Research in Transportation Economics 43 (2013) 112–122

Contents lists available at SciVerse ScienceDirect

Research in Transportation Economics

journal homepage: www.elsevier.com/locate/retrec

Determinants of motor vehicle crash fatalities using Bayesian model selection methods

Gail Blattenberger [a], Richard Fowles [a,*], Peter D. Loeb [b]

[a] Department of Economics, 260 S. Central Campus Drive, Orson Spencer Hall 343, Salt Lake City, UT 84112, USA
[b] Department of Economics, Rutgers University, 360 Dr. Martin Luther King Jr. Blvd, Hill Hall 325, Newark, NJ 07102, USA

A R T I C L E I N F O

Article history:
Available online 5 March 2013

JEL codes:
C11
R41

Keywords:
Motor vehicle crashes
Bayesian econometrics
Cell phones
Vehicle safety

A B S T R A C T

Motor vehicle crashes continue to result in large numbers of fatalities each year and represent the leading cause of death for young persons. In 2006, for example, there were over 42,700 fatalities associated with these crashes. Understanding the causes of these crashes and methods to reduce them continues to be of great interest to economists, public health officials, and policy makers. We present in this paper statistical models using a rich set of panel data covering the period 1980 to 2007 by state and the District of Columbia. Our choice of variables is based on an extensive literature highlighting the importance of policy, safety, demographic, and economic determinants of fatality rates.

The estimation techniques used in this paper takes cognizance that standard econometric inference focuses on parameter uncertainty. Models are estimated conditional on the assumption that the model to be estimated and reported is the “true” model. Tests are then made on a multitude of alternative models, each sequentially assumed to be the “true” model. Model uncertainty is manifested in this procedure, but it is often ignored in practice. Recent Bayesian statistical methods speak directly to the issue of both model choice and variable selection. This paper utilizes three Bayesian techniques: Extreme Bounds Analysis, Bayesian Model Averaging, and Stochastic Search Variable Selection to address model and parameter uncertainty in models estimating the determinants of motor vehicle crash fatalities.

1. Introduction

The determinants of fatalities due to motor vehicle crashes continue to be of major interest to public health officials, economists, statisticians, policy makers, among others. Fatalities hit a high point in 1972 with 54,589 deaths. Between that year and 2007, deaths due to crashes were generally above forty thousand per year. More recently the numbers have diminished significantly. Between 2007 and 2010, deaths have fallen from 41,059 to 32,885.[1] This precipitous drop in deaths lately may be due to things other than the prior trend. For example, some of this decline may be due to economic events, i.e., the Great Recession as well as a change in tastes among the public. That is, there seems to be a preference, especially among the youth, to move from the suburbs to the cities where there is a greater reliance on public transportation and walking. Furthermore, while the baby boomers had a strong desire to obtain powerful cars while in their teens, the current population of youths may be more inclined to desire powerful cell phones and electronic equipment. Regardless of these changes in preferences, there remains a significant number of crash related fatalities in the U.S. which scientist attempt to explain.

Many factors thought to contribute to motor vehicle crashes and crash fatalities have been examined over the last two decades. These include, motor vehicle speed, speed variance, alcohol, speed limits, vehicle miles traveled, measures of income and wealth, unemployment rates, advances in technology, the age of the motor vehicle fleet, population characteristics, insurance effects, seat belts and seat belt legislation, the deregulatory climate of the 1980's, among others. In general we can classify these factors into three categories: those associated with vehicles such as technology and design characteristics; those associated with roadways such as speed limits; and those associated with drivers themselves such as alcohol consumption and seat belt usage. Most recently, there has been an interest in the effects of cell phone usage on motor vehicle crashes and fatalities along with the effects of other socioeconomic and technology factors such as the age of the fleet, education, and suicidal propensities.

* Corresponding author. Tel.: +1 (801) 581 45771; fax: +1 (801) 585 5649.
E-mail address: richard.fowles@economics.utah.edu (R. Fowles).

[1] See the National Highway Traffic Safety Administration (NHTSA), Fatality Analysis Reporting System (FARS) for data on crash fatalities.

0739-8859/$ – see front matter © 2012 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.retrec.2012.12.004

We provide a background for our research in Section 2. Our data is described in Sections 3 and 4 lays out the general Bayesian framework underlying all the procedures we use in this analysis; it then introduces Extreme Bounds Analysis (EBA), and finally uses EBA on models of motor vehicle fatality rates. Section 5 discusses and implements Stochastic Search Variable Selection (SSVS) developed by George and McCulloch (1993) as another way of exploring models. Section 6 explores the problem further using Bayesian Model Averaging (BMA) as discussed by Raftery, Madigan, and Hoeting (1997). Section 7 highlights agreement among EBA, SSVS, and BMA methods as applied to the motor vehicle crash fatality rate models. Section 8 provides a summary of our findings and some concluding comments and policy recommendations.

2. Background

Most of the factors considered to be important determinants of crash fatalities have been investigated using econometric models based on classical methods. These models have attempted to examine whether various factors had significant statistical effects on crash fatalities and to provide a measure of their marginal effects. Many of these determinants are reviewed in Loeb, Talley, and Zlatoper (1994). The models have often followed the classic approach suggested by Peltzman (1975). His model is of particular interest given that he was concerned with the potential offsetting behavior of drivers as they adjusted their driving behavior in the face of various imposed regulations by the state. For example, one can recognize that drivers (as well as other members of society) have a given risk tolerance. If seat belts were required by law, risk imposed on the driver might be reduced, all else equal. However, drivers might then increase other risk behaviors, such as driving faster, which might not only affect their vehicle, but impose additional risks on, e.g., pedestrians. In any case, many factors have been evaluated by econometric models and not all of them have provided consistent results which might be expected from economic theory, public heath experiences and various statistical studies viewing the same determinants. Differences between studies may be due to different models, different estimation techniques, data differences, and changes which occur over time. As suggested above, the examined factors effecting safety are numerous. Some of the more significant ones (although not all of them) are reviewed below.

Motor vehicle inspection has been imposed in different degrees by many states over the last several decades. The catalyst for these regulations stems from the Highway Safety Act of 1966 which set standards for inspection and used the threat of withholding federal highway funds for noncompliance. In 1976, Congress relaxed its position regarding the imposition of state inspections. Numerous studies were conducted over time on the effectiveness of inspection on safety resulting in varying conclusions. Crain (1980), for example did not find strong statistical results suggesting the effectiveness of inspection. Garbacz and Kelly (1987) also did not find reason to support vehicle inspection. These results were countered by other investigators including, Loeb (1985, 1988, 1990), Loeb and Gilad (1984), among others.[2] The reason for such different results may be due to model uncertainty and different time regimes under investigation. To address some of this, Loeb often made use of specification error analysis to minimize the likelihood of model misspecification. In more recent work using Bayesian methods, Blattenberger, Fowles, Loeb, and Clarke (2012) found the effects of inspection to be fragile. However, the efficacy of inspection may also have changed over time as the age of the fleet, and hence the technologies available, changed. Keeler (1994) finds some evidence for this where inspection is found efficacious using data for the period 1970 but not so using data for 1980.

Speed and later speed variance were considered major factors contributing to crash related injuries and deaths. Speed adds to utility by diminishing travel time and, at least for some, provides thrills. Yet, it is argued that speed comes at a price of increasing the probability of crashes and deaths. This has been found to be the case in papers, e.g., by Peltzman (1975), Forester et al. (1984), Zlatoper (1984), Sommers (1985), and Loeb (1987, 1988), among others. However, Lave (1985) has argued that it is primarily speed variance as opposed to vehicle speed itself which is the speed related factor contributing to fatalities. Levy and Asch (1989) and Snyder (1989) found some evidence for this as well, while Fowles and Loeb (1989) using Bayesian methods found support for both speed and speed variance.

Speed limits have also been investigated as contributors to crashes especially after the Arab Oil Embargo in 1973. Statistical results varied among studies depending on model specifications and data used. Contributing effects have been found by Forester et al. (1984), Loeb (1991), among others. However, speed limits have been found also to reduce measures of crash fatalities by Garbacz and Kelly (1987), and Loeb (1990). More variable results are found, for example, in Keeler (1994), Blattenberger et al. (2012), and Fowles, Loeb, and Clarke (2010).

The effect of alcohol use has almost uniformly been found to have significant effects on motor vehicle crashes in recent research. This result is found both using classical as well as Bayesian methods as seen in Loeb, Clarke, and Anderson (2009), Fowles et al. (2010), Blattenberger et al. (2012) among others.[3] The effect of the minimum legal drinking age has also been investigated with varying results. For example, Sommers (1985) found a negative relationship between the minimum legal drinking age and fatality rates, while more recently, Blattenberger et al. (2012) and Fowles et al. (2010) find fragile results regarding the effect of the Minimum Legal Drinking Age on crash related fatalities.[4]

Related to alcohol consumption itself has been an analysis of the use of varying blood alcohol thresholds to determine if a driver is operating a vehicle under the influence. Recently, some evidence has been found by Loeb et al. (2009) indicating more severe limits on blood alcohol concentration (BAC) to designate driving while impaired reduced vehicle fatalities.

