Should we strengthen seat belt laws?

A (semi-) Bayesian analysis

Sam Harper

McGill | Department of Epidemiology, Biostatistics and Occupational Health

2021-03-09

Outline

Orientation

Background

Methods

Results

Outline

Orientation

Background

Methods

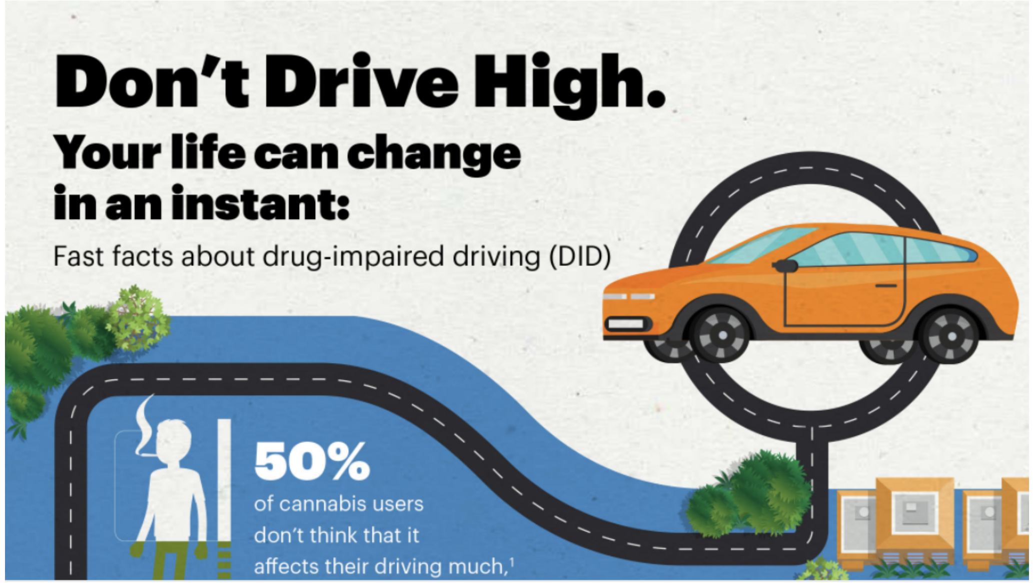
Results

Orientation

Bayesian framework for inference

Thought experiment

Suppose you are interested in the effect of exposure to cannabis on the risk of a motor vehicle



See Hamra et al. (2013) for intuitive description of Bayesian inference, which I modified to this example. Image credit: Government of Canada

Before doing your study: what do you know?

Describe your prior beliefs

- Many studies have found excess risks for those consuming cannabis $(X = 1)$ compared to those who don't $(X = 0)$.
- You think a harmful effect is likely and a protective effect

Quantify your prior beliefs

- Translate your beliefs into probability statements.
- E.g., you're 95% sure that exposure to cannabis increases the risk of a crash between 0% and 40%

*Prior mean = average of limits = (0 + .4)/2 = 0.20; prior SD = width of interval divided by width of SD units = (|0 - 0.4| / (2x1.96)) = 0.1.

Visualizing prior information

- Summary of our prior knowledge about the risk difference.

Conduct your study: what did you find?

Describe your evidence

- Small pilot RCT using a driving simulator (n=100).
- You find 10/50 crashed when exposed to cannabis $(X = 1)$ and 5/50 crashed among those unexposed $(X = 0)$

Image credit: Toronto Star 2015-11-26.

Adding the likelihood based on observed data

- Plot of the likelihood based on observed data.
- Note the peak of the blue curve is the maximum likelihood

Inference: update your beliefs

- Can combine these two pieces of evidence, since both come from normal distributions.
- Inverse-variance weighting to reflect the amount of *information* given by the prior and likelihood.
- Produces a summary estimate.

Note: weights are 1/0.01 = 100 and 1/0.0051 = 196, respectively. So posterior mean = [100(0.20) + 196(0.10)]/(100 + 196) = 0.13. Variance = 1/(100 + 196) = 0.0034.

Combined plot of the posterior distribution

- Product of the prior probability and the information in the likelihood.
- Revised estimate of the probability

