Elsevier Editorial System(tm) for Journal of Safety Research
Manuscript Draft

Manuscript Number:

Title: National Evaluation of the Effect of Graduated Driver Licensing Laws on Teenager Fatality and Injury Crashes

Article Type: Full Length Feature Article

Keywords: Graduate Drivers Licensing, Teenager, Crash, Regression Analysis

Corresponding Author: Dr. Rong Pan, Ph.D.

Corresponding Author's Institution: Arizona State University

First Author: Joshua D Lyon

Order of Authors: Joshua D Lyon; Rong Pan, Ph.D.; Jing Li, Ph.D.

Abstract: Problem: Automobile crashes remain a prominent cause of death and injury for teenagers in the United States. While it is generally agreed that graduated drivers licensing (GDL) influences crash rates, it is unclear which components have the strongest effect on any specific types of crashes. Method: We analyze the relative effect of different stages of GDL on teenage fatal and injury crash risk via a negative binomial generalized linear model with random state effects. Overall, nighttime, and crashes with multiple teenage passengers are considered. Results: The strongest effects are seen by 16-year-olds, for which a strict permit stage is associated with a 58% reduction in fatal crash risk over a lenient permit stage. Similar reductions are seen for injury crashes. The intermediate stage, involving nighttime and passenger restrictions, is associated with a 44% reduction in fatalities but has relatively little effect on injury crashes. The strongest effects are generally seen for passenger crashes, followed by nighttime, and then overall crashes. Impact on Industry: This study identifies stronger relationships between GDL and crash risk than has previously been discovered and captures the relative effects of permit and intermediate licensing restrictions, two high-level components of GDL which differ in intent and implementation.

Ira A. Fulton Schools of Engineering
ARIZONA STATE UNIVERSITY

School of Computing, Informatics
and Decision Systems Engineering
Box 878809
Tempe, AZ 85287-8809

480-965-3190
FAX: 480.965.2751

August 3, 2011

Dear Editor of Journal of Safety Research:

We submit the manuscript entitled "National evaluation of the effect of graduated driver licensing laws on teenager fatality and injury crashes" for the consideration of publication on Journal of Safety Research.

In this paper, we analyze the crash rate of teenagers of different age groups over the past 15 years and the relative effect of different stages of graduated driver licensing (GDL) laws on the teenage crash risk. Both fatality and injury types of crashes are analyzed. In addition, we discuss the specific crash types involving nighttime driving and passenger restrictions, which are targeted by the restrictions in the intemediate stage of GDL.

We believe the paper is well suited to the audience of Journal of Safety Research. We appreciate your time and effort on handling the review of our paper. For future correpondence, please address to Rong Pan at rong.pan@asu.edu. Thank you.

Sincerely,

Rong Pan
Joshua Lyon
Jing Li
Arizona State University
Tempe, Arizona

*Title page with author details

National Evaluation of the Effect of Graduated Driver Licensing Laws on Teenager Fatality and Injury Crashes

Joshua D. Lyon[a], Rong Pan[a,*], Jing Li[a],
[a] *Arizona State University School of Computing, Informatics, and Decision Systems Engineering, 699 S. Mill Avenue, Tempe, AZ 85281, United States*

*Corresponding author. Tel: +1 480 965 4259. Fax: +1 480 965 2751
E-mail address: rong.pan@asu.edu

Author e-mail addresses:
joshua.lyon@asu.edu
rong.pan@asu.edu
jing.li.8@asu.edu

Abstract

Problem: Automobile crashes remain a prominent cause of death and injury for teenagers in the United States. While it is generally agreed that graduated drivers licensing (GDL) influences crash rates, it is unclear which components have the strongest effect on any specific types of crashes. *Method*: We analyze the relative effect of different *stages* of GDL on teenage fatal and injury crash risk via a negative binomial generalized linear model with random state effects. Overall, nighttime, and crashes with multiple teenage passengers are considered. *Results*: The strongest effects are seen by 16-year-olds, for which a strict permit stage is associated with a 58% reduction in fatal crash risk over a lenient permit stage. Similar reductions are seen for injury crashes. The intermediate stage, involving nighttime and passenger restrictions, is associated with a 44% reduction in fatalities but has relatively little effect on injury crashes. The strongest effects are generally seen for passenger crashes, followed by nighttime, and then overall crashes. *Impact on Industry*: This study identifies stronger relationships between GDL and crash risk than has previously been discovered and captures the relative effects of permit and intermediate licensing restrictions, two high-level components of GDL which differ in intent and implementation.

Keywords: Graduate Drivers Licensing, Teenager, Crash, Regression Analysis

***Highlights**

Highlights of the paper:

- This nationwide study evaluates the relationship of licensing laws to teen fatal and injury crash rates.
- The model delineates between permit and intermediate licensing.
- Results show stronger relationships than previously uncovered.
- The permit stage is consistently effective across ages and crash types.
- The intermediate stage has less effect on injury than it does fatal crash rates.

National Evaluation of the Effect of Graduated Driver Licensing Laws on Teenager Fatality and Injury Crashes

Abstract

Problem: Automobile crashes remain a prominent cause of death and injury for teenagers in the United States. While it is generally agreed that graduated drivers licensing (GDL) influences crash rates, it is unclear which components have the strongest effect on any specific types of crashes. *Method*: We analyze the relative effect of different *stages* of GDL on teenage fatal and injury crash risk via a negative binomial generalized linear model with random state effects. Overall, nighttime, and crashes with multiple teenage passengers are considered. *Results*: The strongest effects are seen by 16-year-olds, for which a strict permit stage is associated with a 58% reduction in fatal crash risk over a lenient permit stage. Similar reductions are seen for injury crashes. The intermediate stage, involving nighttime and passenger restrictions, is associated with a 44% reduction in fatalities but has relatively little effect on injury crashes. The strongest effects are generally seen for passenger crashes, followed by nighttime, and then overall crashes. *Impact on Industry*: This study identifies stronger relationships between GDL and crash risk than has previously been discovered and captures the relative effects of permit and intermediate licensing restrictions, two high-level components of GDL which differ in intent and implementation.

Keywords: Graduate Drivers Licensing, Teenager, Crash, Regression Analysis

1. Introduction

Car crashes are the number one cause of death for teenagers over age 15 (Heron, 2006). In 2009, 1,659 drivers aged 15-17 in the United States (US) were involved in a fatal car crash and about 100 times more were reported to be involved in a crash where there was at least a possible injury. Licensed teen drivers are more likely to be involved in a serious crash than those who are marginally older or middle aged.

Fortunately, national crash rates amongst all age groups have been steadily declining over time, indicating that conditions are becoming safer as a whole. In addition to the trend of the general population, teenagers have seen an even bigger improvement over the past 15 years. Fig.1 plots the risk for a teenager to be involved as a driver in a fatal or injury crash relative to a person aged 25-44. Here, risk is calculated as the average number of crashes per 1,000 persons in the population, the data for which was gathered as described in Section 2.1. Relative risk is defined as the risk for the test group divided by the risk for the control group. It is apparent from Fig.1 that injury trends are similar to those of fatality but also that the sheer volume of injury crashes results in considerably less noise. These cleaner trends make it easier to identify underlying changes in risk and thus provide one of the principal benefits of including injury crashes in this study.

