The Effects of Automobiles Recalls on the Severity of Accidents[§]

Yong-Kyun Bae[†]
Hood College

Hugo Benítez-Silva[‡]
SUNY-Stony Brook

May 25, 2010

Abstract

The number of automobile recalls in the U.S. has substantially increased over the last two decades, and after a record of over 30 million cars recalled in 2004, in the last few years it has consistently reached between 15 and 17 million, and in 2009 alone 16.4 million cars were recalled. Toyota's recall crisis in 2010 illustrates how recalls can affect a large number of American drivers and the defects connected to them can result in loss of life and serious accidents. However, in spite of the increase in public concern over recalls and the loss of property and life attached to them, there is no empirical evidence of the effect of vehicle recalls on safety. This paper investigates whether vehicle recalls reduce accidental harm measured by the severity of injuries in vehicle accidents. The results of our analysis show that if a recall for a new-year model is issued, then the severity of injuries of accidents continuously diminishes during the first year after the recall, something we do not find among cars not subject to recalls. This is because defects are repaired over time but also because drivers react by driving more carefully until the defects are fixed. To minimize the losses attached to having dangerously defective cars on our roads, both quick and timely recall issuance are needed and more detailed information on defects should be delivered to owners of defective vehicles. The latter can be made possible through simple but important policy changes by the U.S. government regarding recall information sharing with drivers and insurance companies.

Keywords: Automobile Recalls, Safety Regulation, Vehicle Defects, Car Accidents
JEL Classification : L51, L62

[§] We are grateful to John Hause, Chris Swann, Chris Snyder, Sam Peltzman, Ioana Marinescu, Adam Copeland, Katja Seim, Selcuk Eren, Na Yin, Huan Ni, Caroline Austin, four anonymous referees, a co-editor, and the editor-in-chief for their useful comments and suggestions. We also thank the participants of the Annual International Industrial Organization Conference held in Chicago, and the applied micro workshops at Florida State University, Drexel University, Cornell, Hood College, and the IAE in Barcelona. Benítez-Silva is also grateful to the Economics Departments at the University of Maryland and UPF for their hospitality during the completion of this paper. We bear sole responsibility for any remaining errors.

[†] Department of Economics and Management, Hood College. 401 Rosemont Avenue, Frederick, MD 21701. Phone: (301) 696-3692, fax: (301) 696-3771, e-mail: bae@hood.edu

[‡] Corresponding Author: Economics Department, SUNY-Stony Brook, Stony Brook, NY 11794-4384. Phone: (631) 632-7551, fax: (631) 632-7516, e-mail: hugo.benitez-silva@sunysb.edu

1. Introduction

The National Highway Traffic Safety Administration(NHTSA) reports in its 2009 annual report that in 2008, 37,261 people were killed in the 5,811,000 police-reported motor vehicle traffic crashes. More than 2.35 million people were injured, and 4,146,000 crashes involved property damage only (NHTSA, *Traffic Safety Facts*, 2009). Most accidents were caused by drivers' mistakes or misbehavior. However, it cannot be underestimated that vehicle defects may play a role in causing accidents. With 16.4 million cars recalled during 2009 alone, the recent massive recalls by Toyota, and with the number of recalls consistently above 15 million during the last decade and a half, understanding the effects of recalls have become an important research and policy question.[1]

A safety defect is defined as "a problem that exists in a motor vehicle or item of a motor vehicle equipment that possesses a risk to motor vehicle safety, and may exist in a group of vehicles of the same design or manufacture, or items of equipment of the same type and manufacture."[2] To remove defective cars from the roads, the NHTSA requests vehicle manufacturers to recall them and fix the defects at no cost to the owners of the vehicles. The number of automobile recalls has sharply increased over the last two decades. Vehicle manufacturers issued 588 separate recalls involving 14.8 million vehicles in 2007, and the industry set a record of 30.8 million vehicles in 2004. Then, the number has decreased and increased again. Toyota's recall crisis in 2010 has ignited sharp disputes over appropriate public safety policy, and the NHTSA has been criticized by lawmakers and car-safety groups for acting too slowly on complaints, while Toyota has been widely criticized by lawmakers and the public for hiding information on its vehicle defects. We believe that this puts the focus on the effectiveness of recall regulation: Are recalls effective in reducing accidental harm? If so, should our society spend more resources on this regulation? However, without quantitative evidence, any attempt to change public safety policies related to recalls is hard to defend. Clearly, if we find that recalls reduce accidental harm we can infer that manufacturers are putting on the road vehicles that are dangerous to the public, and that their systems to control the quality of their vehicles are passing the responsibility to find defects to the American drivers, something clearly unacceptable when happening in the numbers that we are seeing in the last years.

While vehicle recalls have been the object of study over the past two decades, most of that research has focused on particular aspects of recalls, such as their effect on demand, vehicle resale prices, firm valuation, liability verdicts, and initiation of recalls, rather than safety.[3] There is no quantitative evidence of the number of vehicle accidents caused by vehicle defects and how many potentially danger-

[1]While it is possible to find cases of large recalls which were linked to few complaints and even fewer accidents or fatalities, this should not be taken as evidence that these problems are not serious. If we naively believe that the probability of recall-related accidents are too low, and we leave in the hands of the manufacturers the mandate of not putting unsafe vehicles on the road, the price to pay will likely be preventable accidents, injuries, and lives. At the very least the profession should seriously tackle the question of understanding the effects of recalls, this paper is a rare step in this direction, and we hope our work encourages others to research this issue.

[2]Motor Vehicle Defects and Safety Recalls: What Every Vehicle Owner Should Know, NHTSA, 2009.

[3]These papers include Jarrell and Peltzman (1985), Crafton, Hoffer, and Reilly (1981), Hartman (1987), Hoffer, Pruitt, and Reilly (1994), Huble and Arndt (1996), Marino (1997), Rupp and Taylor (2002), Rhee and Haunschild (2003), and Bates et al. (2007).

ous vehicles are on the U.S. roads everyday.[4] Furthermore, we do not know whether vehicle recalls are effective in reducing accidental harm. So far, there has been no empirical assessment of vehicle recall regulation related to safety.

There are a number of reasons why there are relatively few studies of the effect of recalls on safety. First, there is no direct link between recall, vehicle, and accident data, therefore researchers need to do this by themselves, and to analyze the effectiveness of recalls researchers should be able to identify potentially dangerous cars. Second, vehicles may have multiple defects, and recalls are issued over time. Therefore, even if an accident has been caused by a particular defect, we do not know which defect causes the accident. Third, defects have different levels of risks. Thus, it is very difficult to measure potential risks accurately and compare them. Fourth, there is no way to find out whether particular car owners have returned their cars to manufacturers to be fixed. Fifth, drivers' behavioral change is not observable before and after recalls. Sixth, the entire recall process is time-consuming and very complicated.[5] Recently, Bae and Benítez-Silva (2010) have developed a synthetic panel model to investigate the effectiveness of recall regulation in terms of the number of accidents on the road. The basic argument of that work is that if the defects, which lead to recalls, directly or indirectly affect accidents, then one would expect recalls would lead to a reduction in the number of accidents. Their finding that indeed recalls reduce the number of accidents, complements our findings on the severity of those accidents.

This paper empirically investigates how recalls affect accident severity using publicly available accident and recall data from the NHTSA. There are two arguments suggesting that there would be a decrease in accident severity. First, behavioral response by drivers' learning about the defect can make them more aware of the risks of driving the vehicle, and may induce greater care in driving, which reduces the harm if the defect leads to an accident and greater care may also reduce the total number of accidents. The changed behavior persists until the driver has the defect repaired (if that is the case). Second, if the defect is a hazardous one, elimination of the defect may reduce the expected accident severity if a vehicle with the defect repaired is subsequently involved in an accident (even if the driver changes back to the level of care he exercised before learning about the defect through the announced recall).

The results of our analysis show that recalls reduce the severity of injuries in vehicle crashes. In particular, we find that if a recall for a new-year model is issued, then the severity of injuries of accidents continuously diminishes during the first year after the recall, something we do not find among cars not subject to major recalls. This is because defects are removed over time as drivers take corrective actions and possibly drivers increase their level of care until the defects are fixed. One clear policy implication is that both quick and timely recall actions should be taken and more detailed information on defects should be delivered to the owners of defective vehicles. Additionally, the government should make an

[4]It is hard to identify the source of an accident given that both human error and vehicle defects can play a role, what seems clear from our analysis is that the severity of accidents after recalls is significantly decreasing over time, which points to the effectiveness of recalls.

[5]The recall regulation and details about the recall process can be found on the website of the NHTSA: "http://www.nhtsa.gov".

effort to allow insurance companies to link, through the release of the Vehicle Identification Numbers (VIN) of recalled cars, the information on recalls to insured vehicles, prompting drivers to fix the problems on their cars to avoid higher premiums, and can also follow suggestions by some policy makers to include recall information on the registration renewal information that DMVs send to drivers. More broadly, given the large number of recalls and our findings, the government should consider reviewing the quality control system of car manufacturers who are putting potentially dangerous products on the roads in record numbers.

Section 2 of the paper presents a theoretical discussion on how recalls can reduce accidental harm, and discusses our identification strategy. Section 3 introduces the recall system and recent recall trends. The estimation strategy, using an ordered probit model, is discussed in section 4, along with a discussion of our data and some summary statistics. Section 5 discusses estimation results, and our concluding remarks are presented in the final section of the paper.

2. Vehicle Recalls, Safety, and Model Identification

This section studies two connected questions, and how they relate to our empirical analysis. First, we need to analyze the conditions under which recalls are efficient from the point of view of social welfare. Second, we need to analyze whether recalls reduce accidental harm and how recalls affect the distribution of injury severity of accidents.

The efficiency question needs to connect the marginal benefits of recalls with their marginal costs, where the former are measured by the monetized value of savings thanks to the possible reduction of the severity and the number of accidents resulting from an additional recall, while the marginal costs are the extra monetary costs of the additional recall. Recall costs can be comparatively easy to compute (even though costs linked to the loss of reputation and the consequences on future sales can be more difficult to calculate), however, the benefits side is more complex since the value of life is often involved in the calculation (Posnerwer(1998), Kaplow and Shavell (1999), and Ashenfelter and Greenstone(2002)).

