Summary

Roadside Inspections and Traffic Enforcement Model

Philip Situmorang

1 Introduction

The basic idea in program evaluation is to compare a sample set which was exposed to the program and a sample set which was not, and evaluate if the set exposed to the program is more strongly associated with the program's intended effect. As an equal alternative, we may compare a set which experienced a higher *exposure* of the program with a set which experienced less, and assess if the high exposure group is observed to experience the program's intended effect more strongly than the low exposure group.

For example, in a test evaluating the effectiveness of a drug to lower blood pressure, we may compare a group which received the drug with a group which received a placebo, and see if the treatment group see a significantly greater reduction in blood pressure than the placebo group. Alternatively, we may give one group a higher dose of the drug and the other a lower dose, and see if the high dose group see significantly greater drop in blood pressure. For both frameworks, the two groups are referred to as *treatment* and *control* groups, respectively.

Both approaches may be applied in evaluating the effectiveness of roadside inspection and traffic enforcement in improving road safety. For example, we may compare carriers which received inspections versus carriers which do not in a given time period, and see thereafter if the inspected carriers perform more safely on the road. Similarly, we may compare carriers with a higher *rate* of inspections with carriers that experienced a lower rate, and see if the high rate carriers, on average, perform more safely on the road relative to the low rate carriers during an evaluation period.

The former approach has been shown to be less practical, however, as the inspection program is ubiquitous and most carriers will experience at least one inspection per year, leaving very few eligible to serve as a control group. The issue

Version June 19, 2024

becomes more prominent when comparing among large carriers, where a case of one receiving zero inspection is extremely unlikely. The latter approach is more feasible in assessing the effectiveness of roadside inspections and traffic enforcement, as the variance in inspection rates among carriers allows partitioning of the carriers analyzed such that both treatment and control groups have sufficient samples.

This project is an implementation of the latter approach. It attempts to evaluate the effectiveness of roadside inspections and traffic enforcement by comparing a group of carriers with *high* inspection rates and a group with *low* inspection rates, and see if the high rate carriers, on average, perform more safely on the road. Specifically, we want to research available data to assess if the high inspection rate carriers, on average, experience fewer crashes relative to the size of their fleets in a period of evaluation after the inspections are performed.

1.1 A NOTE ON CAUSAL INFERENCE

Any discovery of association between high inspection rates with greater negative change in crash rates may serve as evidence of a causal relationship between a treatment and an effect variable, and thereby, the effectiveness of roadside inspections and traffic enforcement. *Evidence* stands in contrast to *proof*, associated in the scientific sense most closely with mathematical proof, which guarantees a conclusion. Correlation does not prove causation, however it is the standard byproduct of researches exploring possible causal relationships between variables. For example, a research may show that smokers are more strongly associated with developing lung cancer than non-smokers. The quantitative measure of the strength of association between smoking and cancer does not prove that one causes the other, but rather serves as correlational evidence indicating causality.

Naturally, some evidence are more compelling than others. Generally, the greater the contrast in the observed effect between the treatment and control groups, the stronger the evidence for a causal relationship. However, more important than the final indication alone are the data and the analysis process which underlies it. The strength of any correlational evidence hinges directly on the data quality and methodologies used to arrive at that indication. The goal of the roadside inspection and traffic enforcement model, therefore, is to implement decisions on selecting data source and methodologies available such that any result indicating the effectiveness of roadside inspections program may be used as evidence for program effectiveness.

1.2 DATA CONSIDERATIONS

One way to obtain data is through a controlled study where measurements of effect are taken before and after the application of a certain treatment or program. For example, drug manufacturers often study the effectiveness of their product by selecting participants, dividing them into control and treatment groups, applying the medication, and measuring changes before and after the treatment while minimizing the likelihood that any measured effect is due to factors other than the treatment.

This rigorous approach is widely considered to be the gold standard in optimizing data quality. Such controlled study, however, is expensive. In the case of roadside inspections and traffic enforcement, implementing this approach would require steps such as selecting drivers or carriers to include in the study, standardizing the inspection process, and monitoring the change for a period of time. While conceivable, the cost associated with controlling the sample and collecting data in this manner is often prohibitive.

Another approach is the study of existing data. In this approach, we forego additional data collection and define from available data variables which best measure the treatment and effect of interest. We may additionally partition our sample set to control for possible confounding factors, provided that our dataset enables us to do so. We then obtain numeric indications of how strongly a particular treatment is associated with a certain effect. This analysis implements this data approach, by using data available in the Motor Carrier Management Information Systems (MCMIS) database.

1.3 METHODOLOGY CONSIDERATIONS

The decisions on how to process and interpret data impacts the value of a research. Below are the primary considerations that drive the analysis decisions made in this project.

Industry Standard. Implementing a commonly used methodology facilitates ease of grasp and minimizes likelihood of misinterpretation of the methodology. This analysis uses a statistical framework most commonly used for causal inference to foster ease of communication, and thus, confidence in the results of the analysis.

