The Annual Cannabis Holiday and Fatal Traffic Crashes

Sam Harper, PhD (McGill University)

Adam Palayew (McGill University)

January 18, 2019

Corresponding author: Sam Harper, PhD. Department of Epidemiology, Biostatistics & Occupational Health, McGill University, Montreal QC H3A 1A2. (sam.harper@mcgill.ca)

Funding: This work was not supported by any specific funding.

Competing interests: None.

Word count, excluding title page, abstract, references, figures and tables: 2351

This article has been accepted for publication in *Injury Prevention* (2019) following peer review, and the Version of Record can be accessed online at <http://dx.doi.org/10.1136/injuryprev-2018-043068>.

Abstract

Background: Cannabis use has been linked to impaired driving and fatal accidents. Prior evidence suggests the potential for population-wide effects of the annual cannabis celebration on April 20th (“4/20”), but evidence to date is limited.

Methods: We used data from the Fatal Analysis Reporting System for the years 1975-2016 to estimate the impact of “4/20” on drivers involved in fatal traffic crashes occurring between 1620h and 2359h in the United States. We compared the effects of 4/20 to those for other major holidays, and evaluated whether the impact of “4/20” had changed in recent years.

Results: Between 1992-2016 “4/20” was associated with an increase in the number of drivers involved in fatal crashes (Incidence rate ratio [IRR]: 1.12, 95% CI 0.97 to 1.28) relative to control days one week before and after, but not when compared with control days one and two weeks before and after (IRR 1.05, 95% CI 0.92 to 1.28) or all days of the year (IRR 0.98, 95% CI 0.88 to 1.10). Across all years we found little evidence to distinguish excess drivers involved in fatal crashes on 4/20 from routine daily variations.

Conclusions: There is little evidence to suggest population-wide effects of the annual cannabis holiday on the number of drivers involved in fatal traffic crashes.

What is already known on this subject

- Cannabis use is associated with moderate excess motor vehicle crash risk.
- There are reported increases in “drugged driving”.
- Prior work has suggested an increase in crashes on April 20, the annual “cannabis holiday”.

What this study adds

- Comparison of cannabis holiday with other known risky days for motor vehicle crashes.
- Evaluation of time trends in the impact of the cannabis holiday on drivers involved in fatal crashes.
- There is little evidence of excess numbers of drivers involved in fatal crashes on April 20 relative to normal daily variations.

Introduction

The potential role that cannabis use and impairment may play in traffic crashes is the subject of limited evidence. Given legislation to legalize recreational cannabis use in some US states and legalization nationwide Canada in 2018, understanding the potential impact of cannabis consumption on traffic crash risks is needed. Traffic crashes already create a substantial burden of death and injury, so reported increases in drugged driving^{1,2} are concerning. Prior studies have suggested that, like alcohol, cannabis use impairs driving in simulated environments.³

Another strand of research has focused on population-level risks, and systematic reviews⁴ have suggested that cannabis use is associated with an approximate doubling of fatal crash risk, though there is important heterogeneity by study design and quality.⁵ On the other hand, two separate studies found that US states that legalized medical cannabis had reduced rates of fatal traffic crashes after passing legislation.^{6,7} These results have been interpreted as supporting the idea that marijuana and alcohol may be substitutes,⁸ particularly for younger people, or perhaps that cannabis users are more likely to consume it at home and are therefore less likely to drive.

However, most studies of cannabis-impaired driving have either been conducted in a laboratory setting or in local field trials, so population level evidence is limited. Recently, however, Staples and Redelmeier reported that the annual “4/20” cannabis celebration that occurs on April 20th each year was associated with a 12% increase in the number of drivers involved in fatal traffic crashes in the United States.⁹ In this paper we aim to further this evidence by comparing the impact of 4/20 to other risky days identified in prior work,^{10–12} as well as evaluating whether the impact of 4/20 has changed over time.

Methods

Data

We used publicly available data from the Fatal Analysis Reporting System,^{13,14} which covers all crashes in which at least 1 person died (within a month of the crash) in the United States. The data cover the years 1975 to 2016, but we primarily focus on the years 1992–2016 since 4/20 is a recent phenomenon. Similar to prior work, we restricted our analysis to drivers, who are in control of the vehicle, rather than all passengers.⁹ Since prior work suggested an association between mass cannabis consumption at 4:20pm and fatal crashes, we also restricted our analysis to the period from 1620h to 2359h.