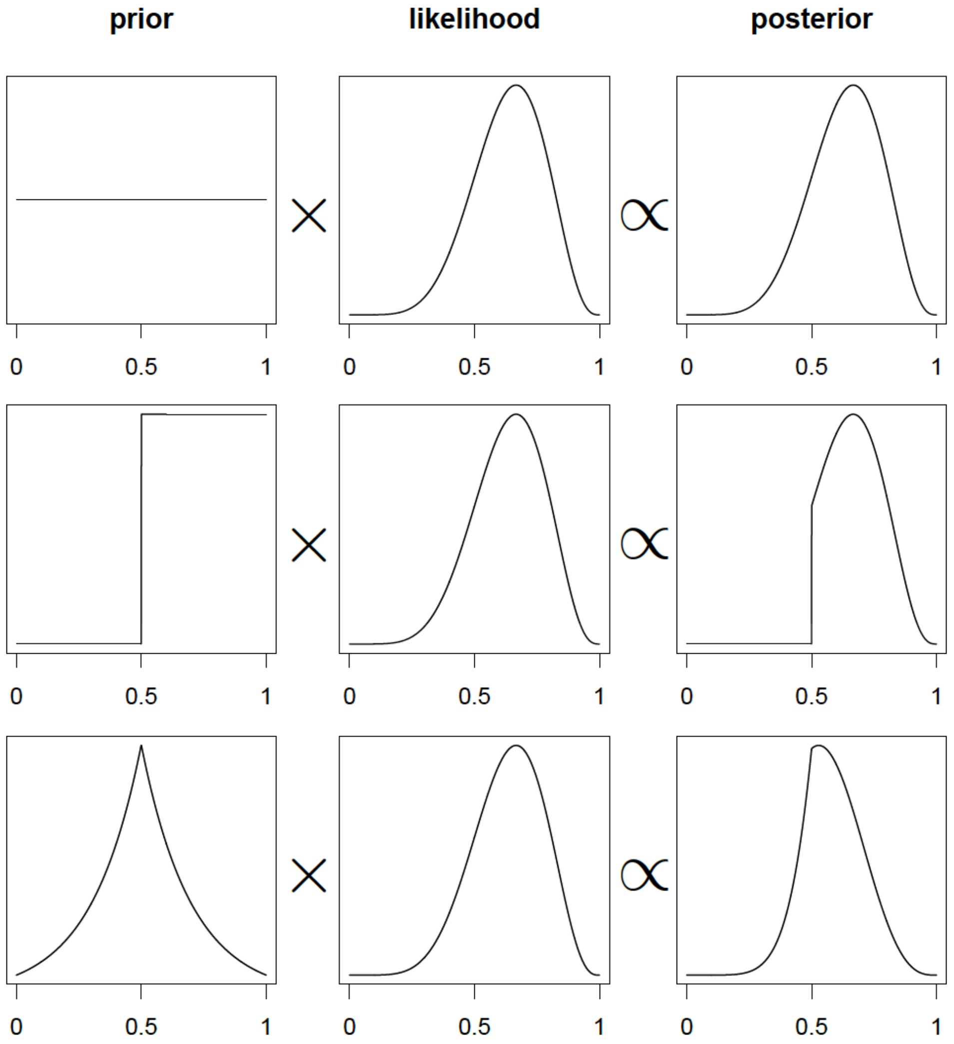
- Normal prior not necessary

- In general, the posterior $P(H|data)$ is equal to

$$\frac{P(data|H) \times P(H)}{P(data)}$$

- The posterior is proportional to the

McElreath (2020)



Interpretation

Frequentist inference

- No probability statements about hypothesis of interest.
- p-value gives evidence of compatibility of observed data with the null, given the model and assuming the null hypothesis is true.
- 95%CI: If we took a large number of repeated samples

Bayesian inference:

- Direct statements about probability of hypothesis from the posterior distribution.
- Can say 95% probability that parameter is within the limits of the interval.
- Is a learning process and is updated with new information.
- But priors require subjectivity...

Summary

Honest uncertainty

Quantitatively incorporating prior knowledge.

Priors need defending

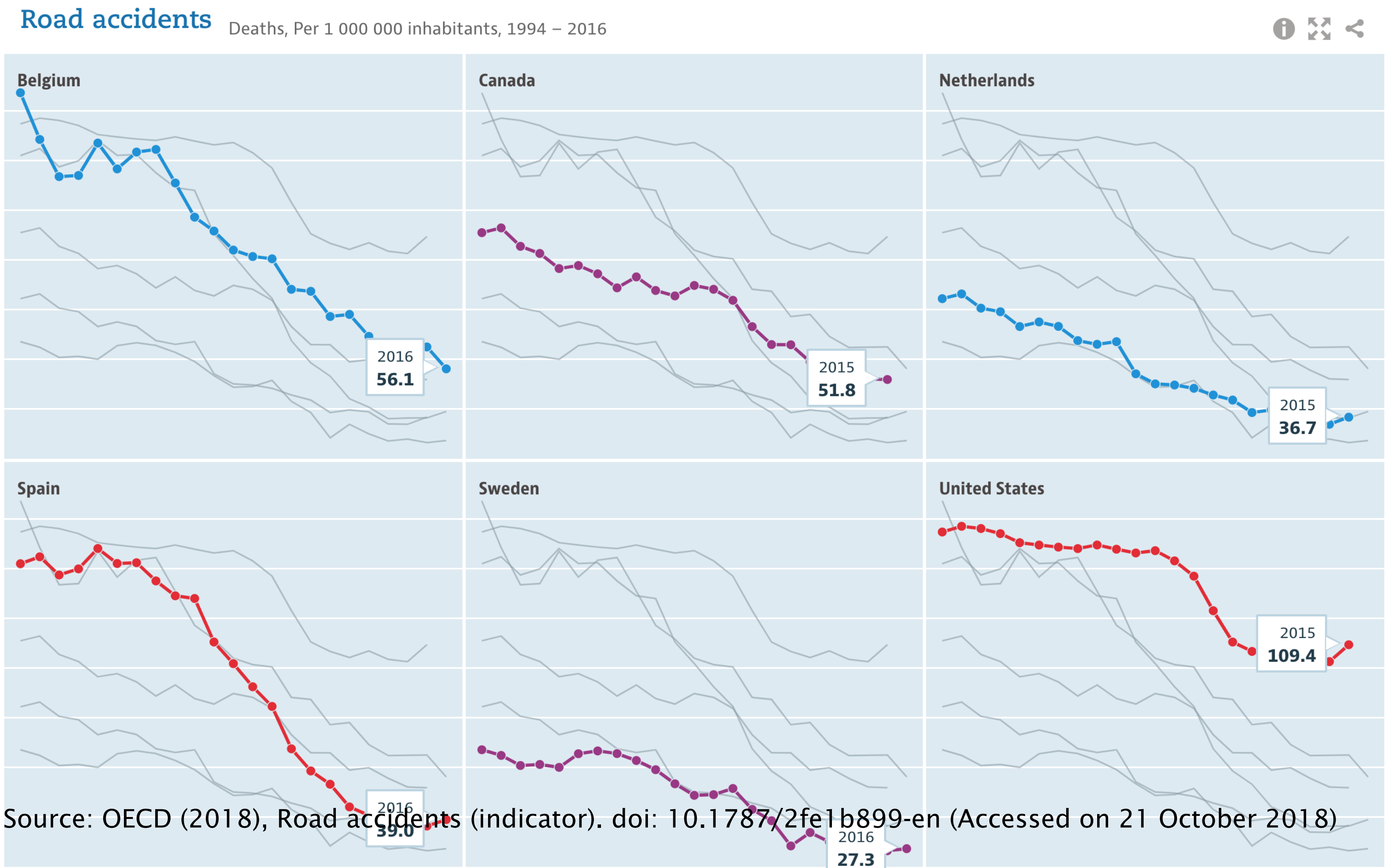
Based on substance. Non-informative are rarely defensible.