The biggest reduction in relative risk has been seen by 16-year-olds, dropping from about 1.2 times that of the control group to less than 0.8 over the past 15 years. While a relative risk less than one is encouraging news for the safety of the teen *population*, the number of crashes per *mile driven* remains high amongst teenagers, motivating a closer look as to how teenage driver safety may be improved.

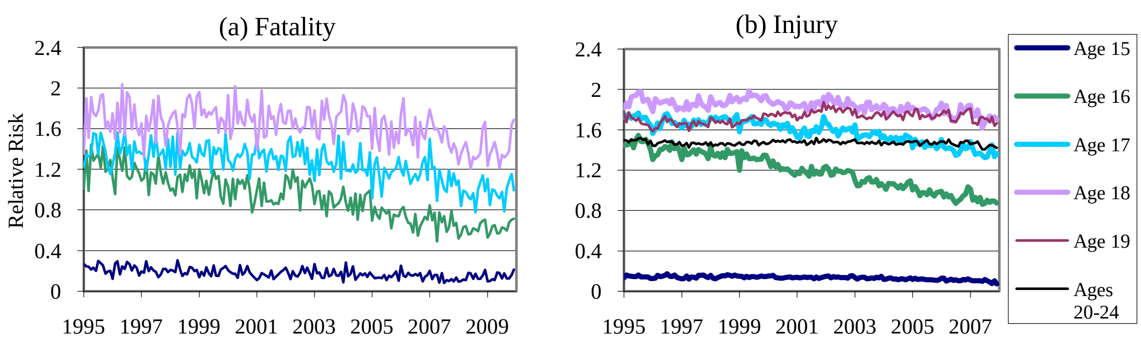

Figure 1. Nationwide crash relative risk adjusted for monthly seasonality. The control group is the age group 25-44. Risk is low for 15-year-olds due to lower licensing rates. Injury trends for ages 19 and 20-24 are shown to illustrate how changes in crash rates for post teens have remained close to the control group. Trends are similar for both crash types.

Over the same period, graduated driver licensing (GDL) laws were introduced and have become increasingly prevalent in the United States (Simpson, 2003). GDL describes state-implemented licensing laws that involve restrictions that progressively become more lenient as the driver works towards a full license. As implemented in the US, GDL laws focus on young drivers and split licensure into three stages. The first (permit) stage allows the young person to drive only when supervised by an appropriate licensed driver. In the second (intermediate) stage, the teenager may drive unsupervised except during certain high risk conditions. Namely, driving at night or with other young passengers has been identified as high risk (Rice et al., 2003). The intermediate GDL stage includes restrictions on these very situations. The third stage of licensure is considered to be when the young driver graduates to a full license, beginning once intermediate restrictions have been lifted and the driver is subject to all the normal driving laws in his/her respective state.

1.1. Literature Review

As teenage licensing laws have become stricter, safety is improved by encouraging learning and reducing exposure to dangerous situations. In recent years, researchers have tried to explain the reduction in young driver crash risk by evaluating licensing restrictions. Regression models, in particular, have uncovered strong correlation of stricter GDL laws and the reduction in young driver risk. Shope (2007) and Williams and Shults (2010) provide excellent literature reviews of GDL and teenager crash studies. There have been a variety of analyses on individual states and relatively few nationwide and international studies. The larger studies are able to provide more holistic insights because a wide range of scenarios can be considered in the same model.

Relatively early on, Eisenberg (2003) briefly considered the effect of GDL using weighted least squares regression. He estimated that GDL programs may reduce crash rates for drivers under 21 by 9.2%. Soon after, others linked GDL to greater reductions of teen driver crashes. Both research by Dee et al. (2005) and Morrisey et al. (2006) associated GDL systems with a score of "good" under the IIHS point system as having about 19% lower crash rates for 15-17 year-olds than those with a score of "poor". The latter study also looked at night and passenger crashes, whose results did not differ significantly (with the probability of a type I error, α, less than 0.05) from those of overall crashes. Nevertheless teenage passengers did show the strongest signs of fatality reductions in the presence of stronger GDL programs. Chen et al. (2006) looked at seven common GDL restrictions and scored states' systems by the number of restrictions above a predefined threshold. In their results, at least five "strict" restrictions were necessary in order to find a statistically significant improvement over a single strict restriction. Six or more strict restrictions were associated with a 21% reduction in risk for 16-year-olds. In the following year the same authors used higher thresholds to show a 38% reduction in risk for states with five strict restrictions (Baker et al., 2007). The latter was also the first nationwide GDL study to consider injury crashes, whose results did not significantly differ from those of fatal crashes.

The above studies were the earliest to show that comprehensive GDL programs are associated with reduced crash rates; more recent papers have found drawn the same conclusion (Pressley et al., 2009; Vannlar et al., 2009; HLDI, 2009; McCartt et al., 2010). Beyond learning that GDL works, it is desirable to understand which aspects are the most important. To this end, two different research studies have built models which simultaneously evaluated individual GDL restrictions. Vannlar et al. (2009) analyzed data from the US and 11 Canadian jurisdictions against several predictors. Their study uncovered limited insights. For 16-year-olds, they only found a significant effect for passenger restrictions which include restrictions against family members.

Using US data, McCartt et al. (2010) identified as significant the state's permit age, intermediate licensing age, and night and passenger restrictions. An important aspect of this study was the decision not to directly control for the state where the crash occurred. It has otherwise been the standard for nationwide studies to control for state. Eisenberg (2003), Dee et al. (2005), Morrisey et al. (2006), and Chen et al (2006) all included fixed state effect estimates in their models. Baker et al (2007) accounted for within state correlations using generalized estimating equations (GEE); Vannlar et al. (2009) did so using random state effects. Such practices are common in longitudinal studies in order to ensure that a statistically valid model is achieved. McCartt et al. (2010) justified not directly controlling for state by using the crash rates of a control age group (30-59 year-olds) as a covariate, arguing that any differences between states should be reflected by the crash rates of the control group. We describe an analysis of the alternatives in Section 2.3.4.

1.2. Contributions

Our research makes two contributions to the field by:

1. Analyzing the *relative* effects of GDL stages one and two on fatal and injury crashes.
2. Modeling the response of night and passenger injury crashes.

The historical tendency for GDL components to simultaneously become stricter makes it challenging to isolate the strength of their individual effects. As a result, estimates on individual restrictions are never very robust. We moderate this pitfall by categorizing permit (stage one) and intermediate (stage two) licensing components. In contrast to modeling individual restrictions, it turns out that this characterization actually helps us to uncover *stronger* relationships between GDL components and crash risk than has been suggested by other national studies.

Decomposing GDL by stage is a natural step for testing its effectiveness because stages one and two differ considerably in their intent and implementation. Stage one serves to increase supervised learning experience and delay the point in which a teenager may drive unsupervised. Meanwhile, stage two reduces the exposure of unsupervised drivers to high risk scenarios. This design allows us to effectively deal with the concerns of Baker et al. (2007) whom caution against the use of regression to fit and interpret individual licensing restrictions, arguing that the results could be misleading since GDL programs are generally designed to work as a whole.