Statistical analysis

The primary challenge in isolating the effect of any specific day on fatal crashes is noise in daily crash rates, which are affected by large-scale secular trends, economic conditions, safety legislation, road and vehicle engineering, seasonal weather patterns, weekends, and holidays.

Prior studies have used crash days in the same year one week before and after as control days.^{9,12,15} This usefully accounts for day of the week and potentially seasonal effects, but also requires accounting for the matched/clustered nature of the data (days clustered within a given year) to obtain valid standard errors. We used several regression-based approaches to account for extra-Poisson variation in daily fatal crash rates:¹⁶ negative binomial regression, re-scaling standard errors using the Pearson chi-squared statistic, cluster-robust variance (clustered by year), and bootstrapping.

Although using control days one week before and one week after helps to adjust for potential confounding by day of the week and seasonal factors, the choice of two days a week apart is arbitrary. To assess robustness we also estimated the increase in potential fatal crashes on 4/20 relative to control days one and two weeks before and after, as well as to all other days of the year. For comparative purposes, we also conducted the same analyses for a day with established excess risk: US Independence Day (July 4th).¹⁰⁻¹² To investigate heterogeneity in daily drivers involved in fatal crashes we also repeated the same analysis for each day of the year in a separate negative binomial regression, using two control days one week before and after each day. As a sensitivity analysis we also repeated these analyses for the number of fatal crashes rather than the total number of drivers involved in fatal crashes.

Since public celebrations of 4/20 are a recent phenomenon, to see whether we could detect increasing excess numbers of fatal crashes over time that may be attributable to 4/20 we also fit a multi-level Poisson model using all of the FARS data from 1975-2016. We included random effects for each year as well as random slopes for 4/20 and a handful of other notable days: July 4th, the Saturday before Labor Day, the day before Thanksgiving, Christmas Day, and New Year's Day. We fit this model using weakly informative priors¹⁷ and used 4000 posterior samples to generate model predictions of the excess number of drivers involved in fatal crashes on each day, year-by-year. More details on the models are given in the Appendix. We used Stata (version 14), R (version 3.4.3), and Stan software¹⁸ to conduct our analysis. The data and code to replicate our findings are posted in a public repository (<https://osf.io/qnrg6/>).

Ethics approval

This study used publicly-available data and did not require ethics review.

Results

Summary statistics showing average number of drivers involved in fatal crashes by day of the week, month, and year over the entire sample are given in Appendix Tables 1-3. The number of drivers involved in fatal crashes between 1620h and 2359h has decreased from around 66 per day in 1975 to 56 in 2016. Figure 1 shows estimates from our series of regression models for 4/20 and July 4th for the years 1992-2016. The first estimate in the upper panel shows a 12% increase in the number of drivers involved in fatal crashes on April 20 relative to control days one week before and after (IRR=1.12, 95%CI 1.04, 1.19). The additional estimates in

Figure 1: Estimates of the effect of 4/20 (April 20th) and 7/4 (July 4th) on excess number of drivers involved in fatal crashes between 1620h and 2359h using various control periods, 1992-2016.

the upper panel demonstrate that the Poisson model underestimates uncertainty by ignoring extra-Poisson variation in daily crashes. The middle panel of Figure 1 shows that comparing drivers in fatal crashes on 4/20 and July 4th to the number occurring on the same day of the week one and two weeks before and after attenuates the effect for 4/20 but not July 4th, and the bottom panel shows evidence consistent with null effects for 4/20 when using fatal crashes on all other days as the reference group. The number of drivers involved in fatal crashes increases by roughly 40% on July 4th, regardless of which control period was used. Estimates from each regression are provided in Appendix Tables 4 and 5. Appendix Tables 6 and 7 show similar results when the outcome is the number of crashes rather than the number of drivers involved in crashes.

Figure 2 shows estimates from separate negative binomial regressions for each day of the year relative to control days one week before and one week after, pooled over the period 1992-2016. Again, there is some evidence for elevated relative rates on 4/20 (shown in red), but the magnitude of the estimate is consistent with the bulk of daily variations for many other days. In Appendix Table 8 we list 16 days with higher rates and 15 days with lower

Figure 2: Estimates of the effect of each day on the number of drivers involved in fatal crashes between 1620h and 2359h relative to control days +/- one week apart, from separate negative binomial regressions, 1992-2016. Filled circles are point estimates and lines are 95% confidence intervals. Estimates for 4/20 are in red.

rates relative to the population average that are also inconsistent with the null. However, some more notable days stand out as particularly risky (July 4th, New Year's Eve).