Parsimony. Minimizing the number and complexity of analysis variables facilitates ease of understanding of the relationship between those variables. For example, the number of medicine tablets a person takes is a simpler variable to measure and understand compared to the rate at which the person takes them.

Objectivity and Transparency. Decisions made throughout the analysis process impact the objectivity of the research. Steps such as using standard ways of analyzing data, making the research data publicly accessible, maximizing code readability, and facilitating peer review fosters objectivity, mitigates confirmation bias, and ultimately improves the overall value of the research product.

2 Implementation

This analysis investigates if carriers inspected at a higher rate, on average, experience relatively fewer crashes after the inspections. To enable quantitative measurement, we define a carrier's **inspection rate** as the number of inspections divided by the number of the carrier's power units, and **crash rate** the number of crashes divided by the number of power units.

$$
\begin{aligned}
&Inspection\,Rate \;=\; Number\,of\,Inspections \;\div Number\,of\,Power\,Units\;(PUs)\\
&Crash\,Rate \;=\; Number\,of\,Crashes \;\div Number\,of\,Power\,Units\;(PUs)
\end{aligned}
$$

The rates are computed on an annual basis. For example, a carrier with 100 power units who, in a given year, received 20 inspections and experienced 3 crashes will have an inspection rate of 0.20 and a crash rate of 0.03 for that year.

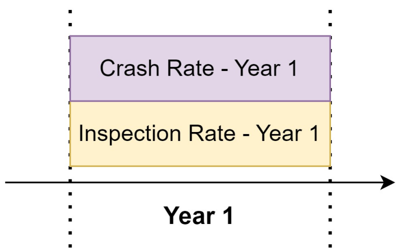

Figure 1: Carriers' inspection and crash rates are computed on a yearly basis

To evaluate if inspections are effective at reducing crashes, we may intuitively assess if higher inspection rates are associated with lower crash rates. We may, for example, plot a graph to see if carriers who are inspected at a higher rate in year 1, on average, experience lower crash rates in that same year. However, in this approach the distribution of when the inspections and crashes occurred could be such that a carrier's crash rate cannot be reasonably attributed to the inspections it received. For example, a carrier with a relatively low crash rate cannot reasonably attribute the rate to inspections if the inspections occur after the crashes in the year. In other words, to logically attribute the effect to the treatment, the inspections would have to occur before the crashes.

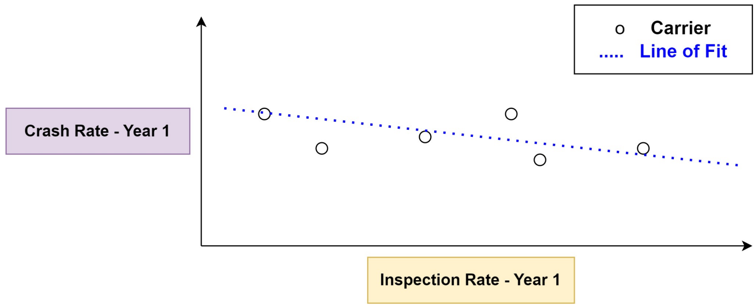

Figure 2: A hypothetical investigation of six carriers. In year 1, carriers with higher inspection rate are shown to have lower crash rates. However, it is possible that the inspections occur after the crashes in that year, in which case the lower crash rates cannot be reasonably attributed to the higher inspection rates.

To address this we compare inspection rates not with crash rates but with the *change* in crash rates. Specifically, we compare if higher inspections in year 1 is negatively associated with change in crash rate from year 1 to year 2. An example best illustrates this idea. Suppose we have two carriers with 100 power units each, and that their inspection and crash rates are as described below.

	Year 1	Year 1	Year 2	
Carrier	Inspections	Crash Rate	Crash Rate	CR Change
1	20	0.03	0.02	- 0.01
2	5	0.03	0.04	+ 0.01

In this example, carrier one, which received twenty inspections, saw a decline in crash rates in year 2. Carrier two, who received only five inspections, saw instead an increase in crash rate from year 1 to year 2. Note that as each has 100 PUs, their inspection rates are 20/100 and 5/100. From a sample of two carriers, we observe that the carrier with higher rate of inspections experience relatively fewer crashes in a period after the inspections. Figure 4 illustrates this relationship.

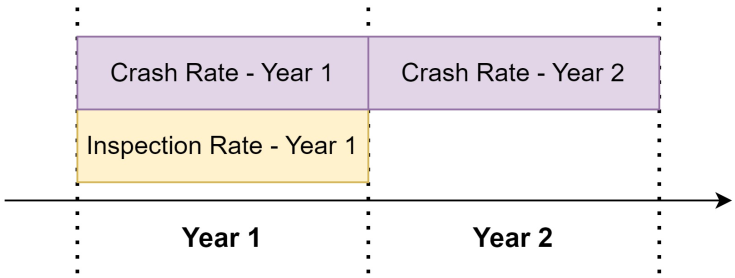

Figure 3: Instead of comparing inspection rates against crash rates in year 1, we compare year 1 inspection rate with the *change* in crash rate from year 1 to year 2.