An interesting observation was found by Fowles and Loeb (1992) when examining the effects of alcohol on motor vehicle related crashes. They found evidence that altitude intensifies the adverse effect of alcohol on highway safety. This may be due to the fact that at higher altitudes, oxygen intake is less than at lower altitudes and may adversely impact on reaction time.[5]

Seat belts (and airbags) have been shown to have life-saving and injury reducing attributes. Researchers have estimated that seat belts have the potential to reduce fatalities by 40 percent or more.[6] Seat belt laws, both primary and secondary, have been imposed so as to induce the driving public to wear belts. New York was the first state to impose a seat belt law and currently 32 states and the District of Columbia have primary seat belt laws and 17 states have secondary laws. The laws vary further from state to state by who must wear a belt (front seat versus all seats) and the fine structures imposed for violating the law. Numerous studies have been conducted regarding the efficacy

[2] See Loeb et al. (1994) for a more complete review of the literature.

[3] See Loeb et al. (1994) for a review of earlier work including those resulting in opposite or insignificant results.

[4] See Loeb et al. (1994) for additional reviews.

[5] See Newman (1949) and Mazess, Picon-Reategui, Thomas, and Little (1968).

[6] See Partyka (1988) and Evans (1991).

of these laws. Most early studies found seat belt laws reduced various measures of fatalities and injuries. These include studies by: Campbell and Campbell (1986), Campbell, Stewart, and Campbell (1986, 1987), Hatfield and Hinshaw (1987), Hoxie and Skinner (1987), Lund, Pollner, and Williams (1987), Reinfurt, Campbell, Stewart, and Stutts (1988), Skinner and Hoxie (1988), Streff, Schultz, and Wagenaar (1989), and Loeb (1991). Womble (1989), however, did not find necessarily a uniform reduction in injury rates due to seat belt laws depending on the types of injuries investigated. More recently, Cohen and Einav (2003) found additional evidence of the efficaciousness of mandatory seat belt laws and, in particular, the benefits of primary laws. Additional evidence on the effectiveness of seat belt laws is provided from a classical probability perspective by Fowles et al. (2010) where a statistically significant effect is found. However, from a Bayesian perspective, both Blattenberger et al. (2012) and Fowles et al. (2010) found fragile results.

Measures of income have been of particular interest to economists and public policy makers. Assuming that driving intensity and safety are both normal goods, then the demand for each should increase with income. Peltzman (1975) argued that income would have an ambiguous impact on crashes given their offsetting effects.[7] The net effect of income would then depend on the relative strengths of these offsetting effects. In addition, Peltzman argued that different types of income would have different effects on crashes, e.g., transitory income should have a lesser life-saving effect than an equal amount of permanent income. In addition, one might expect different results from models based on time-series data (possibly portraying short-run effects) as opposed to models based on cross-section data (possibly portraying long-run effects).[8] The bottom line is then an empirical matter. Recently, Blattenberger et al. (2012) found a positive and significant effect of real income per household on fatality rates with additional support from several Bayesian methods.

Time trends have also been incorporated in many models to adjust for changes in technology and to proxy permanent income.

Additional normalizing variables included in past studies have been: population characteristics, measures of poverty and unemployment, and measures of education. Regarding population, researchers are concerned not only with the size of the population, which may be highly correlated with other factors such as income and a time trend, but the age distribution of the population as well. One might speculate that youthful drivers are more likely to be involved in crashes due to a lack of experience. Asch and Levy (1987), Garbacz (1990), Loeb (1990), Saffer and Grossman (1987a, 1987b) are examples of studies finding a relationship between youthful drivers and measures of fatalities. However, McCarthy (1992), and Loeb (1985) find a significant negative relationship between youthful driving and fatality or injury measures.

Miles driven and types of roadways have also been examined by various researchers. Loeb et al. (1994) review many of the early statistical studies where evidence suggests safety associated with urban interstate mileage. However, some more recent studies, e.g., by Loeb et al. (2009), do not find a statistically significant effect on total vehicle fatalities due to interstate highway mileage.

Weather and daylight conditions have been examined as well. Most notably, Coate and Markowitz (2004) find that moving to daylight savings time is associated with a reduction in both pedestrian and occupant fatalities.

Hospital access, the cost of a crash, insurance attributes, and geographical locations have also been examined.[9] More notably, numerous studies have been conducted regarding the deregulatory effect of the 1980s on safety. Even airline deregulation has been shown to have an impact on motor vehicle fatalities. Bylow and Savage (1991), for example, have shown that deregulation of the airlines led to a reduction in motor vehicle deaths between 1978 and 1988. Perhaps of more interest was the impact of the deregulation of the trucking industry by the Motor Carrier Act of 1980. Concern arose regarding the potential increase in crashes associated with this act. Statistical models, such as those reported by Loeb and Clarke (2007) found no significant impact of deregulation on truck safety. In addition, there was concern that the deregulation of the railroad industry by the Staggers Act of 1980 might have resulted in more fatalities, including those at grade crossings involving motor vehicles. Clarke and Loeb (2005) did not find statistical evidence supporting this using their econometric models.[10]

More recently, safety researchers have been addressing the impact of cell phone use and availability on motor vehicle crashes. 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.[11] 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. As such, ten states plus the District of Columbia have banned the use of hand-held phones by drivers (California, Connecticut, Delaware, Maryland, Nevada, New Jersey, New York, Oregon, Utah, and Washington).[12]

Statistical evaluations of the effects of cell phones on crashes have not always provided consistent results. Perhaps the most well-known early study of the effect of cell phones on motor vehicle crashes was by Redelmeier and Tibshirani (1997). They attribute a four-fold increase in property-only crashes to the use of cell phones. Violanti (1998) attributes a nine-fold increase in fatalities to the use of cell phones and McEvoy et al. (2005) also find an increase in risk from cell phones. Neyens and Boyle (2007) found that cell phones increased the likelihood of rear-end collisions relative to fixed-object collisions amongst teenage drivers. In addition, Consiglio, Driscoll, Witte, and Berg (2003) using a laboratory environment found that brake reaction time was increased when cell phones were in use, regardless of whether they were hand-held or hands-free devices. Similarly, Beede and Kass (2006), also using a laboratory environment, found that hands-free devices adversely effected driving performance. These results are what one would expect a priori. However, not all empirical results are consistent with the above.

Laberge-Nadeau et al. (2003) using logistic regression and Canadian data initially found a relationship between cell phone use and crashes. However, this risk diminished as their basic models were extended. The life-taking effect of cell phones were countered as well by Chapman and Schoefield (1998) who argued that cell phones should be credited with saving lives instead of taking them. They found that, “Over one in eight current mobile phone users

[7] See Peltzman (1975) and Loeb et al. (1994) for a discussion.

[8] See Loeb et al. (1994) for a review of this literature.

[9] See Loeb et al. (1994) for a review of some of this literature.

[10] See Loeb et al. (1994) for reviews of other studies on truck, air, and railroad deregulation.

[11] See CTIA (2011).

[12] 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).

have used their phones to report a road accident."[13] They attribute this beneficial effect of cell phones to the "golden hour," the time period crucial for survivorship from various medical emergencies and crashes which cell phones could affect. Hence, given a crash, the probability of survivorship is influenced by the speed which help can be obtained. Sufficient cell phones in the hands of the public increases the likelihood of medical help arriving at the scene of the crash promptly. Sullman and Baas (2004) added to these investigations using survey data pertaining to crashes of all levels of severity. They did not find a significant correlation between cell phone use and crash involvement after normalizing for demographics and other factors. Similarly, Poysti, Rajalin, and Summala (2005) claim that, "phone-related accidents have not increased in line with the growth of the mobile phone industry."[14]

These convoluting results led to a study by Loeb et al. (2009) using classical econometrics where cell phones were found to have a nonlinear effect on crash related fatalities: initially associated with an increase in such fatalities with low volumes of cell phone subscribers, a decrease in fatalities as cell phone subscribers rose over time, and then an increase of fatalities as cell phone subscriptions rose substantially again later on.

Blattenberger et al. (2012) and Fowles et al. (2010), using Bayesian methods demonstrated a relationship between cell phones and crash fatalities as well.

Finally, education levels, crime rates, and suicide rates were also examined with regard to an association with crash fatalities. One might argue that higher levels of education are associated with greater stocks of human capital which would be expected to be inversely related to crash fatalities. This has been investigated by various researchers, e.g., Blattenberger et al. (2012), indicating a negative association between college education rates and crash fatalities.