Figure 3 shows the average number of excess drivers involved in fatal crashes each year between 1620h and 2359h on 4/20 as well several other notable days (averages across the entire period and 95% posterior intervals are given in Appendix Table 9). There is little evidence of any additional drivers involved in fatal crashes on 4/20 across the entire range of data (average annual excess crashes of -3.5, 95% posterior interval -8.0, 1.4), or in recent years. However, there is a consistent excess every year on July 4th and the days before Labor Day and Thanksgiving, as well as fewer drivers involved in fatal crashes on Christmas and New Year's Day.

Discussion

We found little evidence that the number of drivers involved in fatal traffic crashes are elevated on April 20 relative to any control period. This is not because daily crash rates are too noisy to detect any signal. On the contrary, we find important, systematic, and meaningful variation in the daily number of drivers involved in fatal traffic crashes across the period from 1975 to 2016. We find consistent evidence for increases in the number of drivers involved in fatal crashes on July 4 and the days prior to Labor Day and American Thanksgiving, as well

Figure 3: Model predictions of the excess number of drivers involved in fatal crashes between 1620h and 2359h on selected days, derived from a single model for data from 1975-2016. Each panel highlights a specific day, shaded regions are 95% prediction intervals calculated from 4000 posterior samples.

as systematically fewer drivers involved in fatal crashes on Christmas and New Year’s Day. Similar patterns have also been observed in Canada, with the riskiest periods being weekends before major holidays such as Canada Day, Thanksgiving and Victoria Day.¹⁹

With respect to prior studies, our results are consistent with existing work that has evaluated daily fluctuations in fatal crashes, including elevated risks on July 4, weekends before Labor Day and Christmas, and New Year’s Eve.^{10–12} We were also able to reproduce the previously reported⁹ 12% increase in the national rate of fatal crashes among drivers on 4/20. However, we found that this estimate was sensitive to methods for calculating its variance, as well as to how the control group was constructed, with diminished and largely null associations when crashes on 4/20 are compared to control days 1 and 2 weeks before or after, or to the rest of the days of the year. In contrast, the excess number of drivers in fatal crashes on July 4th was robust to these alternative control periods. We also found little variation in the annual impact of 4/20 over time. If recent celebrations of 4/20 were generating excess fatal crashes we would expect to see a greater excess in recent years.

To generate a reliable increase in the national number of drivers in fatal crashes on any given day requires either the presence of a substantially increased relative risk, population-wide exposure to a moderately increased risk, or some combination of both. In the case of alcohol, there is evidence that increasing impairment increases the relative risk of a fatal crash exponentially,²⁰ as well as demonstrated increases in population-wide exposure to alcohol and drunk driving around major holidays. This likely generates the observed excess risk and is reflected in the higher proportion of fatal crashes involving alcohol on those days.¹⁰

Less is known about how widely 4/20 is celebrated, or about drugged driving more generally. Both American and Canadian reports suggest a similar prevalence (~12%) of past-year cannabis use among adults.^{21,22} Past month use in the US is closer to 8%, and daily use nearer to 2.5%.²³ Estimates of the proportion of the population that may engage in drugged driving are less common. Roadside tests in Canada indicate 4-6% of drivers reported having used cannabis, and roughly 20% of the 12% of past-year cannabis users admitting to driving within 2 hours of cannabis use.²⁴ Given the localized celebration of 4/20, the small fraction of individuals that report driving after using cannabis, as well as the moderate relative risk of a fatal crash for cannabis use,⁵ it seems unlikely that enough of the population is exposed to drugged driving on 4/20 to generate substantial population-wide excesses in fatal crashes.

Our results also have implications for other analyses of study designs focused on transient risks that have used the “double-pair control” design.¹⁵ Conceptually, it seems reasonable to consider a set of control days one week apart from the index date since this provides good control for any day of the week or (less clearly) seasonal factors. However, because the data generating processes resulting in daily counts are noisy and tend to reflect both strong stochastic and deterministic components, failing to model or incorporate this variation leads to degraded inference.

