Empirical Economics (2021) 60:2063–2081
https://doi.org/10.1007/s00181-020-01826-2

A sturdy values analysis of motor vehicle fatalities

Richard Fowles[1] · **Peter D. Loeb**[2]

Received: 23 May 2019 / Accepted: 23 January 2020 / Published online: 3 February 2020

Abstract
Understanding the major determinants of crash fatalities continues to be an important topic of investigation for safety researchers. Regression models using a vast number of explanatory variables are often used which result in a huge array of specifications. Results often vary among studies based on size of estimated coefficients and significance levels. To address this, we explore both significance and model sturdiness in regression models using Leamer's *s*-values. This Bayesian technique allows us to address estimation uncertainty and model ambiguity over all possible subset regressions so as to evaluate the effect of key variables which we focus on as contributors to crash fatalities. These include cell phone use, fleet modernization, suicidal behavior, alcohol use, and speed limits.

Keywords Motor vehicle fatalities, motor vehicle crashes, sturdy inference · *S*-values · Bayesian econometrics · Cell phones, vehicle safety · Suicide, median car age, speed, alcohol

An earlier version of this paper was presented at the meetings of the American Economic Association and the Transportation and Public Utilities Group in Atlanta, GA, on January 6, 2019. The authors are grateful to Wayne Talley and two anonymous referees for insightful comments.

✉ Richard Fowles
richard.fowles@utah.edu

[1] Department of Economics, University of Utah, 260 S. Central Campus Drive, Gardner Commons 4100, Salt Lake City, UT 84112, USA

[2] Department of Economics, Rutgers University, 360 Dr. Martin Luther King Jr. Blvd, Hill Hall 813, Newark, NJ 07102, USA

1 Introduction

In 1972, there were 54,589 fatalities attributed to motor vehicle crashes. By the year 2014, these numbers had declined to 32,744 deaths. As impressive as this decline is, the number of such deaths is still quite significant. A major reduction in these fatalities occurred between 2007 and 2010 when fatalities fell by 20%.[1] This precipitous drop in deaths may be due to things other than, or in addition to, what contributed to the prior trend. Some of this recent decline may be due to economic events such as the Great Recession as well as to changes in tastes among the public. For example, youth may have preferences to move from the suburbs to the cities where there is a greater reliance on public transportation and walking. While the baby boomers had a strong desire to obtain powerful cars when in their teens, the current population of youths may be more inclined to desire powerful cell phones. Independent of these changes in preferences, there remains a significant number of crash-related fatalities in the USA which scientists attempt to explain.

Most troubling has been the increase in traffic-related fatalities since 2010. Deaths have increased by over 13.5% from 32,999 in 2010 to 37,461 in 2016.[2] Understanding the causes of these deaths is a matter of concern to economists, public health scientists and policy makers. Determining the significance of the contributing factors of these crashes resulting in fatalities can lead to policy and technical developments that may mitigate these losses.

Many factors thought to contribute to motor vehicle crashes and crash fatalities have been examined over the last few decades. These factors can be classified into three categories: those associated with vehicles, those associated with drivers (and pedestrians), and those associated with roadways. The potential contributing variables that have been investigated in the past include: motor vehicle speed, speed variance, alcohol consumption, speed limits, vehicle miles travelled, measures of income and wealth, the age of the motor vehicle fleet, seat belt usage and seat belt legislation, and the deregulatory climate of the 1980s, among others.[3] More recently, there has been an interest in the effect of cell phones, the age of the motor vehicle fleet, and the trend in suicides on these fatalities along with continued concern regarding alcohol consumption.[4] Most often, these effects have been investigated using regression models similar to those posed early on by Peltzman (1975). We provide an overview of a myriad of explanatory factors related to fatal crashes in Sect. 2.[5]

That so many explanatory factors are utilized in a regression gives rise to issues of selective model specification, in particular, the choice of factors to include. A common presentation style is to select a few key focus variables and always have

[1] See NHTSA (2016).

[2] See NHTSA (2016).

[3] See Loeb et al. (1994) for a more complete list and discussion of these contributing influences.

[4] See Blattenberger et al. (2013).

[5] An extensive review of the literature regarding the explanatory factors can be found in Loeb et al. (1994) and more recently in Fowles and Loeb (2019) and Blattenberger et al. (2013).

statistics for these variables reported across a few other regressions that include other possible explanatory factors. It is not atypical to find that there are twenty or more candidate regressors that can be moved in or out of the model. Out of millions of possible models, only a few are presented. Early work by Leamer (1978, 1982, 1983) exposed some problems with selective reporting and global sensitivity analysis. Extreme bounds analysis (EBA) was a result of this research which provided a Bayesian approach to address these issues.

In Sect. 3, we discuss what EBA accomplishes and demonstrate some downsides that arise when looking at global bounds in a regression context. As a preview, when using EBA it is possible to devise a model such that any regression parameter could be negative or positive. Fragility (manifest in coefficient sign switching) is an inherent feature of regression modeling and devising sensible ways to bound model estimates but retain the benefits of global sensitivity analysis has been an aspect of ongoing econometric research.[6] Because global bounds are wide, Section 3 introduces a new summary statistic, *s*-values, from Leamer (2016). *s*-values can be used to summarize both meaningful significance and specification sturdiness. Because *s*-values are Bayesian in nature, we discuss how one might view a regression in the context of prior beliefs about the regression problem. This technique can improve our understanding of causal models beyond that provided from purely a classical regression approach.

In Sect. 4, we summarize sets of regressions using *s*-values based on a rich US panel dataset that is largely based on variables discussed in Sect. 2. Our empirical findings provide insight as to how fatality rates in the USA can be falling at the same time as cell phone use among drivers is rising. The results support the nonlinear effect of cell phone use on crash fatalities as initially indicated by Loeb et al. (2009) and Fowles et al. (2010). Results also show that alcohol use, suicidal behaviors, and speed limits are dominant factors relating to vehicle fatalities. In Sect. 5, we highlight the practical significance of important variables. Section 6 concludes with suggestions for policies at the state or federal level.

2 A myriad of possible factors

Models developed to explain motor vehicle crash fatalities have mostly followed the approach suggested by Peltzman (1975) using classical regression. As mentioned, the common list of determining factors in these models include: measures of speed, speed variance, alcohol use, seatbelts and airbags, and laws requiring their use, income, education, population attributes, among others. The rationale for including these variables and their effects on crash fatalities have been reviewed fairly extensively in Loeb et al. (1994) and more recently in Simmons et al. (2016), Fowles and Loeb (2019), Blattenberger et al. (2013), Fowles et al. (2010), and Loeb et al. (2009). What is particularly noteworthy is that the empirical findings as to the effects of some of these factors varied greatly among the studies leading to uncertainty and

[6] See for example, Fowles et al. (2010).

ambiguity with respect to the models and parameter estimates reported. Bayesian methods offer a way to address these problems.