In a thoughtful examination of the existing literature, Williams (2007) concluded that there is strong evidence identifying the end of the permit period, night restrictions, and passenger restrictions as the most important aspects of GDL. He noted that "...although it is not an easy task to sort out the relative contribution of the various GDL components, it is probable that the more major effect comes from the extended learner permit period". Our specifications of stage one and stage two variables aptly serve to test this hypothesis.

Of additional concern is how GDL laws affect injury crashes relative to fatalities. All national studies of which we are aware concern fatalities, but relatively few consider injury crashes. The lesser attention on injury crashes is more likely due to data unavailability than to a belief that injuries are unimportant. While fatal crashes are clearly more severe, injury crashes occur at a much higher rate. Naumen et al. (2010) estimate that in 2005 US traffic injuries of 15-19 year-olds costed about $5 billion, compared to a total of $11 billion for fatalities.

Since removing drivers from high-risk situations at night or with multiple passengers is a significant part of the GDL’s second stage strategy, the relationship between GDL and "night" and "passenger" crashes is evaluated. By doing so, we can more precisely evaluate stage two’s effectiveness at reducing the crash types which it targets. Here, very low fatality counts serve as a detriment to good fit so only injury crashes are considered for model adequacy reasons.

Finally, we obtain improved model adequacy by using a random effects model which assumes that the relative effect of each state varies normally. Random effects are useful for making *inferences* based on the injury model, for which data is not available from every state (see Section 2.1). These state-dependent effects also account for important variables which may not be identified or explicitly included in the regression model. An improved fit across all models contributes to the validity of the results and our confidence in making valid inferences on the relationship between licensing laws and crash rates.

2. Material and Method

2.1. Data Sources

Fatality statistics are available online from the Fatality Analysis Reporting System (FARS, 2010), supported by the National Highway Traffic Safety Administration (NHTSA). The use of this tool is widespread and allows for querying records of all fatal crashes in the US. In order to perform more complicated preprocessing such as classifying "passenger" crashes, we download monthly case listings from every state plus Washington D.C.

Injury statistics are obtained from the State Data System (SDS, 2009) which is also supported by the NHTSA. The SDS is a collection of state crash data which includes a variety of information from many types of crashes which have been recorded via police crash reports in the participating states. Over half of the states provide data to the SDS. However, some do not allow NHTSA to release their information; we obtained permission to use crash statistics from 19 states. Included in our study are all crashes identified to have at least one personal injury.

A number of studies have utilized SDS injury crash data (Campbell et al., 1991; Farmer and Lund, 2002; Farmer and Williams, 2002; Eisenberg, 2004; Karaca-Mandic and Ridgeway, 2010; Cheung and McCartt, 2011). Further uses by the NHTSA are outlined by Rhea and Austin (2006). The fact that the SDS has been used in several research papers still does not guarantee adequate data quality. Accuracy and validity are subject to the individual states' reporting. We perform two checks to validate the data used. First, we compare fatality statistics from the SDS to those pulled from the FARS. We observe that the crash counts are close between the two sources. This verifies that, at least in the case of fatalities, reporting to SDS is consistent with FARS.

Secondly, to evaluate the *injury* data to be used, we plot injury rates for each state by month. By inspection, we find clear quality issues in the data from Texas from 1995 to 1997 and in all available years of one other state. We remove these observations outright before proceeding with the analysis. In the end, data from 18 states are used as summarized in Appendix A. The SDS state-periods used reflect just over 40% of the nationwide exposed population of teenagers aged 15-17 between 1995 and 2007. A geographically diverse set of states are covered with four (of 13) from the West, seven (of 12) from the Midwest, five (of 17) from the South, and two (of nine) states from the Northeast.

In order to interpret changes in crash counts, it is important to take into account exposure. A common approach is to consider the rate of crash counts per population. We take yearly population estimates for each age group from the census bureau and interpolate for quarterly and monthly populations (U.S. Census Bureau, 2010). Another desirable statistic is the number of licensed drivers. Unfortunately, inconsistent reporting between states, and sometimes between years in the same state, creates a data quality issue which is hard to reconcile. See IIHS (2006) and Vannlar et al. (2009) for more details on the issue.

To categorize GDL changes, we utilize data recorded by IIHS between the early 1990's and 2009 on the levels of eight separate restrictions which are commonly applied by different states (IIHS, 2009). This information is used to calculate predictor variables as described in Section 2.3.2. Changes in GDL law which were enacted in the second half of a month are attributed to crashes beginning in the next month.

In many states, GDL restrictions vary based on the teenager's enrollment or completion of state-approved driver education, which often includes public or professional classroom or behind-the-wheel training. Due to a lack of information over time, our model does not explicitly consider such driver education programs. It may be noted that the effectiveness of driver education is debatable and that relaxing GDL restrictions for individuals who complete driver education may very well lead to a net increase in risk (Mayhew, 2007). For the historical tracking of GDL laws which IIHS provides, it does not contain enough information to allow us to include the consideration of a teenager's participation in driver education.

2.2. Definition of night and passenger crashes

To evaluate night crashes, we fit the model to events that occur between 11pm and 6am. These are common times for night restrictions to take affect. Most passenger restrictions allow for one or less teenage passenger. To reflect this, passenger crashes are defined as those where there are two or more passengers between the ages of 15 and 19.

2.3. Regression Model

Crash counts are fit using a generalized linear model (GLM) assuming the negative binomial distribution. A GLM is a generalized version of standard linear regression which can fit data to any probability distribution in the exponential family (Gill, 2001). The negative binomial distribution can be viewed as a generalized version of Poisson distribution, containing an extra parameter that can account for overdispersion (Hilbe, 2007). As such, the negative binomial distribution is useful when variance is not equal to the mean as is assumed with Poisson. We elect to use the negative binomial distribution because it is a natural choice for modeling count data and it demonstrates significantly better fits than Poisson.

With the choice of negative binomial as the response distribution, the GLM includes a linear predictor $\boldsymbol{\beta x}$ and a log link function, where $\boldsymbol{\beta}$ is the vector of coefficients for each predictor and $\boldsymbol{x}$ is the two-dimensional matrix describing the predictor variables for each state-period. With this, the relationship between the predictors and the crash counts $\boldsymbol{y}$ takes the basic form of:

$$y_i \sim \mathrm{NB}(\mu_i, k) \tag{1a}$$
$$\log(\mu_i) = \beta_1 x_{i1} + \beta_2 x_{i2} + \ldots + \beta_n x_{in} \tag{1b}$$

where μ_i is the mean and k is the dispersion parameter. It should be noted that Eq. (1a) represents a re-parameterization of the negative binomial model for convenience of expression using these mean and dispersion parameters; the variance of y_i is equal to $\mu_i + k\mu_i^2$ (Hilbe, 2007). The log link function connects the mean to the predictor $\boldsymbol{x}_i$, with i representing the state-period observation and n the number of predictor variables. SAS is used to carry out the regression analysis. The GLIMMIX procedure allows the user to model exposure (population) using the OFFSET statement to specify the variable log(*population*) with a known (assumed) coefficient of one. In the final model, the exponential and logarithmic terms cancel to give:

$$\mu_i = population_i \times \exp(\beta_1 x_{i1}) \times \exp(\beta_2 x_{i2}) \times \ldots \times \exp(\beta_n x_{in}) \tag{2a}$$
$$\mu_i \,/\, population_i = \exp(\beta_1 x_{i1}) \times \exp(\beta_2 x_{i2}) \times \ldots \times \exp(\beta_n x_{in}) \tag{2b}$$

Therefore, even though we use a count distribution, variable coefficients are actually fitted against a measure of risk: number of crashes per age population.