Assume that the variable (H), representing accidental harm, properly reflects the value of life. Suppose that n is the number of accidents caused by defects, L is the injury severity per accident, and x is the variable that measures recall activity ($0 \leq x < \infty$). For simplicity, assume that the marginal costs of an additional recall activity are constant (c). Then, the social costs of accidents are given by[6]

$$SC = H(x) + cx = n(x) \cdot L(x) + cx.$$

To minimize the social costs of accidents, the marginal benefits should be equal to the marginal costs. Thus, the efficient level (x^*) of recall activity can be found by solving the following equation which results from taking first derivatives:

$$H'(x) = -c \Longrightarrow n'(x) \cdot L(x) + n(x) \cdot L'(x) = -c.$$

[6]We follow the notation of Cooter and Ulen (2000), but the interpretation of x is different in our model.

Given positive recall costs (c), recalls are not welfare-improving unless they reduce accidental harm. This means that there should be a reduction in accidental harm from either the reduction in the number of accidents ($n'(x) < 0$) or (and) the reduction in the severity of injury ($L'(x) < 0$). Bae and Benítez-Silva (2010) show that recalls reduce the number of accidents, and this paper focuses on whether recalls reduce the severity of injuries ($L'(x) < 0$). Notice that a reduction in the number of accidents, and/or a reduction in the severity of those accidents does not guarantee that the level of optimal recalls from a societal perspective will be the actual level we observe, given that social costs are unlikely to be internalized by manufacturers, and the benefits for them are more likely to be linked with reputation and assessment of the consequences on future profits, not with the reduction of accidental harm. This suggests an important role for regulation and policy, which in part already exists, but that has not been at the forefront of the discussion given the lack of quantitative evidence on the benefits of recall activities.

The second question, however, is tricky to answer if the probability that a defect causes accidents is believed to be very small. Many industry people and some experts then believe that recalls do not save lives.[7]

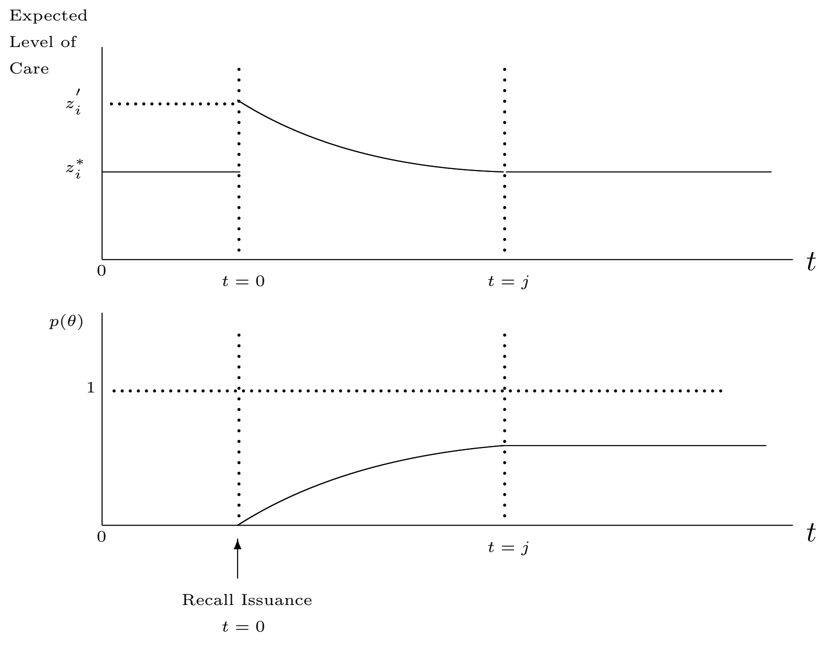


Figure 1: Driver's Level of Care and Corrective Action over Time

To analyze this issue we present a simple discussion which naturally leads to our identification strategy in the empirical analysis. Suppose that there is only one defect for a particular vehicle model.

[7]Of course, we fully understand that drivers' faults are one of the major sources of accidents. For more discussion on this, see Evans (2002).

Also suppose that the manufacturer does not know the potential risk of the defect when it sells the vehicle to the consumer. When the consumer purchases the car, she assumes that the car is safe. Given this safety evaluation, she will choose her optimal levels of care when driving, z_i^*, $(0 \leq z_i \leq \infty)$ such that her utility is maximized. In this situation the consumer will only know about the defect when the manufacturer finds it and issues a recall. Thus, there will be no change in her level of care until that recall is issued. This is illustrated as a horizontal line in the expected level of care before $t = 0$ in Figure 1. Suppose that a recall is announced at $t = 0$. Then, she may take the car to the dealer right away and fix it, or fix later and raise her level of care $(z_i^{'})$ until the defect is being fixed. Once the repair takes place, her level of care will return to the initial level (z_i^*). Therefore, the consumer's expected level of care (z_i) diminishes and the probability that the car is being fixed $(p(\theta))$ increases during the time period between $t = 0$ and $t = j$.[8] The time period could be a year or more than a year, depending on the recall and its risk.

Now, consider accidents that occur after the recall issuance. If the defect had not played any role in causing the accidents, then the severity of injuries would rather increase over time, at least up to $t = j$. This is because drivers reduce their levels of care as their cars are being fixed and they adjust their driving to a potentially hazardous situation. If the recall failed in terms of affecting neither drivers' levels of care nor the likelihood of taking corrective action, then the average severity level would not change between time 0 and time j. If these were true, then there would be no correlation between the severity of injuries and when the accidents occurred over time. If the recall were effective, then the severity of injuries would decrease at least up to time j. We do not know whether drivers raised their levels of care or took any corrective actions before the accidents or when the drivers returned to their original levels of care, therefore the exact length of the process analyzed is a random variable.

This discussion provides an identification strategy given the accident data we analyze. We focus on the accidents of a particular vehicle model. We check when the first recall for a year-model was issued, and then observe when the accident occurred. For instance, we look at the Ford Taurus 2007-year model, that had its first recall on March, 2007. Then, we observe an accident in which this vehicle model was involved, and when the accident occurred. If the accident occurred, say, in May, 2007, then the accident occurred two months after the recall was issued. We observe all the accidents with the same vehicle model and observe when they occurred. Given our previous discussion, in order to argue for a role of recalls in the reduction of the severity of accidents, the average severity of accidents occurred within two months must be greater than the average severity of accidents occurred after three, four, or more months. This is because the probability that the defect of a particular vehicle is eliminated before the accident increases over time.[9] Accordingly, the severity must diminish over time if the recall is effective in reducing accidental harm. Thus, the severity would be negatively associated with the months

[8] For particular recalls and particular vehicles affected, the expected level of care (and the probability of being fixed) need not decrease (increase) in a monotonic fashion. We are showing that case for illustrative purposes. In reality we could observe step functions as waves of owners find out about the recalls over time and adjust behavior and take their cars to be fixed.

[9] The riskier the defect is, the more responsive the drivers are. Thus the level of care diminishes over the relatively short-time period and the correction rate increases at a greater speed.

passed after the recall. Therefore, we can expect that ***the longer the time since a recall issuance is, the lower the severity of injuries of an accident is, at least during a relatively short time (about a year in our empirical application) after the recall***. Notice that this variable is identified through the variation in the sample of the time between the recall and the accident, and that is in separately identified from a vintage control of the number of months between the launching of a model and the time of the accident, given the variation of time between the launching of the model and the recalls. Additionally, we would expect that in a sample of accidents by cars not subject to recalls we should see no significant decrease in the severity of accidents over time. Our empirical analysis tests these hypotheses.

3. The U.S. Recall System and Trends

The NHTSA has the authority to require vehicle manufacturers to issue recalls, whenever their vehicles posses any potential safety-related defect that could cause loss of vehicle control such as steering, braking, tire damage, or repeated stalling. However, the Federal Motor Vehicle Safety Standard regulates the vehicle industry, and manufacturers are asked to issue recalls if serious accidents occur or are expected because of potential defects. However, the manufacturers can initiate them whenever they find critical defects that might cause serious accidental harm. In that case, the manufacturers report detailed information on the corresponding defects to the NHTSA and begin to take actions for recall issuance. In many cases, the NHTSA may find the defects first (maybe because of complaints reported from owners of defective vehicles) and require the manufacturers to issue recalls, followed by its own investigation. The manufacturers may comply with the requirement and issue recalls. If they do not agree with the government's recall decision, they can resolve the disputes in court.

The owners of the vehicles may report complaints to either the manufacturers or the NHTSA. It doesn't matter whether the owners are injured or not in accidents. If the accidents occur and the defective problems are considered as a cause of the accidents, they can file a defect report directly to the NHTSA. Once the recalls are issued, they have to take their vehicles to the places assigned by the manufacturers to be repaired. The entire recall process is lengthy and time-consuming, thus, it may take several years.[10] Once the recalls are announced, the manufacturers send notice letters to their customers and also announce them through the media so that the vehicles should be brought in and the defects fixed. After that, the Recall Management Division, part of the NHTSA, monitors the post-recall process. After the recalls are issued, then it may take several years again to finish all corrective procedures because it depends upon the vehicle owners. If the owners' addresses change and they do not notify of such a change, then there is no way for the manufacturers to send letters. The NHTSA requires manufacturers to submit "Quarterly Recall Reports" that contain detailed corrective actions in accordance with Federal Regulation, 573.6.

The recall system began its operation in 1966. Since then, the number of vehicle recalls has

[10]Recent information indicates that Toyota might have been aware of the problems with sudden acceleration on some its most popular vehicles for more than 7 years before it started to recall them.

increased over time (Figure 2). In 1966, 58 recalls were issued, and by 2008 the number of recalls had increased to 684. The increase has been particularly sharp since the mid 1990s, and in the graph we can see the number of hazardous recalls but only up to 2001 when the government stopped reporting the hazard level of recalls.[11] Since each recall involves a different number of units, we can plot the average number of units per recall over time. Figure 3 shows the annual average units per recall. In 1981 a large number of vehicles were recalled. Other than that, the average number of units per recall has increased during the 1990s and then decreased during the 2000s. For example, more than 58,000 vehicles per recall on average were issued in 1996. Since then, the number has decreased. In 2009, the average unit per recall was 33,000. Additionally, during the late 1990s and the early 2000s, there was a sharp increase in equipment recalls. In the 1990s hazardous recalls tend to be larger than the average recall.