Besides the potential determining factors listed above, several additional variables have been more recently considered of great importance, or renewed importance, in understanding the determinants of crash fatalities. These include: the effect of fleet modernization, cell phone use, and suicidal propensities. The effects of alcohol and speed continue to be of major concern as well. These last five factors are what this study primarily focuses on. The entire list of causal factors which this study concerns itself with is outlined in Table 1 along with their expected effects on crash fatalities. All of these factors have been well reviewed in the past, but three of our five "focus variables" deserve further elaboration given the more recent attention they have received.

Recently, safety researchers have been addressing the impact of cell phones on crash fatalities. It is argued that cell phone usage leads to crashes for several reasons. Cell phones are considered to have a distracting effect on drivers and to diminish attention spans and increase reaction time. Furthermore, the number of cell phone subscribers, and hence usage rate, has increased exponentially over the last two decades. More specifically, the number of cell phone subscribers has increased from about 340 thousand in 1985 to over 310 million in 2010.[7] Not only has the number of cell phones increased dramatically over time, but the propensity of drivers to use them has also increased. Glassbrenner (2005) has estimated that 10% of all drivers are using a cell phone while operating their vehicles during daylight hours. Currently, sixteen states plus the District of Columbia have banned the use of hand-held phones by drivers (California, Connecticut, Delaware, Georgia, Hawaii, Illinois, Maryland, Nevada, New Hampshire, New Jersey, New York, Oregon, Rhode Island, Vermont, Washington, and West Virginia).[8]

Statistical evaluations of the effects of cell phones on crashes have not always provided consistent results. Papers finding cell phone effects contributing to crash fatalities, injuries, or property damage include: Redelmeier and Tibshirani (1997), Violanti (1998), Consiglio et al. (2003), McEvoy et al. (2005), Beede and Kass (2006), Neyens and Boyle (2007), and Welki and Zlatoper (2017). However, others did not find such life-taking effects due to cell phones. These included studies by: Laberge-Nadeau et al. (2003), Sullman and Baas (2004), and Poysti et al. (2005). Chapman and Schoefield (1998) even attribute to cell phones a life-saving effect.[9]

These convoluting results led to studies by Loeb et al. (2009), Loeb and Clarke (2009), and Welki and Zlatoper (2014) which addressed the different results found in prior studies using classical econometrics and a nonlinear model. These models included cell phone usage measured as a polynomial of degree three. The models by Loeb et al. (2009) and Loeb and Clarke (2009) had been subjected to specification

[7] See CTIA (2011). By December 2015, the number of wireless subscriber connections rose to 377.9 million as per CTIA (2016).

[8] Strangely, the bans do not include hands-free devices even though research indicates that such devices have a similar adverse effect. See for example, Consiglio et al. (2003).

[9] See Fowles and Loeb (2019) and Blattenberger et al. (2013) for a full review of these papers.

Table 1 Explanatory variables

Name	Description	Expected sign
MCMY	Median car model year divided by 1000	–
CELLPOP	Number of cell phone subscribers per capita	+
CELLSQ	The square of CELLPOP	–
CELLCUBE	The cube of CELLPOP	+
SPEED	Maximum posted speed limit, urban interstate highways (MPH)	+
PERSE	Dummy variable indicating the existence of a law defining intoxication of a driver in terms of blood alcohol concentration (BAC). PERSE = 1 indicates the existence of such a law, and PERSE = 0 indicates the absence of such a law. (More precisely, PERSE = 1 when the BAC indicating driving under the influence is 0.1 or lower.)	–
BEER	Per capita beer consumption (in gal) per year	+
MLDA	Minimum legal drinking age (years)	–
BELT	Dummy variable for presence of a legislated seat belt law (1 = presence)	–
EDHS	Percent of persons with a high school diploma	–
EDCOLL	Percent of persons with a college degree	–
REALINC	Real per household income in 2000 dollars	+/–
YOUNG	Percentage of males (16–24) relative to population of age 16 and over	+
SUICIDE	Suicide rate (suicides per 100,000 population)	+/–

For data sources, see "Appendix 1" and for descriptive statistics, see "Appendix 2"

error tests, including the regression specification error test (RESET) developed by Ramsey and Zarembka (1971) and Ramsey (1974) which tests for misspecification of the structural form of the regressors. RESET supported the nonlinear specification. Strong evidence was found that cell phone use was associated with a life-taking effect in vehicle crashes. Blattenberger et al. (2012) and Fowles et al. (2010) also demonstrated a life-taking effect of cell phones on crash fatalities using Bayesian methods.[10] Most notably, Fowles et al. (2010), using Bayesian extreme bounds analysis, found a nonlinear model in terms of cell phone usage to be non-fragile in explaining motor vehicle crash fatalities. As such, this nonlinear specification is the Bayesian "prior" for the current study.

The age of the fleet has also been of renewed interest as of late due to the advances in technology.[11] This is of particular interest with the ongoing development of autonomous vehicles. Should the technology and legal issues be successfully addressed, modernization of the fleet is anticipated to reduce crash-related fatalities. This modernization leading to the overwhelming use of autonomous vehicles may prove particularly beneficial in addressing issues dealing with distracted drivers on cell phones and other electronic devices, drivers under the influence of

[10] See also Fowles et al. (2013).

[11] A time trend was also considered as a potential regressor in the model. However, our measure of the age of the fleet (MCMY) was found to be almost perfectly correlated with a time trend, with a correlation coefficient of 0.99962.

alcohol or other drugs who would otherwise be driving, and those dealing with disabilities, among others. However, prior studies of the effect of fleet modernization have resulted in mixed results regarding their beneficial effects to date.[12]

Suicidal propensities, as measured by suicide rates, are not commonly used as a determinant of crash fatalities. However, there is some evidence that there is a relationship, but these studies give dissimilar results. Those associated with a positive relationship between suicidal propensities and crash fatalities include: Porterfield (1960), Pokorny et al. (1972), Phillips (1977, 1979), Conner et al. (2001), Murray and DeLeo (2007), and Blattenberger et al. (2012). Those studies that did not find a uniformly strong relationship include, e.g., those by: Huffine (1971), Souetre (1988), and Connolly et al. (1995). In addition, Etzerdorfer (1995) indicates the difficulty in determining whether a crash victim was actually a suicide.

The suicide–automobile crash link is controversial with mixed empirical findings. Yet, we consider it of interest not only in serving as a measure of societal risk taking propensities, but also in taking cognizance that it may minimize the stigma associated with attempting suicide since it may appear that such deaths were due to traditional traffic crashes. In addition, suicides committed by motor vehicles may have potential economic effects for the families of the deceased due to insurance benefits.