2.3.1. Time periods

The model is applied over years 1995 to 2009. For fatality crashes, *quarterly* time periods are used. Indeed, more observations can be gained by utilizing shorter periods, however we observe that doing so results in poorer fits due to an excess number of periods with zero counts. Lord et al. (2005) note that this is not unusual for crash types of low frequency, and likewise recommend using a longer time period over which to summarize crashes. As such, the period length used for fatality crashes strikes a balance between, what are in this case, the competing objectives of a high number of observations and model adequacy. Injury crash counts are much higher than fatality and consequently adequate model fit is accomplished using *monthly* periods.

In many instances, a GDL change does not come into full affect for one year; many teenagers are "grandfathered" in. For example, consider an individual who gets a permit one day before the legal age is raised by one year; this teen may still be subject to the prior restriction for 364 days. We exclude these transition periods from the analysis by removing periods that fall within 12 months after a state enacts a GDL change.

2.3.2. Definition of GDL predictor variables

Graduated drivers licensing laws have all generally become stricter over time. This has resulted in significant multicollinearity between restrictions. From 1995-2009, states changed at least one GDL restriction 125 times and in 78 of these instances there were changes to multiple restrictions. In more than one third of all cases three or more restrictions were simultaneously made stricter. With such instances, it is difficult to estimate the individual effect of individual GDL restrictions; effect estimates have wider confidence intervals and tend to change significantly following the inclusion or exclusion of other variables.

Table 1. GDL restriction category

(a) Points attributed to each stage according to strictness of restrictions.

Stage 1: permit license period	**Level**	**Points**
Minimum permit age	16 or older	1
	Less than 16	0
Permit holding period	6 or more months	2
	3-5 months	1
	Less than 3 months	0
Minimum supervised driving	30 or more hours	1
	Less than 30 hours	0
Stage 2: intermediate license period		
Night driving curfew	10pm or earlier	2
	After 10pm	1
	None	0
Maximum number of underage passengers	Zero or 1 passenger	2
	2 passengers	1
	3 or more passengers	0
Duration of night driving restriction	12 months or more	1
	Less than 12 months	0
Duration of passenger restriction	12 months or more	1
	Less than 12 months	0

(b) Levels for each stage based on number of points for restrictions within stage.

	Sum of Points	
Category	Stage 1	Stage 2
A	4	5-6
B	1-3	1-4
C	0	0

To reduce the problems coming from multicollinearity, we use categorical scores to summarize GDL laws. Most nationwide studies have used the IIHS GDL rating system of "good", "fair", "marginal", or "poor" as a model predictor. The IIHS rating system categorizes states' licensing laws by assigning points to different restrictions. To calculate GDL prediction variables, we sum points for each restriction in a similar manner but we sum over two separate stages, creating separate categorical variables for stages one and two. The methodology is summarized in Table 1. Each stage is classified into one of three categories, with *A* being the most strict and *C* the most lenient. The trends of these predictor variables are summarized in Fig.2, from which it is apparent that stage one and stage two are not independent. However, the consequence is mild compared to when *three* or more GDL restrictions are simultaneously considered as separate variables.

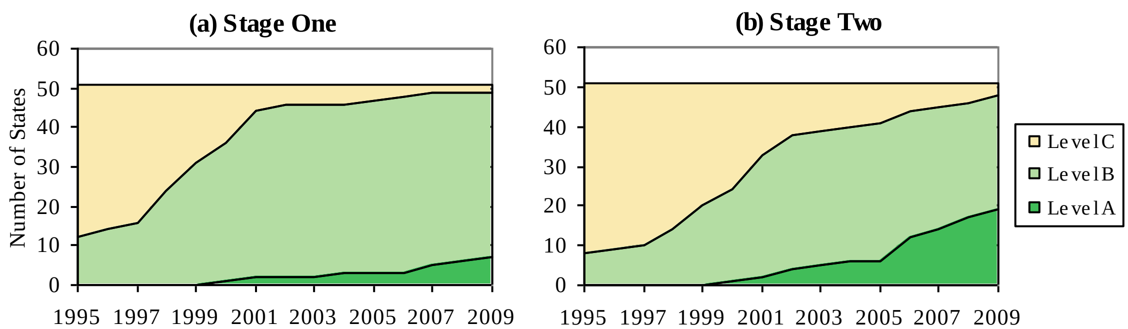

Figure 2. The number of states in each GDL category, as defined in Table 1, by year.

2.3.3. Definition of Non-GDL predictor variables

Besides predictors describing GDL laws, it is appropriate to include variables which help explain other variation in the data even if we are not interested in studying these variables directly. Considering pseudo-AIC statistics provided by SAS, we find it beneficial to include factors describing fatal crash rates for a control age group 25-44, the quarter of the year, and the state in which the crash occurred. Discussion on how state effects are modeled is the focus of Section 2.3.4.

Using a numerical predictor representing the crash rate for a control group (e.g., age group 25-44) shifts the interpretation of the model towards evaluating *relative* risk. The motivation is that it is good to know whether a reduction in crashes can be explained by conditions that make everyone less at risk, regardless of age. Such factors could be unknown or impossible to quantify, and may include the increased prevalence of safer road or vehicles, general cultural shifts, or greater law enforcement. With the inclusion of this control variable, the interpretation of the effect estimate of a GDL predictor shifts towards: how much crash rates for the target age have changed relative to the control group.

A number of reasonable control groups may be specified. A "perfect" control group would: 1) have crash rates which are independent and identically distributed to the marginal distribution of the test group (e.g., the 16-year-old drivers) conditioned on the GDL laws and 2) have a large enough population for a small sampling variance. We elect to use the age group 25-44.

2.3.4. Random State Effects

A central assumption of any regression model is independent observations. In nationwide studies, it is natural to expect that individual states will have different crash rates, due to terrain, road condition, population distribution, etc.; it so happens that, in practice, variation is not sufficiently summarized by the GDL and control variables described in previous sections. The most intuitive way to account for the state effect is to fit a separate intercept for each state such that all responses are measured in relation to the average crash rate for that state. However, this approach substantially increases the number of variables the model must estimate. We avoid such a large increase in degrees of freedom by utilizing a random effects model where the intercept is assumed to vary normally by state. The resulting effects for individual states have the same interpretation as in fixed effect models; however the fitting procedure does not consider these effects to be independent of each other (Burton et al., 1998).

Checking model adequacy of a regression model is important because without it the validity of the results would have to come into question. Residuals are often used as the principal indicator of model adequacy. Among the alternatives, the fixed effects model yields the best residuals but carries the most danger of over-fitting. In testing we find that residuals from the random effects model are negligibly different from the fixed effects model, are somewhat better than generalized estimating equations (GEE), and are much better than the model which does not directly consider state.