Limitations

Some limitations of this analysis should be noted. We do not measure changes in cannabis consumption on 4/20 or any other days, so clearly any estimates of the impact of 4/20 on fatal traffic crashes among drivers must be considered a misclassified estimate of the population impact of cannabis consumption on driving-related deaths. We compared the impact on 4/20 to that on July 4 and other days but those choices are subjective and it would be worthwhile to conduct more detailed investigations of systematic variations in daily fatal traffic crashes. In particular, for comparison with prior research we restricted our analysis to focus on the period from 1600h to 1159h to capture the potential for “mass consumption” of cannabis to have an impact on 4/20. Because the period with the highest crash risk often overlaps 2 days (from 9pm to 3am), using alternative hours could produce different risky days (e.g., New Year’s Day was protective in our analysis but would likely be risky if the hours between midnight and 3am were included).

Conclusion

Population-based evidence on both the prevalence and the relative risk of marijuana-impaired driving is weak, so increased efforts to understand the potential impact that decriminalizing or legalizing marijuana may have on driving risks would be welcome. We conclude that there is limited evidence of any population-wide effects of the annual cannabis holiday on fatal traffic crashes. The policy landscape for addressing marijuana-impaired driving is rapidly changing, and efforts to combat drugged driving are likely to require new initiatives. Given the considerable resources already devoted to heightened enforcement campaigns to reduce the impact of impaired driving,²⁵ there does not seem to be sufficient evidence for including 4/20 among “risky” days.

References

1. Wilson FA, Stimpson JP, Pagán JA. Fatal crashes from drivers testing positive for drugs in the u.s., 1993-2010. *Public Health Rep.* 129(4):342–50.
2. Berning A, Compton R, Wochinger K. Results of the 2013-2014 national roadside survey of alcohol and drug use by drivers. *Journal of Drug Addiction, Education, and Eradication.* 2015;11(1):47.
3. Hartman RL, Huestis MA. Cannabis effects on driving skills. *Clin Chem.* 2013 Mar;59(3):478–92.
4. Asbridge M, Hayden JA, Cartwright JL. Acute cannabis consumption and motor vehicle collision risk: Systematic review of observational studies and meta-analysis. *BMJ.* 2012 Feb;344:e536.
5. Rogeberg O, Elvik R. The effects of cannabis intoxication on motor vehicle collision revisited and revised. *Addiction.* 2016 Aug;111(8):1348–59.
6. Anderson DM, Hansen B, Rees DI. Medical marijuana laws, traffic fatalities, and alcohol consumption. *The Journal of Law and Economics.* 2013;56(2):333–69.
7. Santaella-Tenorio J, Mauro CM, Wall MM, Kim JH, Cerdá M, Keyes KM, et al. US traffic fatalities, 1985-2014, and their relationship to medical marijuana laws. *Am J Public Health.* 2017 Feb;107(2):336–42.
8. Subbaraman MS. Substitution and complementarity of alcohol and cannabis: A review of the literature. *Subst Use Misuse.* 2016 Sep;51(11):1399–414.
9. Staples JA, Redelmeier DA. The April 20 cannabis celebration and fatal traffic crashes in the United States. *JAMA Intern Med.* 2018 Feb;178(4):569–72.
10. Farmer CM, Williams AF. Temporal factors in motor vehicle crash deaths. *Inj Prev.* 2005 Feb;11(1):18–23.
11. Liu C, Chen C-L, Utter D. Trend and pattern analysis of highway crash fatality by month and day. National Highway Traffic Safety Administration; 2005. Report No.: DOT HS 809 855.
12. Weast R. Temporal factors in motor-vehicle crash deaths: Ten years later. *J Safety Res.* 2018 Jun;65:125–31.
13. National Bureau of Economic Research. NHTSA’s fatality analysis reporting system (FARS) data - 1975-2015 [Internet]. Available from: <http://www.nber.org/data/fars.html>
14. National Highway Transportation Safety Association. Fatality Analysis Reporting System (FARS) [Internet]. 2018. Available from: <ftp://ftp.nhtsa.dot.gov/fars/>
15. Redelmeier DA, Tibshirani RJ. Methods for analyzing matched designs with double controls: Excess risk is easily estimated and misinterpreted when evaluating traffic deaths. *J Clin Epidemiol.* 2018 Jun;98:117–22.

