NBER WORKING PAPER SERIES

THE EFFECT OF FUEL ECONOMY STANDARDS ON VEHICLE WEIGHT DISPERSION AND ACCIDENT FATALITIES

Antonio Bento
Kenneth Gillingham
Kevin Roth

Working Paper 23340
http://www.nber.org/papers/w23340

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
April 2017

The authors thank John Graham, Ashley Langer, and Josh Linn; the participants at the NE Workshop on Env. Economics, IIOC, and UCE3; and seminars at CU Boulder, Indiana University, Ifo Institute, UC-Irvine, UC-Berkeley, UCSD, Yale, and Wyoming for helpful comments. The authors are grateful to Yiwei Wang and Paige Weber for excellent research assistance and to UCE3 for funding. The authors also thank Don MacKenzie for the CAFE credit data and Pawel Zal from www.automobile-catalog.com for vehicle characteristics data. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

The Effect of Fuel Economy Standards on Vehicle Weight Dispersion and Accident Fatalities
Antonio Bento, Kenneth Gillingham, and Kevin Roth
NBER Working Paper No. 23340
April 2017
JEL No. H23,I18,Q48,Q58,R41

ABSTRACT

The firm response to regulation is seldom as controversial as in the context of fuel economy standards, a dominant policy to reduce emissions from vehicles worldwide. It has long been argued that such standards lead to vehicle weight changes that increase accident fatalities. Using unconditional quantile regression, we are the first to document the effect of the Corporate Average Fuel Economy (CAFE) standards on the vehicle weight distribution. We find that on net CAFE reduced fatalities, with lowered mean weight dominating increased dispersion. When monetized, this effect suggests positive net benefits from CAFE even with no undervaluation of fuel economy.

Antonio Bento
Sol Price School of Public Policy
and Department of Economics
University of Southern California
Los Angeles, CA 90089-0626
and NBER
abento@usc.edu

Kenneth Gillingham
School of Forestry and Environmental Studies
Yale University
195 Prospect Street
New Haven, CT 06511
and NBER
kenneth.gillingham@yale.edu

Kevin Roth
Department of Economics
University of California, Irvine
kroth1@uci.edu

A online appendix is available at http://www.nber.org/data-appendix/w23340

1 Introduction

For decades economists have pushed policy makers towards more flexible regulations to correct market failures. Flexible regulations allow firms to use the lowest-cost channels, including channels unforeseen by regulators (Stavins, 1998; Carlson et al., 2000). Such unanticipated channels may lower the cost, but in some cases can mean that the policy fails to fully correct the market failure (Auffhammer and Kellogg, 2011) or exacerbates other unpriced externalities.

These issues are particularly notable when regulating the automobile industry. Regulation affecting the average fuel economy of light-duty vehicles—the primary approach used in the United States and many other countries to reduce vehicle oil consumption and emissions—has long faced critiques for leaving open a multitude of channels for firms to meet the regulation (Austin and Dinan, 2005; Bento et al., 2009; Jacobsen, 2013a). Among the channels most important to policymakers is the ability of firms to meet the standard by reducing vehicle weight (Knittel, 2011; Klier and Linn, 2012).

It is well established by the 'arms-race' literature that occupants of lighter-weight vehicles face a higher risk of fatality in multi-vehicle crashes (White, 2004; Anderson and Auffhammer, 2014; Li, 2012). If firms decrease vehicle weight as a strategy for increasing fuel economy, drivers of these vehicles may be at an elevated risk of a fatality when struck in an accident. It has long been argued that Corporate Average Fuel Economy (CAFE) standards increase fatalities based on evidence suggesting that occupants of lighter vehicles are at a greater risk of a fatality and that CAFE standards decreased the mean vehicle weight, and thus must have increased fatalities (Crandall and Graham, 1989). Given typical values of statistical life, even a small change in fatalities due to reduced mean vehicle weight could result in costs even larger than the direct compliance costs of the regulation.

A possible concern with focusing only on mean weight is that both the mean and dispersion of weight can be expected to matter for fatalities (Anderson and Auffhammer, 2014). For example, suppose CAFE standards led to a decrease in the mean weight through a down-weighting of the heaviest vehicles. Then the weight distribution could be compressed by CAFE, reducing weight dispersion and perhaps even lowering accident fatalities. The mean and dispersion of weight in accidents have a clear parallel to the classic arms-race metaphor: an unarmed society will face greater casualties when in a conflict with an armed society, but the total casualties would be lower if both societies disarm.

This paper examines the tension between mean weight and weight dispersion, by taking advantage of new unconditional quantile regression methods. Understanding the full distribution of weight changes is critical to disentangling the implications of CAFE on fatalities, but until recently the tools for such an analysis have been lacking. We examine this question using unconditional quantile regression, an approach largely isolated to the labor literature (Fortin, Lemieux, and Firpo, 2009), most prominently in the examination of unions on wage dispersion (DiNardo, Fortin, and Lemieux, 1996). We investigate the relationship between CAFE regulations and vehicle weight, estimating changes to the full weight distribution. While OLS regression can provide an estimate of the change in mean weight, standard quantile estimates are conditional on the included covariates and are insufficient for understanding the effect on the unconditional distribution of weight. We are particularly interested in the unconditional distribution, since this distribution determines the dispersion of vehicle weights sold on the market, and thus influences accident fatalities. In addition to using the unconditional quantile approach, we also carefully consider the measurement of the CAFE regulation's stringency and develop multiple measures to improve on prior studies and lend credibility to our estimates. We apply our estimates of the effect of CAFE on the weight distribution to a large set of accidents, separating the effects for single and multi-vehicle accidents and simulating fatalities in a counterfactual without CAFE.

To estimate these effects we use two unique datasets. For the first dataset, we bring together two large databases that detail trim-level statistics on vehicles sold in the United States between 1954 and 2005. These data allow us to examine manufacturer offerings not only during the period where CAFE was changing but also before CAFE. We document that CAFE is associated with a lowering of the mean vehicle weight and an increase in the dispersion. While heavy vehicles were down-weighted, already light-weight vehicles saw larger declines in weight, amounting to 40-50 lbs per 1 miles-per-gallon (MPG) increase in the standard. The implication of this heterogeneous response is that CAFE standards increased dispersion in a key part of the domestic vehicle market. One plausible explanation for this result is that firms pursued different strategies for products sold to different types of consumers. Consumers of low-weight vehicles may be more likely to be price sensitive and less sensitive to weight, implying that an optimal strategy for CAFE compliance would be to down-weight these smaller vehicles. Consumers of larger vehicles are likely willing to pay more to preserve attributes, and thus the optimal compliance strategy may have been to add costly technology. Importantly, the heterogeneous response we find is consistent with multiple-product firms having multiple responses to regulation, and in the case of CAFE, these multiple responses are not well represented by the mean statistic.

The second dataset contains 17 million police-reported accidents from 13 states between 1989 and 2005, which allows us to estimate the effect of vehicle weight on fatalities. We find that dispersion is linked with a higher likelihood of a fatality, while lower mean weight reduces the likelihood of a fatality in both single-vehicle and multi-vehicle crashes. When we apply the estimated weight changes to these observed accidents, the simulation indicates that the regulation *decreased* fatalities by 171 to 439 per year depending on the year and assumptions. This result is critical for policy: our illustrative calculations suggest that CAFE standards are more likely to have positive net benefits once we account for the saved lives. Further, CAFE may have positive net benefits *even if* the value of future fuel savings is not included as a benefit. This is particularly relevant because future fuel savings would not be a benefit for a rational consumer perfectly valuing the attribute of fuel economy, and economic evidence is building that consumers only slightly undervalue fuel economy, if at all (Allcott and Wozny, 2014; Busse, Knittel, and Zettelmeyer, 2013; Sallee, West, and Fan, 2016). Although the introduction of the new footprint-based standard limits our ability to predict attribute changes for the recent increases ushered in by Presidents Bush and Obama, concerns that the increases will result in more traffic fatalities are not supported by our empirical analysis.

The rest of the paper is organized as follows. Section 2 provides background on the policy and relevant literature. Section 3 describes the data sources used. Section 4 discusses the methodology used to measure dispersion in the vehicle fleet and section 5 reports the results. Section 6 details the simulation of counterfactual fatalities and the estimation of the fatality risk parameters used in that simulation. Section 7 concludes.

2 Background and Literature Review

The United States first implemented CAFE standards in 1975 in the wake of the 1973 Arab oil embargo.[1] Under the regulation, the harmonic average fuel economy of new vehicles sold by each firm was required to meet or exceed a target that was differentiated by car and truck fleet.[2] For cars the average rose from 17 MPG to 27.5 MPG between 1978 and 1990 and it remained frozen at that level until revised in 2010. For light-trucks the standard was more complex. From 1979 to 1982 firms were required to satisfy separate standards for two-wheel drive (2WD) and four-wheel-drive (4WD) vehicles. From 1982 to 1991 firms could either satisfy the separate 2WD and 4WD standards[3] or a single standard that rose from 17.5 to 20.2 MPG. From 1992 to 1996 firms were required to satisfy a single standard that rose from 20.2 to 20.7 and it remained fixed at 20.7 until 2005.[4] Figure 1 illustrates the pattern of CAFE standards over time.

Under CAFE standards prior to 2011, when the fuel economy of a firm's fleet exceeded

[1] Originally passed as a way of reducing exposure to foreign oil imports, the regulation was administered by the National Highway Traffic Safety Administration but the vehicles were tested by the EPA.

[2] Firms are also required to meet the standard separately for imported and domestically produced vehicles.