Outline

Background

Context of seat belts and motor vehicle crash deaths

- Important declines in motor vehicle crash deaths.
- Many factors contributing to declines.
- Yet, US lags behind



Source: OECD (2018), Road accidents (indicator). doi: 10.1787/2fe1b899-en (Accessed on 21 October 2018)

Flavors of seat belt legislation

Primary enforcement

Allows law enforcement officers to pull over drivers and ticket them if they are not wearing their seatbelts.

Secondary enforcement

Drivers pulled over for a separate violation (speeding, headlight out, etc.) can be ticketed if they

More States Adopt 'Click It or Ticket' Laws; Do They Work?

April 28, 2017 | By Jenni Bergal

SHARE

The Insurance Institute for Highway Safety, a nonprofit research group funded by auto insurance companies, takes a different view of primary enforcement laws. It says they do reduce fatalities and deter motorists from bad behavior.

Chuck Farmer, a vice president at the institute, estimates that moving to primary enforcement reduces traffic deaths by about 7 percent. He said police in primary enforcement states also run successful traffic stop operations such as "Click It or Ticket," in which they target motorists not wearing their seat belts.

"Secondary enforcement laws lack teeth. Cops have to find another reason to pull you over," said Kara Macek, spokeswoman for the Governors Highway Safety Administration. "Primary laws are really the only effective ones."

Too Much Government Intrusion?

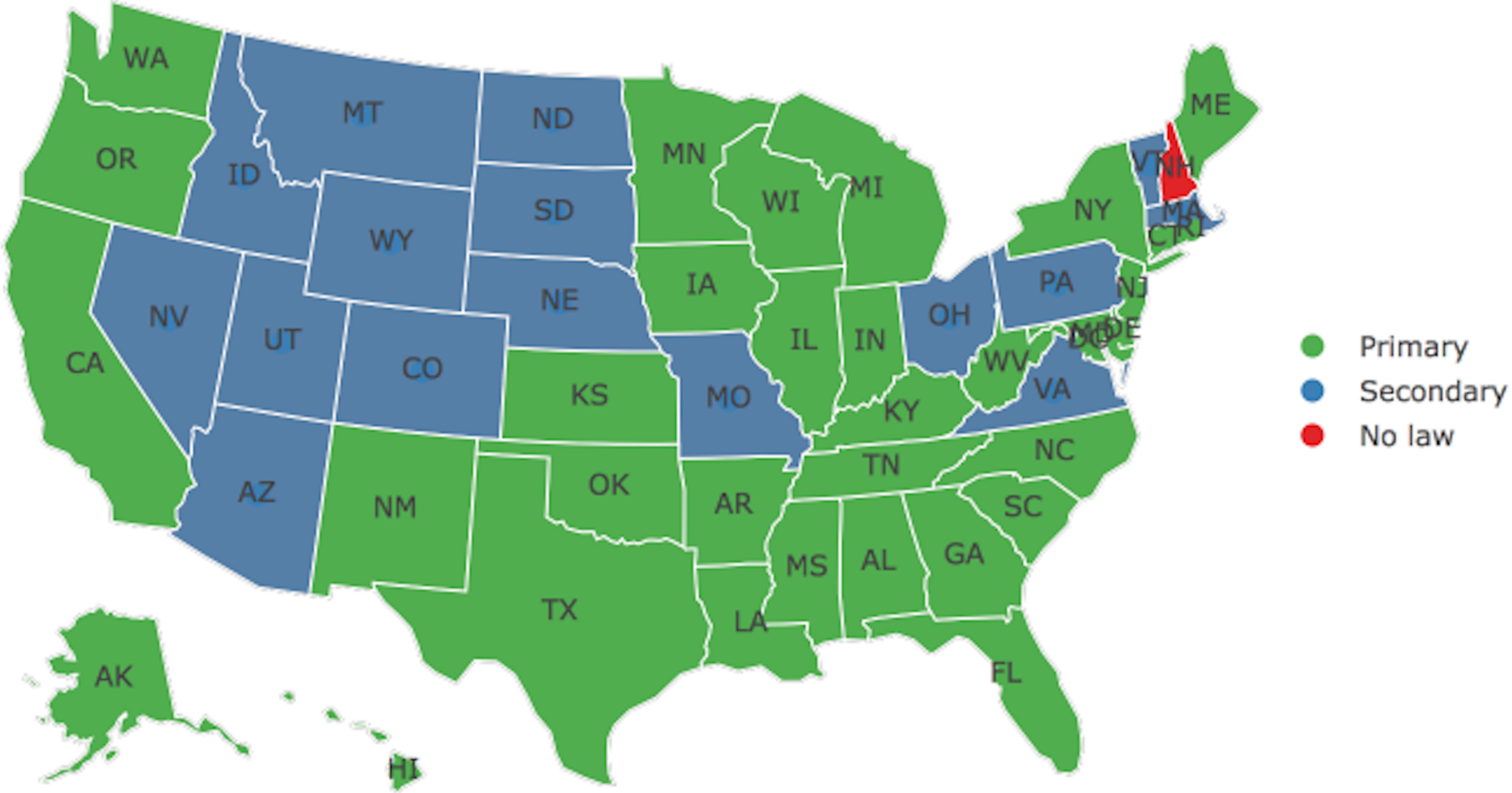
Foes of the tougher laws argue they tread on individual liberties and can be a tool for racial profiling.