Suicide rates have rarely been used as a potential determinant of crash fatalities. Recently, however, Blattenberger et al. (2012) examined the effect of suicide rates on crash fatalities as a way of controlling for factors not already addressed in their models and to potentially account for unexplained societal characteristics over time. They found that suicide rates had a statistically significant effect on fatalities from both a classical and a Bayesian perspective. This possibly unexpected result, from some researchers' perspective, is not completely without prior statistical evidence. For example, Phillips (1979) examined the importance of imitation and suggestion and found that there was a 31% increase in automobile fatalities three days following a publicized suicide.[15] Pokorny, Smith, and Finch (1972) also claim a relationship between suicide and fatal automobile crashes. The strong correlation between suicides and motor vehicle fatalities was noted very early by Porterfield (1960). He summarized that whatever factors play a part in these complex behaviors there is no reason to doubt that aggressive, hazardous driving is likely to be a characteristic of persons who have suicidal tendencies. Along these lines, Conner et al. (2001) utilizing a case–control method observe violence behaviors to be consistent for persons exhibiting suicidal tendencies. However, the association between suicides and automobile crashes is not consistent among studies. Some studies have not found a uniformly strong relationship, e.g., Connolly, Cullen, and McTigue (1995), Huffine (1971), and Souetre (1988). Others, e.g., Etzerdorfer (1995), indicate the difficulty which arises in determining whether a crash victim was indeed a suicide.

The effect of suicides on other modes of transportation has also been investigated to some extent. Some association has been found regarding suicides and aircraft use, but the numbers appear small.[16] Meanwhile, the relationship between suicides and railroad related fatalities with respect to trespassers has been investigated recently by Savage (2007, 2010). Although there is difficulty in determining the exact number of suicides, the author suggests a rather large proportion of trespasser deaths are associated with suicide. Clearly the suicide–automobile crash link is controversial with mixed empirical findings.

Along with suicide rates, there are extraordinarily large numbers of potential contributors to crash fatalities. Many have been suggested by economic theory, some based on conventional wisdom along with statistical studies, as noted above. Some of the contributing factors, such as alcohol, have been consistently found as major determining factors. Others, such as motor vehicle inspection have been found to have fragile effects. These may be due to the types of models investigated, the data sets, and changes over time. Hence, over time, various factors have been added to and others deleted from models of crash fatalities. Importantly, there is no one model. For these reasons, explorations are made over many models with the goal to provide statistically reliable and inferentially meaningful results. Such statistical procedures must account for both model and parameter uncertainty. This is the focus of the current paper which is conducted making use of three Bayesian econometric methods which are ideally suited to deal with the problem of model uncertainly. Model uncertainty is an inferential problem difficult to address via conventional methods which require a single posited model being true that forms the basis of the observed data generating process.

3. Data

We utilize a newly compiled set of data collected on 50 states and Washington, D.C. over the period from 1980 to 2005.[17] The number of fatalities per 100 million vehicle miles traveled is our dependent variable. Our choice of explanatory variables is based on a rich literature (reviewed in Section 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. (2012, 2009), Fowles et al. (2010), and Loeb et al. (2009). These form the basis for the "prior" used in the Bayesian analysis in this paper. Our data covers 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 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 270 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.[18]

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 records the

[13] See Chapman and Schoefield (1998, p. 817).

[14] See Poysti et al. (2005, p. 50).

[15] Phillips (1977) had earlier found a 9.12% average increase in motor vehicle fatalities following front-page newspaper suicide stories in California.

[16] See, for example, Bills, Grabowski, and Li (2005), Cullen (1998), and Lewis, Johnson, Whinnery, and Forster (2007).

[17] We are grateful to Tianda Xing for his assistance in collecting the data.

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

Table 1
Explanatory Variables[a] cross sectional - time series analysis of traffic fatality rates for 50 states and DC from 1980 to 2005.

Name	Description	Expected sign (Priors)
CARYEAR	Median car model year	–
PERSELAW	Dummy variable indicating the existence of a law defining intoxication of a driver in terms of Blood Alcohol Concentration (BAC). PERSELAW = 1 indicates the existence of such a law and PERSELAW = 0 indicates the absence of such a law. (More precisely, PERSELAW = 1 when the BAC indicating driving under the influence is 0.1 or lower.)	–
INSPECTION	Indicator for mandatory annual safety inspection	–
SPEEDMAX	Maximum posted speed limit, urban interstate highways	+
SEATBELT	Indicator for presence of a legislated seat belt law	–
BEER	Per capita beer consumption (in gal) per year	+
MLDA	Minimum legal drinking age	–
YOUNG	Percentage of males (16–24) relative to population of age 16 and over	+
CELLPHONES	Number of cell phone subscribers per capita	+
POVERTY	Poverty rate (percentage)	+
UNEMPLOY	Unemployment rate (percentage)	–
REALINC	Real per household income in 2000 dollars	?
HIGHSCHOOL	Percent of persons with a high school diploma	–
COLLEGE	Percent of persons with a college degree	–
CRIME	Crime rate (reported crimes per 100,000 population)	?
SUICIDE	Suicide rate (suicides per 100,000 population)	?

[a] For data sources, see Appendix.

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.[19] Seat belt legislation varies widely across states. Our data records 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, poverty, the age of the fleet, and suicide on motor vehicle fatality rates. These are of particular interest given the review of the literature in Section 2. The data are organized by geographical coding of states into eleven regions. The variables are defined and described in Table 1 along with their expected effects (priors) on fatality rates.[20]

4. Extreme bounds analysis

Standard econometrics takes a given model as true and summary statistics are provided conditional on that model's truth. Often alternative tests are made on a multitude of competing models, each sequentially assumed to be a true model. Inferences based on sequential search procedures are fraught with problems of doubt regarding the statistical validity of reported summary statistics.[21] Bayesian theory, however, can directly address model uncertainty and in this paper we utilize recent advances in Bayesian research regarding model choice and variable selection as discussed, for example, in Key, Pericchi, and Smith (1999), and Clyde (1999). An early investigator in model uncertainty was Leamer (1978, 1982, 1983, 1985, and 1997) who in a series of articles dealt with specification searches. Extreme Bounds Analysis (EBA) was a product of this work.

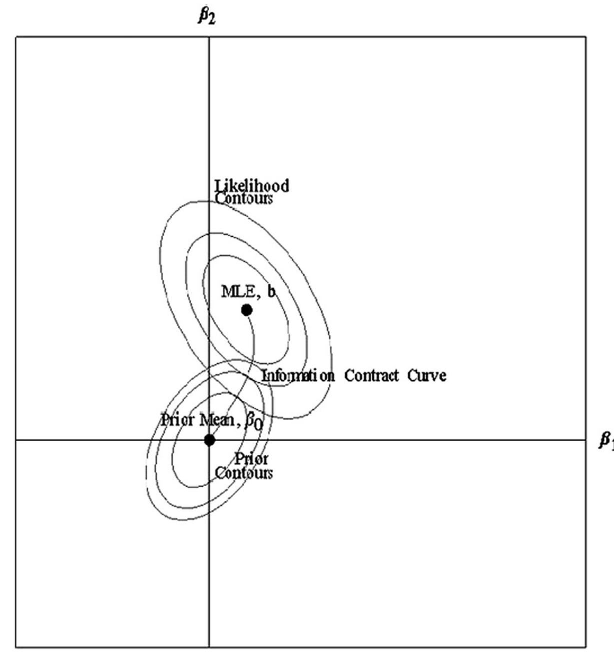

Fig. 1. Likelihood/prior contours: Information contract curve.

We utilize a very simple Bayesian framework that is based on a linear model with a normal – gamma prior. The same model is used in all of the procedures reported. It is

$$\begin{aligned} Y_{it} &= X_{it}\beta + \varepsilon_{it} \\ i &= 1, 2, \ldots, 51 \\ t &= 1, 2, \ldots, 26 \end{aligned} \tag{1}$$

The prior distributions for this model are specified

$$\begin{aligned} &\beta \sim N(\beta_0, \textstyle\sum_0) \\ &\varepsilon \sim N(0, \pi^{-1}) \\ &\pi \sim \Gamma(\nu, \lambda) \end{aligned} \tag{2}$$

Before applying EBA to the motor vehicle fatality data, we introduce the procedure geometrically using a two-dimensional simplification. The basic model is illustrated in Fig. 1. The likelihood contours implied by the data are shown along with the maximum likelihood estimate, b. In addition, the prior contours implied by the prior location, β_0, and the prior precision, Σ_0^{-1}, are shown. The posterior mean for this sample and prior is a matrix weighted average of the sample and prior values:

$$\beta^* = \left(\Sigma_0^{-1} + X'X\right)^{-1}\left(\Sigma_0^{-1}\beta_0 + X'Xb\right). \tag{3}$$

The potential set of posterior means for this prior and likelihood are indicated by the relation labeled the information contract curve. The curve is analogous to the contract curve familiar to economists. The data are indifferent to points on the iso-likelihood contours and the researcher is indifferent to points on the iso-prior probability contours. The exact position of the posterior mean along this curve is a function of the relative sample and prior precisions. The more precise the data, the closer the posterior mean is to the maximum likelihood estimate, b. EBA is illustrated in the context of this model.[22]

[19] 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.