As mentioned, aside from our focus variables, there are many other potential contributors to crash fatalities. A few factors, such as alcohol, have been consistently found to be major determining factors. Others, such as income and suicidal behavior, have been found to have fragile effects. Hence, over time, various factors have been added to and others deleted from models of crash fatalities. Table 1 provides a reasonable set of variables to consider in models of crash fatalities based on economic theory, previous investigations, or intrinsic Bayesian priors. Many other factors could be explored as well in what may be posited as the "true model" of crash fatalities

Importantly, there is no one model. How one selects a model, or a group of models, as true, is often ambiguous at best. For these reasons, explorations are made in this paper over many models with the goal to provide statistically reliable and inferentially meaningful results. Such a statistical procedure must account for both model ambiguity and the uncertainty associated with parameter estimation. This is the focus of the current paper which is conducted making use of a new Bayesian econometric method developed by Leamer (2016) which is ideally suited to deal with the problem of model ambiguity and parameter uncertainly. The next section provides an overview of this method.

3 Bayesian global bounds and *s*-values

Although it is common to indicate regression results for a variety of model specifications, reported statistics are valid on the presumption of a given model's truth. In practice, alternative tests are made on competing models, each sequentially assumed

[12] See Blattenberger et al. (2013) and Loeb et al. (1994) on this issue.

Table 2 Coefficients, bounds, and *s*-values

	Extreme bounds wide		Extreme bounds pessimistic		Extreme bounds optimistic		Wide	Pessimistic	Optimistic	OLS simple	OLS all	*t* stat all
	Lower	Upper	Lower	Upper	Lower	Upper	*s*-value	*s*-value	*s*-value	Coefficient	Coefficient	*t*
MCMY	**−0.904**	**−0.390**	**−0.809**	**−0.396**	**−0.858**	**−0.757**	**−2.515**	**−2.921**	**−15.922**	**−0.749**	**−0.968**	**−14.953**
PERSE	**−0.090**	**−0.059**	**−0.089**	**−0.062**	**−0.073**	**−0.069**	**−4.725**	**−5.656**	**−29.750**	**−0.424**	**−0.068**	**−5.425**
SPEED	**0.033**	**0.076**	**0.035**	**0.071**	**0.056**	**0.062**	**2.509**	**2.912**	**17.281**	−0.122	0.064	**4.770**
BELT	−0.037	0.040	−0.036	0.028	**0.014**	**0.028**	0.041	−0.117	**3.015**	−0.591	0.040	**2.302**
BEER	**0.113**	**0.152**	**0.117**	**0.150**	**0.127**	**0.133**	**6.699**	**8.008**	**42.228**	**0.365**	**0.126**	**9.799**
MLDA	−0.048	0.004	**−0.046**	**−0.004**	**−0.016**	**−0.007**	−0.857	**−1.171**	**−2.558**	−0.480	−0.001	−0.036
YOUNG	−0.026	0.078	−0.009	0.076	−0.011	0.008	0.498	0.780	−0.164	0.687	−0.028	−1.374
CELLPOP	**0.176**	**0.628**	**0.181**	**0.545**	**0.498**	**0.587**	**1.775**	**1.999**	**12.163**	−0.641	0.684	**11.137**
CELLSQ	**−0.091**	**−0.065**	**−0.089**	**−0.067**	**−0.083**	**−0.080**	**−6.041**	**−7.053**	**−42.557**	**−0.511**	**−0.084**	**−3.080**
CELLCUBE	**0.006**	**0.020**	**0.008**	**0.020**	**0.008**	**0.011**	**1.800**	**2.400**	**6.813**	−0.344	0.005	0.954
REALINC	−0.153	0.019	−0.137	0.008	**−0.085**	**−0.057**	−0.776	−0.894	**−5.135**	**−0.763**	**−0.080**	**−2.601**
EDHS	**−0.395**	**−0.284**	**−0.386**	**−0.292**	**−0.356**	**−0.340**	**−6.152**	**−7.271**	**−41.225**	**−0.775**	**−0.349**	**−14.515**
EDCOL	**−0.196**	**−0.063**	**−0.186**	**−0.074**	**−0.137**	**−0.116**	**−1.947**	**−2.329**	**−12.110**	**−0.728**	**−0.123**	**−4.767**
SUICIDE	**0.186**	**0.227**	**0.188**	**0.222**	**0.210**	**0.216**	**10.254**	**12.030**	**70.812**	**0.237**	**0.216**	**16.164**

s-values are obtained by simply taking the ratio of sum of the lower and upper bounds to the difference between them. Boldface type indicates non-fragile bounds, conformity between the OLS simple and OLS all coefficients, *s*-values greater than one in absolute value, and *t* statistics greater than two in absolute value

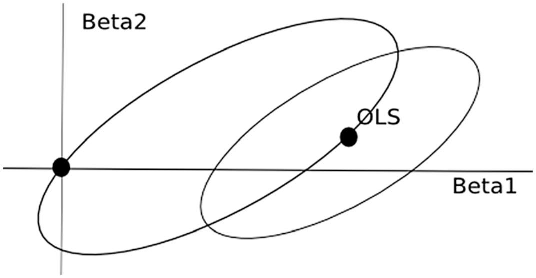

Fig. 1 Illusration of Bayesian global bounds

to be true. Inferences based on sequential search procedures are fraught with problems regarding the statistical validity of models' reported summary statistics. Bayesian theory, however, can directly address both estimation uncertainty and model ambiguity. In a book and series of articles, Leamer (1978, 1982, 1983) explored specification searches and introduced extreme bounds analysis (EBA). It is a methodology of global sensitivity analysis that computes the possible maximum and minimum values for Bayesian posterior means in the context of linear regression models of the form: $Y = X\beta + \mu$.[13] The matrix X contains the data on the fourteen variables listed in Table 2 along with the eleven regional variables.

The region of possible posterior means with a prior location of zero for the β parameters is a function of the prior variance. Under very minimal assumptions for prior variance, the possible region of posterior means is bounded. These global bounds are illustrated in Fig. 1 for a two variable regression model. A typical likelihood ellipse is centered at the OLS estimate. The other ellipse is the same shape and passes through the origin (the prior mean) and the OLS estimate. It contains the set of possible posterior means that could be obtained for all prior variance matrices that are positive definite. This larger ellipse is called the feasible ellipse and highlights a main drawback of EBA: that bounds cover zero. In this example, only the second quadrant (negative Beta1 and positive Beta2) is excluded as a joint region that could contain the posterior mean. Marginally, both Beta1 and Beta2 are fragile in the sense that there exist prior variance matrices that could result in negative or positive posterior estimates for either variable.

That the coefficients for all variables in a regression are necessarily fragile from a global EBA perspective highlights the importance of the prior variance. To narrow these bounds, we incorporate a new perspective on the prior variance developed by Leamer (2014, 2016). *S*-values reveal aspects of parameter fragility for minimally specified prior variance matrices.