16. Hilbe JM. Negative binomial regression. 2nd ed. Cambridge, UK: Cambridge University Press; 2011.
17. McElreath R. Statistical rethinking: A Bayesian course with examples in R and Stan. Vol. 122. Boca Raton, FL: CRC Press; 2016.
18. Stan Development Team. RStan: The R interface to Stan. R package version 2.17.3. [Internet]. 2018. Available from: <http://mc-stan.org>
19. Perreault S. Impaired driving in Canada, 2011 [Internet]. Juristat—Component of Statistics Canada Catalogue no. 85-002-x. Statistics Canada; 2013. Available from: <https://www150.statcan.gc.ca/n1/pub/85-002-x/2013001/article/11739-eng.htm>
20. Zador PL, Krawchuk SA, Voas RB. Alcohol-related relative risk of driver fatalities and driver involvement in fatal crashes in relation to driver age and gender: An update using 1996 data. *J Stud Alcohol*. 2000 May;61(3):387–95.
21. Rotermann M, Macdonald R. Analysis of trends in the prevalence of cannabis use in canada, 1985 to 2015. *Health Rep*. 2018 Feb;29(2):10–20.
22. Hasin DS. US epidemiology of cannabis use and associated problems. *Neuropsychopharmacology*. 2018 Jan;43(1):195–212.
23. Azofeifa A, Mattson ME, Schauer G, McAfee T, Grant A, Lyerla R. National estimates of marijuana use and related indicators - national survey on drug use and health, united states, 2002-2014. *MMWR Surveill Summ*. 2016 Sep;65(11):1–28.
24. Capler R, Bilsker D, Van Pelt K, MacPherson D. Cannabis use and driving: Evidence review [Internet]. Canadian Drug Policy Coalition (CDPC) Simon Fraser University; Available from: https://drugpolicy.ca/wp-content/uploads/2017/02/CDPC_Cannabis-and-Driving_Evidence-Review_FINALV2_March27-2017.pdf
25. Solomon MG, Hedlund JH, Chaffe RH, Cosgrove LA, Group PR, others. The 2006 national labor day impaired driving enforcement crackdown: Drunk driving. Over the limit. Under arrest. United States. National Highway Traffic Safety Administration; 2008.

Appendix: The Annual Cannabis Holiday and Fatal Traffic Crashes

Sam Harper (McGill University)
Adam Palayew (McGill University)

January 18, 2019

Corresponding author: Sam Harper, PhD. Department of Epidemiology, Biostatistics & Occupational Health, McGill University, Montreal QC H3A 1A2. (sam.harper@mcgill.ca)

Funding: This work was not supported by any specific funding.

Competing interests: None.

This article has been accepted for publication in *Injury Prevention* (2019) following peer review, and the Version of Record can be accessed online at <http://dx.doi.org/10.1136/injuryprev-2018-043068>.

Models

Negative binomial model for specific days

Given that the daily number of drivers involved in fatal crashes exhibit overdispersion (i.e., the conditional variance exceeds the mean), some mechanism to account for the excess variance is required to generate valid standard errors.¹ We generally obtained similar estimates of precision using either negative binomial regression, rescaling standard errors using the Pearson dispersion statistic, Poission regression with clustered standard errors (clustered by year), and bootstrapping. Our basic negative binomial regression model assumes that the number of drivers involved in fatal crashes on any given day i follow a Poisson distribution, $y_i \sim \text{Poisson}(\mu_i)$:

$$\ln(\mu_i) = \alpha + \beta \times 420_i$$

where α represents the intercept, and β gives the expected difference in the log number of drivers involved in fatal crashes on 4/20 relative to other days. For the first set of models the control days were restricted to be the same day of the week for the week before and week after 4/20 (April 13th and April 27th). These models include 75 observations (1 “exposed” day + 2 control days +/- 1 week apart x 25 years between 1992-2016). We then expanded the control group to include 2 additional control days +/- 2 weeks apart for the period surrounding 4/20 (1 “exposed” day + 4 control days x 25 years, n=125), and finally we used the entire rest of the year as the control period (n=9132, including leap days).

