The Relationship between Freight Train Length and the Risk of Derailment

Peter M. Madsen[1], Robin L. Dillon[2]*, Konstantinos P. Triantis[3], Joseph A. Bradley[1]

Abstract

In recent years, longer and heavier trains have become more common, primarily driven by efficiency and cost-saving measures in the railroad industry. Regulation of train length is currently under consideration in the U.S. at both the federal and state levels, because of concerns that longer trains may have a higher risk of derailment, but the relationship between train length and risk of derailment is not yet well understood. In this study, we use data on freight train accidents during the 2013-2022 period from the Federal Railroad Administration (FRA) Rail Equipment Accident and Highway-Rail Grade Crossing Accident databases to estimate the relationship between freight train length and the risk of derailment. We determine that longer trains do have a greater risk of derailment. Based on our analysis, running 100-car trains is associated with 1.11 (95% confidence interval: 1.10 to 1.12) times the derailment odds of running 50-car trains (or a 11% increase), even accounting for the fact that only half as many 100-car trains would need to run. For 200-car trains, the odds increase by 24% (odds ratio 1.24, 95% confidence interval: 1.20 to 1.28), again accounting for the need for fewer trains. Understanding derailment risk is an important component for evaluating the overall safety of the rail system and for the future development and regulation of freight rail transportation. Given the limitations of the current data on freight train length, this study provides an important step toward such an understanding.

KEY WORDS: Rail transportation; Freight train length; Derailment; Transportation safety; Quasi-induced exposure

[1] Brigham Young University, Provo, UT, USA
[2] Georgetown University, Washington, DC, USA
[3] Virginia Tech, Falls Church, VA, USA
*Address correspondence to: Robin Dillon, McDonough School of Business, Georgetown University, Washington, DC, 20057, 202-687-5398, rld9@georgetown.edu

1. INTRODUCTION

On February 3, 2023, 38 cars from a 151-car, 9,300-foot-long freight train derailed in East Palestine, Ohio, leading to the release of hazardous materials that required the evacuation of more than 2,000 residents. In response to this event and concerns that the length of the train may have contributed to the derailment, U.S. Senator Sherrod Brown introduced the Railway Safety Act of 2023, which if enacted would require the development of regulations regarding freight train length, among other things (Congress, 2023). Additionally, several U.S. states are currently considering state-level regulations regarding freight train length (Bernton, 2023; CBS, 2023). The major freight rail industry association, the Association of American Railroads, expresses the industry's opposition to regulation of freight train length, arguing that "'Long trains' have operated safely for decades, and the industry's safety record has dramatically improved during that period" (AAR, 2023). However, general improvement in safety over time coinciding with increases in train length cannot be seen as evidence for a lack of association between the two. Moreover, identifying the relationship between freight train length and the risk of derailment is challenging and evidence with respect to this relationship is sparce in the current research literature. Although a number of related questions, such as how train length relates to the severity of a given derailment, have been examined in the literature, data availability challenges have made direct investigation of the relationship between train length and the likelihood of derailment difficult (Dick et al., 2021; Liu et al., 2017). Indeed, a 2019 U.S. Government Accountability Office report to Congress concludes that the safety implications of train length is poorly understood and that more study is needed to assess it (GAO, 2019).

The system-wide safety implications of utilizing longer freight trains are further complicated by the fact that when longer freight trains are used, fewer trains, on average, are needed to transport the same amount of freight. Thus, even if longer freight trains are at an increased risk of derailment, that increased risk must be balanced against the benefits of using fewer freight trains in reducing overall exposure to derailment risk. Moreover, longer freight trains also have additional benefits for the U.S. rail transportation system including improved system-wide fuel efficiency, lower system-wide emissions, and lower overall operating costs (GAO, 2019; Muller et al., 2022). Consequently, a better understanding of how freight train length impacts system-wide derailment risk can help policymakers and rail industry decision makers to better optimize freight train length policies and decisions (Ghofrani, Sun, & He, 2022).