Figure 2 illustrates how *s*-values are obtained for a two variable regression problem. As in Fig. 1, we plot typical likelihood ellipses that are centered at the OLS estimate. There are also two circles centered at the origin that represent two iso-prior probability contours that would result from a prior that is centered at zero

[13] The target of interest in a Bayesian regression setting is the posterior means for the β parameters. Mathematical developments for EBA are found in Leamer (1982).

with spherically symmetric prior variances. The points of tangencies trace the posterior mean from zero to the OLS point. From a non-Bayesian perspective, this is exactly the ridge regression trace (Hoerl and Kennard 1970). If the prior variance increases, the posterior mean will fall closer to the OLS point and if the prior variance decreases, the mean falls closer to the origin. Two middle points are associated with two values of the prior variance. These values translate to prior *R*-squares (variances).[14] The larger prior *R*-square gives more weight to the explanatory variables in the model, and thus, the trace is closer to the OLS point. In Fig. 2, there is also a shaded ellipse that contains the possible posterior means associated with all linear combination of the two explanatory variables. Here, notice that the limits for Beta2 are fragile, but that the limits for Beta1 are unambiguously positive. The extreme values for means within such an ellipse form the basis for *s*-values which are computed as the midpoint of the extremes divided by half their length.

As suggested by Leamer (2014, 2016), useful prior *R*-squares are associated with values of 0.1–1 (wide), 0.1–0.5 (pessimistic), and 0.5–1 (optimistic). A pessimistic belief is that the explanatory variables would not account for much of the variation in the dependent variable, whereas an optimistic belief is that they do and thus the prior defers to the data.[15] In the next section, we see how these priors are used in regards to US vehicle fatality rates.

4 Statistical results using *s*-values

Here, we estimate *s*-values using a newly created set of data collected on 50 states and Washington, DC over the period from 1980 to 2014. The number of fatalities per 100 million vehicle miles travelled is our dependent variable. Our choice of explanatory variables is based on literature reviewed in Sect. 2, highlighting the importance of policy, safety, demographic, and economic determinants of fatality rates. Issues related to the choice of these variables, as well as the general form of the models, are well described in Blattenberger et al. (2009, 2012, 2013), Fowles et al. (2010), and Loeb et al. (2009). These form the basis for the regression model in this paper. All numeric variables are used in standardized form for two reasons: standardized regression coefficients are preferred in the field of statistics to aid in interpretations (Gelman, (2008)) and standardized data provide a mechanism to develop a meaningful Bayesian prior as discussed in the previous section.

Our data cover years during which there were significant changes in several important variables that are a priori plausible predictors of fatalities. A new series includes the median age of cars in states to account for the fact that some fleet ages are older (for example in Montana) and some ages are newer (for example in

[14] The Bayesian natural conjugate model sets the prior variance for the β's = var(β) = $v^2 I_{k\times k}$ where $I_{k\times k}$ is the k by k identity matrix. Bounds are obtained via the scalar v^2 which is set to the minimum or maximum expected *R*-square divided by k, see Leamer (2014, 2016) for details. Calculations are performed in the software R (R Development Core 2016). Also see Fowles (1988).

[15] A super pessimist prior is to exclude a variable from a regression, so the prior mean is at zero and the variance is zero as well (prior *R*-square zero). In this paper, we do not consider this kind of strict prior.

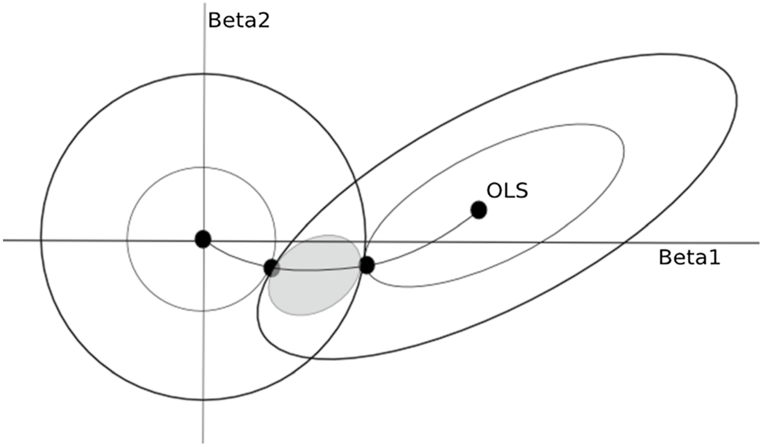

Fig. 2 Illustration of s-value bounds

Connecticut). Over the span of our data, this series is designed to quantify improvements in automobile safety. The data also capture the explosive growth in cell phone subscriptions from effectively zero to over 300 million. Annual subscription data at the state level were only available beginning in year 2000. For the earlier years, we used national level data and imputed state-level subscriptions to be proportional to state population proportions for the prior years.[16] Another major change observed in the data relates to changes in Federal law that allowed individual states to modify the 55 mile per hour speed limit on their interstate highways. Our data record the highest posted urban interstate speed limit that was in effect during the year for each state. Within the data, per se blood alcohol concentration (BAC) laws vary widely, even though by 2005 all states and the District of Columbia had mandated a 0.08 BAC illegal per se law.[17] Seat belt legislation varies widely across states. Our data record the years in which a state mandatory primary or secondary seat belt law came into effect. Additionally, the data allow us to examine the effects of education, income, and youth on motor vehicle fatality rates. These are of particular interest given the literature overview in Sect. 2. However, we concentrate, in particular, on four variables: cell phones, age of the fleet as measured by the median car model year, and suicidal propensities, along with alcohol consumption. The data are organized by geographic coding of states into eleven regions. (These eleven regions are US Federal standardly defined geographic areas with the exception of our separating Alaska and Hawaii since they are non-contiguous states.[18]) The variables are defined and described in Table 1 along with their expected effects on fatality rates.[19]

[16] Our method of imputing cell phone subscriptions correlates with the actual data with a correlation coefficient of 0.9943.

[17] The per se law refers to legislation that makes it illegal to drive a vehicle at a blood alcohol level at or above the specified BAC level. BAC is measured in grams per deciliter.

[18] See Office of the Federal Register (1980).

[19] See Blattenberger et al. (2009).

Table 2 summarizes the findings for our variables of interest for a linear model based on standardized data.[20] The column "Simple" (column 11) regresses the fatality rate on only the one specified explanatory variable and measures the pairwise correlation between the two variables. Leamer argues that these simple correlations "are a feature of the data, while the 'partial' regression coefficients are cooked up by the analyst when he or she selects the control variables."[21] A different sign in the simple correlation and the partial correlation (column 12) then, "requires scrutiny."[22] It is here that *s*-values are particularly useful.