Figure 4 shows the proportion of all total vehicles recalled initiated by domestic manufacturers, which up to the mid 1990s had fluctuated considerably every year moving from highs above 90% to as low as 50%, since then we can observe a sharp decline in the proportion of units recalled initiated by domestic manufacturers, trend that reached a low point just above 40% in the last few years. In part this reflects the growing importance of foreign manufacturers in the U.S. car market, but also their push towards massive production which seems to be coming with a lowering of their quality, reaching a pinnacle with the ongoing recalls by Toyota. It is interesting to emphasize that a larger proportion of the recalls issued by foreign manufacturers are considered hazardous compared with domestic manufacturers. It is also important to emphasize that in Figure 4 we are not accounting for some very small manufacturers that recall very small number of vehicles, but the recalls represented in the figure account for almost all the vehicles actually recalled.

Figure 5 shows the evolution of the proportion of recalled vehicles that are mandatory, meaning that they have been initiated by the government and not the manufacturers.[12] This proportion has fluctuated considerably over time, and since early 1990s, has consistently reached 60% of the vehicles recalled. Recalls can be quite costly for manufacturers and in some cases they seem reluctant to initiate the recall as they balance the cost and benefits of delaying action. Interestingly, if we just look at the proportion of recalls irrespective of number of vehicles involved (figure not shown), the proportion of mandatory recalls is never above 40%, and usually fluctuates between 20% and 30%, which indicates that mandatory recalls are usually also those with large number of units involved.

4. Estimation Strategy and Data

4.1. Empirical Model Setup

[11] Recalls can be of different types: vehicle, equipment, tire, and component recalls. In most cases when recalls are mentioned, it usually refers only to vehicle recalls. However, other types of recalls often contain hazard parts and (or) defective equipment. Recently, the recalls related to defective equipment have sharply increased. Recalls have different hazard ratings according to the potential risk levels. This information was available to the public in the past, but it is no longer the case. For additional discussions on these issues, see Bae and Benítez-Silva(2010).

[12] For more details, see Rupp and Taylor (2002), who emphasize that the recalls of new vehicle models are more likely to be initiated by the manufacturers.

Given the nature of the data that we have access to, we can only analyze the effects of recalls if accidents of particular models actually occur. In our analysis we use Police Accident Reports (PAR) provided by the NHTSA through the General Estimates System (GES) which began operations in 1988. The GES obtains its data from a nationally representative sample selected from the estimated 6 million police-reported crashes which occur annually.[13] A nice feature of this data is that the system contains information on the severity of injuries of the people involved in the accidents. At the same time, it also contains information on vehicle models, and vehicle-year models, that are involved in the accidents. The latter information is the key that allows us to connect accidents data to the independently collected recall information, since recalls are specific to vehicle models.[14]

We want to see if recalls affect a person's injury level, conditional on an accident occurring. Therefore, the person i's injury level is y_i^*, which is the dependent variable in our model. In fact, we do not have information on the exact injury level for the person. Fortunately, the GES has a categorical variable, y_i, indicating the severity of injury of a person in an accident. The variable is based on an unobserved continuous variable, $y_i^* \in \mathbf{R}$. The variable appears as an ordered rating scale, and we use this information to measure accidental harm. Since the dependent variable is discrete and ordered, the ordered probit model is used for our analysis.

A number of reasons can explain the differences in the severity of injuries of individuals involved in accidents. We categorize the variables into four major determinants, and write

$$y_i^* = f(D_i, V_i, E_i, R_i \mid A_i = 1) \tag{1}$$

where, D_i = represents Driver characteristics, V_i = refers to Vehicle characteristics, E_i = indicates Environmental factors, and R_i = stands for Regulation factors including vehicle recalls. With this function being conditional on the accident that has happened, $A_i = 1$. The dependent variable takes five values (0, 1, 2, 3, 4) corresponding to no injury, possible injury, non-incapacitating injury, incapacitating injury, and fatal injury (of drivers, passengers or pedestrians). The ordered probit model is constructed as follows:[15]

$$y* = \beta' X + \epsilon, \qquad \varepsilon | X \sim N(0, 1) \tag{2}$$

[13]Motor Vehicle Defects and Safety Recalls, NHTSA. Notice that the fact that we use data on reported accidents is another reason to restrict our analysis to newer models, given that the reporting behavior of owners of older vehicles might be quite different, and they might be less likely to file accident reports and notify the police about an event, especially if the damage is not too extensive. Another related issue has to do with whether recalled vehicles might be more likely to appear in the police reports than other vehicles, maybe because owners, assuming they know about the recall, are more likely to call the police after an accident. If this was the case we would be in the presence of a selection problem that could bias our coefficient of interest. While this is plausible, although unlikely to be a major issue, we cannot really control for this issue without access to the universe of data on all cars on the road.

[14]There is no specific data on the accidents caused by recalls. There are owners' accident reports submitted to manufacturers and (or) the NHTSA, but we have no access to them.

[15]Following the discussion in Wooldridge(2002), Greene (2000), and Cameron and Trivedi (1998).

$$y_i = \beta' X_i + \epsilon_i \tag{3}$$

Let $\kappa_1 < \kappa_2 \cdots < \kappa_J$ be unknown cut points and define

$$\begin{array}{lll} y = 0 & \text{if} & y* \leq \kappa_1 \\ y = 1 & \text{if} & \kappa_1 < y* \leq \kappa_2 \\ y = 2 & \text{if} & \kappa_2 < y* \leq \kappa_3 \\ y = 3 & \text{if} & \kappa_3 < y* \leq \kappa_4 \\ y = 4 & \text{if} & y* > \kappa_4 \end{array} \tag{4}$$

ϵ is normally distributed across observations. The conditional distribution of y given X is

$$\begin{array}{l} P(y = 0|X) = P(y^* \leq \kappa_1|X) = \Phi(\kappa_1 - \beta' X) \\ P(y = 1|X) = P(\kappa_1 < y^* \leq \kappa_2|X) = \Phi(\kappa_2 - \beta' X) - \Phi(\kappa_1 - \beta' X) \\ P(y = 2|X) = P(\kappa_2 < y^* \leq \kappa_3|X) = \Phi(\kappa_3 - \beta' X) - \Phi(\kappa_2 - \beta' X) \\ P(y = 3|X) = P(\kappa_3 < y^* \leq \kappa_4|X) = \Phi(\kappa_4 - \beta' X) - \Phi(\kappa_3 - \beta' X) \\ P(y = 4|X) = P(y^* > \kappa_4|X) = 1 - \Phi(\kappa_4 - \beta' X) \\ \kappa_1 < \kappa_2 < \kappa_3 < \kappa_4 \end{array} \tag{5}$$

The parameters κ and β are estimated by maximum likelihood (Gould and Sribney (1999)). For each observation i, the log-likelihood function is

$$\begin{aligned} L_i(\kappa, \beta) = 1[y_i = 0] log[\Phi(\kappa_1 - \beta' X)] + \\ 1[y_i = 1] log[\Phi(\kappa_2 - \beta' X) - \Phi(\kappa_1 - \beta' X)] + \\ \cdots\cdots \\ 1[y_i = 4] log[1 - \Phi(\kappa_5 - \beta' X)] \end{aligned} \tag{6}$$

4.2. Recall Variable and Data Structure

We use three years of accident data from the GES, with all the detailed information on the accidents occurred in 2005, 2006, and 2007 on the U.S. roads. Each observation is a person who is involved in an accident, either a driver, a passenger, or a non-occupant, such as a pedestrian. Each observation shows the injury severity of that person, and contains information on the vehicle related with the person. If the person is a driver, then the vehicle is her car. If she is a pedestrian, then the information is about the vehicle that injured her. However, the accident data does not contain any recall information, and there is no direct link between accidents and recalls. Therefore, we need to create a recall variable using data from the NHTSA.[16] We obtain information on each vehicle model's recall history, including the date, month, and year of each recall announcement for the particular vehicle year-model.[17] Since

[16] All the detailed information on recalls are available in the NHTSA website. See "http://www-odi.nhtsa.dot.gov".

[17] It is important to emphasize that the recall information does not take the risk levels into account. In particular,

the cross-sectional accident data include the date that the accident occurred and each accident has information on the vehicle maker and the vehicle model involved, we can find when a recall for the vehicle model involved in the accident was issued.

We define the main recall variable of interest as "**the months that have passed after a recall issuance**", which appears with the name, **RECALL**, in the tables. For instance, if a recall for a particular year-model is issued on January and a vehicle of this year-model is involved in an accident next month, then the value of the variable is one. If the recall is issued on January and the accident occurs on March, then it is two. We measure the months elapsed during the first year after the issuance of the recall, and the maximum value of the recall variable is therefore 12. Having this definition, we see the relationship between the months elapsed and the severity of injury of the accident.

Vehicle recalls have a number of characteristics that make any empirical analysis very complex and require a number of important simplifications. For instance, each vehicle model contains multiple recalls, and some recalls are issued before and some after accidents occur. Furthermore, we do not know if a particular vehicle involved in the accident is the vehicle included in the recall. Thus, our estimation strategy requires us to eliminate the accidents that we cannot link to a particular recall. We enumerate here the main problems and the criteria we use to include or not the accident in our sample:

- Many vehicles may contain multiple defects, and manufacturers often issue more than one recall for a particular vehicle model. If two recalls were issued in a relatively short-time period, it is hard to know which recall affected the accident, if it did, but since this happens quite often, if there are two or more recalls issued for a particular vehicle model within 3 months, we consider them as a single recall. Otherwise, we remove accidents with vehicle models that have multiple recalls within 4 and 18 months after the first recall, unless one of the recalls is minor (includes less than 10,000 units of that year-model) and with some other exceptions that we will discuss below.
- We also require that the owners of the corresponding vehicle models should be able to respond to the issuance, but it is much harder for used car owners to know about recalls given that only if they are in the same State and do not change addresses will they be able to be linked. Furthermore, even for new car owners, if they change their addresses, then the recall notice might not reach them. Thus, we include only vehicle models launched in the 2005 to 2007 period we analyze in the data set, to all but guarantee that owners will receive the information.
- Given that we include only new-year vehicle models, the first recall should be issued within the first year, and the accident can occur up to a year after the recall was issued.
- The hazard levels of different recalls can be quite different, but for recalls after 2001 we do not have this information. Thus, we include accidents with vehicle models whose correction rates

the effects of hazard recalls on severity could show a more solid relationship between recalls and severity. The information was available to the public in the past, but it is no longer available.

are at least 60%. The higher correction rate indicates that the defect is considered a riskier one. Notice that this correction rate is for vehicles already sold to individuals, it does not include vehicles not delivered to individuals and therefore maybe already fixed by the manufacturer or the dealers. This justifies the use of the correction rate as a selection criteria. We discuss the sensitivity of our results to this criteria. It should not be surprising that if we take all recalls independent of correction rates the effects of recalls on severity will be smaller since many drivers will not know about the recalls yet and will not have taken them to be fixed. This selection criteria is in fact connected with the fact that we use new vehicle models since these vehicles are the ones more likely to get safety recalls corrected (See Rupp and Taylor, 2002, and Hoffer, Pruitt, and Reilly, 1994)

- Recalls often contain very different number of units. Some recalls contain a few hundreds, while others contain millions. In our analysis we include only recalls of more than 10,000 units.