3. Results

We fit separate models for ages 16, 17, and 15-17 and observe the incidence rate ratios (IRRs) of each effect. The IRR is calculated by applying the exponential function on the effect estimate. Taken in the context of Eq. (2), it can be seen that IRR represents the percentage change in the expected crash rate per unit change in the predictor. For the categorical stage one and stage two variables, these percentages are defined as the risk relative to the least strict level *C*. All results are summarized in Appendix B. All results mentioned in the text below are statistically significant with $\alpha = 0.05$.

3.1. Overall Crashes

Consider the IRRs of fatal and injury crashes plotted in Fig.3. Effect estimates for *stage one* are similar for each crash type while *stage two* has a weaker relationship with injury crashes than it does with fatalities. Despite including data from fewer states, the injury models yield better residual fits and parameter estimates with tighter confidence intervals.

Among fatal crashes (Fig.3(a)) the effects of stages one and two are reasonably big. For the wider age group 15-17, they also have similar magnitudes, the strictest category for each corresponding to 38% and 35% reductions in risk, respectively. The biggest effect is by stage one on 16-year-olds, corresponding to a 58% reduction in risk. Unsurprisingly given their intentions, stage one has a significantly bigger effect on age 16 than it does on 17-year-olds and vice versa for stage two.

Among injury crashes (Fig.3(b)) stage one effects are very similar to those on fatalities. However, a large disparity between the injury and fatal models is seen for stage two; a stricter stage two is not associated with a steeper drop in injury crashes. In fact, the strictest level, stage two = *A*, is correlated with an increase in injury crash rates among ages 16 and 17 relative to older drivers.

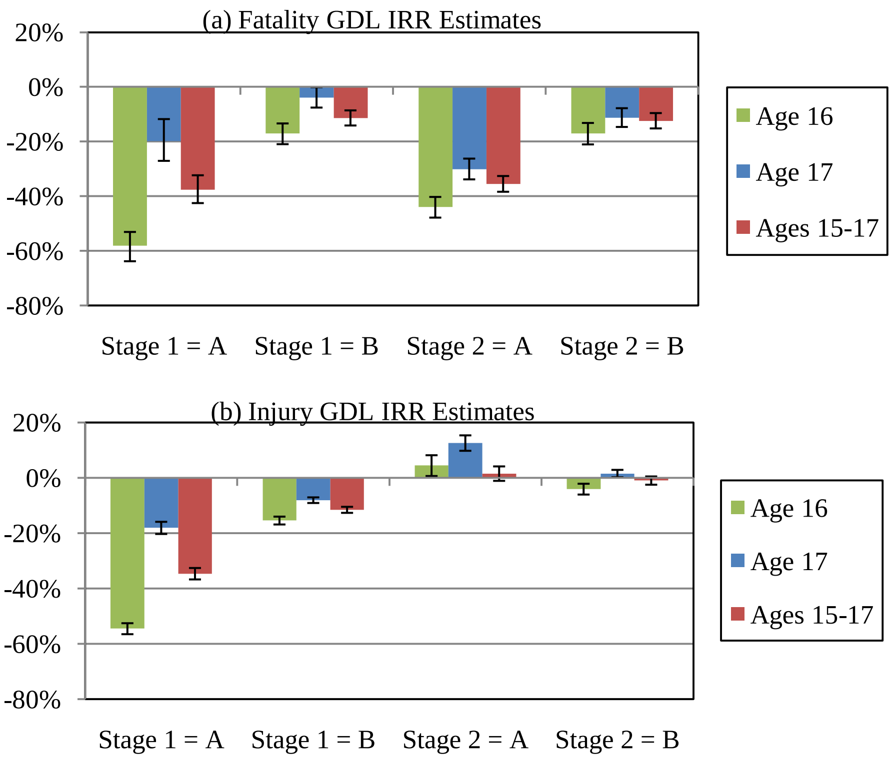

Figure 3. GDL incidence rate ratios (IIR) for (a) fatality and (b) injury crashes. The IIR for each is the risk relative to the most lenient level = *C*. Error bars indicate the 90% confidence intervals.

From model to model, the estimates of the *non-GDL* covariates are relatively consistent. The period from January to March has the lowest crash rates and July through September has the highest. The biggest difference between models on different ages is the variation between states; the random intercept in the age 16 model has an estimated variance of 0.138 compared to 0.028 for the age 17 model. This indicates that the younger age, especially, is more affected by variables absent from the model than the control group is.

3.2. Night and Passenger Injury Crashes

Night and passenger *fatality* crashes may also be analyzed, however doing so results in poor residuals. This is likely due to small crash counts: night and passenger crashes each represent less than one fifth of the all fatal crashes. As outlined in Section 2.3.1, in order to achieve model adequacy it is more appropriate to apply our model to the higher-count injury response.

The resulting GDL effect estimates on 16-year-old injury crashes are summarized in Fig.4. For both 16 and 17-year-olds, intermediate restrictions have a significantly stronger effect on night crashes than on overall crashes. For both ages, night crashes are most strongly related to a strict stage one (*A*). As can be observed in Appendix B, for 17-year-olds a strict stage one (A) is estimated to have a 17% stronger effect on night crashes than on overall crashes.

Among all crash types, the effect of GDL is strongest on passenger crashes. For each age tested estimates for strict GDL are about 10-20% lower than for overall crashes. Even stage one is more strongly related to fewer passenger crashes than overall crashes, with a 69% versus a 54% reduction for age 16.

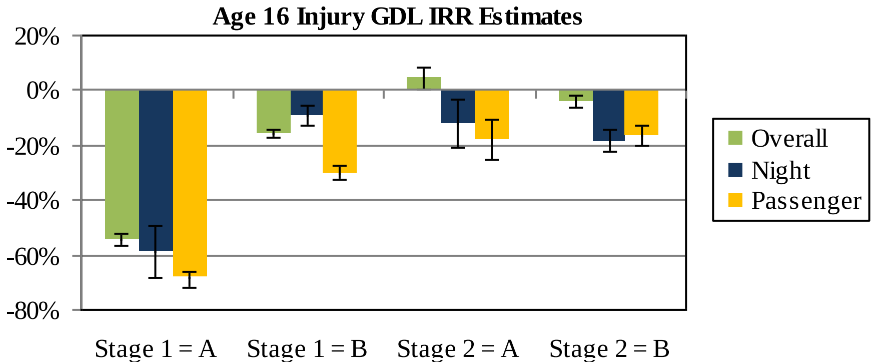

Figure 4. GDL incidence rate ratios for age 16 overall injury crashes compared to those that occur between 11pm and 6am (night) or with more than two passengers between the ages of 15 and 19 (passenger). Error bars indicate the 90% confidence intervals.

4. Discussion

Significant effect estimates for stage one and stage two variables reinforce the consistent finding that GDL laws do have a strong correlation with the reduction in teenager crashes. Variable definitions are based on the IIHS (2010) rating system of GDL laws. This rating system specifies points for different restrictions, and both the stage one and stage two predictor variables used here rely on the assumption that the IIHS points system is credible. The strength of association in our results suggests that their point system is valid. However, any assertion as to why this is so can only be made post hoc; it is unclear if it is the *best* way of characterizing effective GDL laws. Ultimately, an optimal rating statistic should describe the greatest reductions in crash risk. Our results do not allow us to assert the rating system's adequacy in this respect; however we do observe that the magnitude of effects for stages one and two are relatively balanced for fatal crashes. That is, the effect of a strongly scored stage one is similar to a strongly scored stage two.