[3] The 2WD standard rose from 17.2 to 20.7 MPG between 1979 and 1991 while the 4WD standard rose from 15.8 to 19.1 MPG during that same time period.

[4] CAFE applies to all vehicles with gross vehicle weight rating (GVWR) below 8,500 lbs although prior to 1980 the maximum was 6,000 lbs GVWR. The GVWR is the maximum weight of the vehicle when loaded, but we focus our analysis on the curb weight, which is the weight of the vehicle without cargo.

the target, firms generated credits that could be carried forward or back by three years but could not be traded between firms or between fleets. Any firm that did not meet the standard and failed to generate credits within three years was considered out of compliance and fined $55 per vehicle for every MPG below the standard on average in the fleet.

Automobile manufacturers have different product offerings, leading to heterogeneous effects of the regulation on different manufacturers. The domestic firms (Ford, GM, and Chrysler) were required to improve the fuel economy of their fleet to meet the standards, the Asian automakers already exceeded the standards, and the European automakers largely opted to pay fines rather than completely comply (Jacobsen, 2013a). In this paper, we examine all fleets, but focus primarily on the domestic car fleet because it covers a large share of the vehicles on the road during all time periods and the vast majority of vehicles at the inception of CAFE standards.

There are several channels automakers can use to meet CAFE standards. They can add more-efficient technology. They can shift large cars from a passenger vehicle chassis to a light truck chassis (e.g., a cross-over) so that a less-stringent standard applies.[5] They can increase prices on inefficient vehicles and decrease prices on efficient vehicles to push sales towards more fuel-efficient vehicles (Jacobsen, 2013a). Finally, they can alter horsepower or weight to improve fuel economy without using new technology (Klier and Linn, 2012; Knittel, 2011).[6]

As discussed above, the possibility that automakers might lower vehicle weight in response to CAFE has generated substantial concern and has been a crucial point of opposition to raising the standard since the earliest years of the policy.[7] The literature has taken two approaches to quantify the fatalities from CAFE standards. Early work used time series

[5]A similar conversion is possible between 2WD and 4WD vehicles during the earliest years of the truck standard.

[6]There were also short-term loopholes that allowed manufacturers to meet the standard with sales of flex-fuel vehicles at various points in time (Anderson and Sallee, 2011).

[7]An attempt in 1991 to further raise the standard was opposed over concerns that it would "deprive the consumer of a full-sized family vehicle" and "have adverse effects on vehicle safety." See the opening statements by Senator Bryan (D-Nev) on the Motor Vehicle Fuel Efficiency Act February 21, 1991.

variation to directly estimate the effect of mean vehicle weight on annual fatalities. For example, Crandall and Graham (1989) showed that CAFE was associated with a 500 lbs decrease in mean vehicle weight and that vehicles that are 500 lbs lighter are associated with a 27% higher likelihood of a fatality. This would suggest that CAFE increased fatalities by 2,000 to 3,000 annually. Several follow-on papers used the same variation but different model specifications to come to different conclusions (Noland, 2004; Ahmad and Greene, 2005). More recently, the focus has moved to vehicle weight dispersion and in this literature CAFE is modeled as changing relative prices in order to shift drivers towards more efficient vehicle classes (White, 2004; Jacobsen, 2013b). These studies also show increases in fatalities from CAFE standards, but make the restrictive assumption that manufacturers hold vehicle weight constant. In contrast, other recent literature suggests that over a medium-run time frame of 3-5 years, attribute changes are a much more cost-effective way to comply with the standards (Klier and Linn, 2012).

Going forward, the effect of CAFE standards is likely to be all the more complicated due to the switch to footprint-based standards. Both weight and footprint are known to influence vehicle fatalities, but in opposite directions (Wenzel, 2013). The footprint standard allows firms to down-weight vehicles to improve fuel economy, as in the past, but incentivizes them to hold constant or increase the vehicle footprint, which in theory may reduce fatalities. This is in contrast to weight-based standards; for example, Ito and Sallee (2014) show that weight-based standards in Japan exacerbate accident externalities.

3 Data Description

3.1 Data Overview

To analyze the effect of CAFE on vehicle weight and fatalities, we collected and combined several unique datasets. These include a large trim-level dataset on vehicles from 1954 to 2005 that allows us to examine the effect of manufacturer decisions on the attributes of vehicles produced each model year. From these data we generate estimates that allow us to simulate vehicle weights in a counterfatual without the CAFE standard. We connect these to accident outcomes using data on police-reported accidents that describe the vehicles and fatality outcomes of 17 million crashes from 1989 to 2005.

3.2 Vehicle Data

Data on the weight of vehicles comes from two data sources, Ward's Automotive Yearbooks (Ward's) and Automobile Catalog (AC).[8] AC is a database of automobile characteristics collected from various sources, including back issues of Popular Mechanics and manufacturer documentation. These data include trim level statistics on vehicles sold in the United States between 1954 and 2005. We focus our attention on cars and light trucks below 8,500 lbs gross-vehicle-weight-rating (GVWR), the maximum regulated by CAFE, and below 6,000 lbs GVWR before 1980.[9]

While Ward's provides detailed data on trim-level specifications after 1981, there is less detail between 1971 and 1981, and no data are available before 1971. Fortunately, AC data provide detail on vehicles going back to 1954 and help to fill in the gaps in the critical ramp-up period of CAFE between 1971 and 1981 and to examine pre-CAFE trends in fuel economy. These data are used to create one of our measures of CAFE stringency: automaker-specific predictions of average fuel economy in a world without CAFE standards. The AC data include all domestic manufacturers as well as a few foreign manufacturers, including Audi, Honda, and Mazda.

In our analysis we focus on trim-level data rather than sales-level data, since we are most interested in the weight of the vehicles that the manufacturers chose to produce and how those decisions influenced accidents. We then directly capture the prevalence of those trim-level vehicles in crashes using the police-reported accident data. We are not aware of

[8]This data source is maintained by Pawel Zal at www.automobile-catalog.com.

[9]We generally focus on the curb weight of the vehicle, which is the weight of the empty vehicle while GVWR is the maximum operating weight of the vehicle.

any source that records sales data before 1975 and such data is sparse up to 1985, limiting the usefulness of a sales-weighted analysis. Furthermore, the model-level sales data that are available are more highly aggregated than the trim-level data, which is problematic for our approach because individual quantile estimation requires substantial data to precisely estimate coefficients.[10]

Table 1 shows the mean and standard deviation of new vehicle weight and observation counts for various manufacturers and product segments. These are given for 5-year periods except for 1975-1977 and 1978-1979, which are split at the introductory year of CAFE. This table demonstrates the well-known drop in vehicle weight that occurred during the introduction of CAFE: the mean vehicle weight of cars dropped from 3,755 lbs in 1970 to 2,963 lbs in 1985. The standard deviation of vehicle weight was at a maximum between 1975 and 1977, a time period when the standards initially ramped up and when existing on-road vehicles were still relatively heavy at 3,600 lbs. The minimum dispersion occurred in 1985, the year where car weight is at the lightest in our data. These simple summary statistics already indicate that lowering the mean may not increase dispersion.

3.3 Accident and Fatality Data

We use the State Data System (SDS) dataset maintained by the National Highway Traffic Safety Administration (NHTSA) to recover the effect of vehicle weight on total fatalities and to simulate counterfactual fatality outcomes. Although 21 states are available through the SDS, only 13 states had sufficient information to be used in our analysis. These states are Florida, Georgia, Illinois, Kansas, Michigan, Missouri, Nebraska, New Mexico, New York, North Carolina, Pennsylvania, Virginia, and Washington.[11]

[10] We perform robustness checks in the appendix using sales-weighted regressions; the results are similar but far less precise due to the significant aggregation. The similarity in point estimates is perhaps unsurprising as popular models of vehicles seem to result in a proliferation of different trim options for those models. Conditional on model year and class, sales explain more than 37% of the variation in the count of trims where sales data exists.

[11] Our data differs somewhat from that used by Anderson and Auffhammer (2014). Kentucky and Maryland no longer allow access to their data. Ohio and Wyoming did not have VINs to allow us to decode. Our

We obtained permission from each state to use the data, which consists of all police-reported accidents. Each crash in the data is given an identifier linking all vehicles and individuals involved in the crash. There is considerable variation by state and year in the variables recorded for each crash, but our simplest regressions require only the number of vehicles involved, the VINs for each vehicle, and the presence of fatalities. The requirement of a decodable VIN is the strongest restriction on our data. VINs are a unique serial number for all motor vehicles and the first nine digits allow us to decode the year, model, make, and vehicle weight of each automobile involved in a crash. Of the crashes, 38% are dropped because one more more vehicles do not have a usable VIN. Some common reasons a VIN cannot be decoded are that the vehicle was produced before 1981, it is a heavy-duty vehicle, moped or motorcycle, or because of a recording error.[12] Some regressions also make use of variables that are not present in all states such as the county of accident, seatbelt use, or vehicle speed.

For our regressions and simulations we use only crashes involving one, two, or three vehicles. We omit all crashes that involve pedestrians, motorcycles, or bicycles and we assume that CAFE had no effect on these fatality outcomes. We also assume that fatality outcomes for crashes involving more than three vehicles, which represent less than 5% of crashes, are also unaffected by CAFE. These restrictions eliminate 43% of total crashes, but a larger fraction of fatalities, 56%, because accidents with pedestrians, bicycles and motorcycles are deadlier on average.