Even after eliminating vehicle models following the criteria discussed above, approximately 700 observations are included in our sample.

4.3. Summary Statistics

Table 1 shows the severity levels of individuals involved in accidents over the three years we analyze. The first column shows the severity of accidents for new year-models (2005, 2006, and 2007) that have their first recalls over these three years. Each accident occurred after its first recall was issued. More than 65 % of individuals are not injured in accidents. Possible injury accounts for 13.55% of the accidents. Non-incapacitating and incapacitating injury represents 13.12 % of the accidents. Less than one percent of individuals experiences fatal accidents. The second set of columns shows the severity levels of the accidents, after excluding major recalls. Thus, the first set represents what we call recall accidents, while the second set shows non-recall accidents.[18] There are no significant differences in the distribution of severities. Thus, we do not have a selection problem from the sample restrictions explained in the previous subsection. However, given the small number of fatalities, the results regarding this event should be taken with caution.

Other independent variables are included along with the recall variable in order to control for other sources of variation that can explain the severity of accidents. Drivers' behavior (or characteristics) is one of the most important factors that might affect the severity of injuries. To control for this, D_i includes sex, age, whether any safety equipment was used, and whether alcohol was involve at the time of the crash. V_i includes the variables that are not related to the vehicle defects, such as vehicle size, the vehicle role on the crash, and whether there were any contributing factors related to the vehicle.

[18]The second set includes accidents by vehicles whose recalls were issued, but the units involved for those particular vehicle models are less than 10,000. The first set has vehicles whose recalls were major ones and each recall was issued in the first year. So, in the second set, both other major recall-related accidents, unmatched to the selection criteria, and 701 accidents are excluded.

E_i includes environmental factors surrounding the accident. R_i includes recalls, and speed limits. The definitions of the variables and their descriptions can be found in Table 2.

It is interesting to discuss some of these variables of interest. $RESTRAINT_i$ encodes what is documented on the PAR (Police Accident Report) regarding occupant use of available vehicle restraints (i.e., lap and (or) shoulder belt, child safety seat, and motorcycle helmet). About 12 % of persons do not use any devices, while 88 % of them wear a lap or shoulder seat-belt.[19] With this information we construct a dummy variable, which takes the value of one if the person uses any one of the protective restraints. Therefore we expect a negative relationship between the dependent variable and this variable. $ALCOHOL_i$ is a binary indicator that indicates whether alcohol was used by the person involved in the crash. $MOTORIST_i$ indicates role of this person in the accident. If the person is a motorist, then it is one. If she is a non-motorist, then it is zero. Assuming that pedestrians are injured more severely conditional on being involved in accidents, a negative relationship is expected. To account for vehicle factors, the binary indicator $AUTOMOBILE_i$ is included because the degree of severity might be different, not because of vehicle defects, but because of vehicle non-defective factors, so it is important to distinguish between automobiles, motorcycles, light trucks, etc. $FACTOR_i$ indicates whether there were any vehicle factors that might have been the cause of the crash. The possible factors include tires, brake system, steering system, suspension, power train, exhaust system, and so on. The PAR data show whether one of these factors could have caused each accident. However, we do not know if they are coming from defects or just simply maintenance problems.

$SPEED\_LIMIT_i$ shows actual posted speed limit in miles per hour. Half of the accident reports do not contain information on vehicles' travel speed at the time of crash, therefore instead of missing many observations we include the speed limit, the maximum being 75 miles per hour and the minimum is 5. Given the nature of accidents that occurred in the parking lots or alleys we remove them from our sample. Regarding environmental factors, $ROADWAY_i$ identifies the location of the first harmful event. This is a dummy variable, and if the vehicle is on a roadway it takes the value one. If the vehicle is off a roadway (or shoulder) or parking lane, then the value is zero. $STRAIGHT\_ROAD_i$ identifies the horizontal alignment of the roadway in the immediate vicinity of the first harmful event. If the roadway is straight, then the variable is 1, otherwise it is zero. $GOOD\_WEATHER_i$ measures the general atmospheric conditions. If there are no adverse conditions, then the value is 1. If it is rainy, sleety, snowy, foggy, or a similar poor condition, then the variable has the value of zero.

An important variable is $VINTAGE_i$, which measures the months between the October of the previous year in which the model was initially launched in the market and the time of the accident. For instance, if an accident occurred on January, the involved vehicle must have been purchased within 3 months.[20] If the accident occurred on December next year, then it must have been purchased during the last 15 months. The variable can take a value as high as 27 if a vehicle was recalled in December of the year in which it was launched, and the accident happened in the following December.

[19] This is consistent with the NHTSA' data on seat belt usage. As of 2008, 83% of occupants wear their seat belts, regardless of new and used cars. Thus, the effect of new cars on wearing seat belts could be sizeable.

[20] Manufacturers begin to sell their their new-year models from September in previous year. We do not know when particular owners purchased their cars and had accidents.

We present the summary statistics in Tables 3 and 4. Table 3 presents the sample of accidents linked to recalled vehicles, and Table 4 for those linked to vehicle models not subject to major recalls. In Table 3 the total number of observations is 701. The dependent variable, $SEVERITY_i$, has a mean of 0.6505, this means that the large majority of accidents has neither severe injuries nor fatal injuries. The mean statistics for the explanatory variables reveal that the average age in our sample of the individuals involved in accidents is 37.55. The number of males is larger than the number of females and most people have used some kind of restraint system. We also see in the table that only about 6.5% of accidents involve alcohol, so in a large majority of accidents intoxication was not a factor. Although not reported in the table, 69.8% of people were the drivers, while 27.92% were the passengers, and almost 97% are either drivers or passengers. Regarding the vehicle type, 51% of the accidents involved automobiles. Vehicle contributing factors are reported by policemen when they thought there were the possible causes of the accidents. However, we do not know whether the factors really come from the defective parts that brought about recall issuance. Only 3.1% was considered to have had possible contributing factors.

The key recall variable, $RECALL_i$, has an average of 7.06 months. Therefore, the accidents in the sample occurred on average 7 months after the first recalls were issued, and given the value of the $VINTAGE_i$ variable the accidents occurred on average 15 months after they were launched. Regarding speed limit, the average speed limit was 45 miles per hour. This is because the accidents in the sample occurred on the local streets as well as highways. Many environmental factors are considered. Most accidents occurred on local roads, on straight roadways, and when there were no weather related problems. The average number of occupants in cars involved in the accidents in the sample was 1.48, which indicates that most drivers were alone in their vehicles at the time of crash.

In Table 4 we show the summary of the observations linked to vehicles not subject to major recalls. The number of observations is much higher at over 12,300. The main differences that we observe between the two samples is that the non-recalled sample is composed of a lower percentage of automobiles and more SUVs, light trucks, and other trucks, and there is a slightly longer time between the launching of the model and the accidents.

5. Estimation Results

The first set of estimation results are shown in Table 5. We present four models, *MODEL 1* controls for the $RECALL_i$ measure but omits the variable $VINTAGE_i$, while *MODEL 2* includes it because we are interested in assessing the consequences of controlling for the number of months the vehicle has been on the road, which could potentially explain the changes in the severity of the accidents. The two variables are separately identified as we explained above because of the variation in the sample across vehicle models from the time of the launching of the car to the time of the accident, and the independent variation from the time of the recall to the time of the accident. We also present a third model *MODEL 3* which only controls for the vintage measure and leaves out our main recall control, to explore the possibility that the vintage effect could overwhelm any effect of our main variable of interest. We also show another set of results, labeled *MODEL 4*, using the very different sample of accidents not linked to vehicle models subject to major recalls. For this set of accidents we can only control for the vintage effect. As we hypothesized earlier, if we find a significant and negative vintage effect with this sample any similar results in the recall sample would suggest that the recall variable was picking up a vintage effect and therefore we would not be able to convincingly argue that the effect on severity was linked to recalls.

We compare the fit of the models using a pseudo-R^2 measure and the Bayesian Information Criteria (BIC), a scalar measure of fit, which is widely used to compare non-nested models (J.S. Long & J. Freese, 2001). The pseudo-R^2 weakly supports *MODEL 2*, while the BIC is strongly in favor of the model without the vintage variable. Given their significance, and interpretation we prefer to show all results.[21]

Most of the signs across specifications are what we expected.[22] In *Model 1*, the severity of injuries decreases over time after recall issuance, and the coefficient is significantly different from zero at the 5% significance level. Therefore, the more months pass after a recall issuance, the lower the severity of injuries is.[23] With the second model, we reach the same conclusion and the recall variable is still significant although a bit smaller in magnitude, even though the variable, $VINTAGE$, is also negatively associated with the dependent variable but the effect is small and not statistically significant. Thus, after controlling for the vintage effect we still see the decrease in the severity of injuries over the first year. This reinforces our finding that recalls reduce the severity of injuries in the time period we analyze.