[20] See Blattenberger, Fowles, Loeb, and Clarke (2009).

[21] See, for example, Leamer (1978).

[22] Mathematical developments are found in Leamer (1982) and implemented in Fowles (1988).

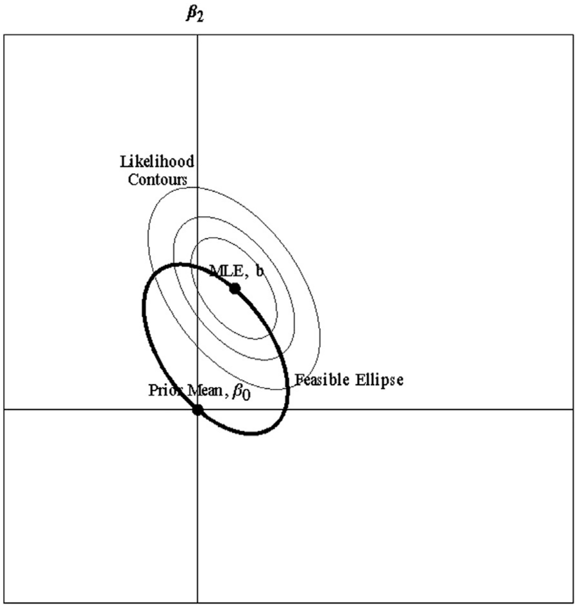

Fig. 2. Feasible ellipse.

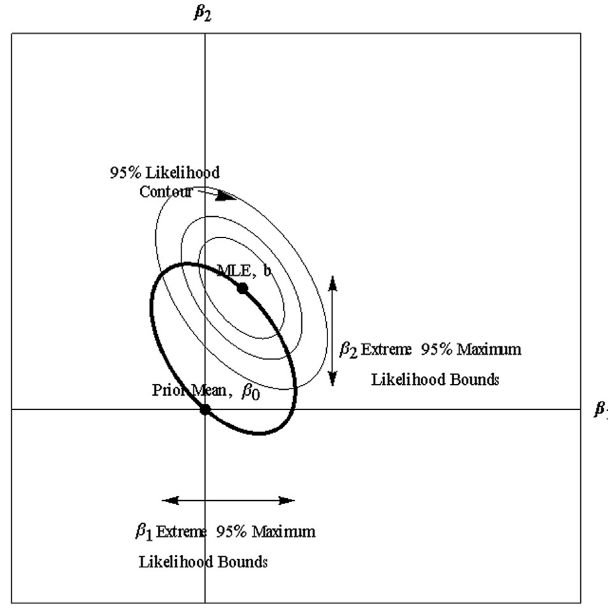

Fig. 3. Extreme bounds within 95% likelihood contour: All variables doubtful.

EBA is a methodology of global sensitivity analysis that computes the maximum and minimum values for Bayesian posterior means in the context of linear regression models. The potential values are those that could be estimated via maximum likelihood estimation when all possible linear combinations of the explanatory variables are considered under a wide assortment of model specifications. This method is rather draconian in the sense that all possible specifications are considered and that very few hypotheses survive a full EBA analysis.[23] Lack of survivability is seen when the extreme bounds of posterior means cover zero. Such variables are called fragile even though associated parameter estimates obtained via classical estimation might be seen to be statistically significant.

The analysis depicted in Fig. 1 requires specifying both a prior mean β_0, and a prior precision, Σ_0^{-1}. The posterior might encompass a wide range of values even if the prior location is restricted to the origin. In a remarkable result of Chamberlain and Leamer (1976), posterior means fall within a feasible ellipse under minimal assumptions regarding the nature of the precision matrix,Σ_0^{-1}. The feasible ellipse is illustrated in Fig. 2. Nonetheless, as indicated by this figure, the range of potential posterior values encompasses zero. Without further restrictions all variables are fragile.

Several options are used to further limit the values under consideration. First one might limit the values to those that are highly likely for the sample, for example those falling within the 95% confidence ellipsoid. Bounds within the 95% ellipsoid are referred to as being data favored. Fig. 3 illustrates the implication of this restriction for the extreme bounds on the parameter values. In this illustration the extreme bounds on β_2 are positive and X_2 is not a fragile variable while bounds on β_1 cover zero and thus X_1 is fragile.

A major advantage in using EBA is that prior distributions only have to be specified for certain sets of variables, yet bounds can be computed for all variables in the model. Following Leamer (1982) we specify a natural conjugate prior for a set of doubtful variables, or those variables which could plausibly be dropped from a specification.[24] In this paper, these are the regional binary variables; the remaining variables, called free variables, are not linked to a proper prior specification. Free variables are associated with a diffuse prior. The effects of the free/doubtful distinction on the information contract curve for our two dimensional example are illustrated in Fig. 4. In this figure, β_1 bounds do not cover zero and X_1 is not considered fragile.

Table 2 reports the maximum and minimum bounds for the posterior means for the non-dummy variables with the widest possible bounds. Column 1 reports the maximum likelihood estimates (ordinary least squares) for the entire model. Columns 2 and 3 report the unconstrained EBA maximum and minimum values when regional dummy variables are doubtful and all other variables free.[25] The last two columns show the EBA maximum and minimum values that lie within a 95% confidence ellipsoid with all variables specified as doubtful.

When the regional variables are considered doubtful, non-fragile inferences are obtained for eleven explanatory variables: CARYEAR, PERSELAW, SPEEDMAX, BEER, MLDA, CELLPHONES, POVERTY, UNEMPLOY, REALINC, COLLEGE, and SUICIDE. These cells are indicated with boldface type in Table 2.[26] When all variables are doubtful, EBA bounds necessarily cover zero. However, the data favored extreme bounds are non-fragile for four variables: CARYEAR, CELLPHONES, POVERTY, and COLLEGE (these are highlighted in boldface type in Table 2).

Although EBA as discussed in this paper provides insight into the range of values that the posterior means can take, it does not pay direct attention to the posterior probabilities of the corresponding models. Models that correspond with extreme posterior means may be quite implausible. The next two procedures address this issue.

[23] See Mayer (2007).

[24] Dropping a variable forces a very strong prior belief that the coefficient is exactly equal to zero with perfect precision.

[25] Table 2 EBA results are based on the inclusion of regional dummy variables but the results are not shown for them or for the constant term which was included as a free variable.

[26] In Table 2, precision is truncated at 4 decimal places.

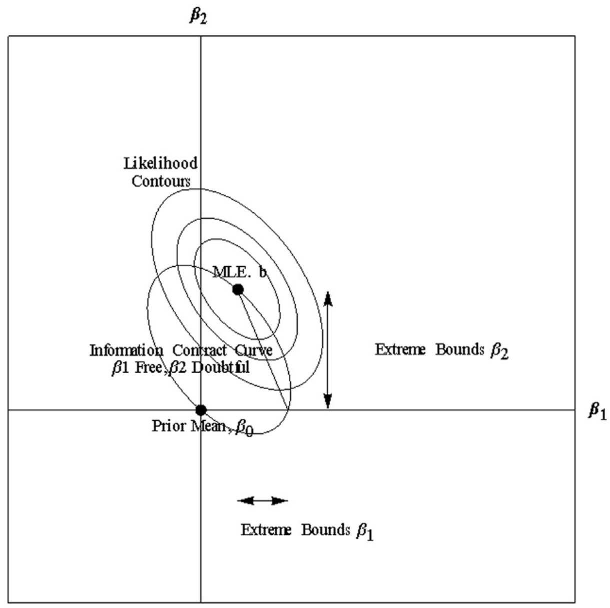

Fig. 4. Information contract curve: β_1 free, β_2 doubtful.

5. Stochastic Search Variable Selection (SSVS)

Stochastic Search Variable Selection was introduced by George and McCulloch (1993) to address the issue of what variables to include in a regression specification. Because it is computationally burdensome, it is one of many recent procedures in Bayesian analysis that takes advantage of the ability to integrate over multidimensional spaces using Markov Chain Monte Carlo Methods (MCMC) typically found when dealing with analyses of the posterior density. This is done with a Gibbs sampler.