3.4 Other Data Sources

Nominal annual gasoline prices from 1954 to 2005 were obtained from the Energy Information Administration and were converted to real values using the Consumer Price Index (CPI) from the Bureau of Labor Statistics. Nominal annual GDP from 1954 to 2004 was ob-

sample includes IL, KS, MI, NE, NM, NY, NC, PE, and VA, which were not available for their study.

[12]These errors are sometimes apparent because a VIN will include a letter not allowable or are an incorrect length. As shown in appendix tables A2 and A3, vehicles with unusable VINs are more prevalent for the earliest years of data and improve with time.

tained from the Bureau of Economic Analysis. In our counterfactual simulations, we project fatality rates by state based on fleet characteristics from the National Household Transportation Survey (NHTS) linked to our vehicle data. This projection also uses county-level data on population and total fatalities. Population data comes from the U.S. Census via the National Cancer Institute that adjusts for some known issues such as population shifts between census waves due to natural disasters. County-level fatalities are from the Fatality Analysis Reporting System (FARS) maintained by the NHTSA. Unlike our police-reported data, FARS only records fatal crashes but exists for all counties nationwide.

4 Dispersion Methodology

4.1 Motivation

Our empirical strategy is motivated by the concern that the true risk of increasing the stringency of CAFE depends on changing not only the mean but also the dispersion of vehicle weight. We employ an unconditional quantile regression technique used for studying wage dispersion developed by Fortin, Lemieux, and Firpo (2009). This technique, known as Recentered Influence Function (RIF) regression, is ideal for our setting where CAFE can be satisfied through several channels, including lowering the weight of some vehicles.

The effect of a policy on dispersion can either be summarized using a statistic of the distribution or in a disaggregated fashion by examining individual quantiles. Common summary measures of dispersion include the Gini coefficient (Borenstein and Rose, 1994), variance (Dahlby and West, 1986), coefficient of variation (Pratt, Wise, and Zeckhauser, 1979; Sorensen, 2000; Lach, 2002), and range (Sorensen, 2000). Although RIF regression can be run on any statistic of a distribution, we focus on quantiles, which allow us to examine how automakers are adjusting the entire weight distribution.

Regression techniques that move beyond the mean were advanced by Koenker and Bassett (1978) with the development of *conditional* quantile regression. But unlike the mean, one cannot use the law of iterated expectations to move from the conditional to the unconditional estimates for quantiles.[13] In some cases the researcher may be interested in the effect of a treatment on the τ-quantile conditional on covariates; in other cases the inability to interpret the coefficients as unconditional quantiles may be problematic.[14] The method developed by Fortin, Lemieux, and Firpo (2009) is ideally suited to our question because we wish to control for variables that are known to affect automaker decisions but ultimately want to know how CAFE affects an arbitrary unconditional quantile τ of the equilibrium weight distribution for the fleet.[15] Other approaches, such as conditional quantile regression, or kernel-weighted ordinary least squares, would improperly characterize how the full distribution of the fleet changes. Appendix B provides an illustrative example to demonstrates the value of unconditional quantile regression for our research question.

RIF regression makes use of an influence function (IF), which represents the influence of adding an infinitesimally small amount of contamination at point y to a distribution F_Y on a distributional statistic ν. Examples of ν include the τth quantile, the mean, the variance, and the Gini coefficient. As first shown in Hampel (1974), the influence function is formally defined as $IF(y; \nu, F) = lim_{\epsilon \to 0}(\nu(F_\epsilon) - \nu(F))/\epsilon$, where $F_\epsilon(y) = (1-\epsilon)F + \epsilon\delta_y$, $0 \leq \epsilon \leq 1$, and δ_y is a distribution that puts all mass at the value of y.

To construct the RIF, the statistic of interest is then added back to the IF:

$$RIF(y; \nu, F_Y) = \nu(F_Y) + IF(y; \nu, F_Y). \tag{1}$$

[13]The linearity of the mean allows for equivalence between the conditional mean $E(Y|X) = X\beta$ and the unconditional mean $E(Y) = E(X)\beta$, but this does not hold for a nonlinear statistic.

[14]As an example, it is possible that a treatment causes a large down-weighting at the 10th quantile of the products produced by only one automaker (the quantile conditional on the automaker), but that particular automaker may contribute very few vehicles to the 10th quantile of the total population, implying that the conditional quantile is very different than the unconditional quantile.

[15]It is also possible to recover the unconditional values by calculating every conditional quantile using the approach of Machado and Mata (2005); however, the method developed by Fortin, Lemieux, and Firpo (2009) is much less computationally burdensome.

For the specific case of the τth-quantile (q_τ), we have

$$RIF(y; q_\tau) = q_\tau + \frac{\tau - \mathbb{1}(y \leq q_\tau)}{f_Y(q_\tau)} \tag{2}$$

where f_Y is the marginal density function of Y and $\mathbb{1}(\cdot)$ is an indicator function taking the value of 1 if the condition holds.

Assuming that the explanatory variables, X, enter linearly, this allows for an OLS regression of Y transformed by the RIF on X

$$E[RIF(Y; \nu)|X] = X\beta_\nu \tag{3}$$

where β_ν is the effect of X on the distributional statistic ν, which in our case is the τth-unconditional quantile.[16]

4.2 Model Specification

For our initial analysis we focus on the impact of CAFE standards on the weight of domestic cars, which made up the vast majority of the fleet at the inception of CAFE.

We model the τ-quantile of vehicle weight, wt, as:

$$E[RIF(wt_{f,i,t}; q_\tau)|S_{f,t-1}, Z_{f,i,t-1}] = \beta_\tau S_{f,t-1} + Z'_{f,i,t-1}\gamma_\tau. \tag{4}$$

Here, f refers to a firm-fleet combination, i a trim, and t a model year. S is a measure of CAFE stringency. Z is a matrix of other controls. In Z we include time trends, lagged gasoline prices, and lagged GDP. We omit variables such as fuel economy and horsepower because we are interested in understanding the influence of CAFE on vehicle weight allowing for changes in these other attributes. After all, the primary reason for down-weighting is to

[16] In the case of the the quantile, the density function f_Y in the RIF must be estimated and the standard errors adjusted by bootstrapping. In the case of other statistics, such as the variance, the RIF is not estimated.

improve fuel economy. Bootstrapped standard errors are computed using 200 replications.

To estimate the effect of CAFE standards on the decisions of domestic automakers we need a measure of CAFE stringency on each automaker's fleet in each year. Ideally we would have a measure of stringency year-by-year, for each automaker, and each fleet. Anderson and Sallee (2011) are able to measure this stringency using a loophole in the CAFE regulation, but unfortunately this loophole only lasted for several years. In lieu of such a measure, prior work has generally used the CAFE level itself (Crandall and Graham, 1989; Ahmad and Greene, 2005; Knittel, 2011) or the difference between the CAFE level and a counterfactual average fuel economy without CAFE. Developing this counterfactual has involved a range of assumptions, with some researchers using cross-sectional variation in the pre-CAFE level (Klier and Linn, 2016), others introducing some structural assumptions by predicting a counterfactual using a reduced form regression (Small and Van Dender, 2007), and others using a full structural estimation (Klier and Linn, 2012; Jacobsen, 2013b,a).

While using only the CAFE level and omitting the counterfactual level of fuel economy is a transparent approach using exogenous variation, there are two drawbacks to the use of this method. The first is that firms had quite different average fleet fuel economy levels before the policy, so it is a relatively poor measure of stringency. The second arises because our ultimate goal is to set the stringency equal to zero and simulate fatalities without CAFE standards. In the absence of CAFE, the average fuel economy of vehicles is not zero but is determined by factors such as fuel prices, GDP, and technology trends.

Because of these concerns we develop a counterfactual fuel economy level using an approach most similar to Small and Van Dender (2007). We define our measure of stringency as:

$$S_{f,t-1} = l(CAFE_{t-1}) - l(\widehat{MPG_{f,t-1}}) \tag{5}$$

where $l(\widehat{MPG_{f,t-1}})$ is the log of the level of fuel economy for firm f in the absence of CAFE. We predict the time path of fuel economy for each firm's fleet $l(\widehat{MPG_{f,t-1}})$ by first estimating a reduced-form equation explaining the fuel economy of vehicles produced between 1952 and

1977. Explanatory variables for fuel economy include gasoline price, GDP, and a time trend.[17] We then use these estimated coefficients to predict fuel economy in the absence of CAFE for the period after 1978. The primary benefit of this measure is that, although the variation is still largely driven by the imposed CAFE schedule, it acknowledges that firms are not starting from identical places and allows for sensible counterfactual simulations by setting this measure of stringency to zero.

Because vehicles are only redesigned every few years, the most relevant measure of stringency for a given model year will be based on the stringency in prior years as well as anticipated future levels. For this reason we average the stringency measure over the three proceeding years, the current year, and one future year, but we also examine the effect using only the prior year of stringency as a robustness check.

There are some shortcomings to capturing stringency using this method.[18] The first is that the variation identifying the coefficient on stringency largely comes from the increase in the CAFE standard between 1978 and 1990. The response of automakers may be different today. A related concern is that our prediction of counterfactual fuel economy is less reliable over time. The automobile market is extremely dynamic with constant technological change. Therefore we stress that our estimates and simulation results are most valid in examining the original CAFE standards and caution should be used in assuming that similar dynamics will prevail under the new standards. We nevertheless believe our results are of first order importance because criticism of the CAFE standard's effect on fatalities has largely been based on the lessons learned during this early time period. Evidence from this early period generated a reluctance to increase the standard and informed the decision to shift to a footprint-based standard to prevent the decreases in vehicle size that happened during that time.

[17]We also examine the inclusion of additional covariates for this prediction in appendix table A6.