For each crash type observed, relative to a lenient stage (*C*) the effect estimate for a moderate stage (*B*) is more muted than that of a strict stage (*A*). By the beginning of 2009 only seven states had implemented what we designate to be a strict stage one and only eighteen had a strict stage two. Furthermore, at present[1] only one state still has a lenient (*C*) stage one or two. While this study has been designed to evaluate the effects of GDL retrospectively, the results indicate that many states can still benefit from reduced teen driver crash rates by moving to a stricter GDL program.

Reduction of injury crashes, as far as they can be attributed to GDL, is almost exclusively associated with a stricter stage one. Even though the intermediate restrictions are less effective at reducing *overall* crashes, they do appear to help reduce the especially *severe* (fatal) ones. The aggregations of night and passenger restrictions together are apparently ineffective at reducing the total number of injury crashes.

That the stage associated with passenger restrictions has less effect on injury crashes is unsurprising. About 20% of fatality crashes with 16-17 year-old drivers have multiple teenage passengers, compared to about 5% of injury crashes. The question remains, though, as to why the relative injury crash rates for 16 and 17-year-olds could increase in the presence of stricter passenger restrictions. A credible explanation for this phenomenon exists. As a result of fewer

[1] As of Summer 2011, North Dakota has no state-wide night or passenger restrictions.

teens being legally allowed to carpool, it is natural to expect many of them to take different vehicles. Teenagers taking different vehicles instead of carpooling can increase the total number of miles driven. As such, reductions in injury crashes with multiple passengers (per Fig.4) may be cancelled out by an overall increase in exposure. Our hypothesis is that an exchange is made between fewer crashes with multiple passengers and more crashes without any passengers. The significance of this effect has been pondered before (Preusser et al., 1998) and tested for fatality crashes (Chen et al., 2001), but as far as we are aware the impact of passenger restrictions on non-passenger *injury* crashes has not previously been examined. In addition, we observe that the permit stage, which does not have direct passenger restrictions, is also closely correlated with reductions in crashes where multiple teenage passengers aged 15-19 are onboard. The reason may be hypothesized that, on average, there exist fewer teenage passengers in supervised vehicles and that the result is related to the delay of intermediate licensure.

For all ages tested, the intermediate stage has a significant reduction in nighttime crashes as the laws intent to do. We also find that for 17-year-olds, a strict stage one has a stronger effect on night crashes than on overall crashes. Again, this result may be influenced by permit restrictions reducing exposure or producing safer drivers. Newly intermediate licensed teenagers *may* also drive less at night.

In most states, 16-year-olds are directly subject to both stage one and stage two restrictions[2]. For this age group, the strictest stage one (delayed permit age, longer permit period, and more supervised driving) has a stronger effect than stage two. We find the opposite is true for 17-year-olds, the strictest stage two has a stronger effect on fatal crashes than stage one. In any case, crash reduction may be influenced by both reduced exposure to unsafe conditions and affording more experience before the young driver takes to the streets unsupervised.

4.1. Limitations

The introduction of young people into the driving system is full of complex and ever changing dynamics. Variance among state effect estimates underline that there exist other significant variables which are not directly accounted for in our model. The observed variance in state effects, which is larger for 16 than for 17-year-olds, could follow from different teenage cultures, road conditions, driver education programs, or levels of enforcement between states. Continuously striving towards a better understanding of the important variables and how they interact with GDL laws is necessary to make the most effective decisions for improving driver safety.

The reasons for GDL's effectiveness are of practical concern. Both stages one and two have the potential to reduce crash risk by either increasing teens' driving aptitude or decreasing their exposure to dangerous situations. Which root cause has the stronger effect cannot be answered by our analysis. The reader may be interested in recent research by Karaca-Mandic and Ridgeway (2010) which suggests that exposure is the stronger driver of GDL effectiveness.

Additional usable data could lead to a greater understanding of the effects of GDL. More trustworthy license statistics would allow for the examination of crashes per licensed driver, leading to further insights as to the effect of GDL on reducing exposure vs. increasing driver safety. Of particular note to this study, injury data is limited to certain states and certain time periods. Although the data collected is geographically disperse, it covers less than one third of all

[2] As of Summer 2011, New Jersey is the only state with a minimum intermediate licensing age of 17 years.

restriction changes occurring within states since 1995. The evaluation of more comprehensive data would increase our confidence in making nationwide inferences. Given the significant costs stemming from injury crashes, efforts to increase the injury data availability would be worthwhile for nationwide studies.

5. Conclusion

The problem of improving teen driver safety is a complex one which may be waged on many fronts. Recently, increasingly prevalent graduated driver licensing laws in the US have corresponded to reductions in crashes for young drivers relative to their older counterparts. Gaining more insight as to "how" and "what" elements of GDL work can only improve our ability to make more intelligent developments in the future. This regression analysis represents an evolution over previous research centered on the "what" question: what is the magnitude of the relationship between GDL and teenage crash risk. From this knowledge, better insights as to the "how" may be gained.

Due to the challenge of isolating the effect of different GDL components, most studies have evaluated categorical variables which describe the state of the overall system. Those who have attempted to evaluate individual GDL restrictions have been disadvantaged by the confounding effects of multiple restrictions becoming stricter at the same time. This increases the standard error of the effect estimates and waters down the insights which can be achieved. In this study we propose a compromise; by grouping restrictions together but categorizing them by stage, the effects of multiple restrictions are pooled together. The advantage is that we obtain a *much* stronger signal as to which aspects of GDL are the most effective.

5.1. Impact on Industry

A strong stage one, regarding stricter permit laws, appears to have had a strong impact on all crash types. Intermediate licensing, taking the form of night or passenger restrictions, have also been significantly associated with a reduction in fatality crashes and nighttime and multiple passenger injury crashes. Results suggest, however, that intermediate restrictions are not effective at reducing the number of injury crashes overall. There is evidence that the number of the crashes with one or less teenage passengers may actually increase, possibly because passenger restrictions encourage teenagers to drive separately putting more vehicles on the road. In other words, the modes of injury crashes have shifted due to these restrictions.

This analysis has also served to validate the findings of other studies. A very strong relationship with GDL is seen by 16-year-olds, with the effect trailing off for older drivers. GDL laws are most associated with reducing injury crashes with multiple teenage passengers, which make up about 5% of the total injury crashes. The results indicate that night restrictions influence injury crash rates, as intermediate GDL restrictions have a stronger effect on crashes which occur between 11pm and 6am than during the rest of the day.

References:

Baker, S.P., Chen, L., Li, G., 2007. Nationwide review of graduated driver licensing. AAA Foundation for Traffic Safety.

Burton, P., Gurrin, L., Sly, P., 1998. Tutorial in Biostatistics: Extending the simple linear regression model to account for correlated responses: an introduction to generalized estimating equations and multi-level mixed modeling. Statist Med 17, 1261-1291.