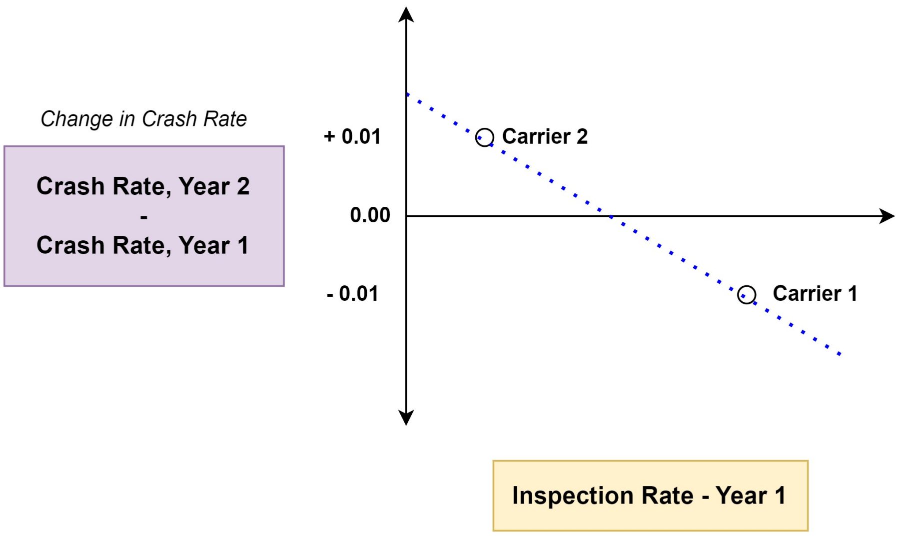

Figure 4: Carrier one, which received more inspections, see relatively fewer crashes in year 2. As indicated by the downward slope of the line of best fit, greater inspection rate is negatively associated with change in crash rate. This negative slope indicates effectiveness. However, because the sample size is small, the difference could be solely due to chance.

The more carriers we analyze, the greater the likelihood that any relationship discovered between inspection rates and change in crash rate is not due to randomness. Sample size drives the level of confidence of the results, thus in this analysis we rely on the volume of data available in MCMIS to discover the relationship between inspection rates and change in crash rates. We assess the strength of evidence for the effectiveness of roadside inspections by asking two questions:

- Are greater inspection rates negatively associated with change in crash rates?
- If true, is the relationship statistically significant?

2.1 Periods of Evaluation

To draw on data available in MCMIS we first define the time range by which we aggregate inspection and crash rates. In this analysis we select our period of aggregation to be a period of one year. To increase rigor, the analysis evaluates two periods of interest and perform two separate inquiries into how inspection rates are related to change in crash rates.

The two periods were chosen arbitrarily. For the first period, year 1 is defined as between July 1st, 2021 and June 30th, 2022. Year 2 is defined as the year thereafter. The second period is analogous to the first, with the time ranges shifted backwards by one year.

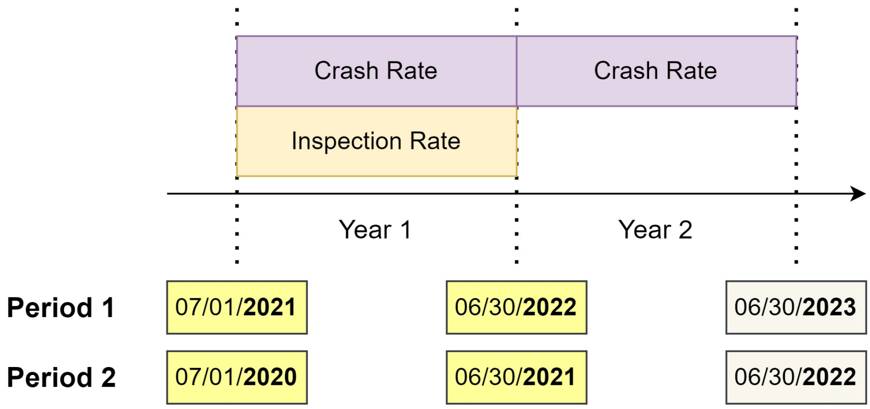

Figure 5: The analysis investigates the relationship between inspection rates and change in crash rates in two time periods.

2.2 Segmentation of Carriers

At this point we may begin our analysis into the relationship between inspection rates and change in crash rate, for all carriers inspected in year 1. However, carriers vary in the sizes of their fleets and their safety risk levels, and these may impact how each responds to inspections. To account for any difference in the effect of inspections due to variance in sizes and safety risk levels, we segment carriers according to the number of power units in their fleets and their assigned Inspection Selection System (ISS) risk groups.