Columns 2 through 7 provide lower and upper extreme values for the three specified prior variances. Columns 2 and 3 provide the extreme values for the wide prior (prior *R*-squared from 0.1 to 1); columns 4 and 5 the pessimistic prior (prior *R*-squared from 0.1 to 0.5); and columns 6 and 7 the optimistic prior (prior *R*-squared from 0.5 to 1). *s*-values are provided in columns 8 through 10. Column 8 provides *s*-values associated with the wide prior; column 9 the values associated with the pessimistic prior; and column 10 with the optimistic prior. Column 11 provides the "simple" regression result, while columns 12 and 13 provide the OLS coefficients and associated *t* values.

The bold cells in Table 2 highlight aspects of model and parameter uncertainty/ambiguity.[23] There are seven variables, i.e., MCMY, PERSE, BEER, CELLSQ, EDHS, EDCOLL, and SUICIDE for which all statistics are in conformity. For these variables, the signs of parameters are always the same, the absolute values of the *s*-values are always greater than one (in absolute value), and the absolute values of the *t* statistics are greater than two. These seven variables exhibit the highest level of sturdiness. CELLPOP and CELLSQ show sturdiness on the basis of *s*-values, conformity of *s*-values with *t* statistics, as well as non-fragility of all bounds. In addition, with respect to CELLSQ, there is no sign switching when examining the simple correlation and coefficient of the full model. However, there is sign switching when viewing the simple correlation and the coefficient in the full model when examining CELLPOP, with OLS simple having a negative sign. This result is due to an aspect of falling fatality rates when cell phones became popular. Again, when other control variables are introduced, the coefficient associated with CELLPOP is positive and, as such, is regarded as a sturdy variable. In the case of CELLCUBE, coefficients are sturdy, but the *t* statistic is small for the full model and there is sign switching between the simple correlation and the coefficient associated with the full model. The sign switching may be once again countered by the consistency of the bounds and *s*-value results across all prior variance assumptions. This would lead to CELLCUBE being viewed as sturdy, and these results are consistent with Loeb et al.

[20] Regional dummies were included as explanatory variables, but results are not shown in Table 2.

[21] See Leamer (2014).

[22] See Leamer (2014).

[23] Leamer's decision rule in the selection of models requires the absolute values of *s*-values to be greater than 1 and the absolute value of *t* statistics for the full OLS model to be greater than 2. Conformity between the sign of *t* simple and *t* all is an additional factor in model selection although not necessarily required. However, when signs differ between these two values, additional information should be considered.

(2009) and Fowles et al. (2010).[24] Also, SPEED may be considered sturdy in spite of the sign switch when comparing the simple correlation and the coefficient in the full model given the stability of all bounds, the *s*-values being greater than 1 in absolute value, the sign associated with its coefficient in the full model, and its large *t* value (in absolute value). Hence, there are up to ten variables which may be considered as sturdy.

An important application of Bayesian *s*-values can be seen when examining variables such as REALINC. If one is dubious that REALINC is an important explanatory variable, then one might focus on the fragile bounds and *s*-values found with the wide and pessimistic priors. However, the opposite position would be supported with an optimistic prior which also conforms with no sign switching between OLS simple and OLS all. In addition, this position is supported by the *t* statistic. A useful feature of this reporting style is that each reader can come to the table with his or her own attitude toward the importance of the variable shown.

These relationships from Table 2 are illustrated in Fig. 3 with horizontal lines at ± 2 (significant) and vertical lines at ± 1 (sturdy).[25] Variables in the northeast and southwest quadrants are associated with more certain and sturdy estimates. We can note that most variables are both significant and sturdy. Figure 3 also demonstrates that there is agreement between calculated *s*-values and *t* statistics.[26]

5 Insights on the relative importance of key explanatory variables

By using standardized data, we can assess the relative importance of each explanatory variable as shown in Fig. 4.[27] It is important to note that the most important variables selected, i.e., those with the greatest impact, include those which we have considered as our focus variables. Once again, they include measures of the newness of the fleet, cell phone usage, suicidal propensities, and alcohol use. Furthermore, we find these factors are selected from a Bayesian *s*-values perspective as well.

The foremost variable in terms of classical regression is MCMY which highlights how improvements in automobile safety have had a pronounced effect on lowering fatality rates. The statistical results show the unambiguous effect of vehicle modernization on reducing fatality rates.

The next most important variable in terms of magnitudinal effects on fatality rates is cell phone use as measured by CELLPOP. One clearly expects the distraction effect of cell phone use to result in additional fatalities as demonstrated by Loeb

[24] It is important to note that all three cell phone variables are sturdy regardless as to one's prior whether pessimistic, optimistic, or not having a preconceived notion at all. The results in Table 2 regarding cell phone subscriptions reveal an overall positive effect of cell phones on fatality rates, but the marginal effect is diminishing. (A pictorial representation of this is available from the authors.) One should remember that the nonlinear specification is the prior proposed by previous studies.

[25] Again, regional dummy variables are included in the analysis, but not shown in Fig. 3.

[26] For *t* statistics and *s*-values, the correlation is .89.

[27] All variables are standardized to mean 0 and variance 1 so as to allow for unit-free comparisons.

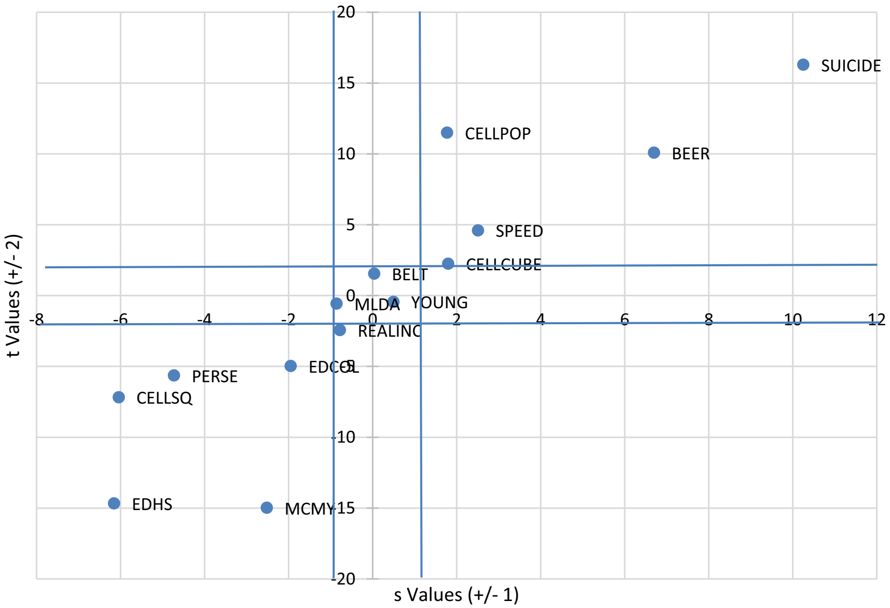

Fig. 3 *t* statistics and *s*-values (standardized data, wide bounds)

et al. (2009) and Fowles et al. (2010). The relative size of this effect, as demonstrated in Fig. 4, might be indicative of distracting influences in general.