Campbell, B.J, Stewert, J.R., Reinfurt, D.W., 1991. Change in injuries associated with safety belt laws. Accid Anal & Prev 23 (1), 87-93.

Chen, L., Baker, S.P., Li, G., 2006. Graduated driver licensing programs and fatal crashes of 16-year-old drivers: a national evaluation. Pediatrics 118, 56-62.

Chen, L.H, Braver, E.R, Baker, S.P, Li, G., 2001. Potential benefits of restrictions on the transport of teenage passengers by 16 and 17 year old drivers. Inj Prev 7, 129-134.

Cheung, I., McCartt, A.T., 2011. Declines in fatal crashes of older drivers: Changes in crash risk and survivability. Accid Anal & Prev 43, 666-674.

Dee, T.S., Grabowski, D.C., Morrisey, M.A., 2005. Graduated driver licensing and teen traffic fatalities. J Health Econ 24 (3), 571-589.

Eisenberg, D., 2003. Evaluating the effectiveness of policies related to drunk driving. J Policy Anal Manage 22 (2), 249-274.

Eisenberg, D., 2004. The mixed effects of precipitation on traffic crashes. Accid Anal & Prev 36, 637-647.

Farmer, C.M, Lund, A.K., 2002. Rollover risk of cars and light trucks after accounting for driver and environmental factors. Accid Anal & Prev 34, 163-173.

Fatality Analysis Reporting System, 2010. Fatality estimates, see: http://www-fars.nhtsa.dot.gov.

Gill, J., 2001. Generalized Linear Models: A Unified Approach. Sage Publishing, Thousand Oaks, Ca.

Heron, M., 2010. Deaths: Leading Causes for 2006. NVSS 58 (14).

Highway Loss Data Institute, 2009. Graduated driver licensing laws and insurance collision claim frequencies of teenage drivers

Hilbe, J.M., 2007. Negative Binomial Regression. Cambridge University Press, Cambridge, New York.

Insurance Institute for Highway Safety, 2006. Comment to the Federal Highway Administration concerning motor vehicle registration and licensed driver information; docket no. FHWA-05-22706.

Insurance Institute of Highway Safety. 2009. Effective dates of US graduated driver licensing laws, see: http://www.iihs.org/laws/pdf/gdl_effective_dates.pdf

Insurance Institute of Highway Safety, 2010. Summary of all US graduated driver licensing laws, see: http://www.iihs.org/laws/pdf/us_licensing_systems.pdf.

Karaca-Mandic, P., Ridgeway, G., 2010. Behavioral impact of graduated drivers licensing on teenage driving risk and exposure. J Health Econ 29 (1), 48-61.

Lin, D.Y., Wei, L.J., Ying, Z., 2002. Model-checking techniques based on cumulative residuals. Biometrics 58, 1-12.

Lord, D., Washington, S.P., Ivan, J.N. 2005. Poisson, Poisson-gamma and zero-inflated regression models of motor vehicle crashes: balancing statistical fit and theory. Accid Anal & Prev 37 (1), 35-46.

Mayhew, D.R., 2007. Driver education and graduated licensing in North America: Past, present, and future. J Safety Res 38, 229-235.

McCartt, A.T., Teoh, E.R., Fields, M., Braitman, K.A., Hellinga, L.A., 2010. Graduated licensing laws and fatal crashes of teenage drivers: a national study. Traffic Inj Prev 11 (3), 240-248.

Morrisey, M.A., Grabowski, D.C., Dee, T.S., Campbell, C., 2006. The strength of graduated drivers license programs and fatalities among teen drivers and passengers. Accid Anal & Prev 38 (1), 135-141.

Naumen, R.B., Dellinger, A.M., Zaloshnja, E., Lawrence, B.A., Miller, T.R., 2010. Incidence and total lifetime costs of motor vehicle-related fatal and nonfatal injury by road user type, United States, 2005. Traffic Inj Prev 11, 353-360.

Pressley, J.C., Benedicto, C.B., Trieu, L., Kendig, T., Barlow, B., 2009. Motor vehicle injury, mortality, and hospital charges by strength of graduated drivers licensing laws in 36 states. J Trauma 67 (1), S43-S53.

Preusser, D.F., Ferguson, S.A., Williams, A.F., 1998. The effect of teenage passengers on the fatal crash risk of teenage drivers. Accid Anal & Prev 30 (2), 217-222.

Rhea, B., Austin, R., 2006. How SDS Data Have Been Used, see: http://www.nhtsa.gov/DOT/NHTSA/NCSA/Content/PDF/SDS_usage.pdf

Rice, T.M., Peek-Asa, C., Kraus, J.F., 2003. Nighttime Driving, Passenger Transport, and Injury Crash Rates of Young Drivers. Inj Prev 9, 245-250.

Shope, J.T., 2007. Graduated driver licensing: Review of evaluation results since 2002. J Safety Res 38 (2), 165-175.

Simpson, H.M., 2003. The evolution and effectiveness of graduated licensing. J Safety Res 34 (1), 25-34.

State Data System, 2009. Injury estimates, see: http://www.nhtsa.gov/Data/State+Data+Program+&+CODES/SDS+Overview.

U.S. Census Bureau, 2010. Population estimates, see: http://www.census.gov.ezproxy1.lib.asu.edu/popest/states/states.html.

Vannlar, W., Mayhew, D., Marcoux, K., Wets, G., Brijs, T., Shope, J., 2009. An evaluation of graduated driver licensing programs in North America. Traffic Injury Research Foundation.

Williams, A.F., 2007. Contribution of the components of graduated licensing to crash reductions. J Safety Res 38, 177-184.

Williams, A.F., Shults, R.A., 2010. Graduated driver licensing research, 2007-present: A review and commentary. J Safety Res 41, 77-84.

Appendix A: Injury data used

Table A1
State periods acquired from SDS for injury data. "A" indicates inclusion in all models: overall, night, and passenger; "O" is overall model only; "O, N" is overall and night models only. Certain crash types were excluded from Illinois and Minnesota due to failure to pass the quality checks outline in section 2.1.

State \ Year	95	96	97	98	99	00	01	02	03	04	05	06	07
Arkansas				A	A	A	A	A	A	A	A	A	A
California	A	A	A	A	A	A	A	A	A	A	A	A	
Florida	A	A	A	A	A	A	A	A	A	A	A	A	A
Illinois	A		A	A	A	A	A	O, N	O, N	A	A	A	A
Kansas	A	A	A	A	A	A	A	A	A	A	A	A	
Kentucky			A	A	A	A	A	A	A	A	A	A	A
Maryland	A	A	A	A	A	A	A	A	A	A	A	A	A
Michigan	A	A	A	A	A	A	A	A	A	A	A	A	A
Minnesota	O	O	O	O	O	O		O		O	O	O	O
Missouri	A	A	A	A	A	A	A	A	A	A	A	A	
Nebraska					A	A	A	A	A	A	A	A	
New Mexico	A	A	A	A	A	A	A	A	A	A	A	A	
Ohio	A	A	A	A	A	A	A	A	A	A	A	A	
Pennsylvania	A	A	A	A	A	A	A		A	A	A		
South Carolina			A	A	A	A	A	A	A	A	A	A	
Virginia	A	A	A	A	A	A	A	A	A	A			
Washington								A	A	A	A	A	
Wyoming				A	A	A	A	A	A	A	A	A	A

Appendix B: Incidence rate ratio estimates

The tables below display the incidence rate ratios (IRR) for the model variables.