[18]In the appendix we also examine the CAFE credit balance of the firms. See appendix table A7. Where fluctuations in the credit balance of a firm are due to unanticipated shocks, firms will face unanticipated changes to stringency that are plausibly exogenous. However, this variation is so limited that the effects, while generally consistent with our preferred results, are imprecisely estimated.

Another concern is that the choice of explanatory variables in predicting $l(\widehat{MPG_{f,t-1}})$ may seem somewhat arbitrary. Additionally one may be concerned that the measure, $S_{f,t-1}$, is not capturing stringency but rather is capturing the effect of the variables included in creating the prediction, which are likely to influence both weight and fuel economy. For example, it may be the case that the stringency measure is in part capturing the direct effect of gasoline prices and GDP (note that we also make sure to include these covariates in our weight regressions).

We address these concerns in several ways. We examine the robustness of our stringency measure to these included regressors. We also examine the results using the CAFE level as our measure of stringency. Using only the variation imposed by the policy removes any concerns related to the prediction of $l(\widehat{MPG_{f,t-1}})$ but does not allow for the counterfactual simulation as noted earlier. Nevertheless to the extent that this variation in the CAFE level drives our estimates, it suggests that our decisions about how how to develop $l(\widehat{MPG_{f,t-1}})$ do not overly influence our estimates. Although no measure of stringency provides a 'silver bullet' for identification, we argue that the measure chosen is the most defensible and, based on robustness checks, its weaknesses are unlikely to be driving the results.

In a second set of regressions we examine the competitive effects of CAFE on the unconstrained firms. In particular, we hypothesize that the Asian automakers may offset the weight changes that CAFE may induce in the domestic firms. For this analysis we develop a composite measure of stringency on the domestic firms by sales-weighted averaging of the measure across the domestic firms.

5 Results

5.1 Empirical Results

This section reports estimates of the effects of CAFE on the quantiles of vehicle weight for domestic cars and trucks, and Asian cars. Because presenting point estimates for 100 quantiles is not feasible, we first present estimated results for the mean and a few key quantiles on the domestic car sample. We then we use a graphical approach that plots all coefficients for various robustness tests and estimates using other fleets.

5.1.1 Results for the Mean and Key Quantiles

Table 2 presents our main results for the mean and five key quantiles. Panel A displays our estimates of the effect of our preferred stringency measure on weight. The OLS regression shows that a 10% increase in the standard will reduce the mean vehicle weight by 2.99%. This implies that a 1 MPG increase in the standard will reduce mean vehicle weight by 32 lbs. A decrease in the mean vehicle weight does not reveal where in the product space the down-weighting occurred or if it is associated with a compression or dispersion of vehicle weight on the road.

Columns 2 through 6 present estimates from RIF regressions on the 10th, 25th, 50th, 75th and 90th quantile. These regressions reveal that the source of the down-weighting observed in the mean was concentrated among already low-weight vehicles. We see the largest decreases in weight were at the 25th quantile where a 1 MPG increase in the standard would result in a 51.9 pound decrease on a vehicle weighting roughly 2,500 lbs. The effects for higher weight vehicles are closer to 30 lbs, a much smaller decrease. These point estimates suggest that for much of the domestic fleet distribution, CAFE generates dispersion in vehicle weight whereby already low weight vehicles are down-weighted faster than heavier ones.

Panel B presents results using only the log of the CAFE level as the stringency measure. There are a few key differences in the results. While the magnitudes are often similar, the largest effects are concentrated in the center of the distribution with the median experiencing a decrease in vehicle weight of 67.8 lbs for a 1 MPG increase in the standard. This measure suggests slightly more compression may occur for lower-weight vehicles, but that CAFE continues to cause dispersion across much of the product space. These results are informative but are limited to only a few coefficients, while a graphical approach provides a more complete picture.

5.1.2 Graphical Display of Full Coefficient Estimates

Table 3 plots the coefficients and 95% confidence intervals for several fleets of vehicles.[19] For ease of interpretation and comparison across fleets, the x-axis plots the weight corresponding to the quantiles.

The figure in row 1 plots the estimates for all quantiles from our preferred specification for the domestic car fleet, including those from Panel A of Table 2. Compression of the distribution occurs when the line connecting the coefficients takes on a negative slope, while dispersion is characterized by a positive slope. As suggested in Table 2, the point estimates here indicate that some compression occurred for low weight vehicles but that dispersion occurred over most of the product space.

Row 2 replaces our preferred measure of stringency with the log of the CAFE standard, as in Panel B of Table 2. These results suggest that a somewhat larger part of the distribution experienced compression than with our preferred measure, but that dispersion continues to occur from the median up to the heaviest cars produced by US firms.

In our discussion of the stringency measure we noted that firms likely consider several years of stringency when making design decisions but it is also possible that firms mostly react to stringency in the prior year. Row 3 shows that our results change only minimally when we examine the stingency measure based only on the prior year.

Row 4 examines quantiles based on the sales of vehicles sold at the model-level, rather than trims produced. This approach greatly reduces the sample size. Although the point estimates are consistent with our previous results in that they continue to slope upward, the results are imprecisely estimated.

[19] Because quantile techniques do not perform well at the edges of the distribution as shown with a Monte Carlo simulation in the appendix, we omit the top and bottom 3 quantiles from all graphs. At the ends, the plotted points are further apart as fewer vehicles are produced in this product space and the distance between quantiles is larger. In the middle of the density, where more vehicles are produced, the points are closer together.

Estimates for the domestic light-truck fleet are presented in row 5. We see the largest down-weighting for the lightest vehicles, and some evidence that up-weighting may occur for the heaviest vehicles. In the appendix, we consider the robustness of our truck fleet results and find that down-weighting of low-weight vehicles is consistent across all specifications but up-weighting of the heaviest vehicles is more fragile. Nevertheless, there are several reasons that the results for the heaviest light-duty trucks may differ from that of heavy cars. Prior to 1992 firms could choose to satisfy separate standards for 2WD and 4WD vehicles. One way to comply with the standards was to redesign smaller 2WD models into larger, and heavier, 4WD models that faced a lower standard. Another possible reason for the difference between cars and light-duty trucks is the presence of unregulated heavy-duty trucks. Because vehicles over 8,500 GVWR are not regulated by CAFE, several common heavy duty trucks are unregulated, including the Ford F-250, Chevrolet Silverado 2500HD, GMC Sierra 2500, and Dodge Ram 2500. The presence of these unregulated substitutes may create different incentives for firms resulting in a different response to the regulation than found in the car segment.[20] Nevertheless, the results for the domestic truck fleet indicate that increased dispersion occurred.

Overall the data provide evidence that the use of down-weighting by firms was not uniform across the weight distribution and that increased dispersion occurred for much of the domestic fleet. We view the result that firms tended to pursue down-weighting on already lighter-weight vehicles as economically plausible, for lighter-weight vehicles might be expected to be sold more price-sensitive customers who are less willing to pay for new technology but are accepting of lighter vehicles.

[20]Our heaviest quantile is significantly less than 8,500 lbs because our data are for curb weight, not GVWR, which is the maximum allowable weight in the vehicle.

5.1.3 Competitive Effects

To examine the effect of the standards on the Asian car segment we begin with the same specification as above using the predicted fuel economy measure of stringency for each firm.[21] These estimates are presented in Row 6 of Table 3. What is immediately apparent from the figure is that these firms seem to react very little to a tightening of 'own-firm' stringency. The marginal effects implied by these coefficients are also small compared with those of the domestic firms, usually less than 3 lbs. A minimal response makes economic sense because the vehicles of these firms are already sufficiently fuel-efficient that these firms did not face a binding standard.

CAFE may still have implications for these firms. Row 7 examines the possibility that increased stringency on the domestic manufacturers allows the Asian manufacturers to increase weight and move into a product space previously occupied by domestic firms. We replace 'own-firm' stringency with the sales-weighted average stringency of the domestic firms. This figure shows that as the stringency on domestic firms is increased, Asian firms will increase the weight of their lighter vehicles, possibly to take advantage of the demand for a type of vehicle that is no longer provided by domestic firms. The precision of these regressions is lower than those for the domestic firms because the stringency measure is no longer firm-specific, but many coefficients are statistically significant and positive. These results suggest that the Asian firms may be capitalizing on the down-weighting by domestic firms, which could partially offset the increased dispersion due to responses by domestic automakers.

[21] The only manufacturers included in this analysis are Mazda and Honda, which have sufficient pre-CAFE observations from which to generate these predicted fuel economy levels. We also lack data during this time period for the European firms and thus we cannot examine European firms using this stringency measure. Estimates using alternative stringency measures are presented in the appendix.

6 Simulation

6.1 Overview

Our results suggest that CAFE standards increased dispersion for a substantial portion of the vehicles on the road. But we cannot conclude from this that CAFE increased fatalities due to two potentially mitigating factors: the change in mean vehicle weight and the sorting of vehicles by location. The mean weight in accidents would be expected have effects independent of weight dispersion, following the arms-race metaphor. Similarly, if small vehicles sort to cities, they may rarely encounter the large trucks often found in rural areas. In this section, we examine the net fatalities accounting for all of these factors.

To model fatalities, we use an approach closely following Anderson and Auffhammer (2014), who find that increased dispersion of vehicle weights increases accident fatalities. Their results also suggest that a lower average vehicle weight might decrease fatalities, but this result is not statistically significant from zero. We expand upon their results by adding further states to the analysis, using a consistent framework that is amenable to counterfactual simulation, and performing extensive robustness checks.