Pew Trusts, Stateline, April 28, 2017. http://bit.ly/2pFTFPp

> The [Centers] for Disease Control and Prevention's systematic review of 13 high-quality studies ([Shults et al., 2004]) found that primary laws increase belt use by about 14 percentage

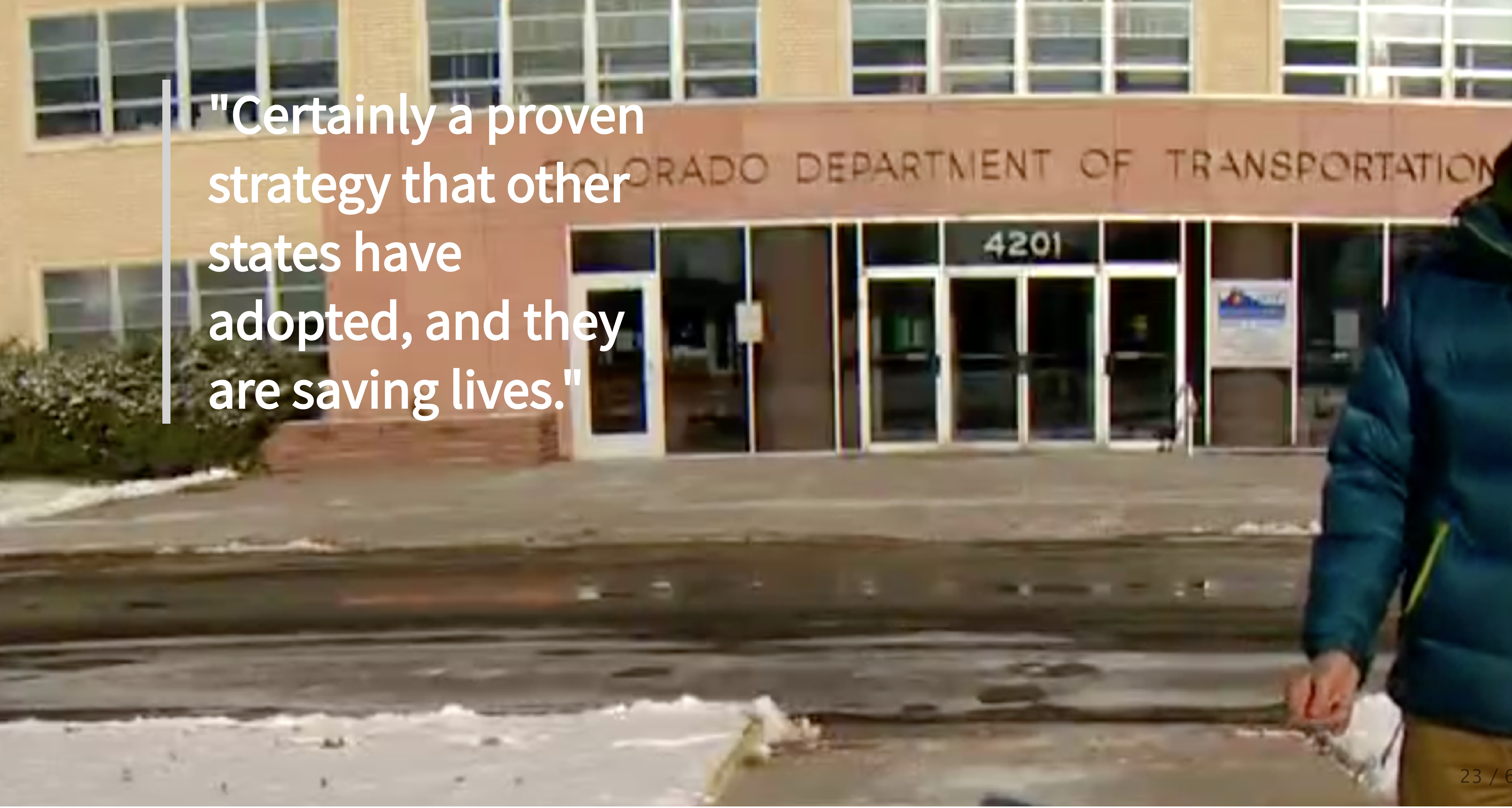
Source: CDC Retreived 2021-04-09.

Current laws in existence in 2017



Source: Insurance Institute for Highway Safety

Should Colorado upgrade to primary?



"Certainly a proven strategy that other states have adopted, and they are saving lives."

Hang on...

> CDOT should be concerned with building and maintaining our roads. Not joining the nanny crowd trying to protect us from our own bad decisions.
>
> Darwinism exists for a reason. If people want to meet their windshields at 60 miles and hour and remove themselves from

http://www.9news.com, Dec 8, 2016.

Potential unintended consequences

ACLU

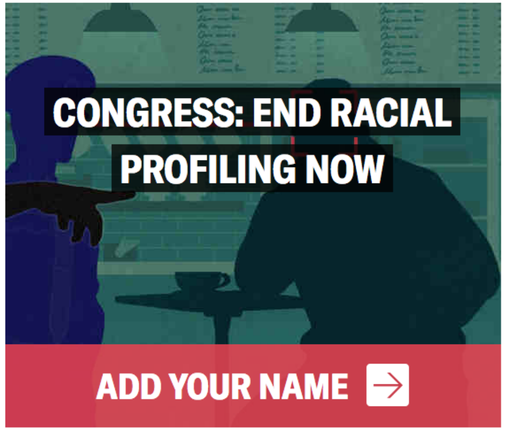
Email address | ZIP code | GET UPDATES

BECOME A MEMBER / RENEW / TAKE ACTION / DONATE

ISSUES | KNOW YOUR RIGHTS | DEFENDING OUR RIGHTS | BLOGS | ABOUT | SHOP

RACIAL DISPARITIES IN FLORIDA SAFETY BELT LAW ENFORCEMENT

Black motorists in Florida are stopped and ticketed for seatbelt violations in far greater numbers than white motorists – nearly twice as often statewide and up to four times as often in certain counties – according to a new report from the American Civil Liberties Union. The ACLU is calling for investigation by the Florida Attorney General's Office of Civil Rights and County Commissions charged with oversight over specific law enforcement agencies.



https://www.aclu.org/report/racial-disparities-florida-safety-belt-law-enforcement

Rationale for Bayesian evaluation

- Prior evidence on primary laws:
    - Strong evidence they reduce deaths and increase seat belt use;
    - This evidence is dated (1990s, early 2000s).
- 16 states have upgraded to primary since 2000.