However, we could still wonder whether the vintage and the recall effects are separately identified, even if we have argued that independent sources of variation suggest theoretical identification of the model. That is why we present the results from models three and four. In *MODEL 3* we only include the vintage effect, and while it becomes negative and significant at the 10% level, the coefficient decreases in magnitude considerably, and so does the fit of the model, suggesting that the recall variable certainly belongs in the model and captures variation that the vintage measure misses. In fact the vintage variable could be spuriously significant in this specification because of its correlation with the recall variable. To check this conjecture we estimate *MODEL 4* using the sample of accidents not linked to major recalls. Here the vintage effect all but disappears, and while still negative the coefficient is very small and highly insignificant, showing that the recall effect on severity is not a vintage effect but a true relationship that

[21]BIC is defined as $BIC_k = D(M_k) - df_k lnN$, where k is a model, and $D(M_k)$ is the deviance of the model M_k. df_k is the degrees of freedom associated with the deviance. The difference in the BICs between models indicates which model is more likely to have generated the observed data. The BIC values are -3110.445 without $VINTAGE$ and -3103.993 with it, respectively. The difference (6.452) between them indicates **strong** support for the model without the variable. Thus, our main results are linked to the specification shown in *MODEL 1*.

[22]These β parameters, by themselves are of limited interest, and here we can only discuss their signs and statistical significance. Also we need to note that the explanatory variables only affect the predicted probabilities. Therefore, the interpretation of the tables should be understood as the effects of the explanatory variables on the likelihood of severity.

[23]To check for a possible non-linear relationship, we included the variable, $RECALL\_SQ$, along with $RECALL$ in the equation. The coefficients on the quadratic terms were not significant.

links recalls with a decreased severity of the accidents of the vehicles.[24]

Regarding the other explanatory variables, the older the person is, the higher the severity is, while gender does not affect the severity of injuries. If a person uses any protective restraints, then she is injured less. Non-drunken persons are injured less. Regarding the variable, $MOTORIST$, motorists are less injured, while non-motorists, such as pedestrians, are more severely injured. Vehicle characteristics do not seem to affect severity much. While this might not be true in general as shown by the significant effects in *MODEL 4*, but when only new-year models are involved in accidents, this is plausible. Most environmental factors, except $GOOD\_WEATHER$, appear not to be significant, this is probably because the data is constructed with only new-year models. When there are good weather conditions, people are injured less.

Given the difficulty of interpreting the coefficients discussed above, Table 6 reports the marginal effects for all explanatory variables in *MODEL 1.* The marginal effects for the categorical variables, such as the recall variable, explain the slopes of the probability curves at the point intersection with the vertical line of the average months (7.05) after recall issuance in the sample. This implies that for the average person as one more month passes after a recall issuance, the predicted probability of no injury increases by 1.31%. It also decreases the predicted probability of an incapacitating injury by 0.37%, and the probability of fatal injury by 0.03%. The marginal effects for the dummy variables are explained by the change in the predicted probability for a change in x_k from 0 to 1. For instance, wearing protective devices, such as seat belts, increases the probability of no injury by 33% and decreases the predicted probability of an incapacitating injury by 13.8%, and of fatal injury by 2.06%. Thus, wearing seat belts has substantial impacts on safety, which is clearly consistent with most of the literature on seat belts laws.

It is interesting to analyze the individual predicted probabilities with the different months after a recall, we show this in Tables 7-a and 7-b corresponding to model specifications one and four. In Table 7-a, if an accident occurred this month and the first recall was issued in the previous month, then the predicted probability that the person does not have any injury is 0.5741. For the person who is involved in an accident this month and her vehicle had a recall twelve months ago, the probability is 0.7187. The probability that the accident results in a non-incapacitating injury goes down from 17% to 10.77%, and the probability of resulting in an incapacitating injury goes down from 7.97% to 3.73% during the year after the recall. Finally, the probability of resulting in fatal injury also goes down (0.0054 vs 0.0016). The last two results, however, should be taken with caution, given the small number of observations with these major injuries that we have in the data set of accidents linked to major recalls. Table 7-b presents the same exercise but using *MODEL 4* and the effect of the vintage variable, and not surprisingly all the probabilities are essentially unchanged as the cars involved in accidents grow older. We show these results graphically in Figures 6 to 8, which bring home the point that the recall effect is very strong,

[24] The sample of accidents of non-recalled vehicles has a smaller proportion of automobiles, and more sports utility vehicles and light/heavy trucks, we could wonder then whether the lack of significance of the vintage effect in *MODEL 4* is due to the different composition of vehicles in the larger sample. To study this possibility, we re-estimated both models using only the sub-sample of automobiles, the results were essentially unchanged, with the recall effect becoming somewhat stronger and the vintage effect even smaller.

and the vintage effect essentially non-existent.

6. Conclusions and Policy Implications

This paper analyzes the effects of recalls on vehicle safety in the U.S., and in particular, we investigate whether recalls affect accidental harm measured by the degree of severity of the accidents. Our results show that recalls reduce the severity of injuries of individuals involved in motor vehicle accidents. We find that if a recall for a new-year model is issued, then the severity of injuries of accidents continuously diminishes during the first year after the recall. This can be because defects are removed over time but also (and) because drivers may increase their level of care until the defects are fixed. Our quantitative findings show that after recall issuance as time goes by, the predicted probability that a person does not have any injury increases, and the probability that a person has significant or fatal injuries becomes smaller, as the number of months passed increases. Due to data limitations, we have only included new vehicle models with one valid recall in the year and a half after the original recall, but given this evidence, we can say that recalls are effective for new-year models. All this shows that defective cars linked to serious accidents are put on the road every year by manufacturers and that an overhauling of the recall system and the quality control of new automobiles should be considered by policy makers.[25].

We have found among readers of our work some researchers who are rather skeptical of our results. They seem to believe that recalls are unlikely to be an important explanation behind accidents and their severity. Of course, we admit that drivers' mistakes are one of the main causes of many accidents. While it might be true that a small proportion of accidents and their increased severity might be linked to recalls, our findings are clear, they are effective, even if we are not able to identify the direct effect from the likely behavioral changes that recall notices have on the driving attitudes of individuals. More importantly, if recalls were not effective and if they were not to blame for accidents, why would companies spend millions (even billions) of dollars on them. They do, among other reasons, because they are afraid of being sued, because they know they are putting faulty cars on the road, and not even they know whether those mechanical problems could be linked to accidents, but know that car owners could easily make that argument in court. [26] However, the system is broken, when in a given year more cars are recalled than sold in the United States it means we have to turn our attention to the production processes, and the effectiveness of the current recall system.

Given the need to know about a recall in order to act on it, every effort should be made to make owners aware of the possible defects of their vehicles. For example, in the year 2001, the IIHS (Insurance Institute for Highway Safety) submitted a petition to the NHTSA to ask for more information release.

[25]In the last weeks a bill is being considered in Congress that would allow the NHTSA to speed up the process of mandatory recalls which can now take months, and they are even considering requiring car manufacturers to install devices that would record what happened during an accident, similar to what is in place in airplanes. See "Highway safety agency wants more auto recall power," AP Press, May 6, 2010

[26]Among others, companies may have enough incentives to put their faulty cars in the market in haste, because of severe competition among them. For more discussion, see Bae (2010).

The petition was for inclusion of VINs (Vehicle Identification Numbers) in defect and noncompliance information reports. The institute insisted that if NHTSA were provided with the VINs of recalled vehicles at the beginning of recall issuance, drivers would know immediately whether they were driving potentially unsafe vehicles, and the information could even be taken into account by insurance companies when pricing insurance. This is especially true when customers purchase used cars, given that people do not know whether the used cars that they want to purchase were subject to recalls. Table 8 shows two recalls issued in 2005 and their correction rates quarterly reported by two major car makers. As this example illustrates, correction rates are much higher when recalls contain highly risky defects and they are issued within a year. Thus, both appropriate risk evaluations on risks and responsive actions are very important.

For a long time manufacturers have resisted any change that could lead to more awareness regarding recalls or affect correction rates (like asking the DMVs to include recall notices with the registration renewals of vehicles), arguing that they would incur in a substantial administrative burden, but with unclear effects on safety. However, producing safer and higher quality cars should be much more important even to them. Over the last months of 2009 and the beginning of 2010, Toyota has recalled more than 6 million vehicles in the U.S. (9 million vehicles worldwide) because of acceleration problems and braking flaws. Regulators have linked 52 deaths to crashes allegedly caused by accelerator problems. According to a report,[27] Toyota recalls could cost up to \$ 2 billion, with \$ 1.1 billion for repairs and \$ 0.9 billion in lost sales. As of March, 2010, at least 89 class-action lawsuits have been filed against Toyota, which could also cost it at least \$ 3 billion. If Toyota had taken recall actions in a more timely fashion and the important recall information had been immediately released to the public, then the costs would be much smaller. This is especially true when recalls are effective in reducing accidental harm.

Along with the work by Bae and Benítez-Silva (2010), that finds that recalls reduce the number of accidents of particular models driven by particular types of drivers, we provide a first step towards understanding the effects of recalls on safety, the importance of which has increased given the prominence, size, and costs of vehicle recalls in the last years. Further research is needed to draw a more complete picture of the effects of recalls on safety, through analyzing the benefit-cost analysis of recalls, the effects of recalls on vehicle demand, measurement of accidental loss and the value of life, and the effects of recalls on manufacturers' reputation.

[27] "Toyota counts rising cost of recall woes," AP Press, March 16, 2010.