To briefly summarize, we consider a set of $\bar{N}$ accidents involving three or fewer vehicles indexed by i.[22] Let $\mathbf{wt}$ be a vector of vehicle weights, $\mu_i(\mathbf{wt})$ be the mean weight of vehicles in the crash, and $\sigma_i(\mathbf{wt})$ a measure of dispersion where $\sigma_i = 0$ for 1-vehicle crashes. Let the probability of a fatality be defined by

$$f_i(\mathbf{wt}) = f(\mu_i(\mathbf{wt}), \sigma_i(\mathbf{wt}), V, D, T) \tag{6}$$

where V are other characteristics of the vehicles involved[23], D are driver characteristics

[22] As we document below, most crashes involve 3 or fewer vehicles and those with more have relatively complicated dynamics.

[23] It is possible that many vehicle attributes are a function of vehicle weight but one important attribute for fatalities is the size, or footprint, of the vehicle measured as length times the width of the wheelbase. Prior to the current footprint-based standard, firms tended to change footprint in tandem with vehicle weight (Wenzel, 2013). In appendix Table A22 we consider the implications of undoing footprint changes to vehicles.

or behaviors, and T are time or vehicle model year controls. Appendix F contains a detailed discussion of the accident fatality estimation, including the exact model specifications, and the full results.

When a fuel economy standard is implemented, automakers change vehicle weights to $\widehat{\mathbf{wt}}$. CAFE may affect accidents through many channels. For example, more efficient vehicles cost less per mile to drive, so consumers will drive more (the rebound effect), which generates additional fatalities (Gillingham, Rapson, and Wagner, 2015). Alternatively, CAFE provides an incentive for automakers to use a light truck chassis, arguably even giving rise to the SUV class, which has different crash risks than large cars. We restrict our analysis to asking if the weight changes caused by CAFE caused crashes to become more fatal, so we preserve $\bar{N}$, V, and D. We also omit any crashes involving pedestrians, motorcycles, bicycles, heavy-duty trucks and crashes involving more than three vehicles as these crashes, as noted in the summary statistics.[24] We discuss these and other potentially omitted factors in more detail in section 6.4.

Our simulation proceeds in several steps. First, we use our police-reported accident data to estimate $f_i(\mathbf{wt})$. We perform this estimation separately for 1-, 2-, and 3-vehicle accidents as is discussed in Appendix F. Next, we simulate fatalities in a counterfactual world by removing CAFE and adjusting the vehicle weight according to our RIF estimates from section 5 to generate counterfactual weights $\widehat{\mathbf{wt}}$. Specifically, we adjust by the estimated treatment effect for each quantile of the equilibrium weight distribution to generate the counterfactual equilibrium weight distribution without CAFE. We perform this analysis for the domestic car and truck fleet, and in a second simulation for all fleets, domestic and foreign. We then apply the new counterfactual vehicle weights to the existing accidents in the SDS data and the accident estimation results in order to estimate fatalities $\widehat{f}_i = f(\mu_i(\widehat{\mathbf{wt}}), \sigma_i(\widehat{\mathbf{wt}}), V, D)$.

Finally we consider how these results might extend to the entire United States. States

[24]Ballesteros, Dischinger, and Langenberg (2004) study the effect of vehicle weight on pedestrian and bicycle accidents and find that heavier vehicles are more deadly to these agents but suggest that the correlation may be due to design elements of the front of the vehicle and hood.

not represented in the SDS crash data differ along some key dimensions. Most importantly, states in the West and South have more Asian vehicles and are under-represented in the SDS data. To scale our results up and account for heterogeneity and sorting of the fleet, we extend our predictions using imputation. To impute fatalities, we calculate the percent change in fatalities for each county in our SDS data, $p_c = \sum_{i \in c} \widehat{f}_i / \sum_{i \in c}^{\bar{N}} f_i$.[25] We then regress the percent changes for observed SDS counties on county and state characteristics and predict the percent change in fatalities in all counties. Finally, these percent changes are multiplied by observed fatalities by county nationwide from the FARS dataset.

6.2 Summary of the Effect of Vehicle Weight on Fatalities

Our accident fatality results (in Appendix F) confirm the results in Anderson and Auffhammer (2014): increased dispersion increases fatalities and decreased mean weight *decreases* fatalities. These two results together suggest that down-weighting low-weight vehicles will both lower the total weight, reducing fatalities, and increase dispersion, increasing fatalities. The net effect on fatalities depends on the characteristics of the fleet and the effect of CAFE on the full weight distribution.

There is a clear physical explanation for the positive relationship between weight and fatalities. Larger vehicles carry more kinetic energy. This places higher demands on braking to stop a vehicle, and in the event of a crash, increases the amount of energy the vehicle structure needs to absorb (Wenzel, 2013). Thus, it is only when a vehicle encounters a movable object does weight provide a protective effect. For such crashes, a heavier vehicle is likely to displace the smaller vehicle or movable object, reducing the deceleration experienced by an occupant.

A final useful result is that vehicles with a larger footprint have lower fatalities. This again corresponds with the engineering safety literature (Wenzel, 2013).

[25] We include all fatalities, not only those for 1-, 2-, and 3-vehicle accidents, in the denominator of this percent change calculation since FARS fatality counts are not limited to crashes involving 3 or fewer vehicles.

6.3 Simulation Results

To simulate the effect of removing CAFE, we first create a counterfactual fleet weight distribution by using the RIF estimates from section 5 to adjust the weights of vehicles.[26] Then we calculate the percent change in fatalities for the calendar year 2005 to illustrate the effects.[27] Because this is roughly two decades after the the largest changes in CAFE standards, these results for 2005 can best be thought of as 'steady state' results in the sense that the full fleet has been adjusted by CAFE by this point. We note that the transition to this steady state—when only the newer vehicles in the fleet have been affected by CAFE standards—may have different effects on fatalities.

Table 4 shows the results of these simulations. Panel A shows these effects on the domestic fleet for the states in our SDS sample. The simulation indicates that removing CAFE would have resulted in 171 additional fatalities. In other words, CAFE *saved* lives. This result can largely be attributed to 1-vehicle accidents: down-weighting reduced 1-vehicle accident fatalities by 141 lives. Obviously 1-vehicle accidents are only affected by the down-weighting and not dispersion. The additional dispersion increased fatalities for multi-vehicle accidents, but this effect appears to be offset by the lower mean vehicle weight in those accidents.

One important effect noted in section 5 is that the Asian automakers tended to up-weight their vehicles. Panel B considers weight changes for the entire light duty vehicle fleet, including both domestic and foreign manufacturers. We see that the up-weighting by the Asian automakers largely offsets the changes in mean weight across all categories of accident. At the same time, the up-weighting reduces dispersion, saving a considerable number of lives

[26] To calculate the weight of vehicles in the absence of CAFE we multiply the estimated coefficient for a vehicle's weight quantile by the stringency in the model year and the weight of the vehicle. This gives the total weight change induced by CAFE, which is then added back to the vehicle weight for a counterfactual vehicle weight. For fleets that are a small share of the market, such as Asian light trucks or European vehicles we use estimates from those fleets that are presented in Appendix Table A15.

[27] We also reweight the observations to match the frequency in the initial data. If a given fraction of vehicles are missing information, this will eliminate more multi-vehicle accidents than single vehicle accidents. We reweight each category of crash (1,2 or 3-vehicle) to match the number of crashes present in each category in the initial SDS accident files before deletions. Although the simulation is for the set of accidents in 2005, it sets stringency of CAFE to zero for all years current and past.

for multi-vehicle crashes. On net, this scenario indicates that CAFE saved 171 lives, due to the effect of decreased dispersion dominating.

Panel C extends the analysis nationally for only the domestic fleet. The county and state characteristics used for the imputation include county-level population and annual traffic fatalities, state level average vehicle weight, age, share light truck, share US and Asian manufactured vehicles, and interactions between county-level population and fatalities with fraction US and Asian.[28] As expected, the results are larger in magnitude than the effects in Panel A since a larger geographic area is covered. The results indicate that CAFE saved 439 lives.

Panel D extends the analysis nationally for the entire light-duty vehicle fleet. The aggregate results largely align with other simulations and indicate that 393 lives were saved nationally. There are however some interesting differences visible in the decomposition. For example, consider 1-vehicle accidents. In the SDS states CAFE saved 20 lives (Panel B), while nationally CAFE generated nearly 21 additional fatalities (Panel D). This difference stems from differences in the vehicle fleet between the SDS states and the nation. In particular, outside of the SDS states (especially on the West Coast) there are many more vehicles from Asian automakers. CAFE led to these vehicles being up-weighted, which generated more deadly 1-vehicle crashes. There are even larger effects for multi-vehicle accidents than 1-vehicle accidents. For 2-car and 3-car accidents, the change in dispersion due to CAFE led to fewer deaths, and in fact, this effect overwhelms the increased deaths due to changing of the mean weight.[29]

One of the primary benefits of our county-by-county simulations is that it captures the sorting of vehicles by location, which can amplify, mute, or even reverse the patterns that emerge from the RIF-regressions. Automarkers are required to meet CAFE at a national level, which can lead to significant local heterogeneity in the effect of CAFE. Given sorting,

[28] See Appendix Table 20 for a full listing of the controls.