Our aims:

1. Evaluate recent policy changes.*
2. Combine the evaluation of recent data with prior evidence.
3. Provide updated evidence on the impact of upgrading to primary enforcement.

*Harper and Strumpf, 2017

Intuition for Bayesian analysis

Purpose

We will generate new empirical evidence on the impact of policy changes since 2000.

Outcomes

Deaths per mile traveled and % of deaths wearing seat belts

Intuition for Bayesian analysis

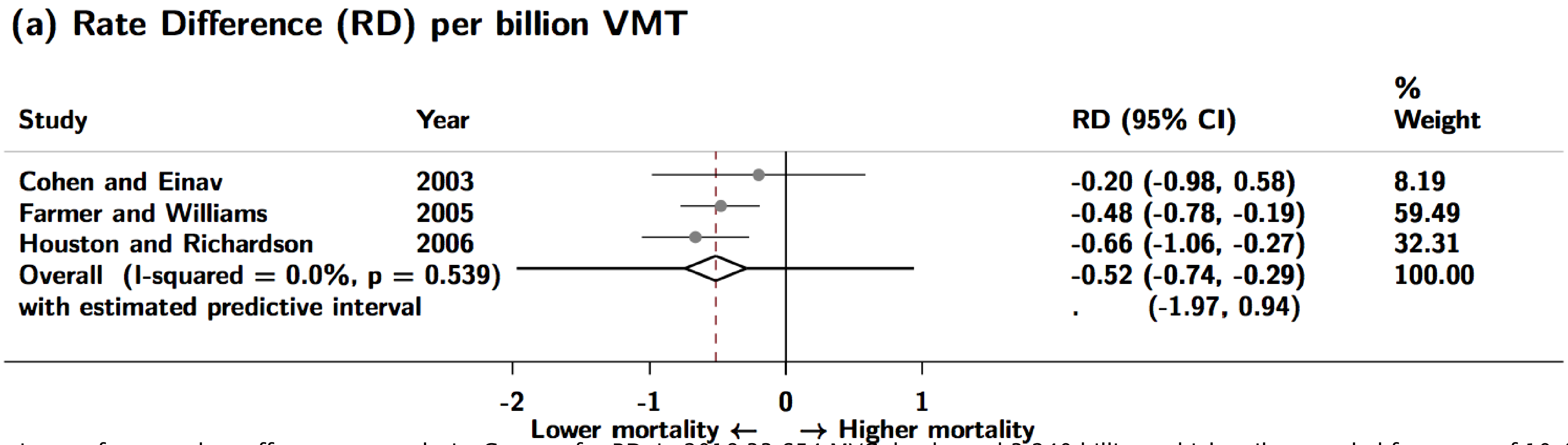
- Frequentist analysis ignores background information.
- Equivalent to belief that primary laws just as likely to:
  - Decrease death rates by a factor of 100 or 10.
  - *Increase* death rates by a factor 10 or 100.

- Use priors to encode existing information.
- Bayesian inference explicitly incorporates prior information to estimate the posterior probability distribution:

$$\underbrace{P(\theta|D)}_{\text{posterior}} \propto \underbrace{P(D|\theta)}_{\text{likelihood}} \times \underbrace{P(\theta)}_{\text{prior}}$$

But where do we get our priors?

Prior empirical evidence on upgrades to primary enforcement



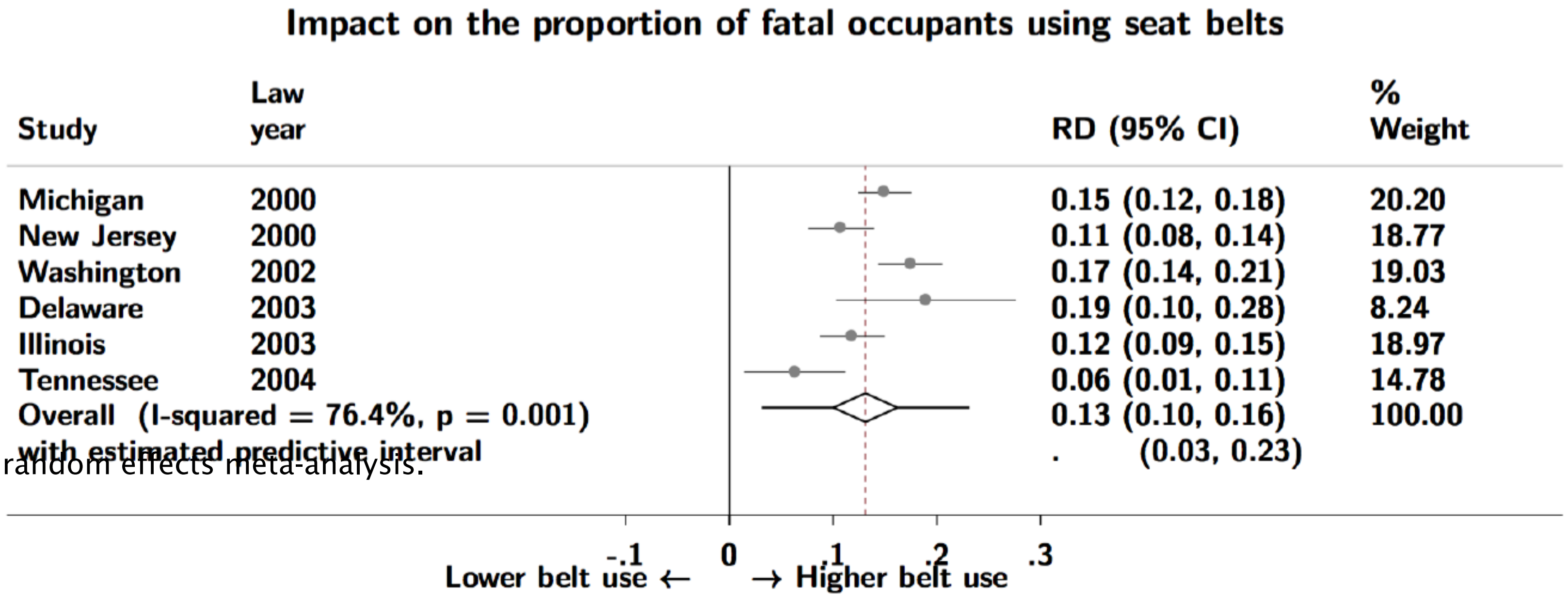
Estimates from random effects meta-analysis. Context for RD: in 2018 33,654 MVC deaths and 3,240 billion vehicle miles traveled for a rate of 10.4 (NHTSA data)