Figure 2: Number of Recalls Issued

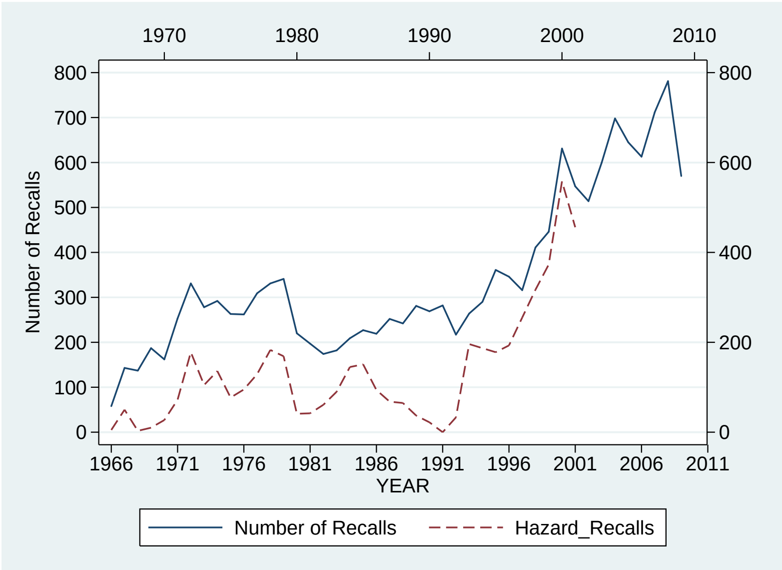


Figure 3: Average Units Per Recall

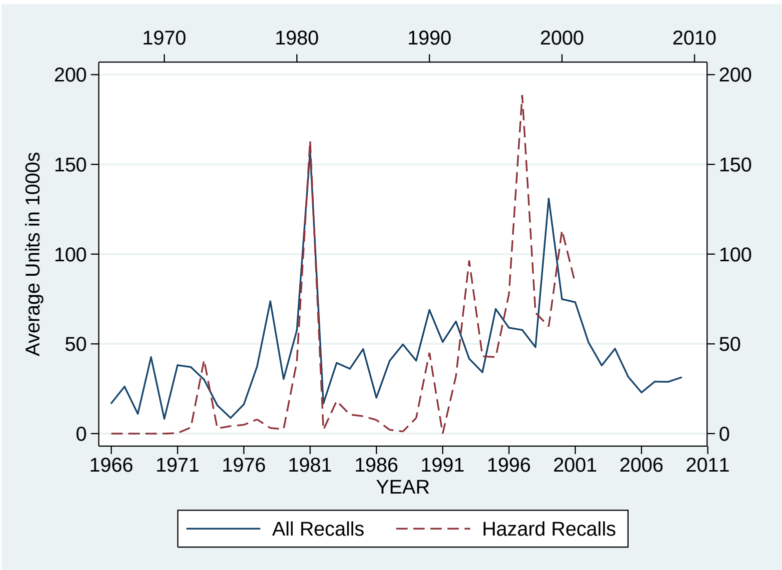


Figure 4: Ratio of Domestic Recalls to Total Recalls

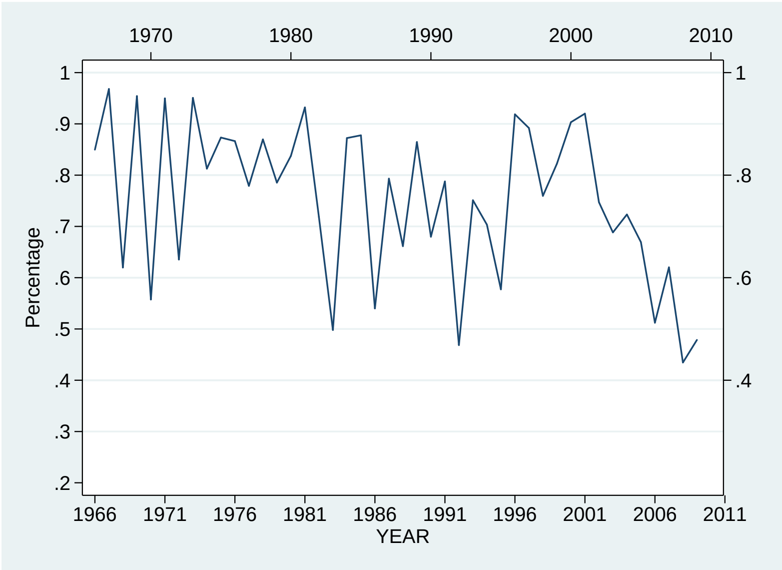


Figure 5: Ratio of Mandatory Recalls to Total Recalls

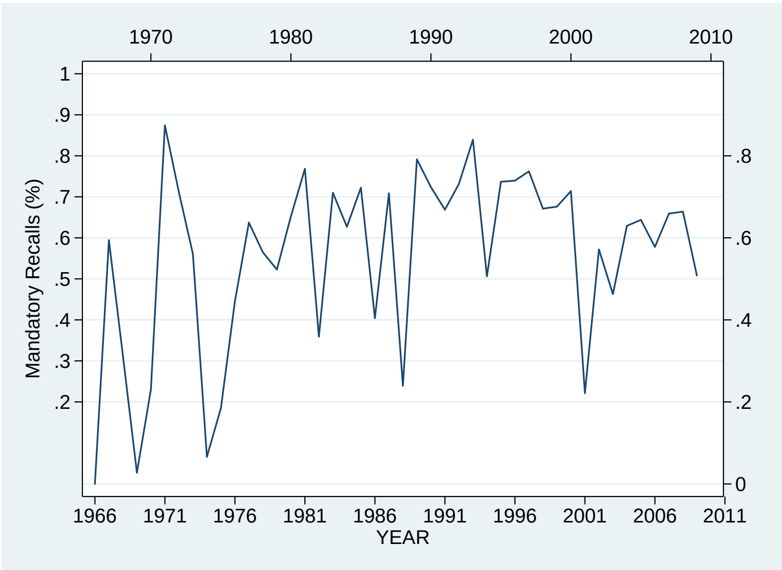


Table 1: Severity Level of Accidents Occurred in the Years 2005, 2006, and 2007

Severity level	Freq	Percent	Cum	Freq	Percent	Cum
No injury (=0)	457	65.19	65.19	8,022	65.12	65.12
Possible injury (=1)	95	13.55	78.74	1,596	12.96	78.08
Non-incapacitating injury (=2)	92	13.12	91.86	1,502	12.19	90.27
Incapacitating injury (=3)	51	7.28	99.14	1,124	9.12	99.40
Fatal injury (=4)	6	0.86	100.00	75	0.61	100.00
Total	701§	100.00	100.00	12,319†	100.00	100.00

Note : The data set comes from the GES.

Only new-year vehicle models are included.

Since the GES data are from a probability sample of police-reported traffic crashes, nationally representative estimates can be obtained from these data. Refer to "NASS GES Analytical Users Manual, 1988 - 2009" regarding the methodology to do that.

§ The total number of observations that include accidents linked to recalled vehicles.

† The total number of observations that include accidents not linked to recalled vehicles.

Table 2: Description of Variables

Variable	Description	Dummy
Dependent variable:		
SEVERITY	Injury categories	N
Personal characteristics		
AGE	Age of the person (years)	N
MALE	Gender: 1 if male, 0 if female	Y
RESTRAINT	Use of any protective device: 1 if use, 0 if no use	Y
ALCOHOL	Police-reported alcohol involvement:	Y
	1 if drunk, 0 if no drunk	
MOTORIST	The role of the person in the accident:	Y
	1 if motorist, 0 if non-motorist	
Vehicle characteristics		
AUTOMOBILE	Body type of the vehicle: 1 if automobile, 0 if other	Y
VEHICLE_FACTOR	Vehicle factors: 1 if any factor, 0 if no contributing factor	Y
VEHICLE_ROLE	Vehicle role: 1 if striking vehicle, 0 otherwise	Y
Regulation factors		
RECALL	Months that have passed since a recall issuance	N
RECALL_SQ	Square of RECALL	N
VINTAGE	Months that have passed since October in the previous year	N
VINTAGE_SQ	Square of MONTH_VINTAGE	N
VINTAGE_2	Months that have passed since October in the previous year	N
	(For non-recall accidents)	N
SPEED_LIMIT	Actual posted speed limit (in miles per hour)	N
Environmental factors		
HIGHWAY	Inter-state highway 1 if highway, 0 if local	Y
ROADWAY	The location of the first harmful event:	Y
	1 if the vehicle is on the roadway, 0 if it is off the roadway	
STRAIGHT_ROAD	1 if straight road, 0 if curved road	Y
GOOD_WEATHER	The general weather conditions:	
	1 if it is good, 0 if there was any adverse condition	Y
OCCUPANT_INVOLVED	Number of occupants involved	N
ROAD_LANES	Number of road lanes	N
WEEKDAY	Weekday: 1 if weekdays, 0 if weekends	Y

Table 3: Summary Statistics for Recall Accidents

Variable	Obs	Mean	SD	Min	Max	Acronym
Dependent variable:						
Severity	701	.6505	1.0152	0	4	SEVERITY
Personal characteristics						
Age	701	37.5578	16.1820	11	91	AGE
Male	701	.6220	.4852	0	1	MALE
Restraint system use	701	.8873	.3164	0	1	RESTRAINT
Alcohol involvement	701	.0656	.2478	0	1	ALCOHOL
Motorist	701	.9672	.1783	0	1	MOTORIST
Vehicle characteristics						
Body type	701	.5136	.5002	0	1	AUTOMOBILE
Vehicle contributing factor	701	.0314	.1745	0	1	VEHICLE_FACTOR
Vehicle role	701	.5635	.4963	0	1	VEHICLE_ROLE
Regulation factors						
Recall in months	701	7.0599	3.3414	1	12	RECALL
Recall squared	701	60.9914	46.4534	1	144	RECALL_SQ
Vintage in months	701	15.2810	4.3507	3	27	VINTAGE
Vintage squared	701	252.4108	139.1169	9	729	VINTAGE_SQ
Speed limit	701	45.0927	13.6915	5	75	SPEED_LIMIT
Environmental factors						
Highway	701	.1697	.3757	0	1	HIGHWAY
Relation to roadway	701	.9415	.2348	0	1	ROADWAY
Straight road	701	.8959	.3057	0	1	STRAIGHT_ROAD
Weather Conditions	701	.8274	.3782	0	1	GOOD_WEATHER
Occupants involved	701	1.4836	.9358	0	9	OCCUPANT_INVOLVED
Number of travel lanes	701	2.9601	1.2495	1	7	ROAD_LANES
Weekday	701	.7575	.4289	0	1	WEEKDAY