[29] In appendix table A22 we consider alternative scenarios where footprint is adjusted along with vehicle weight and the effects in calendar year 1995 when the fleet was somewhat different than 2005. Across all scenarios CAFE saves lives.

the durability of vehicles, and the interactions between the car and light truck fleet, nearly any effect is possible at the county level. We illustrate the effects in our simulation in Figure 2. Panels A and B show our results for the the predicted changes in fatalities when CAFE is removed for the observed accidents in our SDS dataset. Most counties saw fewer fatalities, but some experienced more. In panel B, for example, Cook County, IL is better off because of CAFE since the policy saved 17 lives. In Queens County, NY CAFE generated an additional 0.1 fatalities. Panels C and D extend these values nationally, and again there is considerable heterogeneity. Consider some examples from panel D that are outside of the SDS states. In Maricopa County, AZ, which contains Phoenix, CAFE saved nearly 10 lives, while in Los Angeles County, CA it generated an additional 26 fatalities. These results are primarily driven by the market share of domestic automakers, which varies greatly across the country.[30]

6.4 Discussion

Our simulation methodology has several advantages. The first advantage is that unlike randomly drawing vehicles to crash at a national level, we account for sorting over location. This is important because small vehicles tend to sort into large cities and large vehicles into suburban and rural areas, and we only want to interact vehicles that are likely to meet on the road. Jacobsen (2013b) notes that drivers also sort by risk into vehicle class. For example, minivans may attract low-risk drivers. By preserving accidents as observed, we preserve the crash frequency by class. We also do not require assumptions about vehicle lifetime, miles driven, and the population of vehicles on the road.

There are however some caveats to our results. First, we assume that drivers do not switch their vehicle class in response to CAFE.[31] To the extent that drivers buy larger vehicles in

[30]For example, Illinois has a fleet that is 80% domestic and 18% Asian, while California has only 52% domestic and 40% Asian.

[31]Our assumption is in effect the mirror image of Jacobsen (2013b). While he assumes that CAFE preserves vehicle risk but shifts driver risk across vehicle classes, we assume that CAFE changes vehicle risk but does not affect driver risk.

response to down-weighting, we will overestimate the number of lives saved. However, we do not anticipate this dynamic to substantially alter our result. Klier and Linn (2012) suggest that satisfying CAFE by using pricing to alter market shares across classes is an expensive way to satisfy CAFE. Furthermore, we observe that firms find it optimal to change the weight of vehicles to comply with CAFE. While firms could use expensive technology that improves fuel economy without sacrificing attributes, the observed down-weighting suggests that many consumers are more price-sensitive than attribute-sensitive and would be unlikely to shift to larger, expensive vehicles.

Second, we recognize that drivers or firms may respond to changes in risk. Our regressions including speed and seat belt covariates suggest these driver behaviors will have a relatively minor effect on our estimates. Nevertheless, it is possible that drivers of lighter vehicles may drive more cautiously to avoid accidents with large vehicles (Peltzman, 1975).[32] Prior research has not found evidence of such behavior (Cohen and Einav, 2003). There may also be safety adjustments by firms. Even though small cars may be safer for society, drivers may be concerned that they are more dangerous for occupants. Firms may respond by putting safety devices into at-risk vehicles to reassure drivers. If such behavior were the case for either consumers or producers, it would suggest that we are underestimating the lives saved by CAFE standards, reinforcing our result that CAFE standards saved lives.

Lastly, as noted earlier, CAFE may alter fatalities through channels besides vehicle weight if the frequency of accidents changes as a result of more miles driven due to the rebound effect or changes to total vehicle sales. While these effects are important, we see these as the safety implications of the rebound effect or of total vehicle sales and not the safety effects of changing vehicle weights. Both of these effects could be analyzed separately.

[32] Although we find that lower vehicle weight reduces total fatalities in multi-vehicle crashes, this is unlikely to make these agents drive more carelessly because most of these benefits are external and accrue to the driver of the larger vehicle (White, 2004).

7 Conclusions

Despite considerable resistance to CAFE standards based on concerns that down-weighting leads to increased fatalities, we find that CAFE likely saved lives. Ours is the first paper to empirically examine the distributional changes to vehicle weight generated by CAFE. Although we find increased dispersion in vehicle weight in several fleets, these effects are offset by the reduced dispersion due to upweighting of the Asian automaker fleet and the lower fatality risk from lowered mean vehicle weight. This demonstrates the importance of separating the effects due to a change in the mean from effects due to a change in dispersion. The fatality risk of decreasing the weight of all vehicles by 50 lbs is different than the relative risk of a vehicle that is 50 lbs lighter in a multi-vehicle crash. It also demonstrates how heterogeneous responses to regulation can influence the unintended consequences of the regulation.

These results further speak to the potentially large welfare gains of a Pigouvian tax on vehicle weight, but our dispersion results suggest that the optimal tax would have to be contingent on the weight distribution of other vehicles on the road. While such an optimal policy may be politically infeasible, we find that CAFE has been providing safety cobenefits because it encourages down-weighting. Using the Department of Transportation's value of a statistical life, at $9.4 million, the welfare benefits of reduced fatalities, found in our simulations to range between 393 and 439 lives nationally, are more than $3.5 billion annually.

Thus, we can compare the monetized value of our results to the estimates the latest Regulatory Impact Analysis of CAFE/Greenhouse Gas Standards (https://www.epa.gov/regulations-emissions-vehicles-and-engines). We must be cautious with extrapolating to the new footprint-based standards, but we can note that vehicles with larger footprints have even lower fatalities, which has the potential to strengthen the results. The projected compliance cost of the new standard is roughly $1.5 billion annually, with an additional cost of $0.2 billion in driving related externalities.[33] Some prior studies by

[33] These values are for the lifetime of model year 2011 vehicles. While differences may arise when summing

economists have suggested higher compliance costs than those used by the EPA (Jacobsen, 2013a), while others have suggested lower costs (Anderson and Sallee, 2011).[34] Based on EPA's estimate of the compliance cost or even a cost as high as $200 per vehicle annually, our findings suggest that CAFE would pass a cost-benefit test based on benefits from reduced fatalities alone. This is particularly notable because the current EPA analysis does not include benefits or costs from fatalities, but relies heavily on myopic consumers realizing large future fuel savings to pass a cost-benefit test, which recent economic evidence brings into question (Busse, Knittel, and Zettelmeyer, 2013; Allcott and Wozny, 2014; Sallee, West, and Fan, 2016). Our results suggest CAFE standards may have positive net benefits even without any undervaluation of fuel economy.

References

Ahmad, Sanjana and David Greene. 2005. "Effect of Fuel Economy on Automobile Safety: A Reexamination." *Transportation Research Record* 1941:1–7.

Allcott, Hunt and Nathan Wozny. 2014. "Gasoline Prices, Fuel Economy, and the Energy Paradox." *Review of Economics and Statistics* 96 (5):779–795.

Anderson, Michael L and Maximilian Auffhammer. 2014. "Pounds that Kill: The External Costs of Vehicle Weight." *The Review of Economic Studies* 81 (2):535–571.

Anderson, Soren T and James M Sallee. 2011. "Using Loopholes to Reveal the Marginal Cost of Regulation: The Case of Fuel-Economy Standards." *The American Economic Review* 101 (4):1375–1409.

driving related external costs across all model years in one calendar year compared summing driving related external costs across calendar years for one model of vehicle, the compliance costs occur only in the year of production and will be the same under both cost calculations.

[34]For example, Anderson and Sallee (2011) estimate an upper bound of compliance costs of $55 per vehicle. Estimates in Jacobsen (2013a) range from $52 to $438 per vehicle depending on the vehicle fleet. At $55 per vehicle and an illustrative 15 million vehicles sold, the compliance costs are less than $1 billion. At $200 per vehicle, the compliance cots are $3 billion.

Auffhammer, Maximilian and Ryan Kellogg. 2011. "Clearing the Air? The Effects of Gasoline Content Regulation on Air Quality." *The American Economic Review* 101 (6):2687–2722.

Austin, David and Terry 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.

Ballesteros, Michael F, Patricia C Dischinger, and Patricia Langenberg. 2004. "Pedestrian Injuries and Vehicle Type in Maryland, 1995–1999." *Accident Analysis & Prevention* 36 (1):73–81.

Bento, Antonio M, Lawrence H Goulder, Mark R Jacobsen, and Roger H Von Haefen. 2009. "Distributional and Efficiency Impacts of Increased US Gasoline Taxes." *The American Economic Review* :667–699.

Borenstein, Severin and Nancy Rose. 1994. "Competition and Price Dispersion in the US Airline Industry." *Journal of Political Economy* 102 (4):653–83.

Busse, Meghan, Christopher Knittel, and Florian Zettelmeyer. 2013. "Are Consumers Myopic? Evidence from New and Used Car Purchases." *American Economic Review* 103 (1):220–256.

Carlson, Curtis, Dallas Burtraw, Maureen Cropper, and Karen L Palmer. 2000. "Sulfur Dioxide Control by Electric Utilities: What are the Gains from Trade?" *Journal of Political Economy* 108 (6):1292–1326.

Cohen, Alma and Liran Einav. 2003. "The Effects of Mandatory Seat Belt Laws on Driving Behavior and Traffic Fatalities." *Review of Economics and Statistics* 85 (4):828–843.

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.

Dahlby, Bev and Douglas S West. 1986. "Price Dispersion in an Automobile Insurance Market." *The Journal of Political Economy* :418–438.

DiNardo, John, Nicole M. Fortin, and Thomas Lemieux. 1996. "Labor Market Institutions and the Distribution of Wages, 1973-1992: A Semiparametric Approach." *Econometrica* 64 (5):1001–1044.

Fortin, Nicole, Thomas Lemieux, and Sergio Firpo. 2009. "Unconditional Quantile Regression." *Econometrica* 77 (3):953–973.