To do this we first aggregate all carriers which were inspected in year 1. Next, we divide them into four size groups, size group 1 being the group consisting of the smallest carriers (between 1 and 5 power units) and size group 4 being the group consisting of the largest carriers (over 100 power units). This division method is identical to that implemented in FMCSA's Carrier Intervention Effectiveness Model.

Carriers in each of the four size groups are then further divided according to their safety risk levels. The ISS categorizes carriers into three risk groups: Pass, Optional, Inspect. The "inspect" category is used to identify carriers with the highest risk levels, and for ease of comprehension this analysis refers to the three groups as Low, Medium, and High risk groups.

This process yields twelve final sub-groups of carriers segmented by size and safety risk levels (see figure 6). This research performs twelve separate analyses for each of these groups (shaded in red in figure 6), in addition to four analyses of carrier groups divided by size only (groups shaded in blue). In total, sixteen groups from each period were separately analyzed to determine if inspection rates are negatively associated with change in crash rates, and if those associations are statistically significant.

3 Results

The analyses were performed using R and SQL. A separate R Markdown summary output which details the data processing and analysis steps accompanies this document. The results of the analyses for each of the two periods are summarized in the following two sections.

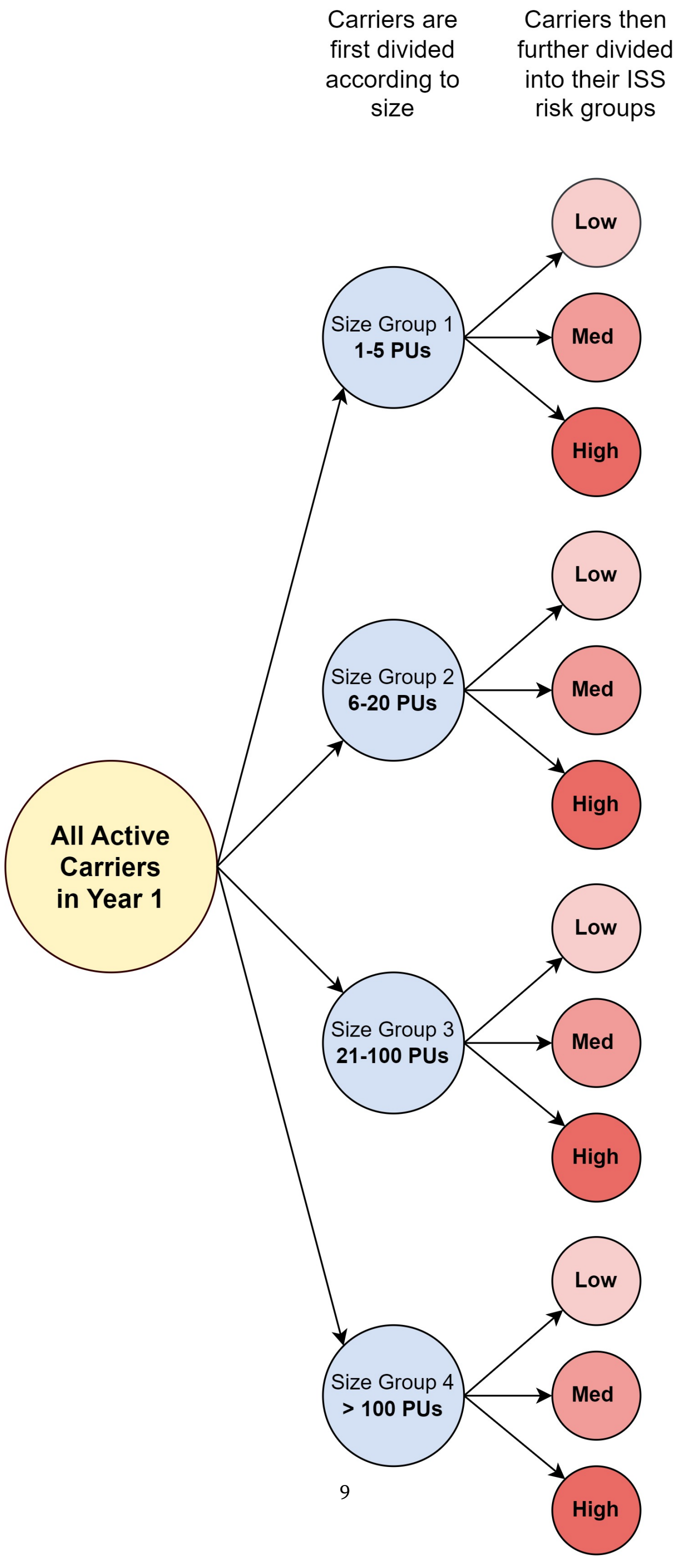

Figure 6: Subdivision of carriers to account for possible variance in the effects of inspection due to variance in carrier sizes and safety risk levels

3.1 Period 1 - Inspections Between 07/01/2021 and 06/30/2022

Period one analysis discovers that higher inspection rates are generally negatively associated with change in crash rate. Analyses of the four size groups (not further divided by ISS risk groups) finds that higher inspection rates are negatively associated with change in crash rates across all groups. The results are less uniform when the groups are divided by their ISS risk groups. Group results in which we discover a statistically significant negative association between higher inspection rates and change in crash rates are shaded in green.