The major challenge to examining the relationship between freight train length and the risk of derailment is the lack of available exposure data. Although the length of trains involved in accidents is included in accident reports that railroads file with the Federal Railroad Administration (FRA), data on the lengths of trains that are not involved in accidents are largely unavailable. This lack of "exposure" data has precluded estimating the derailment risk by freight train length using traditional methods. In the work reported here, we overcome this lack of available exposure data for freight trains by applying the quasi-induced exposure (QIE) approach, a methodology designed to study risks in settings where exposure data are missing (Jiang, Lyles, and Guo, 2014; Keall and Newstead, 2009; Stamatiadis and Deacon, 1997). QIE compensates for the lack of true exposure data by employing data from a control group of accidents whose occurrence is unrelated to the factor of interest to proxy for exposure risk, creating "quasi-induced" exposure data (Jiang et al., 2014; Keall and Newstead, 2009). Using QIE, we analyze data on freight train accidents from the FRA Rail Equipment Accident and Highway-Rail Grade Crossing Accident databases to estimate the relationship between freight train length and the risk of derailment.

2. LITERATURE REVIEW

2.1 Quasi-Induced Exposure (QIE)

QIE has been used extensively in studies of road traffic accidents because data on exposure by vehicle or driver characteristic is usually unavailable, while data on the characteristics of vehicles and drivers involved in accidents are accessible (Jiang et al., 2014; Keall and Newstead, 2009; Stamatiadis and Deacon, 1997). In practice, the QIE approach involves the creation of a combined sample containing both the accidents of interest as well as a set of control accidents whose occurrence is unrelated to the independent variable being analyzed. For example, Leslie et al. (2021) employed QIE to study, among several other questions, the effect of automobile lane departure warning (LDW) systems on the incidence of "lane departure" accidents by automobiles. These authors employed "rear-end struck" accidents, accidents in which a vehicle is rear-ended by a following vehicle as the control type of accidents, given that the likelihood of being rear-ended by another vehicle was assumed to be independent of the presence of LDW. These authors estimated the odds of a lane departure accident for LDW-equipped vs. unequipped vehicles using a logistic regression analysis with a lane departure accident indicator as the dependent variable, an indicator of the presence of LDW on a vehicle as the key independent variable, and a set of controls for vehicle characteristics, driving conditions, and driver characteristics.

QIE analysis uses logistic regression with an indicator variable indicating whether a given event was of the accident type of interest (rather than the control accident type) as the dependent variable (Jiang et al., 2014; Keall and Newstead, 2009; Leslie et al., 2021; Stamatiadis and Deacon, 1997). Continuing the previous example, Leslie et al. (2021) conduct a logit regression analysis with an indicator that an accident is a "lane departure" accident (rather than a "rear-end struck" accident) as the dependent variable and the presence of a LDW system as the independent variable.

Odds ratios estimated using this logistic regression approach are essentially equivalent to the raw odds ratios that could be obtained with proxy (quasi-induced) exposure data, but have the advantage that they may include control variables to account for variables other than the factor of interest that are known or suspected to affect the occurrence of the accident type of interest.

Although, to our knowledge, QIE has not previously been used to study rail accidents, we believe the approach is well suited to study freight train derailments so long as a control type of accident can be identified. The advantages are that train derailment risk as a function of train length can be systematically studied without obtaining train length data for the exposed population of trains (i.e., data that are not available), and this systematic study can include other factors that are known to impact train derailment risk to isolate the effects of train length. The disadvantage of this approach is that it requires the assumption that the control accident is independent of the factor of interest (i.e., train length), but we can demonstrate the impact of this assumption on the analysis results using simulated data (see section 4.1).