Gillingham, Kenneth, David Rapson, and Gernot Wagner. 2015. "The Rebound Effect and Energy Efficiency Policy." *Review of Environmental Economics & Policy* 10 (1):68–88.

Hampel, Frank R. 1974. "The Influence Curve and its Role in Robust Estimation." *Journal of the American Statistical Association* 69 (346):383–393.

Ito, Koichiro and James M Sallee. 2014. "The Economics of Attribute-Based Regulation: Theory and Evidence from Fuel-Economy Standards." Tech. rep., National Bureau of Economic Research.

Jacobsen, Mark R. 2013a. "Evaluating US Fuel Economy Standards in a Model with Producer and Household Heterogeneity." *American Economic Journal: Economic Policy* 5 (2):148–187.

Jacobsen, Mark R. 2013b. "Fuel Economy and Safety: The Influences of Vehicle Class and Driver Behavior." *American Economic Journal: Applied Economics* 5 (3):1–26.

Klier, Thomas and Joshua Linn. 2012. "New-vehicle Characteristics and the Cost of the Corporate Average Fuel Economy standard." *The RAND Journal of Economics* 43 (1):186–213.

———. 2016. "Technological Change, Vehicle Characteristics and the Opportunity Costs of Fuel Economy Standards." *Journal of Public Economics* forthcoming.

Knittel, Christopher R. 2011. "Automobiles on Steroids: Product Attribute Trade-Offs and Technological Progress in the Automobile Sector." *American Economic Review* 2012 (101):3368–3399.

Koenker, Roger and Gilbert Bassett. 1978. "Regression Quantiles." *Econometrica* :33–50.

Lach, Saul. 2002. "Existence and Persistence of Price Dispersion: An Empirical Analysis." *Review of Economics and Statistics* 84 (3):433–444.

Li, Shanjun. 2012. "Traffic Safety and Vehicle Choice: Quantifying the Effects of the Arms Race on American Roads." *Journal of Applied Econometrics* 27 (1):34–62.

Machado, José AF and José Mata. 2005. "Counterfactual Decomposition of Changes in Wage Distributions Using Quantile Regression." *Journal of Applied Econometrics* 20 (4):445–465.

Noland, Robert B. 2004. "Motor Vehicle Fuel Efficiency and Traffic Fatalities." *The Energy Journal* :1–22.

Peltzman, Sam. 1975. "The Effects of Automobile Safety Regulation." *The Journal of Political Economy* :677–725.

Pratt, John W, David A Wise, and Richard Zeckhauser. 1979. "Price Differences in Almost Competitive Markets." *The Quarterly Journal of Economics* :189–211.

Sallee, James, Sarah West, and Wei Fan. 2016. "Do Consumers Recognize the Value of Fuel Economy? Evidence from Used Car Prices and Gasoline Price Fluctuations." *Journal of Public Economics* 135:61–73.

Small, Kenneth A and Kurt Van Dender. 2007. "Fuel Efficiency and Motor Vehicle Travel: The Declining Rebound Effect." *The Energy Journal* :25–51.

Sorensen, Alan T. 2000. "Equilibrium Price Dispersion in Retail Markets for Prescription Drugs." *Journal of Political Economy* 108 (4):833–850.

Stavins, Robert N. 1998. "What Can We Learn from the Grand Policy Experiment? Lessons from SO2 Allowance Trading." *Journal of Economic Perspectives* 12 (3):69–88.

Wenzel, Tom. 2013. "The Estimated Effect of Mass or Footprint Reduction in Recent Light-Duty Vehicles on US Societal Fatality Risk per Vehicle Mile Traveled." *Accident Analysis & Prevention* 59:267–276.

White, Michelle J. 2004. "The 'Arms Race' on American Roads: The Effect of Sport Utility Vehicles and Pickup Trucks on Traffic Safety." *Journal of Law and Economics* 47 (2):333–355.

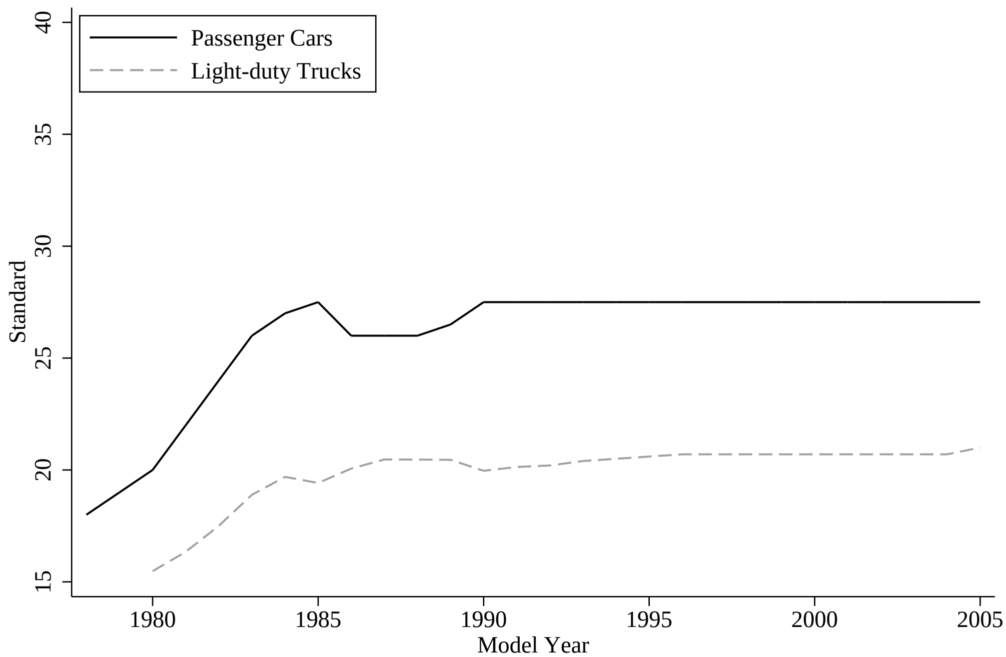
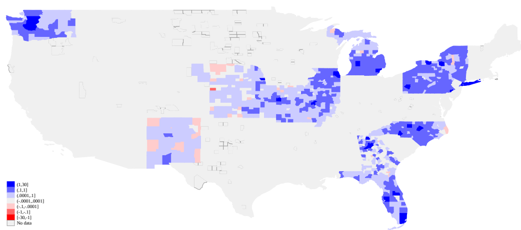


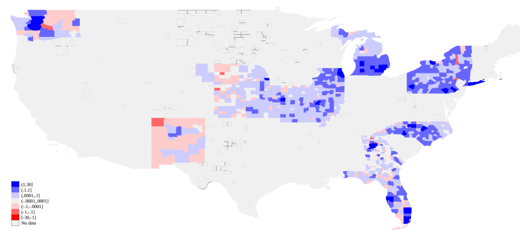
Figure 1: Historical Fuel Economy Standards (miles-per-gallon)

Panel A



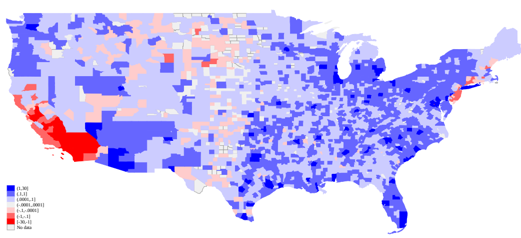
SDS States
Remove CAFE for Domestic Firms