- Mean estimate: -0.5 deaths/billion VMT (-0.7, -0.3) ⇝ 1500 deaths/yr.

Prior empirical evidence on upgrades to primary enforcement

- Average estimate: 0.13 (0.10, 0.16). ⇝ 13 percentage points.
- Prediction interval for new "trial":

6 moderate-quality studies on proportion belted



Hedlund et al. (2008). Estimates from random effects meta-analysis.

Outline

Methods

Design and analysis

Semi-Bayesian analysis via data augmentation

- Data augmentation expresses prior information by adding empirical observations to the observed data.
- Advantages:
  - Avoids cumbersome full Bayesian machinery (resampling, MCMC).

Greenland (2006, 2007), Sullivan and Greenland (2013)

How to execute:

1. Define priors.
2. Encode priors as observations and add to the observed data.
3. Conduct analysis on all data.
4. Regression estimate and 95% CI provide approximations to

Why *semi-* Bayes and not fully Bayes?

Focus

Only interested in the policy effect, not other parameters.

Clustered treatment

Straightforward with conventional software, harder with fully specified Bayesian framework.

How do we augment the observed data?

- One way to think about priors is as bets on the "true" effect.
- Suppose we were 95% sure the true RR is between (0.25, 4).
- What kind of data would correspond to the prior?*
- Imagine we had a small randomized evaluation of upgrading to primary enforcement:

	Primary	Secondary	Estimates
Deaths	4	4	$RR_{prior} = 1.0$
Pop	100,000	100,000	$95\% CI_{prior} = (0.25, 4.0)$

*Higgins & Spiegelhalter (2000); Greenland (2007,2008). Check these numbers yourself in Stata by typing `csi 4 4 99996 99996`

- Add these prior data to our observed data as an additional stratum.

Construction of priors

- Specify prior interval for the RR (effect of law), translated into mean and variance.

3 sets of priors:

1. **Non-informative**: similar to frequentist assumptions.
   $\rightsquigarrow N(ln(1), 100)$

2. **Empirical**: prediction interval from prior meta-analysis.

   - MVC: 95% bet that true RR between 0.83, 1.09
     $\rightsquigarrow N(ln(0.95), 0.005)$

*NHTSA, Primary Enforcement Saves Lives (2006); CDC, Motor Vehicle Prioritizing Interventions.

Summary of prior record construction

- Recall the priors encode **information** that we append to our data.
- Means adding deaths and

(a) Poisson model: Fatalities per vehicle mile traveled (VMT)

Prior	RD	$SD(RD)$	95% limits	Deaths	Implied VMT (billions)
Non-informative	0	10	3e-9, 4e+8	6.25	6.51
Empirical	-0.5	0.44	-1.7, 0.90	125000	129841
Subjective	-1.0	0.26	-2.0, 0.0	208333	215970

(b) Grouped logit model: Proportion of deaths wearing seat belts

Prior	RD	$SD(RD)$	95% limits	Deaths	Implied offset
Non-informative	0	10	3e-9, 4e+8	4	0
Empirical	0.12	0.05	0.03, 0.23	22222	-0.0693
Subjective	0.14	0.02	0.06, 0.24	28571	-0.1099

Data

- Person-level FARS data on fatal crashes and vehicle miles traveled (VMT) from the Fatal Analysis Reporting System (FARS), 2000-2016.

- Dates of primary enforcement upgrades in each state.

- Other time-varying state policies:

  - Speed limits;
  - Graduated driver's license programs;
  - Blood alcohol content laws.

- Other time-varying state covariates:

Data structure (10 ages x 50 states x 17 years ~ n=8500)

How do we add the prior? (empirical version)

- First define the mean and variance for the prior

```
set obs 1          // only adding 1 record for 1 parameter
gen m = ln(0.95) // mean, i.e., RR=0.95
gen v = 0.005      // variance, equiv to 95%CI 0.83, 1.09
gen s = 25         // scaling factor
```

- Next set a variable representing number of prior cases, based on the variance of the prior and the scaling factor to facilitate normality (if using).

```
gen a = round(25^2 / 0.005,0) // num of prior cases
gen c = 0 // set constant to 0 for prior records, 1 for actual data
```

- Now set a variable for which prior is constructed to 1 for the covariate of interest, as well as equivalent person-time based on prior. Save as a dataset.

```
gen cov = 1 / 25
gen pt = a / exp(-0.95 / 25) // person-time for prior record
save "$cdir/data/derived/priorf2", replace
list
 +----------------------------------------------------------------+
 | prior           m      v     s        a   c    cov          pt |
 |----------------------------------------------------------------|
 |     2   -.0512933   .005    25   125000   0    .04    129841.4 |
 +----------------------------------------------------------------+
```