Multilevel model for time-varying daily effects

For the analysis that allowed the impact of 4/20 and other days to vary year-by-year, we used the following model:

$$\begin{aligned}
 Y_i &\sim \text{Poisson}(\lambda) && \text{[likelihood]} \\
 \ln(\lambda_i) &= \alpha_{year[i]} + \beta_{year[i]} * e420 + \gamma_{year[i]} * e704 + \delta_{year[i]} * e900 \\
 &\quad + \epsilon_{year[i]} * e1126 + \zeta_{year[i]} * e1225 + \eta_{year[i]} * e0101 && \text{[linear model for log rate]} \\
 \begin{bmatrix} \alpha_{year} \\ \beta_{year} \\ \vdots \\ \eta_{year} \end{bmatrix} &\sim \text{MVNormal} \left(\begin{pmatrix} \alpha \\ \beta \\ \vdots \\ \eta \end{pmatrix}, \mathbf{S} \right) && \text{[joint prior for varying effects]} \\
 \mathbf{S} &= \begin{pmatrix} \sigma_\alpha & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \sigma_\eta \end{pmatrix} \mathbf{R} \begin{pmatrix} \sigma_\alpha & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \sigma_\eta \end{pmatrix} && \text{[construct covariance matrix]} \\
 \sigma &\sim \text{Normal}(0, 2) && \text{[prior for average year]} \\
 (\beta, \gamma, \delta, \epsilon, \zeta, \eta) &\sim \text{Normal}(0, 2) && \text{[prior for average slope each day]} \\
 (\sigma_\alpha, \sigma_\beta, \dots, \sigma_\eta) &\sim \text{HalfCauchy}(0, 1) && \text{[prior variance for each coefficient]} \\
 \mathbf{R} &\sim \text{LKJcorr}(2) && \text{[prior for correlation matrix]}
 \end{aligned}$$

The likelihood describes the negative binomial model, and the linear portion of the model defines an average intercept that is allowed to vary by year ($\alpha_{year[i]}$), as well as randomly varying slope paramters for several specific days: 4/20 ($e420$), July 4th ($e704$), the Saturday before Labor Day ($e900$), the day before Thanksgiving ($e1126$), Christmas Day ($e1225$), and New Year’s Day ($e0101$). The next line describes the population of varying effects and says that each year has a separate intercept α_{year} and slopes associated with each of the days included in the model (e.g., β_{year}) that have a prior distribution that is multivariate Gaussian with

means $\alpha, \beta, \dots, \theta$ and covariance matrix $\mathbf{S}$, which is comprised of standard deviations for each parameter and correlation matrix $\mathbf{R}$.

The prior for the correlation matrix (*LKJcorr*) is set at 2, which is skeptical of extreme correlations (near -1 or 1) between parameters (*LKJcorr*(1) is a flat prior, i.e., all correlations are equally plausible). We use weakly regularizing priors on the coefficients for each of the slopes, which impose little discipline on the parameters. For example, a prior $\beta \sim N(0, 2)$ would assign low probabilities to absolute values of coefficients on the log(drivers) scale of greater than 4, which still allow for effect sizes rarely seen in most of the literature. We fit the models using the **rethinking** package in *R* by Richard McElreath.²

Results

Table 1. Summary statistics for drivers involved in fatal crashes between 1620h and 1159h by day of week.

	Mean	Std.Dev	Min	Median	Max
Monday	49.77	12.79	13.00	49.00	126.00
Tuesday	49.55	12.57	16.00	49.00	121.00
Wednesday	52.32	14.21	15.00	51.00	158.00
Thursday	56.34	14.32	18.00	55.00	183.00
Friday	80.20	21.33	24.00	78.00	240.00
Saturday	77.18	19.80	23.00	76.00	149.00
Sunday	58.31	15.89	18.00	57.00	128.00

Table 2. Summary statistics for drivers involved in fatal crashes between 1620h and 1159h by month.

	Mean	Std.Dev	Min	Median	Max
January	48.37	15.15	13.00	46.00	114.00
Feburary	50.83	16.22	15.00	48.00	126.00
March	54.30	17.33	17.00	51.00	118.00
April	58.07	18.33	17.00	55.00	126.00
May	62.42	19.93	22.00	59.00	136.00
June	64.82	19.25	25.00	62.00	144.00
July	65.91	19.41	25.00	63.00	149.00
August	66.36	20.19	26.00	63.00	147.00
September	64.52	21.27	22.00	61.00	149.00
October	64.74	21.22	23.00	60.00	144.00
November	63.57	19.24	26.00	60.00	150.00
December	61.81	21.43	23.00	58.00	240.00

Table 3. Summary statistics for drivers involved in fatal crashes between 1620h and 1159h by year.