Panel B



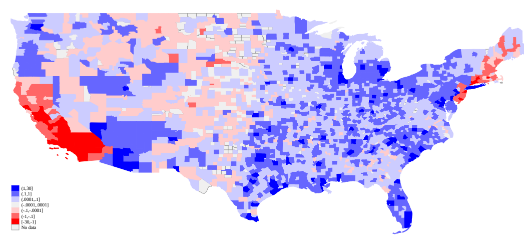
SDS States
Remove CAFE for All Firms

Panel C



Impute Nationally
Remove CAFE for Domestic Firms

Panel D



Impute Nationally
Remove CAFE for All Firms

Figure 2: Predicted Fatality Changes from Removing CAFE in 2005

		Trim Level							
	Years	1970-1974	1975-1977	1978-1979	1980-1984	1985-1989	1990-1994	1995-1999	2000-2004
All Vehicles	Mean	3755.2	3602.3	3294.3	3034.8	2962.7	3273.2	3573.8	4020.1
	St.Dev.	(634.8)	(795.8)	(654.5)	(663.5)	(685.3)	(731.3)	(857.5)	(1065.6)
	N	5803	3649	2857	5339	6680	7291	6741	9014
All Cars	Mean	3753.3	3600.6	3262.3	2952.6	2763.4	2977.4	3085.0	3226.9
	St.Dev.	(634.0)	(802.1)	(640.6)	(624.8)	(533.8)	(544.1)	(543.5)	(513.0)
	N	5608	3378	2611	4580	4870	4614	3515	3597
All Light Trucks	Mean	3811.2	3622.8	3633.6	3530.9	3498.8	3783.0	4106.5	4546.8
	St.Dev.	(658.2)	(714.7)	(704.1)	(674.3)	(756.7)	(731.9)	(818.9)	(1009.8)
	N	195	271	246	759	1810	2677	3226	5417
Domestic Vehicles	Mean	3792.7	3695.2	3394.0	3141.3	3085.8	3437.4	3802.9	4370.9
	St.Dev.	(595.5)	(734.9)	(579.7)	(631.0)	(720.6)	(754.3)	(884.0)	(1114.1)
	N	5641	3362	2553	4420	4496	4539	4288	5594
Ford	Mean	3779.5	3734.8	3437.7	3052.3	3222.9	3493.4	3921.8	4469.3
	St.Dev.	(591.8)	(835.1)	(784.1)	(629.6)	(761.5)	(782.1)	(854.4)	(1214.4)
	N	1564	689	448	1019	992	1265	1324	2168
GM	Mean	3888.3	3615.3	3394.0	3259.6	3137.0	3524.2	3814.6	4366.5
	St.Dev.	(603.9)	(744.2)	(500.7)	(648.2)	(774.8)	(806.1)	(881.4)	(1075.4)
	N	2240	1253	1127	1875	2071	2100	1970	2589
Chrysler	Mean	3779.3	3956.6	3543.5	2954.6	2913.0	3239.5	3574.4	4001.0
	St.Dev.	(577.7)	(653.2)	(568.7)	(657.3)	(574.8)	(562.5)	(759.6)	(806.7)
	N	1339	924	544	827	1259	1280	1119	1326
Asian Vehicles	Mean	2217.9	2222.3	2200.7	2272.2	2594.6	2885.5	3036.9	3360.1
	St.Dev.	(368.2)	(422.3)	(498.2)	(420.0)	(467.1)	(549.4)	(580.7)	(674.4)
	N	90	130	168	512	1568	2077	1827	2335
European Vehicles	Mean	2735.2	2756.2	2773.3	2837.6	3001.2	3361.8	3571.9	3631.4
	St.Dev.	(652.1)	(731.0)	(673.0)	(580.6)	(539.3)	(608.7)	(618.4)	(577.9)
	N	72	157	136	407	616	675	626	1085

Table 1: Summary Statistics for Vehicle Weight (in lbs.) by Year

Panel A: Fuel Economy Regressions[a]

	I	II	III	IV	V	VI
Regressand	OLS	Q10	Q25	Q50	Q75	Q90
$\frac{\sum_{i=-1}^{3} S_{t-i}}{5}$	-0.299***	-0.536***	-0.553***	-0.304***	-0.222***	-0.214***
	(0.027)	(0.049)	(0.059)	(0.054)	(0.048)	(0.044)
$l(Gasoline\ Price_{t-1})$	-0.284***	-0.338***	-0.555***	-0.422***	-0.164***	-0.114***
	(0.014)	(0.023)	(0.026)	(0.028)	(0.027)	(0.024)
$l(GDP_{t-1})$	-0.292***	0.175	-0.209	-0.105	-0.441**	-0.438**
	(0.082)	(0.147)	(0.155)	(0.151)	(0.155)	(0.135)
Weight in Quantile[c]	3,023.4	2,280.0	2,582.0	3,001.0	3,429.0	3,803.0
Weight Change for 1MPG[c]	-32.8	-44.4	-51.9	-33.2	-27.7	-29.6
R-squared	0.132	0.064	0.109	0.126	0.073	0.048
N	15,021	15,021	15,021	15,021	15,021	15,021

Panel B: CAFE Level Regressions[b]

	I	II	III	IV	V	VI
Regressand	OLS	Q10	Q25	Q50	Q75	Q90
$l(CAFE)$	-0.305***	-0.152	-0.436**	-0.624***	-0.341*	-0.183
	(0.075)	(0.141)	(0.147)	(0.142)	(0.135)	(0.109)
$l(Gasoline\ Price_{t-1})$	-0.132***	-0.192***	-0.314***	-0.166**	-0.013	-0.016
	(0.028)	(0.050)	(0.056)	(0.055)	(0.051)	(0.041)
$l(GDP_{t-1})$	-0.183*	0.438***	0.014	-0.048	-0.379*	-0.355**
	(0.082)	(0.123)	(0.152)	(0.170)	(0.155)	(0.132)
Weight in Quantile[c]	3,013.7	2,271.0	2,568.0	2,989.0	3,419.0	3,792.0
Weight Change for 1MPG[c]	-33.4	-12.6	-40.7	-67.8	-42.4	-25.2
R-squared	0.126	0.056	0.103	0.125	0.072	0.047
N	15,021	15,021	15,021	15,021	15,021	15,021

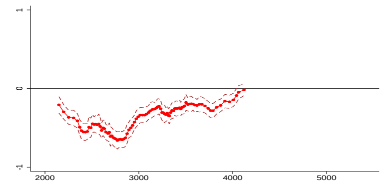
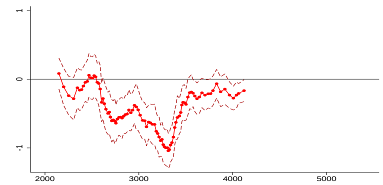
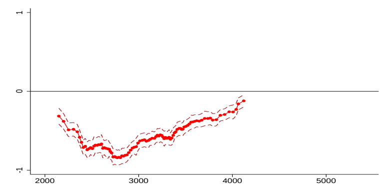
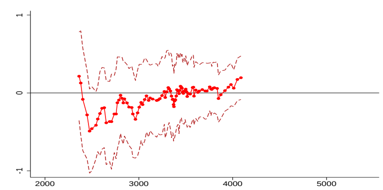
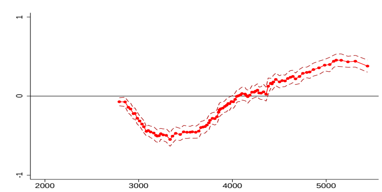
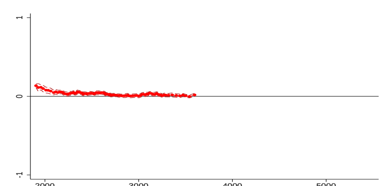
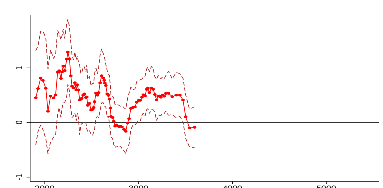
Notes: All regressions include a quadratic trend and automaker fixed effects. Bootstrap standard errors in parentheses with * indicating significance at 5%, ** at 1%, and *** at >1%.

[a] Predicted Fuel Economy Regressions use 1978 to 2005. To predict the fuel economy without CAFE for a given fleet (S_{t-i}) the harmonic mean fuel economy of the fleet produced prior to CAFE is regressed on a trend, the gasoline price, and GDP. These coefficients are then used to predict the fuel economy level after the introduction of the CAFE standard. This is done separately for each automaker and each fleet (cars vs. light trucks).

[b] CAFE Level Regressions use model years 1978 to 2005.

[c] Weight in pounds.

Table 2: Effects for Domestic Cars

	Coefficients	Segment	Model Years	Displayed Coefficient	Sales Weighted
		DOMESTIC FIRMS			
1		Cars	1978-2005	$\frac{\sum_{i=-1}^{3} S_{t-i}}{5}$	No
2		Cars	1978-2005	$l(CAFELevel_t)$	No
3		Cars	1978-2005	S_{t-1}	No
4		Cars	1978-2005	$\frac{\sum_{i=-1}^{3} S_{t-i}}{5}$	Yes
5		Trucks	1978-2005	$\frac{\sum_{i=-1}^{3} S_{t-i}}{5}$	No
		ASIAN FIRMS			
6		Cars	1978-2005	Own $\frac{\sum_{i=-1}^{3} S_{t-i}}{5}$	No
7		Cars	1978-2005	Average US $\frac{\sum_{i=-1}^{3} S_{t-i}}{5}$	No

Notes: All regressions include a quadratic trend and automaker fixed effects. Bootstrap standard errors are calculated using 200 repetitions. 95% confidence intervals are indicated by dashed lines. To predict the fuel economy without CAFE for a given fleet (S_{t-i}) the harmonic mean fuel economy of the fleet produced prior to CAFE is regressed on a trend, the gasoline price, and GDP. These coefficients are then used to predict the fuel economy level after the introduction of the CAFE standard. This is done separately for each automaker and each fleet (cars vs. light trucks).

Table 3: Graphical Regression Results

Panel A: SDS States, Remove CAFE for Domestic Firms				
	1-car	2-car	3-car	Total
Change Dispersion Only	0.0	-46.1	-6.3	-52.4
Change Mean Weight Only	141.3	61.6	20.5	223.5
Total Change	141.3	15.4	14.2	171.0
Panel B: SDS States, Remove CAFE for All Firms				
	1-car	2-car	3-car	Total
Change Dispersion Only	0.0	128.4	28.0	156.4
Change Mean Weight Only	20.4	-3.4	-2.0	15.0
Total Change	20.4	125.0	25.9	171.3
Panel C: Impute to National Level, Remove CAFE for Domestic Firms[a]				
	1-car	2-car	3-car	Total
Change Dispersion Only	0.0	-100.0	-10.5	-110.4
Change Mean Weight Only	342.7	155.3	51.6	549.7
Total Change	342.7	55.4	41.1	439.2
Panel D: Impute to National Level, Remove CAFE for All Firms[a]				
	1-car	2-car	3-car	Total
Change Dispersion Only	0.0	395.4	87.4	482.8
Change Mean Weight Only	-20.7	-49.7	-19.3	-89.7
Total Change	-20.7	345.7	68.1	393.1

Notes: Calculates change in fatalities if CAFE is removed and the stringency is set to zero. Positive values indicate that CAFE saved lives in 2005.
[a] Because the states with SDS accident data tend to be Eastern and Midwest we impute a percent change in fatality rate based on county population, total fatalities, and state level average vehicle weight, age, share light truck, and share US and Asian manufacturer. This imputation tends to reduce lives saved because western counties have low population density and have a larger share Asian manufacturer.

Table 4: Simulation of Fatalities without CAFE 2005