second striking vehicles. Parentheses contain standard errors clustered at the collision level. Effects of a 1,000 lb increase in striking vehicle weight are computed as the average effect of a 1,000 lb increase in the weight of one striking vehicle across all observations included in the regression. All regressions include as right-hand-side variables the weight of each vehicle, indicators for whether each vehicle is a light truck, and year fixed effects. Weather, time, and county fixed effects controls include rain, darkness, day of week (weekday versus weekend), Interstate highway, a quadratic in model year for each vehicle, and year, hour, and county fixed effects. Driver characteristic controls include quadratics in driver age, indicators for drivers under 21 or over 60, indicators for male drivers and young male drivers, and indicators for any seat belt usage in the vehicle. The results indicate that there is a strong positive relationship between striking vehicle weight and struck vehicle fatalities and that this relationship is not sensitive to the inclusion of a large set of controls.

Table 6: Relationship Between Vehicle Weight and NHTSA Crash Test Performance

| Dependent Variable: | HIC | HIC>700 | HIC | HIC>700 |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| Weight of Vehicle | 15.8 | 0.016 | 38.2 | 0.018 |
| | (18.3) | (0.019) | (43.5) | (0.040) |
| Percentage Effect of 1,000 lb Increase | 2.6% | 5.6% | 6.7% | 7.2% |
| Sales Share Weighted | | | Yes | Yes |
| Sample Size | 5,003 | 5,003 | 2,847 | 2,847 |

*Notes:* Each column represents a separate regression. The estimation sample in the first two columns contains all NHTSA vehicle-to-barrier frontal crash test results. The estimation sample in the last two columns contains only crash tests involving vehicles for which we have sales share data. Parentheses contain standard errors clustered by NHTSA crash test. All regressions include the following right-hand-side variables: weight of tested vehicle, a quadratic in model year, a light truck indicator, and a quadratic in collision speed. Sales share weighted regressions are weighted by the tested vehicle's sales share for a given year. The results indicate that heavier vehicles score slightly worse in NHTSA crash test results, but that the relationship is small and statistically insignificant.

Table 7: Effect of Vehicle Weight in Collisions Between Two Equal Weight Vehicles

Dependent Variable: Presence of Fatalities in Struck Vehicle

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Average Vehicle Weight in Collision (1000s of lbs) | 0.00015 | -0.00085 | 0.00020 | -0.00128 |
| | (0.00017) | (0.00104) | (0.00023) | (0.00133) |
| Effect of 1000 lb Increase in Average Weight/ | 0.00015 | -0.00085 | 0.00020 | -0.00128 |
| Percent Increase Over Sample Mean | 7% | -14% | 9% | -24% |
| Max Weight Difference Between Vehicles | 200 lbs | 200 lbs | 100 lbs | 100 lbs |
| Frontal Collisions Only | | Yes | | Yes |
| Sample Size | 491,580 | 38,380 | 263,876 | 20,038 |

*Notes:* Each column represents a separate regression. The estimation sample is limited to collisions in which the difference in weight between the two vehicles is less than 200 lbs (first two columns) or 100 lbs (last two columns). Parentheses contain standard errors clustered at the collision level. Effects of a 1,000 lb increase in striking vehicle weight are computed as the average effect of a 1,000 lb increase in weight across all observations included in the regression. All regressions include the following right-hand-side variables: weight of each vehicle, a quadratic in model year for each vehicle, indicators for whether each vehicle is a light truck, rain, darkness, day of week (weekday versus weekend), Interstate highway, quadratics in driver age, indicators for drivers under 21 or over 60, indicators for male drivers and young male drivers, indicators for any seat belt usage in the vehicle, and year, hour, and city fixed effects. The results indicate that there is a small, statistically insignificant relationship between vehicle weight and fatalities in collisions between two equal weight vehicles.

Table 8: Effect of Vehicle Weight on Fatalities in Single-Vehicle Collisions

Dependent Variable: Presence of Fatalities in Struck Vehicle

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Weight of Vehicle | 0.00093 | 0.00056 | 0.01345 | 0.01330 |
| (1000s of lbs) | (0.00026) | (0.00047) | (0.00626) | (0.01298) |
| Effect of 1000 lb Increase in Vehicle Weight/ | 0.00093 | 0.00056 | 0.00045 | 0.00044 |
| Percent Increase Over Sample Mean | 6% | 4% | 3% | 3% |
| Collision Type | 1 Vehicle | 1 Veh, Frontal | 1 Vehicle | 1 Veh, Frontal |
| Specification | OLS | OLS | Probit | Probit |
| Sample Size | 683,430 | 214,857 | 863,819 | 212,571 |

*Notes:* Each column represents a separate regression. The estimation sample is limited to collisions involving a single vehicle. Parentheses contain robust standard errors. Effects of a 1,000 lb increase in vehicle weight are computed as the average effect of a 1,000 lb increase in weight across all observations included in the regression. All regressions include the following right-hand-side variables: weight of vehicle, a quadratic in model year, indicators for whether a vehicle is a light truck, rain, darkness, day of week (weekday versus weekend), Interstate highway, quadratic in driver age, indicators for drivers under 21 or over 60, indicators for male drivers and young male drivers, indicators for any seat belt usage in the vehicle, and year, hour, and either city fixed effects (OLS) or county fixed effects (probit). The results indicate that there is a very small positive relationship between vehicle weight and fatalities in single-vehicle collisions.

