Traffic Safety Program Evaluation: The Empirical Bayes Model and Mean Reversion Bias

Paul Fisher

Independent Researcher

Justin Gallagher

Montana State University

A Foolproof Traffic Safety Intervention

Traffic Safety Studies and Regression to the Mean Bias

- **Regression to the mean (RTM) bias is a well-known challenge when evaluating traffic safety programs (e.g. USDOT 2023)**
  - Roadway safety interventions often occur following years with an unusually high number of vehicle crashes
  - We would expect crashes to fall in subsequent years regardless of a new safety countermeasure (or "treatment")

- **Empirical Bayes (EB) Model developed to "correct" for RTM bias (Hauer 2002)**
  - Standard, widely used model in traffic safety engineering literature
  - Model is popular because of belief that it corrects for RTM bias:

    e.g. "*the empirical Bayes (EB) methodology has been applied for over 20 years [...] The appeal of the methodology is that it corrects for regression to the mean*" (Persaud and Lyon, 2014, p1)
  - A literature review of peer-reviewed studies in top safety engineering journals shows that the vast majority the studies state that the EB model "corrects", "accounts", "addresses", or "eliminates" RTM bias

EB Model Does Not Correct for RTM Bias

1. EB model never fully eliminates RTM bias
2. As applied in the literature, EB model usually adjusts for very little RTM bias

- **We show this**
  - (1) Analytically using the EB model equations
  - (2) Via a placebo Monte Carlo policy experiment (simulation)
  - (3) Running EB model using all vehicle crashes 2003-2022 (20 years) in San Antonio, TX
- **We provide recommendations to improve reliability of traffic safety studies**
  - (1) Diagnostics to evaluate how likely EB model estimates will suffer from RTM bias
  - (2) Draw a parallel to early program evaluation literature in labor economics which confronted a similar sample selection problem. Proposed solutions:
    - (i) Transparent analysis of the raw data
    - (ii) Recognition that conventional models do not solve RTM bias
    - (iii) Emphasis on careful sample construction

The EB Model leads to Biased Safety Countermeasure Treatment Effects

2 Sources of Bias in the EB Model

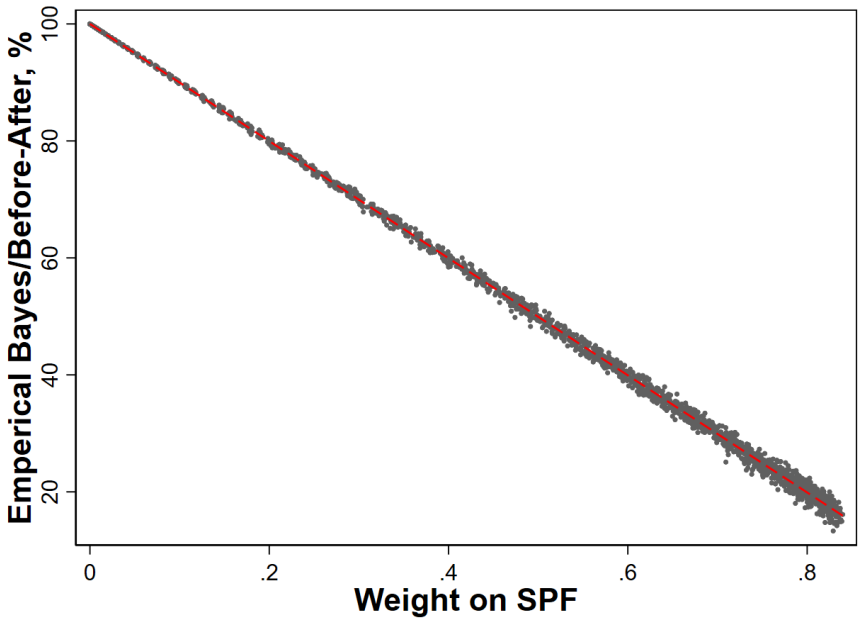
1. Misspecification of SPF in EB Model $\rightarrow$ omitted variable bias $\rightarrow$ RTM Bias
2. Small weights ($\hat{w}_i$) in EB Model $\rightarrow$ RTM Bias