All k of the explanatory variables are included at each iteration of the Markov chain. This implies a full conditional distribution for each variable allowing the application of the Gibbs sampler to the hierarchical Bayesian model. The variable selection choice is enabled by means of a latent variable, γ. Each model is represented by a binary vector $\gamma = (\gamma_1, \gamma_2, \ldots \gamma_k)$ with $\gamma_i = 1$ if the explanatory variable is to be effectively included in the model and $\gamma_i = 0$ if the variable is to be effectively excluded from the model. The prior distributions on the slope parameters (β's) for the explanatory variables are distributed normally with mean zero and variance $c^2\tau_i^2$ when $\gamma_i = 1$, i.e., $\beta \sim N(0, c^2\tau_i^2)$, and normally with mean zero and variance τ_i^2 i.e., $\beta \sim N(0, \tau_i^2)$ when $\gamma_i = 0$ with c greater than 1. This is illustrated in Fig. 5. Effective inclusion implies a diffuse prior analogous to a free variable in EBA. The effective exclusion of the ith variable is imposed by forcing β_i to be close to zero. This is written as:

$$\beta_i|\gamma_i \sim (1-\gamma_i)N\left(0, \tau_i^2\right) + \gamma_i N\left(0, c^2\tau_i^2\right). \tag{4}$$

This framework results in sets of posterior distributions for all vectors γ of dimension k and pays attention to the relatively sharp prior distribution around zero when a variable is effectively excluded in a model compared with a more diffuse prior when a variable is effectively included.[27]

Table 2
Maximum likelihood & extreme bounds analysis for the fatality model specification (Boldface type indicates non-fragile bounds).

Variable name	Maximum likelihood estimate	EBA maximum policy prior	EBA minimum policy prior	EBA maximum 95% likelihood all doubtful	EBA minimum 95% likelihood all doubtful
CARYEAR	−0.1152	**−0.0904**	**−0.1249**	**−0.0642**	**−0.1659**
PERSELAW	−0.2274	**−0.2063**	**−0.2982**	0.0189	−0.4732
INSPECTION	0.0688	0.1789	−0.0265	0.2608	−0.1234
SPEEDMAX	0.0124	**0.0174**	**0.0079**	0.0272	−0.0024
SEATBELT	0.0628	0.1521	−0.0082	0.3328	−0.2074
BEER	0.2900	**0.4094**	**0.1298**	0.6775	−0.0981
MLDA	0.0348	**0.0435**	**0.0079**	0.1424	−0.0728
YOUNG	0.2111	2.4739	−0.2654	3.9486	−3.5268
CELLPHONES	0.0178	**0.0191**	**0.0142**	**0.0297**	**0.0059**
POVERTY	0.0410	**0.0550**	**0.0344**	**0.0729**	**0.0090**
UNEMPLOY	−0.0219	**−0.0039**	**−0.0423**	0.0301	−0.0740
REALINC	0.0000	**0.0000**	**0.0000**	0.0000	−0.0000
HIGHSCHOOL	0.0015	0.0118	−0.0125	0.0268	−0.0239
COLLEGE	−0.0359	**−0.0214**	**−0.0476**	**−0.0055**	**−0.0661**
CRIME	0.0000	0.0000	−0.0000	0.0000	−0.0000
SUICIDE	0.0254	**0.0704**	**0.0158**	0.0676	−0.6895

Each variable is examined in random order at the end of each iteration of the Gibbs sampler to evaluate the marginal effect of effectively including/excluding that variable in the model. Based on this, a probability of including the variable is computed and the value of γ_i for the next iteration is computed stochastically based on this probability. Initial values of the γ_i are all set at 1 and the initial probabilities of inclusion are set at 0.5. Then a stochastic iteration scheme is implemented using Gibbs sampling to search for the models with the highest posterior densities.[28] In particular, the Gibbs sampler begins with initialized parameters $\gamma^{(0)}$, $\beta^{(0)}$, $\sigma^{2(0)}$ and generates the sequence $\gamma^{(1)}$, $\beta^{(1)}$, $\sigma^{2(1)}$, $\gamma^{(2)}$, $\beta^{(2)}$, $\sigma^{2(2)}$, …. This sequence converges to a posterior distribution which supplies the complete posterior $P(\beta, \sigma^2, \gamma|Y)$. Concurrent with the iterative values of the vector γ are iterative values of the vector p, the probability of variable inclusion, enabling us to compute an expected value of the vector β at each iteration.

Table 3 summarizes the findings for SSVS for the linear fatality model based on 100,000 iterations.[29] These results utilize standardized data with boldface type in cells indicating non-fragile bounds, i.e., the maximum and minimum range does not cover zero. The first column (Mean coefficient) gives the weighted average for the sequence of slope coefficients, weighted by the probability of inclusion. The second column (Probability inclusion) is the mean value of the probability of a variable's inclusion in the model. The third and fourth columns are the minimum and maximum values found over the SSVS iterations. It should be noted that the regional dummies were treated like any other variable. Of the three variables with the highest values for inclusion (probability of inclusion > 0.99) in the model, all are EBA non-fragile variables (CARYEAR, CELLPHONES, and POVERTY) as presented in the final two columns of Table 2. The final column in Table 3 provides the mean values of the iterative t-statistics indicating the precision ratio.[30]

[27] c and τ_i are choice variables. In this paper the reported results are for $c = 10$ and $\tau_i = 2\log(c)\,(c^2/(1-c^2))^{-5}\sigma_{\beta i}$ where the parameter $\sigma_{\beta i}$ is the OLS coefficient standard deviation. This choice is consistent with George and McCulloch (1993) and follows their notation.

[28] In this paper, SSVS was implemented via Markov chain Monte Carlo methods using *R*. This code is available on request.

[29] The first 500 iterations were deleted as a break-in period so there were a total of 99,500 iterations employed in the results reported.

[30] A t-statistic is calculated for each iteration and is multiplied by the probability of the variable's inclusion at that iteration. If the mean probability of inclusion is one this is the density of t values. If probabilities vary, then those with low probabilities are devalued.

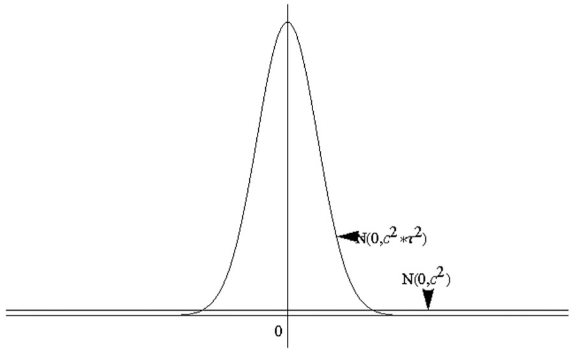

Fig. 5. Concentrated density $N(0,c^{2*}\tau^2)$, diffuse density $N(0,c^2)$.

6. Bayesian Model Averaging (BMA)

Bayesian Model Averaging was addressed extensively by Raftery et al. (1997) following a suggestion by Leamer (1978) and acknowledging George and McCulloch (1993). By averaging across many model specifications, especially among those with high posterior probabilities, BMA is able to explicitly account for model uncertainty as it relates to parameter estimation. As presented in Hoeting, Madigan, Raftery and Volinsky (1999), BMA provides a straightforward method to summarize the effects of explanatory variables as measured by their regression coefficients as they are manifested in assorted models.

Table 4 summarizes BMA analysis for the same model presented above (in Tables 2 and 3), regressing fatality rates on the core set of explanatory variables.[31] These results, as in EBA, are based on raw (non-standardized data). Regional binary variables were included in the analysis but are not reported in Table 4. The column headed "Probability inclusion" gives the posterior probability that the particular variable is included in the model. The "Posterior Mean" column shows the average posterior mean for that variable in the BMA runs and "Posterior Standard Deviation" is the average posterior standard deviation for that variable. The best performing model included 16 explanatory variables. In that particular model, INSPECTION, SEATBELT, UNEMPLOY, MLDA, HIGHSCHOOL, and CRIME were not present. Over all models, BMA never chooses to include SEATBELT, HIGHSCHOOL, or CRIME, and rarely chooses to include MLDA. The procedure always includes CARYEAR, SPEEDMAX, PERSELAW, BEER, CELLPHONES, POVERTY, and COLLEGE. BMA selects SUICIDE 95.8% of the time.[32] Of these eight favored variables, all are non-fragile under EBA when the regional dummy variables are considered doubtful (the policy prior). SPEEDMAX is an explanatory variable that was non-fragile under EBA. BMA selects this variable 100% of the time but has a probability of inclusion of 40.97% under SSVS. Although there are some differences, Tables 2–4 demonstrate considerable agreement among EBA, SSVS, and BMA model choice.

7. Comparing EBA, BMA, and SSVS

Table 5 summarizes the findings from the three Bayesian procedures we have utilized in this research. The cells in boldface type for the three variables CARYEAR, CELLPHONES, and POVERTY indicate agreement between EBA robustness, and SSVS and BMA inclusion certainty. For EBA, the boldface type indicates that the posterior range did not cover zero for either the policy prior where regional variables were doubtful, or for the data favored 95% likelihood bounds when all variables were considered doubtful. For the EBA columns, a "YES" indicates bounds do not cover zero – or are non-fragile. For SSVS and BMA, the shading indicates that both procedures resulted in a 100% probability of variable inclusion. Only these three variables pass the most rigorous tests of EBA non-fragility and 100% inclusion in SSVS and BMA. COLLEGE nearly passes these tests. It is non-fragile, is nearly always included in BMA, and is included just under 65% of the time in SSVS.