Size	ISS Risk	Slope	P-value
1-5 PUs	**All**	**-0.002243**	**2.00E-16**
	Low	-0.0006452	0.0894
	Medium	0.003815	0.0237
	High	-0.0033	0.109
6-20 PUs	**All**	**-0.0056457**	**2.00E-16**
	Low	**-0.0023519**	**0.0000149**
	Medium	-0.001771	0.187
	High	**-0.003138**	**0.0452**
21-100 PUs	**All**	**-0.0055779**	**2.00E-16**
	Low	**-0.0022953**	**0.000254**
	Medium	**-0.005161**	**0.000938**
	High	**-0.004504**	**4.26E-03**
>100 PUs	**All**	**-0.0066261**	**2E-16**
	Low	**-0.0048926**	**2.73E-10**
	Medium	**-0.005364**	**5.25E-03**
	High	**-0.007401**	**0.0000453**

Figure 7: Statistical significance is defined as having a P-value of less than 0.05

In period 1, the negative association does not hold true when the smallest carriers are further divided into their ISS risk groups. In the larger sized carriers, the associations remain after their safety risk level subdivision. Larger carriers generally see steeper slopes, which signifies greater change in crash rates with greater increase in inspection rates. The results are illustrated below.

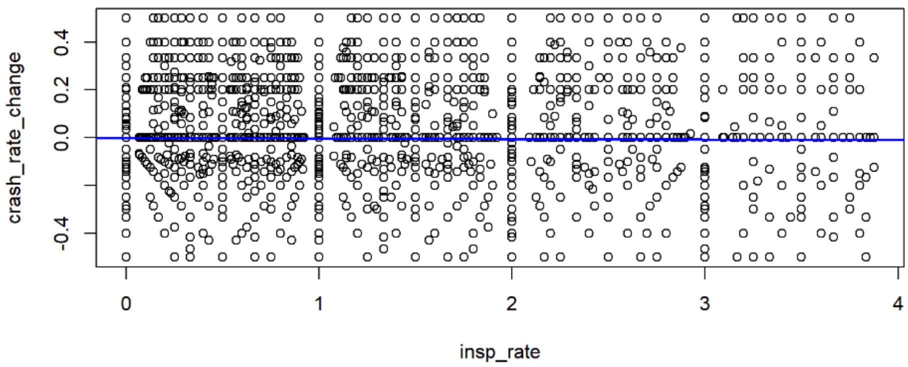

Figure 8: Size Group 1 (1-5 PUs), All Carriers. Slope: -0.002243, P-value = 2e-16

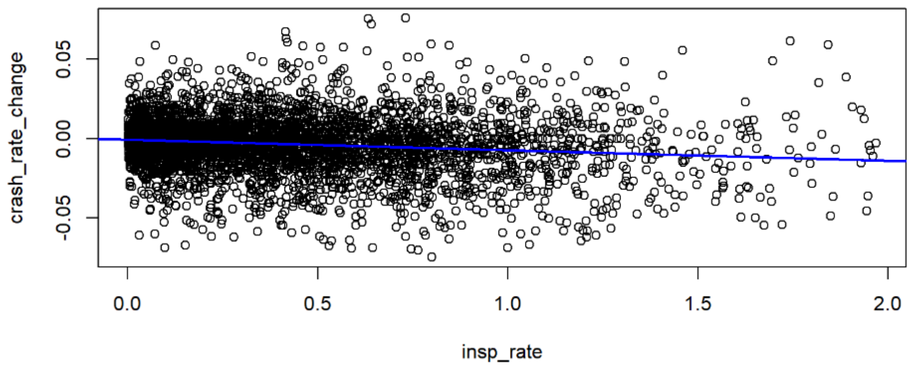

Figure 9: Size Group 4 (> 100 PUs), All Carriers. Slope: -0.006626, P-value = 2e-16