Table 9: Effect of Vehicle Weight on Fatalities Using Alternative Sources of Identification

Dependent Variable: Presence of Fatalities in Struck Vehicle

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Weight of Striking Vehicle/ | 0.00078 | 0.00080 | 0.00064 | 0.00078 |
| Additional Weight in Striking Vehicle (1000s of lbs) | (0.00009) | (0.00017) | (0.00007) | (0.00027) |
| Effect of 1000 lb Increase in Striking Vehicle Weight/ | 0.00078 | 0.00080 | 0.00064 | 0.00078 |
| Percent Increase Over Sample Mean | 50% | 52% | 45% | 54% |
| Specification | OLS | OLS w/Model FEs | OLS | IV |
| Sample Size | 892,970 | 892,970 | 1,318,012 | 1,318,012 |

*Notes:* Each column represents a separate regression. The estimation sample is limited to collisions involving two vehicles. Parentheses contain standard errors clustered at the collision level. Effects of a 1,000 lb increase in striking vehicle weight are computed as the average effect of a 1,000 lb increase in weight across all observations included in the regression. All regressions include the following right-hand-side variables: weight of each vehicle, a quadratic in model year for each vehicle, indicators for whether each vehicle is a light truck, rain, darkness, day of week (weekday versus weekend), Interstate highway, quadratics in driver age, indicators for drivers under 21 or over 60, indicators for male drivers and young male drivers, indicators for any seat belt usage in the vehicle, and year and hour fixed effects. OLS regressions with model fixed effects contain fixed effects for each vehicle model and county fixed effects. IV regressions contain city fixed effects and use the number of occupants in the striking vehicle times 164 lbs per occupant as the instrument for additional weight in the striking vehicle weight. The results indicate that adding model fixed effects or instrumenting for vehicle weight does not change the estimated effects of vehicle weight on fatalities.

Table 10: Valuing the Traffic Fatality Externality

| | Two Vehicle Accidents | | | Two and Three Vehicle Accidents | | |
|---|---|---|---|---|---|---|
| | | iQ Baseline | iQ Baseline + Peds/Mot | Weight Gain Baseline | iQ Baseline | iQ Baseline + Peds/Mot |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Delta | 0.12% | 0.12% | 0.12% | 0.12% | 0.12% | 0.12% |
| P(Death\|Accident) | 0.23% | 0.23% | 0.23% | 0.23% | 0.23% | 0.23% |
| Simulated Vehicle Weight (lbs) | 3,449 | 3,449 | 3,449 | 3,449 | 3,449 | 3,449 |
| Fatalities (2 Vehicle Collisions) | 18,694 | 18,694 | 18,694 | 18,694 | 18,694 | 18,694 |
| Total Fatalities | 18,694 | 18,694 | 18,694 | 20,483 | 20,483 | 20,483 |
| Total Cost (Billlion US$) | 108 | 108 | 108 | 119 | 119 | 119 |
| Baseline Vehicle Weight (lbs) | 2,824 | 1,896 | 1,896 | 2,824 | 1,896 | 1,896 |
| Counterfactual Fatality Probability in Struck Vehicle | 0.15% | 0.04% | 0.04% | 0.15% | 0.04% | 0.04% |
| Sum of Expected Fatalities Across All Vehicles | 6,135 | 15,255 | 15,255 | 6,722 | 16,715 | 16,715 |
| External Costs Across All Vehicles (Billion US$) | 36 | 88 | 138 | 39 | 97 | 146 |
| External Cost per Gallon (US$/gallon) | 0.25 | 0.63 | 0.98 | 0.28 | 0.69 | 1.04 |
| External Cost per VMT (Cents/Mile) | 0.01 | 0.03 | 0.05 | 0.01 | 0.03 | 0.05 |

Notes: The simulation assumes the DOT VSL of $5.8 million. The delta coefficient is the estimated coefficient form our preferred specification (7) in table 2. The value for the P(Death|Accident) is the mean fatality probability for the main regression sample. The simulated vehicle weight of 3,449 is the average weight of all carsin 2005 in our sample. The total external cost from pedestrian and motorcycle deaths is derived from national FARS data for 2005 and mutiplied by the DOT VSL. The number of gallons of gas we use for the pergallon calculation is 1.40E+11, which is obtained from EIA (2011). The total number of VMT for the per mile calculations is obtained from FARS. Finally, in order to determined the total number of two car accidents, we multiply the share of two car accidents in our sample times the total number of national accidents in order to arrive at a number of US total two car accidents.