The model then ranks in importance the percent of the population with a high school education, EDHS. This may be due to the association between income and education as well as an attempt of higher educated individuals to protect their human capital.

Suicide and suicidal propensities have been shown to have an effect on motor vehicle fatality rates. Suicide ranks next in importance with respect to motor vehicle fatality rates. This may be due to an association between suicidal propensities and violent/aggressive tendencies.

Alcohol consumption, as measured by BEER, is the next ranked factor affecting motor vehicle fatality rates. The importance of this factor on motor vehicle fatality rates has long been demonstrated.[28]

The above discussed results, of course, are based on a single model using normalized data. Questions of model ambiguity and parameter uncertainty still need to be addressed. The Bayesian approach using sturdy values adds significantly to this discussion.

[28] See for example, Loeb et al. (2009), Fowles et al. (2010), and Fowles and Loeb (1992).

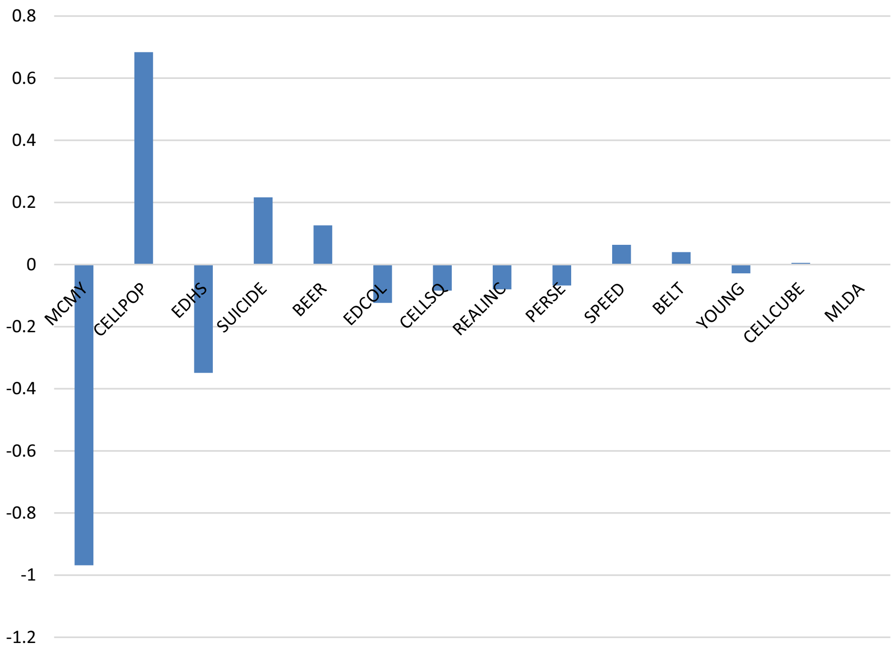

Fig. 4 Standardized OLS regression coefficients

6 Concluding remarks

One of the most important statistical problems today is the task of variable selection in regression model choice.[29] Dealing with both parameter uncertainty and model ambiguity is a challenging endeavor due to the sheer magnitude of the number of models that need to be considered. In this paper, we have looked at a newly developed Bayesian method which addresses these issues with reference to variables thought to be related to motor vehicle fatalities in the USA. With our data, there are millions of potential model specifications. Sturdy values are nicely suited to explore this high dimensional model space to discover plausible model specifications. It allows for different opinions regarding the strength of belief one has on the importance of factors affecting, in our case, motor vehicle fatality rates. The procedure is based on solid probability and statistical theory and provides researchers with inferential tools that are not a part of the non-Bayesian toolkit. As such, this technique adds considerable confidence to the results presented and to the policy measures discussed next.

The *s*-values note the importance of MCMY as a fundamental determinant of crash-related fatalities. This result is consistent with both Bayesian and classical results (Table 2 and Fig. 3). Our estimates suggest that there would be a 64% decline in fatality rates if a state could modernize their cars by 7 years, which is about a 1

[29] See Breiman (2001).

standard deviation change of MCMY in our sample.[30] We note in passing that Wyoming did exhibit a 55% decline in fatality rates coinciding with a fleet modernization of 7 years. Upgrading the fleet of motor vehicles is almost automatic through time as vehicles wear out and are replaced. Modernization is a potentially important policy recommendation. This may suggest a policy of government funded rebates and/or the reduction of taxes on newly purchased vehicles be evaluated with respect to life-saving, injury prevention, and property damage avoidance. The advent of autonomous vehicles will be observed with the modernization of the fleet, and should this technology prove efficacious, one would anticipate a life-saving effect and a reduction in crashes, especially as they reduce crashes due to distracted and inexperienced drivers.

Cell phone usage, as measured with CELLPOP, was found to be non-fragile using sturdy values. We anticipate that autonomous vehicles would reduce crash-related fatalities when they replace drivers who may be disposed to using cell phones or other distracting devices. In addition, policies reducing the use of cell phones by drivers are advisable and are already enacted in sixteen states and the District of Columbia with regard to hand-held phones. The extension of cell phone bans to all states and to include hands-free devices and a complete ban on text-messaging by all drivers across states is worthy of study. Also, evaluating the effect of stricter police enforcement combined with different fine structures and educational policies so as to limit cell phone use by drivers are potentially important areas of investigation.

Alcohol consumption clearly remains unambiguously a major contributor to crash fatalities as seen by s-values. The issue of imposing stricter sanctions against drinking and driving via fines and stricter policing and the use of substance abuse treatment centers are worthy of additional investigation.[31] Once again, we would anticipate that autonomous vehicles may prove to be advantageous in reducing alcohol-related vehicle fatalities if they replace drivers who have been drinking, assuming alternative risk behaviors are not encouraged as drivers attempt to reach a given risk tolerance.

Of particular interest are our results associated with suicides, especially considering that motor vehicle fatalities and suicides are among the leading causes of death among youth in the USA.[32] Our findings replicate earlier studies showing that the high suicide states are also the high motor vehicle fatality states.[33] As noted earlier, the association between these two variables is strong and is not generally well understood. It may well be the case that suicide rates represent changes in the risk taking behavior by individuals and act in a manner similar to a companion variable to account for unobservable factors. A potential avenue of future research may be to investigate the effectiveness of posting phone numbers/help lines for those in need of emotional/psychiatric assistance and/or investing public monies for additional

30 See Blattenberger et al. (2013).

31 See Chaloupka et al. (1993) and Freeborn and McManus (2007).

32 See Centers for Disease Control and Prevention, National Center for Injury Prevention and Control (2012).

33 These are Region 9 states in our model.

psychiatric health care, or policies designed to reduce recklessness and violent behaviors.[34]

S-values also indicate the importance of lowering blood alcohol concentration which indicates driving while impaired due to alcohol consumption. The results found are very stable and unambiguous. Autonomous vehicles may assist in taking the alcohol-affected driver out of the driver's seat should the would-be driver allow the technology to drive on his/her behalf.