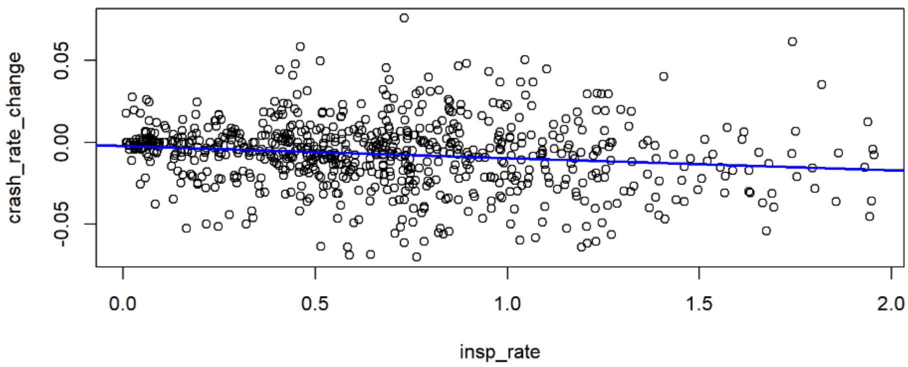

Figure 10: Size Group 4 (> 100 PUs), High Risk. Slope: -0.007426, P-value = 2e-16

3.2 Period 2 - Inspections Between 07/01/2020 and 06/30/2021

Period two analysis discovers similar results, with higher inspection rates generally negatively associated with change in crash rate. Analyses of the four size groups (not further divided by ISS risk groups) finds that higher inspection rates are negatively associated with change in crash rates across all groups. Similar to period 1 results, the indications are less uniform when the groups are divided by their ISS risk groups.

Size	ISS Risk	Slope	P-value
1-5 PUs	**All**	**-0.0005056**	**2.06E-09**
	Low	-0.0006743	0.0808
	Medium	0.003026	0.0712
	High	-0.0005697	0.79888
6-20 PUs	**All**	**-0.0017163**	**2.97E-09**
	Low	0.0010054	0.0639
	Medium	0.0004064	0.758
	High	**-0.006235**	**0.000166**
21-100 PUs	**All**	**-0.0038314**	**2.00E-16**
	Low	0.000605	0.327563
	Medium	**-0.005122**	**0.000541**
	High	**-0.0066807**	**3.16E-05**
>100 PUs	**All**	**-0.0036113**	**9.19E-09**
	Low	0.0001045	0.8943
	Medium	**-0.008221**	**9.76E-06**
	High	**-0.006124**	**0.000411**

Figure 11: Statistical significance is defined as having a P-value of less than 0.05

In period 2, the negative association does not hold true when the smallest carriers are further divided into their ISS risk groups. This is also true in the low risk groups of the larger carriers. Similar to period 1 results, however, larger carriers generally see steeper slopes, which signifies greater change in crash rates as inspection rates increase. The results are illustrated below.

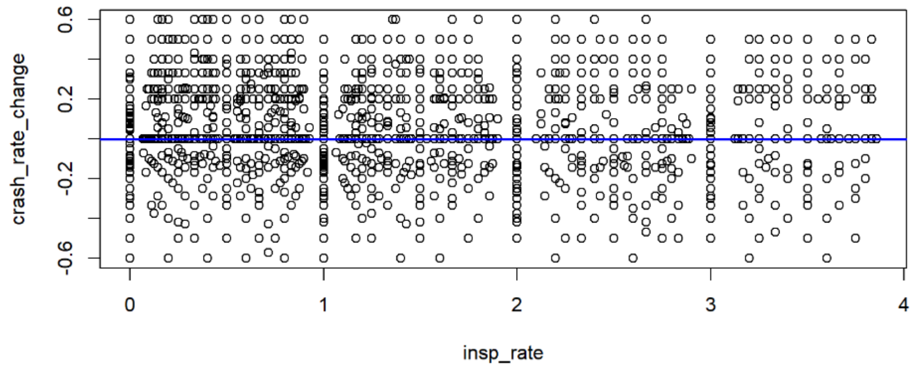

Figure 12: Size Group 1 (1-5 PUs), All Carriers. Slope: -0.000506, P-value = 2.06e-9

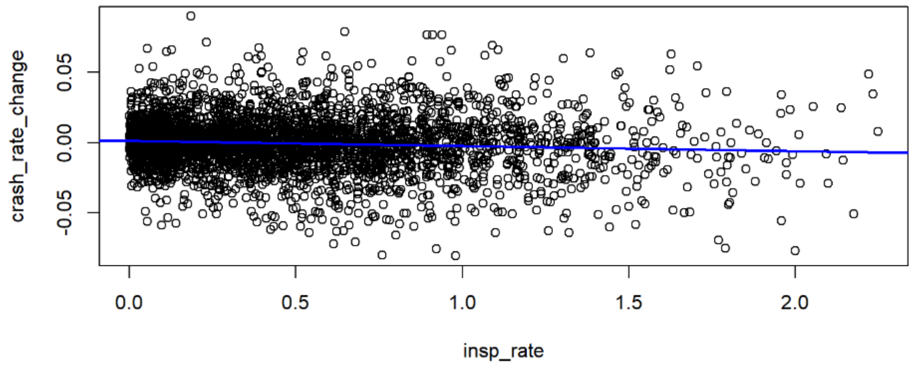

Figure 13: Size Group 4 (> 100 PUs), All Carriers. Slope: -0.003611, P-value = 9.19e-9