SPEEDMAX and BEER are non-fragile under the EBA policy prior and are always selected by BMA. SPEEDMAX has a nearly 41% probability of inclusion in SSVS although BEER has only a 16.7% probability of inclusion via this criterion. An interesting variable is SUICIDE in that it is non-fragile under the policy prior, has nearly a 70% probability of being included in SSVS, and is nearly always included in BMA.

By using standardized data we can assess the relative importance of each explanatory variable. SSVS posterior means and bounds are shown in Fig. 6. The foremost variable is CARYEAR which highlights how improvements in automobile safety have had a pronounced effect on lowering fatality rates. The second tier of most important variables includes CELLPHONES, POVERTY, COLLEGE, and SUICIDE.

Table 3
Stochastic search variable selection for the fatality rate model specification (Boldface type indicates non-fragile bounds).

Variable name	Mean coefficient	Probability inclusion	SSVS MIN	SSVS MAX	Mean *t*-Statistic
CARYEAR	−0.323	1.000	**−0.408**	**−0.228**	−22.257
PERSELAW	−0.009	0.293	**−0.046**	**−0.000**	−2.371
INSPECTION	0.001	0.436	−0.012	0.020	0.514
SPEEDMAX	0.002	0.409	−0.020	0.036	0.521
SEATBELT	0.000	0.443	−0.019	0.019	0.123
BEER	0.013	0.167	**0.001**	**0.122**	1.606
MLDA	0.001	0.438	−0.017	0.026	0.325
YOUNG	0.015	0.331	−0.016	0.095	1.509
CELLPHONES	0.148	1.000	**0.077**	**0.207**	14.374
POVERTY	0.119	1.000	**0.056**	**0.176**	12.206
UNEMPLOY	−0.014	0.337	−0.096	0.015	−1.216
REALINC	0.020	0.390	−0.066	0.174	1.071
HIGHSCHOOL	−0.013	0.378	−0.139	0.061	−0.990
COLLEGE	−0.105	0.645	**−0.260**	**−0.000**	−7.610
CRIME	0.008	0.383	−0.055	0.096	0.757
SUICIDE	0.100	0.691	−0.005	0.233	7.796

Table 4
Bayesian model averages.

Variable name	Probability inclusion	Posterior mean	Posterior standard deviation
CARYEAR	100.00	−0.0924	0.0056
SPEEDMAX	100.00	0.0117	0.0023
INSPECTION	11.60	0.0078	0.0240
PERSELAW	100.00	−0.2247	0.0384
SEATBELT	0.00	0.0000	0.0000
BEER	100.00	0.3258	0.0636
MLDA	0.70	0.0001	0.0023
YOUNG	18.40	0.2290	0.5419
CELLPHONES	100.00	0.0144	0.0016
POVERTY	100.00	0.0306	0.0049
UNEMPLOY	28.10	−0.0052	0.0094
REALINC	63.20	0.0000	0.0000
HIGHSCHOOL	0.00	0.0000	0.0000
COLLEGE	100.00	−0.0309	0.0058
CRIME	0.00	0.0000	0.0000
SUICIDE	95.80	0.0254	0.0088

[31] BMA results were obtained using the bicreg procedure. See Raftery, Hoeting, Volinsky, Painter, and Yeung (2009).

[32] Unlike SSVS, BMA does drop variables over models. SSVS selects CARYEAR, CELLPOP, and POVERTY with an inclusion probability of 100%.

Table 5
Comparison of EBA, SSVS, and BMA results.

Variable	EBA non-fragile policy prior	EBA non-fragile 95% likelihood all doubtful	SSVS probability of inclusion	BMA probability of inclusion
CARYEAR	**YES**	**YES**	**100.0**	**100.0**
PERSELAW	YES		29.30	100.0
INSPECTION			43.6	11.6
SPEEDMAX	YES		40.9	100.0
SEATBELT			44.3	0.0
BEER	YES		16.7	100.0
MLDA	YES		43.8	0.7
YOUNG			33.1	18.4
CELLPHONES	**YES**	**YES**	**100.0**	**100.0**
POVERTY	**YES**	**YES**	**100.0**	**100.0**
UNEMPLOY	YES		33.7	28.1
REALINC	YES		39.0	63.2
HIGHSCHOOL			37.8	0.0
COLLEGE	YES	YES	64.5	100.0
CRIME			38.3	0.0
SUICIDE	YES		69.1	95.8

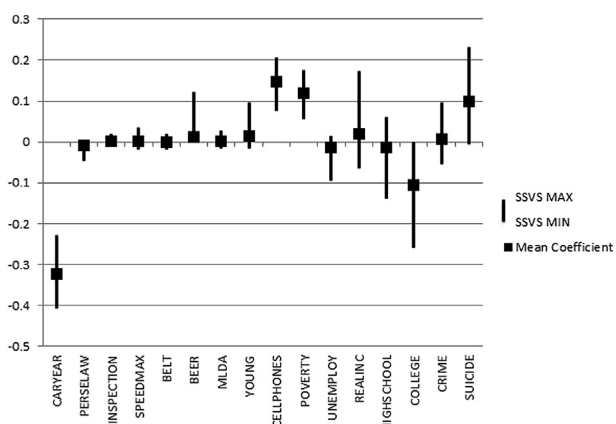

Fig. 6. Posterior means & bounds for the fatality rate model: SSVS procedure.

8. Concluding comments

One of the most important statistical problems today is the task of variable selection in regression model choice.[33] Dealing with both parameter and model uncertainty 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 three Bayesian methods which triangulate on regression models sharing several important variables related to the issue of motor vehicle fatalities in the United States. With our data, there are over 134 million model specifications. EBA, BMA, and SSVS procedures are nicely suited to explore this high dimensional model space and find sets of model specifications that have a posterior probability support. The procedures are based on solid probability and statistical theory and provide researchers with inferential tools that are not a part of the non-Bayesian toolkit.

The recommendations forthcoming from these statistical analyses include reducing the use of cell phones by drivers – a policy enacted already by ten states and the District of Columbia (California, Connecticut, Delaware, Maryland, Nevada, New Jersey, New York, Oregon, Washington, and West Virginia). The extension of cell phone bans to include hands-free devices and a complete ban on text-messaging by all drivers across states is worthy of study. Evaluating the effect of stricter police enforcement combined with different fine structures and educational announcements are potentially important areas of investigation.

The strength of the poverty link is most likely associated with fundamental infrastructural differences between high and low poverty states. Such differences can be manifest in improved highways, traffic enforcement, newer fleet age, and faster emergency response to vehicle crashes.[34]

Alcohol consumption clearly remains a major contributor to crash fatalities. 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.[35]

Of particular interest are results associated with suicides, especially considering that motor vehicle fatalities and suicide are among the leading causes of death among young persons in the United States.[36] Suicides have been found to be an issue in other areas of transportation safety, such as with respect to railroad related fatalities.[37] The strong SUICIDE relationship with motor vehicle fatality rates is non-fragile under the EBA policy prior and is nearly always included in Bayesian Model Averaging. Our findings replicate earlier studies showing that the high suicide states are also the high motor vehicle fatality states.[38] 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 investigating the effectiveness of posting phone numbers/help lines for those in need of emotional/psychiatric assistance and/or investing public monies for additional psychiatric health care, or policies designed to reduce recklessness and violent behaviors.[39]

Finally, we note the importance of CARYEAR as a fundamental determinant of crash related fatalities as seen in Fig. 6. The elasticity of CARYEAR measured with the mean coefficient provided with BMA is the largest (in absolute value) associated with the five major variables of interest by orders of magnitude. For example, our BMA estimates suggest that there would be 64% decline in fatality rates if a state could modernize their cars by 7 years, which is about a 1 standard deviation change of CARYEAR in our sample. Although this might be an overstatement, Wyoming did exhibit a 55% decline in fatality rates coinciding with a fleet modernization of 7 years.[40] Upgrading the fleet of motor vehicles is almost automatic through time as vehicles wear out and are replaced. The force of modernization is a potentially important policy recommendation. More specifically, the efficaciousness of government funded rebates and/or the reduction of taxes on newly purchased vehicles could be evaluated with respect to life-saving, injury-prevention, and property damage avoidance.

[33] See Breiman (2001).

[34] These factors are well described in Cubbin and Smith (2002).
[35] See Chaloupka, Saffer, and Grossman (1993) and Freeborn and McManus (2007).
[36] See Centers for Disease Control and Prevention, National Center for Injury Prevention and Control (2012).
[37] See, for example, Savage (2007).
[38] These are Region 9 states in our model.
[39] See Savage (2007) and Conner et al. (2001).
[40] The elasticities of the variables CELLPHONES, POVERTY, COLLEGE, and SUICIDE are: 0.12, 0.19, −0.32, and 0.15 respectively.