	Mean	Std.Dev	Min	Median	Max
1975	65.71	22.15	18.00	62.00	140.00
1976	65.87	23.91	21.00	61.00	183.00
1977	70.47	25.53	18.00	66.00	178.00
1978	74.70	26.64	27.00	70.00	191.00
1979	75.85	26.02	13.00	70.00	240.00
1980	73.55	25.50	29.00	66.50	158.00
1981	73.64	24.07	23.00	68.00	149.00
1982	62.23	19.77	27.00	58.00	135.00
1983	60.30	19.22	21.00	56.00	123.00
1984	63.10	20.66	20.00	61.00	151.00
1985	65.58	20.12	16.00	63.00	120.00
1986	69.01	21.58	25.00	65.00	143.00
1987	69.72	21.61	28.00	67.00	130.00
1988	69.36	21.94	16.00	65.00	141.00
1989	67.07	19.80	25.00	63.00	132.00
1990	64.76	20.55	27.00	60.00	123.00
1991	59.51	18.73	22.00	55.00	112.00
1992	57.11	18.14	18.00	54.00	127.00
1993	57.10	16.89	18.00	53.00	111.00
1994	57.34	16.67	27.00	55.00	115.00
1995	60.18	17.76	18.00	58.00	124.00
1996	60.60	16.28	23.00	59.00	114.00
1997	59.60	15.76	25.00	58.00	105.00
1998	58.70	15.99	24.00	57.00	126.00
1999	58.12	15.50	24.00	57.00	111.00
2000	58.51	15.16	27.00	57.00	102.00
2001	59.05	16.11	24.00	57.00	110.00
2002	59.99	15.62	26.00	58.00	104.00
2003	60.48	17.07	18.00	58.00	110.00
2004	60.40	16.21	27.00	58.50	107.00
2005	61.42	16.32	28.00	60.00	113.00
2006	59.97	15.26	28.00	58.00	103.00
2007	58.24	16.26	22.00	56.00	138.00
2008	52.62	14.34	25.00	52.00	105.00
2009	48.19	13.55	19.00	47.00	95.00
2010	46.24	13.01	18.00	44.00	84.00
2011	45.73	13.28	15.00	45.00	96.00
2012	48.19	13.31	22.00	46.00	87.00
2013	47.02	12.74	18.00	46.00	93.00
2014	48.22	13.97	19.00	47.00	98.00
2015	52.55	13.97	22.00	51.00	90.00
2016	56.00	13.64	26.00	55.00	95.00

Results for the number of drivers involved in fatal crashes

Table 4: Rate ratio of number of drivers involved in fatal crashes on 4/20

Model	Control period								
	+/- 7 days			+/- 14 days			All other days		
	IRR	95%CI		IRR	95%CI		IRR	95%CI	
Poisson regression	1.12	1.05	1.19	1.05	0.99	1.19	0.98	0.93	1.03
Negative binomial regression	1.12	0.97	1.28	1.05	0.92	1.28	0.98	0.88	1.10
Poisson regression (scaled SEs)	1.12	0.97	1.28	1.05	0.92	1.28	0.98	0.87	1.10
Poisson regression (clustered SEs)	1.12	0.99	1.27	1.05	0.94	1.27	0.98	0.88	1.09
Poisson regression (bootstrapped SEs)	1.12	0.97	1.29	1.05	0.93	1.29	0.98	0.88	1.09

Table 5: Rate ratio of number of drivers involved in fatal crashes on July 4th

Model	Control period								
	+/- 7 days			+/- 14 days			all other days		
	IRR	95%CI		IRR	95%CI		IRR	95%CI	
Poisson regression	1.39	1.32	1.47	1.34	1.28	1.47	1.46	1.40	1.53
Negative binomial regression	1.39	1.24	1.56	1.34	1.21	1.56	1.46	1.31	1.63
Poisson regression (scaled SEs)	1.39	1.25	1.55	1.34	1.22	1.55	1.46	1.33	1.61
Poisson regression (clustered SEs)	1.39	1.29	1.50	1.34	1.26	1.50	1.46	1.38	1.56
Poisson regression (bootstrapped SEs)	1.39	1.25	1.55	1.34	1.24	1.55	1.46	1.38	1.56

Results for the number of fatal crashes (sensitivity analysis)