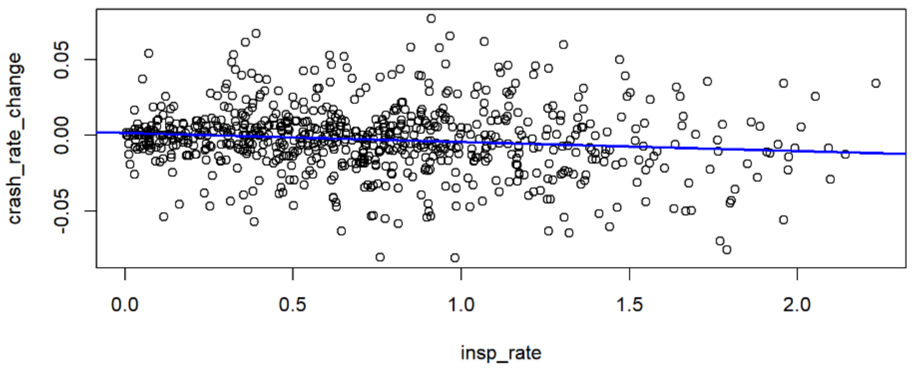

Figure 14: Size Group 4 (> 100 PUs), High Risk. Slope: -0.006124, P-value = 4.11e-4

4 Benefits Estimations

So far the preliminary results imply effectiveness, as indicated by the negative association between inspection rates and change in crash rates. The following sections will discuss ways by which we can interpret the resulting parameters of the linear model and estimate, in more tangible terms, the degree to which inspections impact road safety.

4.1 Interpretation of the Negative Slope

To interpret the negative slope, we will use period one analysis on carriers in size group 3 (21-100 PUs), all ISS risk group results as an example. Below are the indications for the group.

```
Coefficients:
              Estimate Std. Error t value Pr(>|t|)    
(Intercept) -0.0013177  0.0002474  -5.326 1.01e-07 ***
insp_rate   -0.0055779  0.0004057 -13.749  < 2e-16 ***
```

The y-intercept of -0.0013 estimates that, on average, carriers with zero inspections see their crash rates decline by 0.0013 in year 2. The group's slope of -0.0056 indicates that carriers with an inspection rate of 1.0, on average, experience lower crash rate in year 2 by a factor of -0.0056 compared to carriers with an inspection rate of 0.0.

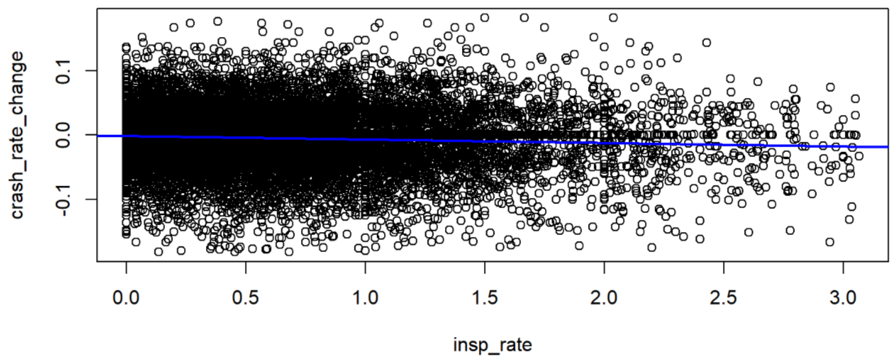

Figure 15: Size Group 3 (21-100 PUs), All Risk. The Y-intercept estimates the change in crash rate when inspection rate is equal to zero. Increasing the inspection rate of a carrier by a factor of one, on average, reduces crash rates by a factor of 0.0056.

To illustrate further, suppose we have three carriers each with 100 power units and a 0.10 crash rate in year 1. Using the linear model above to predict crash rates in year 2, a carrier with inspection rates of 0.0, 1.0, and 2.0 respectively, on average, will see their safety performance unfold as follows.

		Year 1	Year 1	f(IR)	Year 2*
Carrier	PUs	CR	IR	Y-int + IR * Slope	CR
1	100	0.10	**0.0**	- 0.0013 + (**0.0**) * -0.0056	0.0987
2	100	0.10	**1.0**	- 0.0013 + (**1.0**) * -0.0056	0.0931
3	100	0.10	**2.0**	- 0.0013 + (**2.0**) * -0.0056	0.0875

*$Year\ 2\ CR = Year\ 1\ CR + f(IR)$

Carrier 1, which received zero inspection, will see its crash rate decline by 0.0013. Carrier 2, which received 100 inspections (as indicated by an inspection rate of 1.0) will see a negative change in crash rate by a factor equal to the y-intercept *plus* the inspection rate times the slope. To estimate the degree to which 100 inspections reduce crashes, we can use this difference in change in crash rates between carrier 1 and carrier 2.