Acknowledgments

An earlier version of this paper was presented to the Transportation and Public Utilities Group at the 2012 ASSA annual meeting in Chicago. The authors appreciate insightful comments by James Peoples. Loeb acknowledges the research support of a Rutgers University Research Council Grant. We also wish to thank the editor and two anonymous referees for providing valuable comments.

Appendix. Data sources

Name	Data source
FATAL	Highway Statistics (various years), Federal Highway Administration, Traffic Safety Facts (various years), National Highway Traffic Safety Administration
CARYEAR	National Automobile Dealers Association (various years) and the National Household Travel Survey, U.S. Department of Transportation.
PERSELAW	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
INSPECTION	Highway Statistics (various years), Federal Highway Administration
SPEEDMAX	Highway Statistics (various years), Federal Highway Administration
SEATBELT	Traffic Safety Facts (various years), National Highway and Traffic Safety Administration
BEER	U.S. 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, U.S. Census Bureau
YOUNG	State Population Estimates (various years), U.S. Census Bureau http://www.census.gov/population/www/estimates/statepop.html
CELLPHONES	Cellular Telecommunication and Internet Association Wireless Industry Survey, International Association for the Wireless Telecommunications Industry.
POVERTY	Statistical Abstract of the United States (various years), U.S. Census Bureau website http://www.census.gov/hhes/poverty/histpov19.html
UNEMPLOY	Statistical Abstract of the United States (various years), U.S. Census Bureau
REALINC	State Personal Income (various years), Bureau of Economic Analysis website http://www.bea.doc.gov/bea/regional/spi/dpcpi.htm
HIGHSCHOOL	Digest of Education Statistics (various years), National Center for Education Statistics, Educational Attainment in the United States (various years), U.S. Census Bureau
COLLEGE	Digest of Education Statistics (various years), National Center for Education Statistics, Educational Attainment in the United States (various years), U.S. Census Bureau
CRIME	Statistical Abstract of the United States (various years), U.S. Census Bureau
SUICIDE	Statistical Abstract of the United States (various years), U.S. 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

References

Asch, P., & Levy, D. T. (1987). Does the minimum drinking age affect traffic fatalities? *Journal of Policy Analysis and Management, 6*, 180–192.

Beede, K. E., & Kass, S. J. (2006). Engrossed in conversation: the impact of cell phones on simulated driving performance. *Accident Analysis and Prevention, 38*, 415–421.

Bills, C. B., Grabowski, J., & Li, G. (2005). Suicide by aircraft: a comparative analysis. *Aviation, Space, and Environmental Medicine, 76*(8), 715–719.

Blattenberger, G., Fowles, R., Loeb, P. D., Clarke, W. A. (2009). Understanding the cell phone effect on motor vehicle fatalities using classical & Bayesian methods. In *Presented at the Allied Social Sciences Associations Meeting (TPUG), San Francisco, USA January 2009*.

Blattenberger, G., Fowles, R., Loeb, P. D., & Clarke, W. A. (2012). Understanding the cell phone effect on vehicle fatalities: a Bayesian view. *Applied Economics, 44*, 1823–1835.

Breiman, L. (2001). Statistical modeling: the two cultures. *Statistical Science, 16*(3), 199–231.

Bylow, L. F., & Savage, I. (1991). The effect of airline deregulation on automobile fatalities. *Accident Analysis and Prevention, 23*, 443–452.

Campbell, B. J., & Campbell, F. A. (1986). *Seat belt law experience in four foreign countries compared to the United states*. Falls Church, VA: AAA Foundation for Traffic Safety.

Campbell, B. J., Stewart, J. R., & Campbell, F. A. (1986). *Early results of seat belt legislation in the United states of America*. Chapel Hill, NC: University of North Carolina Highway Safety Research Center.

Campbell, B. J., Stewart, J. R., & Campbell, F. A. (1987). *1985–1986 experience with belt laws in the United states*. Chapel Hill, NC: University of North Carolina Highway Safety Research Center.

Centers for Disease Control and Prevention, National Center for Injury Prevention and Control. Web-based injury statistics query and reporting system (WISQARS): [online] Available at: <http://www.cdc.gov/injury/wisqars/> Accessed 07.07.12.

Chaloupka, F. J., Saffer, H., & Grossman, M. (1993). Alcohol-control policies and motor vehicle fatalities. *Journal of Legal Studies, 22*, 161–218.

Chamberlain, G., & Leamer, E. E. (1976). Matrix weighted averages and posterior bounds. *Journal of the Royal Statistical Society Series B, 38*(1), 73–84.

Chapman, S., & Schoefield, W. N. (1998). Lifesavers and samaritans: emergency use of cellular (mobile) phones in Australia. *Accident Analysis and Prevention, 30*, 815–819.

Clarke, W. A., & Loeb, P. D. (2005). The determinants of train fatalities: keeping the model on track. *Transportation Research a Part E, 41*, 145–158.

Clyde, M. A. (1999). Bayesian model averaging and model search strategies. *Bayesian Statistics, 6*, 157–185.

Coate, D., & Markowitz, S. (2004). The effects of daylight and daylight savings time on U.S. pedestrian fatalities and motor vehicle occupant fatalities. *Accident Analysis and Prevention, 36*(3), 351–357.

Cohen, A., & Einav, L. (2003). The effects of mandatory seat belt laws on driving behavior and traffic fatalities. *Review of Economics and Statistics, 85*(4), 828–843.

Conner, K. R., Cox, C., Duberstein, P. R., Tian, L., Nisbet, P. A., & Conwell, Y. (2001). Violence, alcohol, and completed suicide: a case-control study. *American Journal of Psychiatry, 158*, 1701–1705.

Connolly, J. F., Cullen, A., & McTigue, O. (1995). Single road traffic deaths: accident or suicide. *Crisis: The Journal of Crisis Intervention and Suicide Prevention, 16*(2), 85–89.

Consiglio, W., Driscoll, P., Witte, M., & Berg, W. P. (2003). Effect of cellular telephone conversations and other potential interference on reaction time in braking responses. *Accident Analysis and Prevention, 35*, 495–500.

Crain, W. M. (1980). *Vehicle safety inspection systems*. Washington, D.C.: American Enterprise Institute for Public Policy Research.

CTIA. (2011). *The wireless association* [online] Available at http://www.ctia.org Accessed 10.02.11.

Cubbin, C., & Smith, G. S. (2002). Socioeconomic inequalities in injury: critical issues in design and analysis. *Annual Review of Public Health, 23*, 349–375.

Cullen, S. (1998). Aviation suicide: a review of general aviation accidents in the U.K., 1970–96. *Aviation, Space, and Environmental Medicine, 69*(7), 696–698.

Etzerdorfer, E. (1995). Single road traffic deaths: accidents or suicide? Comment. *Crisis: The Journal of Crisis Intervention and Suicide Prevention, 16*(4), 188–189.

Evans, L. E. (1991). *Traffic safety and the driver*. New York: Van Norstrand Reinhold.

Forester, T., McNown, R. F., & Singell, L. D. (1984). A cost benefit analysis of the 55 mph speed limit. *Southern Economic Journal, 50*, 631–641.

Fowles, R. (1988). Micro EBA. *American Statistician, 4*, 274.

Fowles, R., & Loeb, P. D. (1989). Speeding, coordination, and the 55-MPH limit: comment. *American Economic Review, 79*, 916–921.

Fowles, R., & Loeb, P. D. (1992). The interactive effect of alcohol and altitude on traffic fatalities. *Southern Economic Journal, 59*, 108–112.

Fowles, R., Loeb, P. D., & Clarke, W. A. (2010). The cell phone effect on motor vehicle fatality rates: a Bayesian and classical econometric evaluation. *Transportation Research Part E, 46*, 1140–1147.

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

Garbacz, C. (1990). How effective is automobile safety regulation? *Applied Economics, 22*, 1705–1714.

Garbacz, C., & Kelly, J. G. (1987). Automobile safety inspection: new econometric and benefit/cost estimates. *Applied Economics, 19*, 763–771.

George, E. I., & McCulloch, R. R. (1993). Variable selection via Gibbs sampling. *Journal of the American Statistical Association, 88*, 881–889.