Table 6: Rate ratio for number of fatal crashes among drivers on 4/20

Model	Control period								
	+/- 7 days			+/- 14 days			All other days		
	IRR	95%CI		IRR	95%CI		IRR	95%CI	
Poisson regression	1.11	1.02	1.20	1.05	0.97	1.20	0.98	0.92	1.05
Negative binomial regression	1.11	0.98	1.26	1.05	0.93	1.26	0.98	0.88	1.10
Poisson regression (scaled SEs)	1.11	0.97	1.27	1.05	0.93	1.27	0.98	0.88	1.10
Poisson regression (clustered SEs)	1.11	1.00	1.23	1.05	0.96	1.23	0.98	0.90	1.08
Poisson regression (bootstrapped SEs)	1.11	0.99	1.25	1.05	0.94	1.25	0.98	0.89	1.09

Table 7: Rate ratio for number of fatal crashes among drivers on July 4th

Model	Control period								
	+/- 7 days			+/- 14 days			all other days		
	IRR	95%CI		IRR	95%CI		IRR	95%CI	
Poisson regression	1.47	1.38	1.58	1.42	1.33	1.58	1.52	1.45	1.61
Negative binomial regression	1.47	1.32	1.64	1.42	1.28	1.64	1.52	1.38	1.69
Poisson regression (scaled SEs)	1.47	1.32	1.64	1.42	1.29	1.64	1.52	1.39	1.67
Poisson regression (clustered SEs)	1.47	1.36	1.59	1.42	1.32	1.59	1.52	1.44	1.61
Poisson regression (bootstrapped SEs)	1.47	1.35	1.61	1.42	1.32	1.61	1.52	1.44	1.61

Table 8: Selected days with higher or lower relative rates of drivers involved in fatal accidents

More risky days				Less risky days			
Day	IRR	95%CI LL	95%CI UL	Day	IRR	95%CI LL	95%CI UL
May 25	1.13	1.01	1.27	January 1	0.87	0.77	0.98
July 3	1.17	1.04	1.32	January 2	0.83	0.73	0.96
July 4	1.39	1.24	1.56	January 7	0.80	0.70	0.91
July 14	1.15	1.02	1.29	March 25	0.87	0.76	0.99
September 1	1.20	1.04	1.38	April 26	0.87	0.77	0.99
September 2	1.19	1.04	1.35	June 27	0.80	0.70	0.90
September 4	1.14	1.01	1.30	July 5	0.88	0.78	0.99
October 31	1.16	1.02	1.32	July 7	0.86	0.75	0.98
November 21	1.20	1.07	1.35	July 11	0.83	0.74	0.93
November 23	1.16	1.04	1.30	September 8	0.83	0.72	0.96
December 6	1.15	1.01	1.32	November 28	0.84	0.73	0.96
December 18	1.16	1.02	1.31	November 29	0.87	0.77	1.00
December 20	1.19	1.02	1.39	December 16	0.86	0.74	0.99
December 21	1.20	1.07	1.35	December 27	0.85	0.73	0.98
December 23	1.26	1.10	1.45	December 28	0.87	0.77	0.98
December 31	1.17	1.03	1.34				

Table 9: Estimates of the average number of annual excess drivers involved in fatal crashes for specific days from multi-level Poisson model. Statistics based on 4000 samples of the posterior distribution.

Day	Mean	SD	95LL	95UL
4/20	-3.44	2.38	-7.93	1.37
July 4th	33.61	2.78	28.15	39.18
Saturday before Labor Day	26.59	2.19	22.52	31.06
Day before Thanksgiving	23.22	1.96	19.42	27.18
Christmas Day	-12.36	1.86	-15.90	-8.58
New Year's Day	-21.64	1.83	-25.14	-18.01

Plots

Traceplots from the Markov chains from the multi-level model for daily drivers involved in fatal crashes on specific days (model `m1`). Each plot shows the 4000 post-warmup samples (using the `mcmc_trace` function of the `bayesplot` package) that were used for inference. For brevity we only show traceplots for the random effects for each of the specific days (42 plots for each year on each page, 1 page for each specific day with a random coefficient).

References

1. Hilbe JM. Negative binomial regression. 2nd ed. Cambridge, UK: Cambridge University Press; 2011.
2. McElreath R. Statistical rethinking: A Bayesian course with examples in R and Stan. Vol. 122. Boca Raton, FL: CRC Press; 2016.

Figure 1: Traceplots for random effects by year for 4/20

Figure 2: Traceplots for model with random effects for July 4th

Figure 3: Traceplots for model with random effects for Labor Day

Figure 4: Traceplots for model with random effects for Thanksgiving

Figure 5: Traceplots for model with random effects for Christmas

Figure 6: Traceplots for model with random effects for New Year's Eve