Table A1: Effect of Vehicle Weight on Fatalities in Accidents Excluding Light Trucks

Dependent Variable: Presence of Fatalities in Struck Vehicle

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) |
|---|---|---|---|---|---|---|---|---|---|
| Weight of Striking Vehicle (1000s of lbs) | 0.00113 | 0.18790 | 0.00118 | 0.17409 | 0.00129 | 0.18243 | 0.00141 | 0.00070 | 0.00069 |
| | (0.00005) | (0.00643) | (0.00006) | (0.00741) | (0.00007) | (0.00843) | (0.00007) | (0.00008) | (0.00008) |
| Effect of 1000 lb Increase in Striking Vehicle Weight/ | 0.00113 | 0.00104 | 0.00118 | 0.00110 | 0.00129 | 0.00116 | 0.00141 | 0.00070 | 0.00069 |
| Percent Increase Over Sample Mean | 67% | 62% | 58% | 53% | 58% | 51% | 63% | 62% | 61% |
| Weight of Struck Vehicle (1000s of lbs) | -0.00036 | -0.06735 | -0.00052 | -0.08348 | -0.00103 | -0.16845 | -0.00096 | -0.00063 | -0.00068 |
| | (0.00004) | (0.00766) | (0.00005) | (0.00866) | (0.00006) | (0.01030) | (0.00007) | (0.00008) | (0.00008) |
| Specification | OLS | Probit | OLS | Probit | OLS | Probit | OLS | OLS | OLS |
| Weather, Time, and County Fixed Effects | | | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Driver Characteristics | | | | | Yes | Yes | Yes | Yes | Yes |
| City Fixed Effects | | | | | | | Yes | Yes | Yes |
| Occupants and Seat Belt Usage | | | | | | | | | Yes |
| Sample Size | 3,374,741 | 3,374,741 | 2,402,157 | 2,338,923 | 2,069,802 | 2,013,910 | 1,792,640 | 816,026 | 816,026 |

*Notes:* Each column represents a separate regression. The estimation sample is limited to collisions involving two cars – collisions involving light trucks are excluded. Parentheses contain standard errors clustered at the collision level. Effects of a 1,000 lb increase in striking vehicle weight are computed as the average effect of a 1,000 lb increase in weight across all observations included in the regression. All regressions include as right-hand-side variables the weight of each vehicle and year fixed effects. Weather, time, and county fixed effects controls include rain, darkness, day of week (weekday versus weekend), Interstate highway, a quadratic in model year for each vehicle, and year, hour, and county fixed effects. Driver characteristic controls include quadratics in driver age, indicators for drivers under 21 or over 60, indicators for male drivers and young male drivers, and indicators for any seat belt usage in the vehicle. The results indicate that there is a strong positive relationship between striking vehicle weight and struck vehicle fatalities in collisions involving two cars and that this relationship is not sensitive to the inclusion of a large set of controls.

Table A2: Effect of Vehicle Weight on Fatalities for States with High and Low Missing Weight Data

Dependent Variable: Presence of Fatalities in Struck Vehicle

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Weight of Striking Vehicle (1000s of lbs) | 0.00042 | 0.00120 | 0.00041 | 0.00129 |
| | (0.00005) | (0.00005) | (0.00005) | (0.00006) |
| Effect of 1000 lb Increase in Striking Vehicle Weight/ | 0.00042 | 0.00120 | 0.00041 | 0.00129 |
| Percent Increase Over Sample Mean | 52% | 49% | 47% | 49% |
| Percent of Accidents with Missing Weight Data | 27% | 57% | 27% | 57% |
| Weather, Time, Driver, and City Controls | | | Yes | Yes |
| Sample Size | 1,592,656 | 3,173,989 | 1,214,544 | 2,286,700 |

*Notes:* Each column represents a separate regression. The estimation sample is limited to collisions involving two vehicles Columns (1) and (3) are estimated using data from states in which a low proportion of observations are missing weight data (Ohio, Washington, and Wyoming). Columns (2) and (4) are estimated using data from states in which a high proportion of observations are missing weight data (Florida, Kansas, Kentucky, Maryland, and Missouri). Parentheses contain standard errors clustered at the collision level. All regressions include as right-hand-side variables the weight of each vehicle, indicators for whether each vehicle is a light truck, and year fixed effects. Weather, time, driver, and city controls include rain, darkness, day of week (weekday versus weekend), Interstate highway, quadratics in driver age, indicators for drivers under 21 or over 60, indicators for male drivers and young male drivers, indicators for any seat belt usage in the vehicle, and year, hour, and city fixed effects. The results indicate that the relationship between striking vehicle weight and fatalities in the struck vehicle is identical (in percentage terms) in states with high and low proportions of accidents that are missing weight data.