Table 4: Summary Statistics for Non-Recall Accidents

Variable	Obs	Mean	SD	Min	Max	Acronym
Dependent variable:						
Severity	12319	.6715	1.0417	0	4	SEVERITY
Personal characteristics						
Age	12319	37.9306	16.4014	0	102	AGE
Male	12319	.6121	.4873	0	1	MALE
Restraint system use	12319	.8971	.3039	0	1	RESTRAINT
Alcohol involvement	12319	.0472	.2122	0	1	ALCOHOL
Motorist	12319	.9742	.1586	0	1	MOTORIST
Vehicle characteristics						
Body type	12319	.3889	.4875	0	1	AUTOMOBILE
Vehicle contributing factor	12319	.0264	.1603	0	1	VEHICLE_FACTOR
Vehicle role	12319	.6014	.4896	0	1	VEHICLE_ROLE
Regulation factors						
Vintage in Months	12319	17.2785	7.0316	4	27	VINTAGE
Speed limit	12319	45.0889	13.6249	0	75	SPEED_LIMIT
Environmental factors						
Highway	12319	.1670	.3730	0	1	HIGHWAY
Relation to roadway	12319	.9574	.2020	0	1	ROADWAY
Straight road	12319	.8831	.3213	0	1	STRAIGHT_ROAD
Weather Conditions	12319	.8146	.3886	0	1	GOOD_WEATHER
Occupants involved	12319	1.5229	1.1374	0	39	OCCUPANT_INVOLVED
Number of travel lanes	12319	3.0370	1.3485	1	7	ROAD_LANES
Weekday	12319	.7438	.4365	0	1	WEEKDAY

Table 5: Impact of Vehicle Recalls on the Severity of Accidents

	MODEL 1	MODEL 2	MODEL 3	MODEL 4
With	RECALL	RECALL		
With		VINTAGE	VINTAGE	VINTAGE
Personal characteristics				
AGE	.0062 (.0029)**	.0062 (.0029)**	.0062 (.0029)**	.0018 (.0007)**
MALE	.0077 (.0958)	.0075 (.0958)	-.0030 (.0956)	-.2454 (.0225)**
RESTRAINT	-.8589 (.1882)***	-.8564 (.1883)***	-.8565 (.1879)***	-.7254 (.0409)***
ALCOHOL	.8190 (.1803)***	.8201 (.1801)***	.8300 (.1782)***	.7657 (.0505)***
MOTORIST	-.6435 (.2112)***	-.6502 (.2123)***	-.6759 (.2106)***	-.7079 (.0512)***
Vehicle Characteristics				
AUTOMOBILE	.0334 (.0937)	.0295 (.0939)	.0074 (.0934)	-.0535 (.0231)**
VEHICLE_FACTOR	-.1598 (.2844)	-.1608 (.2829)	-.1597 (.2847)	-.0639 (.0669)
VEHICLE_ROLE	-.0374 (.0936)	-.0379 (.0936)	-.0423 (.0935)	-.0882 (.0229)***
Regulation Factors				
RECALL	-.0357 (.0148)**	-.0321 (.0186)*	-	-
VINTAGE	-	-.0044 (.0139)	-.0193 (.0112)*	-.0007 (.0016)
SPEED_LIMIT	.0002 (.0039)	.0002 (.0039)	-.0003 (.0039)	-.0002 (.0010)
Environmental factors				
HIGHWAY	-.2248 (.1400)	-.2247 (.1400)	-.2185 (.1403)	.0913 (.0346)*
ROADWAY	-.0268 (.1933)	-.0294 (.1935)	-.0730 (.1934)	.0136 (.0545)
STRAIGHT_ROAD	.1101 (.1552)	.1083 (.1554)	.1056 (.1568)	.0405 (.0364)
GOOD_WEATHER	-.2377 (.1228)*	-.2284 (.1226)*	-.2321 (.1216)*	.0049 (.0288)
OCCUPANT_INVOLVED	.0206 (.0473)	.0215 (.0475)	.0260 (.0477)	-.0273 (.0108)**
LANES	.0002 (.0387)	-.0001 (.0388)	-.0008 (.0388)	-.0132 (.0085)
WEEKDAY	-.0664 (.1047)	-.0660 (.1049)	-.0801 (.0112)	-.0444 (.0256)
NUM of OBS	701	701	701	12319
Pseudo R2	0.0841	0.0842	0.0821	0.0484

Note : Standard errors are in parentheses. The Huber/White/sandwich estimator of variance is used.
*: Significant at the 10-percent level.
**: Significant at the 5-percent level.
***: Significant at the 1-percent level.

Table 6: Marginal Effects from *MODEL 1*

Variable		Ave_Change	P[y=0]	P[y=1]	P[y=2]	P[y=3]	P[y=4]
AGE	Marginal	.0009	-.0023	.0006	.0010	.0006	.0001
MALE	0 → 1	.0011	-.0028	.0007	.0012	.0008	.0001
RESTRAINT	0 → 1	.1326	.3316	-.0422	-.1309	-.1380	-.0206
ALCOHOL	0 → 1	.1270	-.3175	.0377	.1246	.1347	.0205
MOTORIST	0 → 1	.1004	.2511	-.0359	-.1017	-.1000	-.0135
AUTOMOBILE	0 → 1	.0049	-.0123	.0032	.0053	.0035	.0003
VEHICLE_FACTOR	0 → 1	.0227	.0568	-.0162	-.0246	-.0149	-.0011
VEHICLE_ROLE	0 → 1	.0055	.0138	-.0036	-.0059	-.0039	-.0003
RECALL	Marginal	.0052	.0131	-.0034	-.0057	-.0037	-.0003
SPEED_LIMIT	Marginal	.0000	-.0001	.0000	.0000	.0000	.0000
HIGHWAY	0 → 1	.0320	.0799	-.0228	-.0345	-.0210	-.0016
ROADWAY	0 → 1	.0040	.0099	-.0025	-.0043	-.0029	-.0002
STRAIGHT_ROAD	0 → 1	.0159	-.0398	.0109	.0172	.0108	.0009
GOOD_WEATHER	0 → 1	.0359	.0898	-.0208	-.0385	-.0278	-.0026
OCCUPANT_INVOLVED	Marginal	.0030	-.0076	.0020	.0033	.0022	.0002
LANES	Marginal	.0000	-.0001	.0000	.0000	.0000	.0000
WEEKDAY	0 → 1	.0098	.0246	-.0063	-.0106	-.0071	-.0006

† 0 → 1 indicates that the variable is a dummy. All others are continuous variables.

‡ Standard errors are excluded from the table.

Table 7-a: Predicted Probabilities: *RECALL* variable, indicating months since recall

Months since recall	1	2	3	4	5	6
Pr(y=0|x)	0.5741	0.5881	0.6019	0.6156	0.6291	0.6425
Pr(y=1|x)	0.1709	0.1683	0.1655	0.1625	0.1595	0.1562
Pr(y=2|x)	0.1698	0.1639	0.1580	0.1522	0.1464	0.1406
Pr(y=3|x)	0.0797	0.0748	0.0702	0.0657	0.0615	0.0574
Pr(y=4|x)	0.0054	0.0049	0.0044	0.0040	0.0036	0.0032
Months since recall	**7**	**8**	**9**	**10**	**11**	**12**
Pr(y=0|x)	0.6557	0.6688	0.6816	0.6942	0.7066	0.7187
Pr(y=1|x)	0.1529	0.1494	0.1459	0.1422	0.1385	0.1346
Pr(y=2|x)	0.1349	0.1293	0.1237	0.1183	0.1129	0.1077
Pr(y=3|x)	0.0536	0.0500	0.0465	0.0433	0.0402	0.0373
Pr(y=4|x)	0.0029	0.0026	0.0023	0.0021	0.0018	0.0016

Table 7-b: Predicted Probabilities: *VINTAGE* variable, indicating months since launching of model

Months since launch	1	2	3	4	5	6
Pr(y=0|x)	.6503	.6506	.6508	.6511	.6513	.6516
Pr(y=1|x)	.1398	.1398	.1397	.1397	.1396	.1395
Pr(y=2|x)	.1263	.1262	.1261	.1260	.1259	.1258
Pr(y=3|x)	.0802	.0801	.0800	.0799	.0798	.0797
Pr(y=4|x)	.0033	.0033	.0033	.0033	.0033	.0033
Months since launch	7	8	9	10	11	12
Pr(y=0|x)	.6519	.6521	.6524	.6527	.6529	.6532
Pr(y=1|x)	.1395	.1394	.1394	.1393	.1393	.1392
Pr(y=2|x)	.1257	.1256	.1256	.1255	.1254	.1253
Pr(y=3|x)	.0796	.0795	.0794	.0793	.0792	.0791
Pr(y=4|x)	.0033	.0033	.0033	.0032	.0032	.0032