Finally, the education variables are found to have significant and non-fragile effects on reducing crash-related fatality rates. This may be due to the association between education and income. Individuals with higher investments in education often earn higher incomes than their lower educated counterparts. Such individuals may also be more likely to afford high tech autonomous vehicles as they enter the market which could then reduce fatalities, all else equal.

To reiterate, one can anticipate that the successful introduction of autonomous vehicles will contribute to the reduction of vehicle fatalities by removing the task of driving from those who might otherwise drive while under the influence of alcohol and other drugs, as well as those who might otherwise be distracted by the use of cell phones or other such distracting devices. Autonomous cars may also diminish the risk of crashes due to driving inexperience and alcohol use inexperience associated with youthful drivers or those with diminished reflex actions. We may also notice a decline in crashes and fatalities due to suicidal propensities should these proclivities serve as a proxy for greater risk taking on the part of individuals or society. The technology in driverless vehicles may even prevent aggressive and suicidal drivers from using their vehicles in a reckless manner.

The results using *s*-values are supported by traditional classical methods using standardized data. However, the Bayesian *s*-values afford the reader of statistical evaluations, such as presented here, a way of dealing with various degrees of ambiguity and uncertainty they are willing to entertain as priors to the investigation. Our results over various widely different degrees of uncertainty have found strong evidence on the effects of our focus variables as well as others.[35]

[34] See Conner et al. (2001).

[35] A note is also due regarding the potential shortcomings in the paper. Clearly, alternative models, e.g., double log models among others, could be investigated. One's initial prior establishes the structure of the model, and *s*-values can be viewed then to demonstrate whether they appear reasonable. Quality of data is another matter in which all empirical investigations are subject to. One is left with using the best available at the time, recognizing measurement, and sampling errors are prone to show up. The data used in this study are obtained from well-known sources (as indicated in "Appendix 1") and are sensitive to such errors no more than in other studies.

Appendix 1: Data sources

Name	Data source
FATAL	Highway Statistics (various years), Federal Highway Administration, Traffic Safety Facts (various years), National Highway Traffic Safety Administration
MCMY	National Automobile Dealers Association (various years) and the National Household Travel Survey, US Department of Transportation.
PERSE	Digest of State Alcohol-Highway Safety Related Legislation (various years), Traffic Laws Annotated 1979, Alcohol and Highway Safety Laws: A National Overview 1980, National Highway Traffic Safety Administration
SPEEDHI	Highway Statistics (various years), Federal Highway Administration
BELT	Traffic Safety Facts (various years), National Highway and Traffic Safety Administration
BEER	US Census Bureau, National Institute on Alcohol Abuse and Alcoholism
MLDA	A Digest of State Alcohol-Highway Safety Related Legislation (various years), Traffic Laws Annotated 1979, Alcohol and Highway Safety Laws: A National Overview of 1980, National Highway Traffic Safety Administration, US Census Bureau
YOUNG	State Population Estimates (various years), US Census Bureau http://www.census.gov/population/www/estimates/statepop.html
CELLPOP	Cellular Telecommunication and Internet Association Wireless Industry Survey, International Association for the Wireless Telecommunications Industry
UNEMPLOY	Statistical Abstract of the United States (various years), US Census Bureau
REALINC	State Personal Income (various years), Bureau of Economic Analysis website http://www.bea.doc.gov/bea/regional/spi/dpcpi.htm
EDHS	Digest of Education Statistics (various years), National Center for Education Statistics, Educational Attainment in the United States (various years), US Census Bureau
EDCOLL	Digest of Education Statistics (various years), National Center for Education Statistics, Educational Attainment in the United States (various years), US Census Bureau
SUICIDE	Statistical Abstract of the United States (various years), US Census Bureau
REGION	US States 1: ME, NH, VT; 2: MA, RI, CT; 3: NY, NJ, PA; 4: OH, IN, IL, MI, WI, MN, IA, MO; 5: ND, SD, NE, KS; 6: DE, MD, DC, VA, WV; 7: NC, SC, GA, FL; 8: KY, TN, AL, MS, AR, LA, OK, TX; 9: MT, ID, WY, CO, NM, AZ, UT, NV; 10: WA, OR, CA; 11: AK, HI

Appendix 2: Descriptive statistics (primary variables, raw data)

Name	Mean	Median	Standard deviation
MCMY	1987.00	1987.01	10.10
CELLPOP	360.27	206.56	378.44
SPEED	65.06	65.00	6.65
PERSE	0.91	1.00	0.29
BEER	1.29	1.28	0.23
MLDA	20.74	21.00	0.76
BELT	0.75	1.00	0.43

Name	Mean	Median	Standard deviation
EDHS	0.60	0.60	0.06
EDCOLL	0.16	0.15	0.05
REALINC	28,743.65	28,466.36	7227.34
YOUNG	0.20	0.19	0.03
SUICIDE	13.02	1.00	0.29

References

Beede KE, Kass SJ (2006) Engrossed in conversation: the impact of cell phones on simulated driving performance. Accid Anal Prev 38:415–421

Blattenberger G, Fowles R, Loeb PD, Clarke WA (2009) Understanding the cell phone effect on motor vehicle fatalities using classical and Bayesian methods. Presented at the Allied social sciences associations meeting (TPUG), San Francisco

Blattenberger G, Fowles R, Loeb PD, Clarke WA (2012) Understanding the cell phone effect on vehicle fatalities: a Bayesian view. Appl Econ 44:1823–1835

Blattenberger G, Fowles R, Loeb PD (2013) Determinants of motor vehicle crash fatalities using bayesian model selection methods. Res Transp Econ 43:212–222 **(Special issue on: the economics of transportation safety)**

Breiman L (2001) Statistical modeling: the two cultures. Stat Sci 16(3):199–231

Centers for Disease Control and Prevention, National Center for Injury Prevention and Control. Web-based Injury Statistics Query and Reporting System (WISQARS). http://www.cdc.gov/injury/wisqars/. Accessed 7 July 2012

Chaloupka FJ, Saffer H, Grossman M (1993) Alcohol-control policies and motor vehicle fatalities. J Legal Stud 22:161–186

Chapman S, Schoefield WN (1998) Lifesavers and samaritans: emergency use of cellular (mobile) phones in Australia. Accid Anal Prev 30:815–819