		Year 1	Year 2	Year 2	
Carrier	PUs	IR	CR	Crashes	**Crashes Prevented**
1	100	0.0	0.0987	9.87	**0**
2	100	1.0	0.0931	9.31	|9.31 - 9.87| = **0.56**
3	100	2.0	0.0875	8.75	|8.75 - 9.87| = **1.02**

Crashes prevented is the number of year 2 crashes minus the the year 2 crashes of the carrier with zero inspections, which in this case serves as our control group. This estimate indicates that the inspections performed on carriers 2 and 3 prevented a total of 1.58 crashes.

4.2 Treatment and Control Groups

So far we estimated the benefits of roadside inspections by comparing carriers with high and low inspection rates, where *high* is defined as carriers with inspection rates greater than zero. To estimate the number of crashes prevented, we computed the difference in the change in crash rates between each carrier in the treatment group (carriers 2 and 3) and the control group carrier (carrier 1), and multiplied that difference by the number of power units carriers 2 and 3.

More generally, to estimate the number of crashes prevented, we took each carrier in the treatment group, measure the difference between the *change in crash rate* of the treatment carrier and the *mean change in crash rate* of the control group carriers, and multiply that difference by the number of power units the treatment carrier has. We can simplify this procedure to the following equation.

$$Crashes\,Prevented = \sum_{i=1}^{n} Number\,of\,PU_i * Inspection\,Rate_i * Slope \tag{1}$$

Where n is the number of carriers in the treatment group and i is a unique index number for each carrier in the group. Note that this is only valid when the treatment group consist of all carriers with inspection rate greater than zero.

Using this threshold, however, leads to highly imbalanced control and treatment group sizes especially among large carriers where the vast majority have an inspection rate of greater than zero and thus will fall under the treatment group. For example, in carriers with greater than 100 power units, the Y-intercept is determined solely by the projection of the line of best fit, not by actual carriers with zero inspection rates. Therefore, the control group size is effectively zero.

4.3 Balancing the Control and Treatment Group Sizes

To increase balance in the sizes of the two groups, we can choose a new threshold such as the mean of the inspection rate for a particular carrier group. That is, the treatment group will consist of all carriers with inspection rates greater than this mean rather than zero. By doing this, we effectively treat those carriers with inspection rates below the mean *as though they received no inspections*. This leads to a more conservative estimate and higher balance between control and treatment group sizes.

We may then estimate how much the two groups differ in their change in crash rates. To do this, we compute the mean change in crash rates for each group, measure the difference, and determine if the difference is statistically significant. We can use this difference in means to estimate for the number of crashes prevented using the following equation.

$$Crashes\,Prevented = (\mu_A - \mu_B) * \sum_{i=1}^{n} Number\,of\,PU_i \tag{2}$$

Where:

- μ is the mean change in crash rates for treatment (A) and control (B) groups
- n is the number of carriers in the treatment group
- i is a unique index number for each carrier in the treatment group

4.4 Crashes Prevented Estimations for Periods 1 and 2

Applying the above formula to estimate crashes prevented from the analysis of the two periods, we obtain the following results.

Size Group	Period One	Period Two
1-5 PUs	2,140.94	423.42
6-20 PUs	2,425.61	406.43
21-100 PUs	2,395.84	1,283.29
> 100 PUs	3,102.82	2,075.07
Total Crashes Prevented	**10,065.22**	**4,188.21**

4.5 Injuries Avoided and Lives Saved

To estimate for the number of injuries avoided and lives saved, an estimate of how many injuries and deaths occur per accident is needed. Various ways to estimate these rates exist, and in this analysis the simple method of summing the count of injuries in MCMIS and dividing it by the total count of crashes was chosen. Similarly, to estimate a deaths per accident rate, we aggregate the count of deaths in all crashes in MCMIS and divides it by the count of all crashes available in the database.

$$Injuries\,Per\,Accident = Total\,Count\,of\,Injuries\ \div\ Total\,Count\,of\,Crashes \tag{3}$$

$$Injuries\,Per\,Accident = Total\,Count\,of\,Injuries \div Total\,Count\,of\,Crashes \quad (4)$$

To estimate for number of injuries avoided and lives saved, we multiply the above rates by the number of crashes prevented from each period.

Size Group	Period One	Period Two
Total Crashes Prevented	10,065	4,188
Total Injuries Avoided	**4,308**	**1,792**
Total Lives Saved	**314**	**130**

5 Conclusion

The preliminary analysis indicates that higher inspection rates are associated with negative change in crash rates. Estimating safety benefits using the results of the model requires selecting analysis parameters which ultimately impacts the final estimate. The above methods of estimation were chosen primarily due to their simplicity. More rigorous methods may be applied while balancing the tradeoff between simplicity and complexity.