Each GDL variable is defined per Table 1 where the order of strictness (most to least strict) is *A*, then *B*, then *C*. The IRR is interpreted as the crash risk relative to the reference category; e.g., a 16 year old in a state with stage one = *A* is estimated to be 42% as likely to be involved as a driver in a fatal crash as a 16 year old in a state with stage one = *C*.

Table B1
GDL incidence rate ratios rounded to the nearest percent for (a) fatality (b) and injury crashes. The crash rate for the control group is defined as the number of crashes per 100,000 persons. Quarters refer to the periods of Jan-Mar, Apr-Jun, Jul-Sep, and Oct-Dec.

(a) Fatality

Parameter	Level	Age 16 (95% CI)	Age 17 (95% CI)	Ages 15-17 (95% CI)
Stage 1	A	0.42 (0.36 0.49)	0.80 (0.71 0.91)	0.62 (0.56 0.69)
Stage 1	B	0.83 (0.78 0.88)	0.96 (0.91 1.01)	0.89 (0.85 0.92)
Stage 1	C	1	1	1
Stage 2	A	0.56 (0.51 0.61)	0.70 (0.65 0.75)	0.64 (0.61 0.68)
Stage 2	B	0.83 (0.78 0.88)	0.89 (0.84 0.93)	0.88 (0.84 0.91)
Stage 2	C	1	1	1
rate2544		1.02 (1.02 1.02)	1.02 (1.02 1.02)	1.02 (1.01 1.02)
quarter	1	0.89 (0.84 0.93)	0.86 (0.83 0.90)	0.86 (0.83 0.89)
quarter	2	1.00 (0.96 1.05)	1.03 (0.99 1.07)	1.02 (0.99 1.05)
quarter	3	1.13 (1.08 1.18)	1.08 (1.04 1.13)	1.12 (1.08 1.15)
quarter	4	1	1	1
State covariance		0.138	0.028	0.062

(a) Injury

Parameter	Level	Age 16 (95% CI)	Age 17 (95% CI)	Ages 15-17 (95% CI)
Stage 1	A	0.46 (0.43 0.48)	0.82 (0.79 0.85)	0.65 (0.63 0.68)
Stage 1	B	0.85 (0.83 0.86)	0.92 (0.91 0.93)	0.88 (0.87 0.90)
Stage 1	C	1	1	1
Stage 2	A	1.04 (1.00 1.09)	1.13 (1.09 1.16)	1.02 (0.98 1.05)
Stage 2	B	0.96 (0.93 0.98)	1.02 (1.00 1.03)	0.99 (0.97 1.01)
Stage 2	C	1	1	1
rate2544		1.00 (1.00 1.00)	1.00 (1.00 1.00)	1.00 (1.00 1.00)
quarter	1	0.89 (0.88 0.91)	0.96 (0.95 0.97)	0.92 (0.91 0.93)
quarter	2	1.00 (0.98 1.02)	1.04 (1.02 1.05)	1.02 (1.01 1.04)
quarter	3	1.08 (1.06 1.10)	1.03 (1.02 1.04)	1.06 (1.05 1.08)
quarter	4	1	1	1
State covariance		0.091	0.035	0.054

Table B2

GDL incidence rate ratios rounded to the nearest percent for injury crashes (a) between 11pm and 6am and (b) with more than two passenger between ages 15 and 19. The crash rate for the control group is defined as the number of overall injury crashes per 100,000 persons. Quarters refer to the periods of Jan-Mar, Apr-Jun, Jul-Sep, and Oct-Dec.

(a) Injury night

Parameter	Value	Age 16 (95% CI)	Age 17 (95% CI)	Ages 15-17 (95% CI)
Stage 1	A	0.41 (0.29 0.53)	0.65 (0.56 0.74)	0.57 (0.49 0.65)
Stage 1	B	0.91 (0.86 0.95)	0.93 (0.89 0.97)	0.92 (0.88 0.95)
Stage 1	C	1	1	1
Stage 2	A	0.88 (0.77 0.99)	0.98 (0.89 1.07)	0.91 (0.84 0.98)
Stage 2	B	0.82 (0.77 0.87)	0.91 (0.87 0.95)	0.88 (0.84 0.91)
Stage 2	C	1	1	1
rate2544		1.00 (1.00 1.00)	1.00 (1.00 1.00)	1.00 (1.00 1.00)
quarter	1	0.86 (0.82 0.90)	0.88 (0.85 0.91)	0.87 (0.84 0.90)
quarter	2	1.08 (1.05 1.12)	1.11 (1.08 1.13)	1.09 (1.07 1.12)
quarter	3	1.22 (1.18 1.25)	1.16 (1.14 1.19)	1.19 (1.17 1.22)

quarter	4	1	1	1
State covariance		0.135	0.040	0.059

(b) Injury passenger

Parameter	Level	Age 16 (95% CI)	Age 17 (95% CI)	Ages 15-17 (95% CI)
Stage 1	A	0.31 (0.28 0.35)	0.61 (0.56 0.67)	0.47 (0.43 0.51)
Stage 1	B	0.70 (0.67 0.74)	0.80 (0.77 0.83)	0.76 (0.73 0.78)
Stage 1	C	1	1	1
Stage 2	A	0.82 (0.73 0.93)	0.97 (0.87 1.07)	0.84 (0.77 0.91)
Stage 2	B	0.84 (0.79 0.89)	0.97 (0.92 1.02)	0.91 (0.87 0.95)
Stage 2	C	1	1	1
rate2544		1.01 (1.01 1.01)	1.00 (1.00 1.00)	1.00 (1.00 1.00)
quarter	1	1.00 (0.96 1.05)	1.08 (1.04 1.11)	1.04 (1.01 1.07)
quarter	2	1.06 (1.02 1.11)	1.12 (1.08 1.16)	1.09 (1.05 1.12)
quarter	3	1.07 (1.03 1.11)	1.07 (1.04 1.11)	1.08 (1.05 1.11)
quarter	4	1	1	1
State covariance		0.264	0.260	0.266

***Biographical sketch of all authors**

Joshua Lyon is a Ph.D. student of Industrial Engineering at Arizona State University (ASU). He obtained a B.S.E. in Industrial Engineering in 2009 from ASU. His research interests include optimization in power transmission and statistical analysis and optimization as applied to public health. He is a member of INFORMS, APM, and IIE.

Rong Pan is an associate professor in the School of Computing, Informatics, and Decision Systems Engineering at Arizona State University. He received his doctorate in Industrial Engineering from the Pennsylvania State University in 2002. His research interests include quality and reliability engineering, design of experiments, time series analysis, and statistical learning theory. He is a senior member of ASQ and IIE, and a member of IEEE and SRE.

Jing Li is an assistant professor in Industrial Engineering of Arizona State University. She obtained a M.A in Statistics in 2005 and a Ph.D. in Industrial and Operations Engineering in 2007 from the University of Michigan. Her research interests include quality engineering, industrial statistics, and data mining. She is a member of ASQ, IIE, and INFORMS.