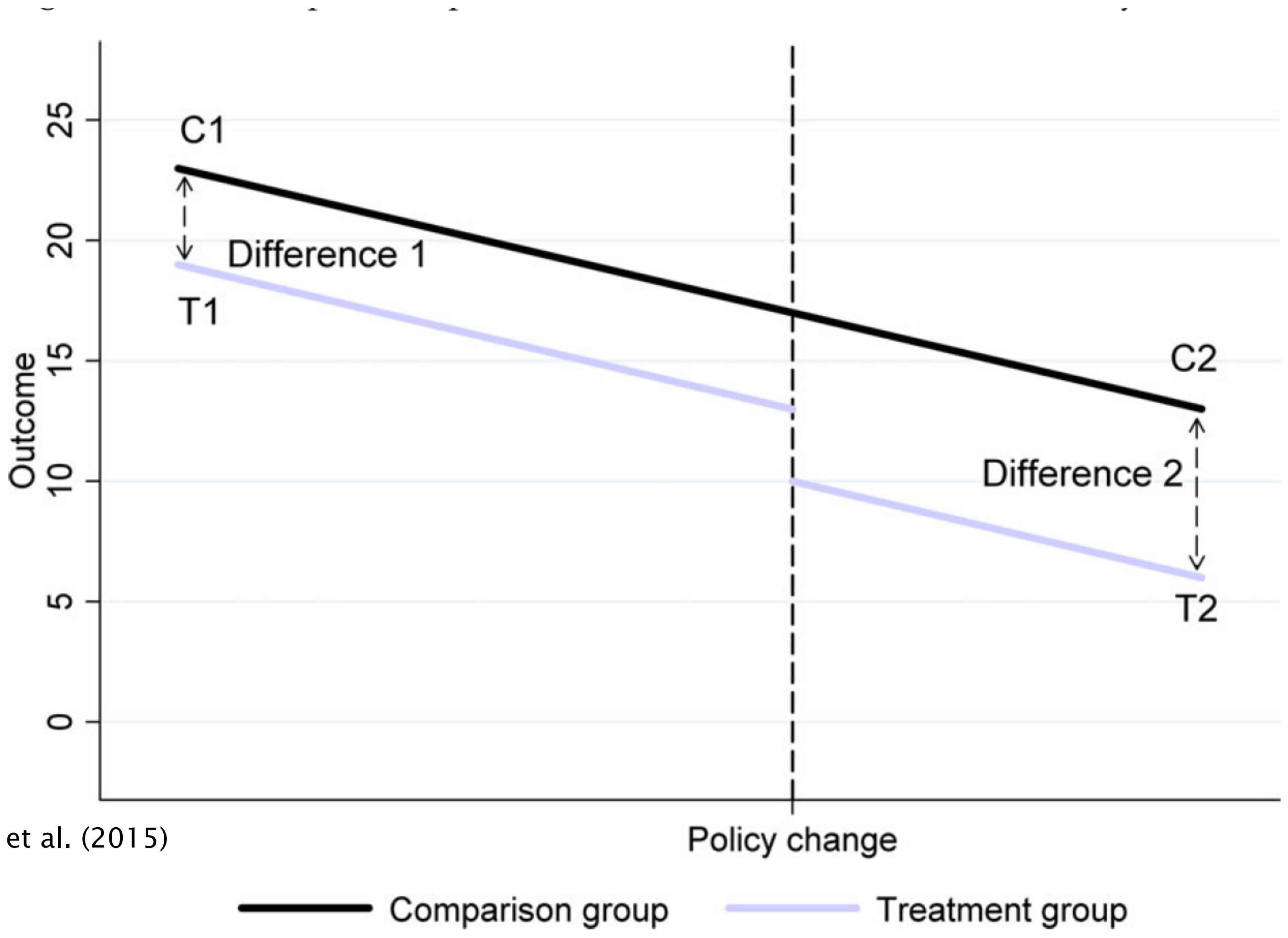
- Append this prior data to the observed data

```
use "$cdir/data/derived/mvc-laws-prior-data.dta", clear
append using "$cdir/data/derived/priorf1"
```

Added priors (n=8503, 1 added record for each prior)

Study design: Difference-in-Differences

Use pre/post data on treated and control groups to estimate



Meyer (1995); Angrist & Pischke (2009); Ryan et al. (2015)

- Identification strategy: timing of legislation.
- Assume that conditional on covariates, precise timing of upgrades is "as-if" random.

Likelihood model specifications

For MVC death rates

- Poisson model, i.e., $y_{ast} \sim Poisson(\mu_{ast})$:

$$ln(\mu_{ast}) = \alpha + \beta \times Primary_{st} + \boldsymbol{\gamma}\mathbf{A}_{ast} + \boldsymbol{\delta}\mathbf{Z}_{st} + \boldsymbol{\sigma}_s + \boldsymbol{\tau}_t + ln\left(VMT_{ast}\right)$$

where:

- y_{ast} is deaths for age group *a*, state *s*, and time *t*.
- $Primary = 1$ when a primary enforcement law is in effect

*speed limit laws, graduated driver's license laws, BAC laws, alcohol consumption per capita, police officers per capita, state median income, police reported alcohol involved, and proportion of crash deaths on rural roads.

Likelihood model specifications

For proportion belted

- Grouped logit model:

$$ln\left(p_{ast}/[1-p_{ast}]\right) = \alpha + \beta \times Primary_{st} + \boldsymbol{\gamma}\mathbf{A}_{ast} + \boldsymbol{\delta}\mathbf{Z}_{st} + \boldsymbol{\sigma}_s + \boldsymbol{\tau}_t$$

where:

- p_{ast} is the proportion of deaths wearing a seat belt in age group *a* and state *s* at time *t*.
- other coefficients as before.

Outline

Orientation

Background

Methods

Results

Some descriptive comparisons

- Important differences in covariates across levels of legislation.
- Also likely for unobserved confounders.