Conner KR, Cox C, Duberstein PR, Tian L, Nisbet PA, Conwell Y (2001) Violence, alcohol, and completed suicide: a case-control study. Am J Psychiatry 158:1701–1705

Connolly JF, Cullen A, McTigue O (1995) Single road traffic deaths: Accident or suicide? Crisis J Crisis Interv Suicide Prev 16(2):85–89

Consiglio W, Driscoll P, Witte M, Berg WP (2003) Effect of cellular telephone conversations and other potential interference on reaction time in braking responses. Accid Anal Prev 35:495–500

CTIA—The Wireless Association (2011). http://www.ctia.org. Accessed 10 Feb 2011

CTIA—The Wireless Association (2016). http://www.ctia.org/industry-data/ctia-annual-wireless-industry-survey. Accessed 7 Nov 2016

Etzerdorfer E (1995) Single road traffic deaths: Accidents or suicide? comment. Crisis J Crisis Interv Suicide Prev 16(4):188–189

Fowles R (1988) Micro EBA. Am Stat 4:274

Fowles R, Loeb PD (1992) The interactive effect of alcohol and altitude on traffic fatalities. South Econ J 59:108–112

Fowles R, Loeb PD (2019) Motor vehicle fatalities from the perspective of sturdy values: the autonomous vehicle effect. Presented at the American economic association and transportation and public utilities group meetings in Atlanta, GA, January 6, 2019

Fowles R, Loeb PD, Clarke WA (2010) The cell phone effect on motor vehicle fatality rates: a Bayesian and classical econometric evaluation. Transp Res Part E 46:1140–1147

Fowles R, Loeb PD, Clarke WA (2013) The cell phone effect on truck accidents: a specification error approach. Transp Res Part E. https://doi.org/10.1016/j.tre.2012.10.002

Freeborn BA, McManus B (2007) Substance abuse treatment and motor vehicle fatalities. College of William and Mary, Department of Economics, Working Paper Number 66

Gelman A (2008) Scaling regression coefficients by dividing by two standard deviations. Stat Med 27:2865–2873

Glassbrenner D (2005) Driver cell phone use in 2005—overall results. Traffic safety facts: research note, NHTSA, DOT HS 809967. http://www-nrd.nhtsa.dot.gov/Pubs/809967.PDF. Accessed 11 Feb 2011

Hoerl AE, Kennard RW (1970) Ridge regression: biased estimation for nonorthogonal problems. Technometrics 12(1):55–67
Huffine CL (1971) Equivocal single-auto traffic fatalities. Life Threat Behav 1(2):83–95
Laberge-Nadeau C, Maag U, Bellavance F, Lapiere SD, Desjardins D, Messier S, Saidi A (2003) Wireless telephones and risk of road crashes. Accid Anal Prev 35:649–660
Leamer EE (1978) Specification searches: ad hoc inference with non-experimental data. Wiley, New York
Leamer EE (1982) Sets of posterior means with bounded variance priors. Econometrica 50(3):725–736
Leamer EE (1983) Let's take the con out of econometrics. Am Econ Rev 73(1):31–43
Leamer EE (2014) S-values and all subsets regressions. University of California at Los Angeles, working paper
Leamer EE (2016) S-values and Bayesian weighted all-subsets regressions. Eur Econ Rev 81:15–31
Loeb PD, Clarke W (2009) The cell phone effect on pedestrian fatalities. Transp Res Part E 45:284–290
Loeb PD, Talley WK, Zlatoper T (1994) Causes and deterrents of transportation accidents: an analysis by mode. Quorum Books, Westport
Loeb PD, Clarke WA, Anderson R (2009) The impact of cell phones on motor vehicle fatalities. Appl Econ 41:2905–2914
McEvoy SP, Stevenson MR, McCartt AT, Woodward M, Haworth C, Palamara P, Cercarelli R (2005) Role of mobile phones in motor vehicle crashes resulting in hospital attendance: a Case-Crossover Study. BMJ 33:428–435
Murray D, DeLeo D (2007) Suicidal behavior by motor vehicle collision. Traffic Injury Prev 8:244–247
Neyens DM, Boyle LN (2007) The effect of distractions on the crash types of teenage drivers. Accid Anal Prev 39:206–212
NHTSA (2016) Traffic safety facts 2016, DOT HS 812554, May 2018. https://crashstats.nhtsa.dot.gov/Api/Public/ViewPublication/812554. Accessed 28 Mar 2018
Office of the Federal Register (1980) National archives and records service, general service administration. The United States government manual: 1980–1981
Peltzman S (1975) The effect of automobile regulation. J Polit Econ 93:677–725
Phillips DP (1977) Motor vehicle fatalities increase just after publicized suicide stories. Science 196:1464–1466
Phillips DP (1979) Suicide, motor vehicle fatalities, and the mass media: evidence toward a theory of suggestion. Am J Sociol 84(5):1150–1174
Pokorny AD, Smith JP, Finch JR (1972) Vehicular suicides. Life Threat Behav 2(2):105–119
Porterfield AL (1960) Traffic fatalities, suicide, and homicide. Am Sociol Rev 25(6):897–901
Poysti L, Rajalin S, Summala H (2005) Factors influencing the use of cellular (mobile) phones during driving and hazards while using it. Accid Anal Prev 37:47–51
R Development Core (2016) R: a language and environment for statistical computing. http://www.R-project.org. Accessed 10 Jan 2016
Ramsey JB (1974) Classical model selection through specification error tests. In: Zarembka P (ed) Frontiers in econometrics. Academic Press, New York, pp 13–47
Ramsey JB, Zarembka P (1971) Specification error tests and the alternative functional form of the aggregate production function. J Am Stat Assoc Appl Sect 57:471–477
Redelmeier DA, Tibshirani RJ (1997) Association between cellular-telephone calls and motor vehicle collisions. N Engl J Med 336:453–458
Simmons WO, Welki A, Zlatoper TJ (2016) The impact of driving knowledge on motor vehicle fatalities. J Transp Res Forum 55(1):17–27
Souetre E (1988) Completed suicides and traffic accidents: longitudinal analysis in france. Acta Psychiatr Scand 77(5):530–534
Sullman MJ, Baas PH (2004) Mobile phone use amongst New Zealand drivers. Transp Res Part F Traffic Psychol Behav 7:95–105
Violanti JM (1998) Cellular phones and fatal traffic collisions. Accid Anal Prev 30:519–524
Welki A, Zlatoper TJ (2014) The effect of cell phones on international motor vehicle fatality rates: a panel-data analysis. Transp Res Part E Log Transp Rev 64:103–109
Welki A, Zlatoper TJ (2017) An analysis of illicit drug use and motor vehicle fatalities using contiguous state-level data. J Transp Res Forum 56(2):5–20