Glassbrenner, D. (2005). *Driver cell phone use in 2005 – overall results* Traffic Safey Facts: Research Note, NHTSA, DOT HS 809967.[online] Available at http://www-nrd.nhtsa.dot.gov/Pubs/809967.PDF Accessed 11.02.11.
Hatfield, N. J., & Hinshaw, W. M. (1987). *Evaluation of the fatality reduction effectiveness of the Texas mandatory safety belt law*. The Texas A&M University System, Texas Transportation Institute.
Hoeting, J., Madigan, D., Raftery, A., & Volinsky, C. (1999). Bayesian model averaging: a tutorial. *Statistical Science, 14*(4), 382–417.
Hoxie, P., & Skinner, D. (1987). *Effects of mandatory seatbelt use laws on highway fatalities in 1985*. Cambridge, MA: U.S. Department of Transportation, Research and Special Programs Administration, Transportation Systems Center.
Huffine, C. L. (1971). Eqivocal single-auto traffic fatalities. *Life-threatening Behavior, 1*(2), 83–95.
Keeler, T. E. (1994). Highway safety, economic behavior, and driving environment. *American Economic Review, 84*(3), 684–693.
Key, J. T., Pericchi, L., & Smith, A. F. M. (1999). Bayesian model choice: what and why? *Bayesian Statistics, 6*, 343–370.
Laberge-Nadeau, C., Maag, U., Bellavance, F., Lapiere, S. D., Desjardins, D., Messier, S., et al. (2003). Wireless telephones and risk of road crashes. *Accident Analysis and Prevention, 35*, 649–660.
Lave, C. A. (1985). Speeding, coordination, and the 55 MPH limit. *American Economic Review, 75*, 1159–1164.
Leamer, E. E. (1978). *Specification searches: Ad Hoc inference with non-experimental data*. New York: Wiley & Sons.
Leamer, E. E. (1982). Sets of posterior means with bounded variance priors. *Econometrica, 50*(3), 725–736.
Leamer, E. E. (1983). Let's take the con out of econometrics. *American Economic Review, 73*(1), 31–43.
Leamer, E. E. (1985). Sensitivity analyses would help. *American Economic Review, 75*(3), 308–313.
Leamer, E. E. (1997). Revisiting Tobin's 1950 study of food expenditure. *Journal of Applied Econometrics, 12*(5), 533–553.
Levy, D. T., & Asch, P. (1989). Speeding, coordination, and the 55-MPH limit: comment. *American Economic Review, 79*, 913–915.
Lewis, R. J., Johnson, R. D., Whinnery, J. E., & Forster, E. M. (2007). Aircraft-assisted pilot suicides in the United States, 1993–2002. *Archives of Suicide Research, 11*(2), 149–161.
Loeb, P. D. (1985). The efficacy and cost-effectiveness of motor vehicle inspection using cross-sectional data – an econometric analysis. *Southern Economic Journal, 52*, 500–509.
Loeb, P. D. (1987). The determinants of motor vehicle accidents with special consideration to policy variables. *Journal of Transport Economics and Policy, 21*, 279–287.
Loeb, P. D. (1988). The determinants of motor vehicle accidents – a specification error analysis. *Logistics and Transportation Review, 24*, 33–48.
Loeb, P. D. (1990). Automobile safety inspection: further econometric evidence. *Applied Economics, 22*, 1697–1704.
Loeb, P. D. (1991). *The effectiveness of seat belt legislation in reducing driver involved injury and fatality rates in Texas, Maryland, and California – Final report*. U.S.D.O.T./NHTSA.
Loeb, P. D., & Clarke, W. A. (2007). The determinants of truck accidents. *Transportation Research Part E, 43*, 442–452.
Loeb, P. D., Clarke, W. A., & Anderson, R. (2009). The impact of cell phones on motor vehicle fatalities. *Applied Economics, 41*, 2905–2914.
Loeb, P. D., & Gilad, B. (1984). The efficacy and cost-effectiveness of vehicle inspection – a state specific analysis using time series data. *Journal of Transport Economics and Policy, 18*, 145–164.
Loeb, P. D., Talley, W. K., & Zlatoper, T. (1994). *Causes and deterrents of transportation accidents: An analysis by mode*. Westport, CT: Quorum Books.
Lund, A. K., Pollner, J., & Williams, A. F. (1987). Preliminary estimates of the effects of mandatory seat belt use laws. *Accident Analysis and Prevention, 17*, 219–223.
McCarthy, P. (1992). Highway safety implications of expanded use of longer combination vehicles (LCVs). In *Paper presented at the 6th World Conference on Transport Research, Lyon, France*.
McEvoy, S. P., Stevenson, M. R., McCartt, A. T., Woodward, M., Haworth, C., Palamara, P., et al. (2005). Role of mobile phones in motor vehicle crashes resulting in hospital attendance: a case-crossover study. *British Medical Journal, 33*, 428–435.
Mayer, T. (2007). The empirical significance of econometric models. In *Measurement in economics: A handbook*. London: Marcel Boumans.
Mazess, R. B., Picon-Reategui, E., Thomas, R. B., & Little, M. A. (1968). Effects of alcohol and altitude on man during rest and work. *Aerospace Medicine, 39*, 403–406.
Newman, H. W. (1949). The effect of altitude on alcohol tolerance. *Quarterly Journal on Studies of Alcohol, 10*, 398–403.
Neyens, D. M., & Boyle, L. N. (2007). The effect of distractions on the crash types of teenage drivers. *Accident Analysis and Prevention, 39*, 206–212.
Partyka, S. C. (1988). *Lives saved by seat belts from 1983 through 1987*. Washington, D.C: NHTSA Technical Report.
Peltzman, S. (1975). The effect of automobile regulation. *Journal of Political Economy, 93*, 677–725.
Phillips, D. P. (1977). Motor vehicle fatalities increase just after publicized suicide stories. *Science, 196*, 1464–1466.
Phillips, D. P. (1979). Suicide, motor vehicle fatalities, and the mass media: evidence toward a theory of suggestion. *American Journal of Sociology, 84*(5), 1150–1174.
Pokorny, A. D., Smith, J. P., & Finch, J. R. (1972). Vehicular suicides. *Life-Threatening Behavior, 2*(2), 105–119.
Porterfield, A. L. (1960). Traffic fatalities, suicide, and homicide. *American Sociological Review, 25*(6), 897–901.
Poysti, L., Rajalin, S., & Summala, H. (2005). Factors influencing the use of cellular (mobile) phone during driving and hazards while using it. *Accident Analysis and Prevention, 37*, 47–51.
Raftery, A., Hoeting, J., Volinsky, C., Painter, I., & Yeung, K. Y. (2009). *BMA: Bayesian model averaging* R package version 3.10. [online]. http://CRAN.R-project.org/package=BMA Accessed 11.0.11.
Raftery, A. E., Madigan, D., & Hoeting, J. A. (1997). Bayesian model averaging for linear regression models. *Journal of the American Statistical Association, 92*, 179–191.
Redelmeier, D. A., & Tibshirani, R. J. (1997). Association between cellular-telephone calls and motor vehicle collisions. *New England Journal of Medicine, 336*, 453–458.
Reinfurt, D. W., Campbell, B. J., Stewart, J. R., & Stutts, J. C. (1988). *North Carolina's occupant restraint law: A three year evaluation*. Chapel Hill, NC: University of North Carolina Highway Safety Research Center.
Saffer, H., & Grossman, M. (1987a). Beer taxes, the legal drinking age, and youth motor vehicle fatalities. *Journal of Legal Studies, 16*, 351–374.
Saffer, H., & Grossman, M. (1987b). Drinking age laws and highway mortality rates: cause and effect. *Economic Inquiry, 25*, 403–417.
Savage, I. (2007). Trespassing on the railroad. In S. M. Dennis, & W. K. Talley (Eds.), *Research in transportation Economics: Railroad economics*. Amsterdam: Elsevier Science.
Savage, I. (2010). Analysis of data on pedestrian/rail safety in Illinois 2005–2009. In *Pedestrian Rail Safety Symposium, June 22. 2010 and Proceedings of the DuPage Rail Safety Council Conference, September 29, 2010*.
Skinner, D., & Hoxie, P. (1988). *Effects of seatbelt laws on highway fatalities: Update – April 1988*. Cambridge, MA: U.S. Department of Transportation, Research and Special Programs Administration, Transportation Systems Center.
Snyder, D. (1989). Speeding, coordination, and the 55-MPH limit. *American Economic Review, 79*, 922–925.
Sommers, P. M. (1985). Drinking age and the 55 MPH speed limit. *Atlantic Economic Journal, 13*, 43–48.
Souetre, E. (1988). Completed suicides and traffic accidents: longitudinal analysis in France. *Acta Psychiatrica Scandinavica, 77*(5), 530–534.
Streff, F. M., Schultz, R. H., & Wagenaar, A. C. (1989). *Changes in police reported injuries associated with Michigan's safety belt law: 1988 update*. Ann Arbor, MI: The University of Michigan, Transportation Research Institute.
Sullman, M. J., & Baas, P. H. (2004). Mobile phone use amongst New Zealand drivers. *Transportation Research Part F: Traffic Psychology and Behavior, 7*, 95–105.
Violanti, J. M. (1998). Cellular phones and fatal traffic collisions. *Accident Analysis and Prevention, 30*, 519–524.
Womble, K. B. (1989). *The estimated effect of safety belt use laws on injury reduction*. Washington, D.C.: U.S. Department of Transportation, NHTSA.
Zlatoper, T. J. (1984). Regression analysis of time series data on motor vehicle deaths in the United States. *Journal of Transport Economics and Policy, 18*, 263–274.