Figure 6: Predicted Probabilities of "No Injury": *RECALL* vs. *VINTAGE*

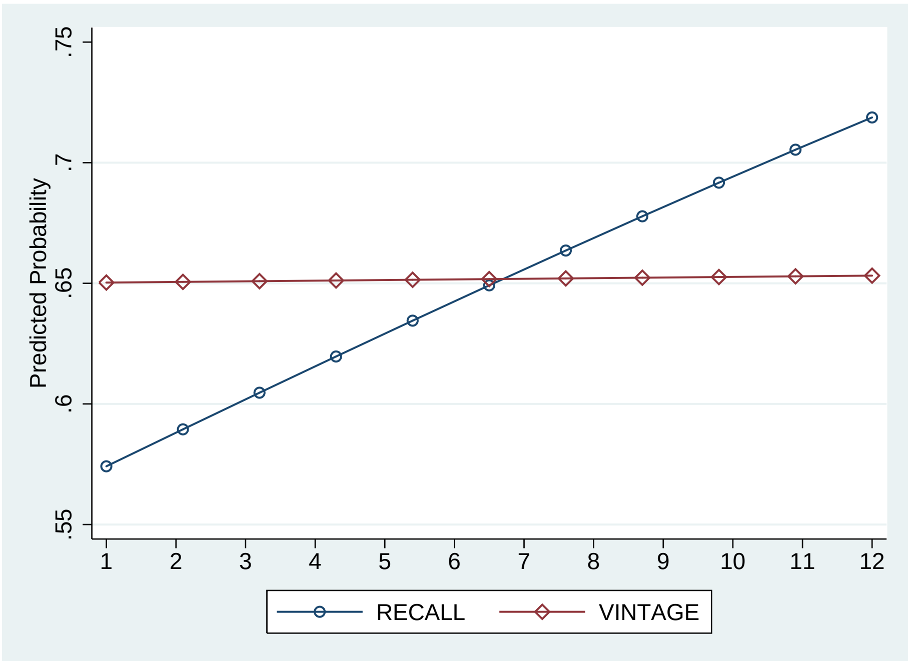
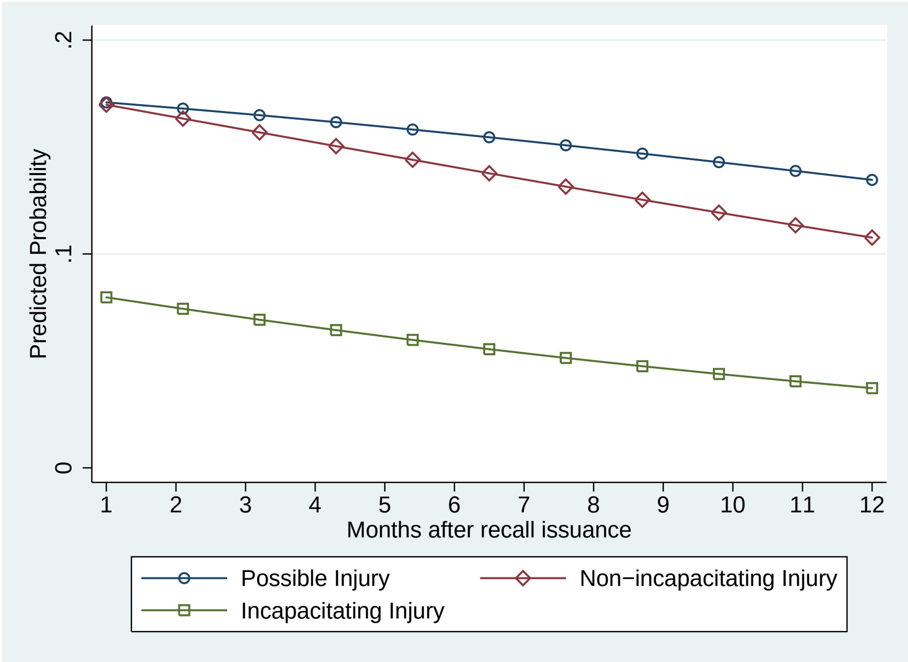
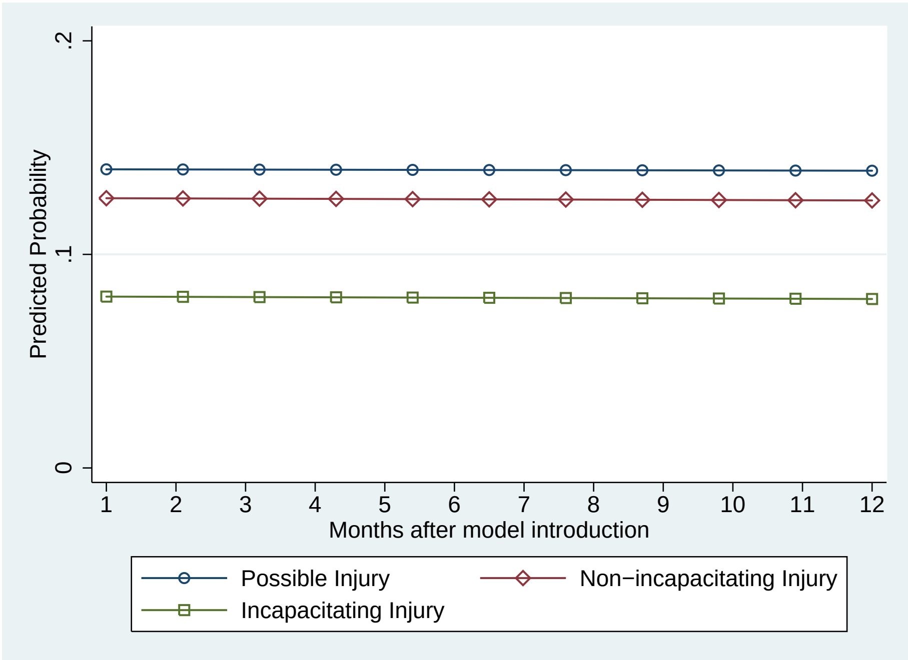


Figure 7: Predicted Probabilities from MODEL 1†: Recall effect



† The predicted probabilities for "No Injury" and "Fatal Injury" are not included.

Figure 8: Predicted Probabilities from MODEL 4‡: Vintage effect



‡ The predicted probabilities for "No Injury" and "Fatal Injury" are not included.

Table 8: Quarterly Correction rates for two 2005 Recalls

Auto-Maker		Ford†			GMC / CHEVROLET‡		
Year Model		2004, 2005			2000, 2001		
Issue Dates		7-22-2005			10-19-2005		
Units Involved		180,113			316,591		
Year	Month	Units Fixed	Units Unreachable	Correction Rate	Units Fixed	Units Unreachable	Correction Rate
2005	11	120,610	2,534	.669	-	-	-
2005	12	-	-	-	-	-	-
2006	1	141,618	2,185	.786	68,937	8,612	.218
2006	2	-	-	-	-	-	-
2006	3	-	-	-	-	-	-
2006	4	-	-	-	110,280	9,194	.348
2006	5	154,140	2,597	.856	-	-	-
2006	6	-	-	-	-	-	-
2006	7	-	-	-	169,186	9,471	.534
2006	8	160,376	2,210	.890	-	-	-
2006	9	-	-	-	-	-	-
2006	10	-	-	-	188,016	9,559	.594
2006	11	165,092	1,860	.917	-	-	-
2006	12	-	-	-	-	-	-
2007	1	-	-	-	206,817	9,461	.653
2007	2	168,647	1,576	.936	-	-	-

Source: "http://www-odi.nhtsa.dot.gov/recalls/recallsearch.cfm"

† Campaign ID #: 05V270000, Initiation: *ODI*, Components: *Electrical system*, *wiring*.

‡ Campaign ID #: 05V155000, Initiation: *ODI*, Components: *Fuel system*, *fuel pump*.

References

[1] Ashenfelter, Orley., Greenstone, Michael: "Using Mandated Speed Limits to Measure the Value of a Statistical Life.", NBER Working Paper No. w9094, (2002)

[2] Bae, Yong-Kyun and Benítez-Silva, Hugo:"Do Vehicle Recalls Reduce the Number of Accidents? The Case of the U.S. Car Market", manuscript. Hood College, MD and Stony Brook University. (2010)

[3] Bae, Yong-Kyun: "Vertical Differentiation with Product Safety Quality", manuscript. Hood College, MD. (2010)

[4] Bates, Hillary., Matthias Holweb, Michael Lewis, and Nick Oliver: "Motor vehicle recalls: Trends, patterns and emerging issues," *Omega: International Journal of Management Science*, vol. 35-2, 202-210, (2007).

[5] Cameron, Colin A., and Trivedi, Pravin K: Regression analysis of count data, Cambridge University Press. (1998).

[6] Crafton, Steven M., Hoffer, George E., and Reilly, Robert J., "Testing the Impact of Recalls on the Demand for Automobiles", *Economic Inquiry.* vol 19:4. (1981): 694-703.

[7] Cooter, Robert., and Ulen, Thomas., Law and Economics, 3rd Ed, Addison Wesley Longman, Inc., (2000).

[8] Evans, Leonard: "Traffic Crashes", *American Scientists*, vol. 90, 244-253, (2002).

[9] Gould, William., and Sribney, William: Maximum Likelihood Estimation with Stata, Stata Corporation, College Station, Texas, 1999.

[10] Greene, William H: Econometric Analysis, Upper Saddle River, New Jursey: Prentice Hall, 2000.

[11] Hartman, Raymond S., "Product Quality and Market Efficiency: The Effect of Product Recalls on Resale Prices and Firm Valuation", *The Review of Economics and Statistics*, vol. 69, No. 2. (1987), 367-372.

[12] Hoffer, George E., Pruitt, Stephen W., and Reilly, Robert J: "When Recalls Matter: Factors Affecting Owner Response to Automotive Recalls", *Journal of Consumer Affairs*, vol 28:1.(1994): 96-106.

[13] Huble, Norma F., Arndt, Mark W: "A Review of Crash Data Analysis in a Defect and Recall Investigation of the General Motors C/K Pickup Trucks", *Accident Analysis and Prevention*, vol 28:1. (1996): 33-42.

[14] Jarrel, Gregg., and Peltzman, Sam: "The Impact of Product Recalls on the Wealth of Sellers." *Journal of Political Economy*, vol 93:3. (1985): 512-536.

[15] Kaplow, Louis., and Shavell, Steven: "Economic Analysis of Law", NBER Working Paper No. 6960, (1999).

[16] Long, Scott J., and Freese, Jeremy: Regression Models for Categorical Dependent Variables using Stata, Stata Corporation, College Station, Texas, 2001.

[17] Marino, Anthony M: "A Model of Product Recalls with Asymmetric Information." *Journal of Regulatory Economics*, vol 12, (1997): 245-265.

[18] NHTSA: Motor Vehicle Defects and Safety Recalls: What Every Vehicle Owner Should Know, 2009.

[19] NHTSA: Traffic Safety Facts, U.S. Department of Transportation, DOT HS 811 172, 2009.

[20] NHTSA: National Automotive Sampling System(NASS), General Estimates System(GES), U.S. Department of Transportation, 1988-2007.

[21] NHTSA: Safety Recall Compendium, Defects and Recall Information Analysis Division, U.S. Department of Transportation, 2001.

[22] Peltzman S: "The Effects of Automobile Safety Regulation", *Journal of Political Economy*, vol. 83, (1975): 677-725.

[23] Posnerwer, Richard A: Economic Analysis of Law, Aspen Law and Business, (1998).

[24] Rhee, Mooweon., and Pamela, Haunschild: "The Liability of Good Reputation: A Study of Product Recalls in the U.S. Automobile Industry," manuscript, Stanford University, (2003).

[25] Rupp, Nicholas G., and Taylor, Curtis R: "Who Initiates Recalls and Who Cares? Evidence from the Automobile Industry." *Journal of Industrial Economics*, vol L, No. 2, (2002): 123-149.

[26] Ward's Automotive Yearbook, Annaul, Detroit, MI.

[27] Wooldridge, Jeffrey M: Econometric Analysis of Cross Section and Panel Data, The MIT Press. (2002).