Variable	Remained Secondary	Remained Primary	Upgrade to Primary
Max speed limit, %			
<70mph	60.3	30.1	35.6
70mph	9.3	48.2	57.2
>70mph	30.5	21.7	7.2
BAC law, %			
<0.10	12.8	4.6	18.3

Visualization of empirical prior for MVC deaths per billion VMT

Visualization of empirical prior for MVC deaths per billion VMT

Add likelihood

Visualization of empirical prior for MVC deaths per billion VMT

Add likelihood

Posterior distribution

Visualization of **subjective** prior for MVC deaths per billion VMT

Add likelihood

Posterior distribution

Results for MVC deaths

Results for proportion belted

Sensitivity analysis with subjective priors (95% CIs)

Fully Bayesian analysis

- Consistent with fully Bayesian estimates
- Clustered standard errors more challenging in full Bayes

*Rstan model bglm2 results above, 4 chains, each with iter = 2000; warmup = 1000; thin = 1; total post-warmup samples = 4000)

Added benefit of full Bayes: estimate anything!

- Creating a full posterior distribution allows lots of flexibility to estimate different quantities of interest (we don't have to be slaves to NHST).
- E.g., what is the probability that the law parameter is < -0.05?

$$P(\beta < -0.05|\theta) = 0.04$$

or < -0.03?

$$P(\beta < -0.03|\theta) = 0.87$$

Outline

Orientation

Background

Methods

Results

Given our empirical priors, model, assumptions, and data, we can revise our inference.

Interpretation: Impact on MVC deaths per billion VMT

Before seeing the data:

We were 95% certain true effect was in the interval (-1.7, 0.9)

After seeing the data:

Now we are 95% certain the true effect is in the interval (-0.89, 0.35).

Interpretation: Impact on proportion belted

Before seeing the data:

We were 95% certain true effect was in the interval (0.03, 0.23)

After seeing the data:

Now we are 95% certain the true effect is in the interval (0.04, 0.07)

Implications

- Rates still declining.
- Other mechanisms?

Some recent evidence for unintended consequences

Conclusions

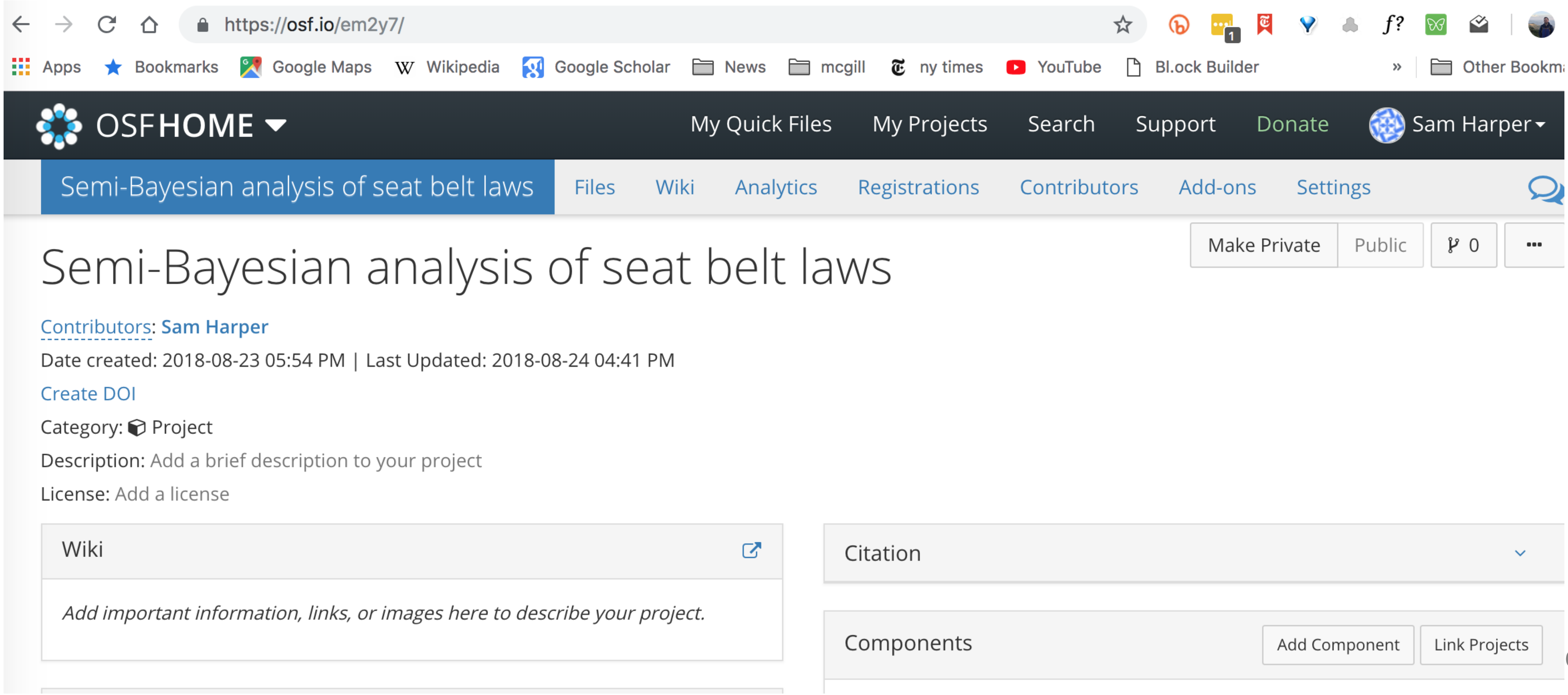
Bayesian interpretation

Newer evidence suggests weaker impact, despite priors.

Other mechanisms

Continued declines in MVC death rates due to improved road and vehicle technology, traffic calming measures, or other changes.

Reproducible materials (including these slides)




sam.harper@mcgill.ca
@sbh4th