Intuition for how the EB Model ostensibly corrects for RTM bias

- Treated roadway locations are often selected due to a large number of recent crashes
- A simple before-after model exaggerates the safety benefit
- The EB model adjusts the baseline (pre-treatment) level of crashes to "correct" for natural variation in the number of crashes
- The baseline crash adjustment is done using a model for crashes estimated out-of-sample (*Safety Performance Function*, *SPF*)
- A new baseline level of crashes at treated locations is calculated using a weighted average (weights, $\hat{w}_i$) of the actual crashes and the predicted level of crashes

Placebo Simulation Experiment

- The average weight is critical in eliminating RTM bias:
  - → The EB model can only eliminate RTM bias if $\hat{w}_i = 1$ for all treated locations
- We simulate data with multiple time periods & no actual safety countermeasure
  - → Placebo experiment so the treatment effect should be 0
  - → Before-After model estimate is -80%
- Follow USDOT: use Before-After model as benchmark to measure effectiveness of EB model at reducing RTM bias

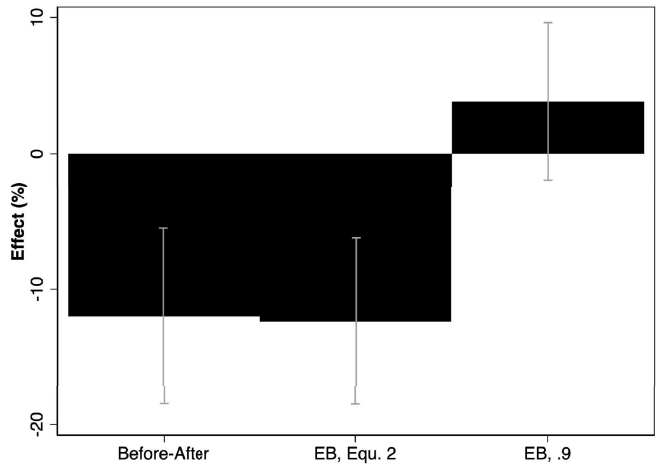


San Antonio Placebo Intersection-level Analysis

- Use GIS software to identify the yearly count of crashes "in or related to" 624 San Antonio, TX intersections
- Estimate the average treatment effect for a fake intersection-level policy we call "placebo red light camera" program
  - 50 intersections with highest avg number of crashes (2003-2005) are "treated"
  - Control group: 44 intersections with similar underlying crash risk (2003-2022)
  - Average Treatment Effect should be 0

Dependent Variable:	Total Crashes
Average Daily Traffic	-0.0000026
	(0.0000067)
Risk Proxy	0.077
	(0.021)
Psuedo R-Squared	0.042
Overdispersion	0.15
Avg weight, treated obs	0.139
Observations	44

(c) SPF Model for San Antonio, TX



(d) EB Model Placebo Treatment Effect

Summary

- EB model is a widely used model to "correct" for RTM bias
- We show:
  1. EB model does not correct for RTM bias
  2. The standard model that chooses the weights to maximize the efficiency of the EB model estimator adjusts for very little RTM bias
- Our recommendations:
  - (1) Researchers should recognize that the EB model does not correct for RTM bias
  - (2) The EB model will do a better job adjusting for RTM bias when:
    - (i) The SPF explains a large amount of the crash variation
    - (ii) A very high weight is selected for all treated intersections: $\hat{w}_i > 0.9$
    - (iii) Careful selection of the out-of-sample control intersections (ideally via randomization)
  - (3) Consider other models that have been shown to be successful in avoiding RTM bias e.g., difference-in-differences model (Gallagher and Fisher, 2020)