2.2. Derailment risk factors

Freight train derailment is an issue of significant interest to railroads, rail regulators, and communities because the impacts of a derailment can be substantial (Li et al., 2018; Liu et al., 2017, Liu et al, 2012; Kaeeni et al., 2018). Indeed, derailment is the most common type of serious train accident in the U.S. (Cao et al., 2020; Liu et al., 2017; Wang et al., 2020; Liu, 2016). Extant research examines several different risk factors for derailment. For example, a number of studies examine the effects of railroad track segment characteristics—such as traffic volume, track class, and method of operation—on risk of derailment (Anderson and Barkan, 2004; Liu et al., 2017; Nayak et al., 1983; Wang et al., 2020). This work has consistently found that derailment risk is lower for higher track classes, which is not surprising because higher track classes are required to be built to more stringent standards and maintained more carefully (Anderson and Barkan, 2004; Liu et al., 2017; Nayak et al., 1983). These studies also find that derailment risk is lower on track segments with higher traffic densities, due to more frequent inspection and maintenance of high-density track (Liu et al., 2017; Wang et al., 2020). Finally, this literature also finds that derailment risk is reduced on track segments where the method of operation is signaled relative to non-signaled segments (Liu et al., 2017).

Other literature has examined the relationship between characteristics of trains and the risk of derailment. For example, studies have examined the derailment risk of loaded relative to unloaded trains, finding significantly increased risk of derailment for loaded trains over unloaded ones (Li et al., 2018; Zhang et al., 2022). Additionally, Zhang et al. (2022) also studied the derailment risk of unit trains (trains composed of only one railcar type carrying only one type of freight) compared to mixed trains (trains with multiple railcar and freight types). They found that derailment risk per railcar-mile and per ton-mile of freight is lower for unit trains than for mixed trains.

2.3 Exposure data in studies of derailment

As noted above, one of the biggest hurdles to studying the relationship between freight train length and derailment risk is lack of exposure data by train length. Traditional accident analysis methods would evaluate the impact of a factor (like train length) on the odds that a certain type of accident (such as derailment) occurs using data on both the number of accidents that occurred when the factor was and was not present and the number of non-accidents occurrences (exposures) when the factor was and was not present. In the context of studying derailment, such exposure data would constitute the data on the number of trains of various lengths that did not derail.

Data on train length for trains involved in a derailment are available from the FRA—with length operationalized as the number of railcars making up the train, a common measure of freight train length (Multer et al., 2022; Zhang et al., 2022). However, data on how many safely completed, non-accident trips have been made by trains of different lengths (exposure data) are not readily available. Some have attempted to solve this challenge by employing high-level, nationwide exposure data. For example, Nayak et al. (1983) estimated the effect of track class on derailment risk using nationwide estimates of rail traffic flows by track class as exposure data. More recently, Liu et al. (2017) studied the effects of track class, method of operation, and rail traffic density on derailment risk using aggregate systemwide traffic data collected from each of the major U.S. railroads as exposure data. While each of these extant approaches was well suited to the research question it was designed to address, none of them would allow us to address our question of interest because the available nationwide or systemwide data on non-accident rail trips are simply not broken down by train length. Because our analysis of interest requires exposure data by train length that does not exist, we required a different type of exposure data than any reported in the extant literature.

3. DATA

3.1. Derailment data

Freight train derailments in the U.S. are tracked by the FRA's Rail Equipment Accident (REA) database, which reports information filed by railroads that experience incidents or accidents that result in rail equipment and/or track damage above a certain threshold (currently $11,500).

This information is reported to the FRA by involved railroads via form FRA F 6180.54. These REA reports contain information on several characteristics of the accident including but not limited to: accident type, accident date and time, accident location, reporting railroad, rail equipment type, visibility condition at the time of the accident, number of loaded and unloaded railcars, number of locomotives, and the number of railcars transporting hazardous materials. The FRA REA database is the primary source of information available on train derailments in the U.S. and has been used in virtually all prior work on derailments for U.S.-based railroads (Liu et al., 2017; Zhang et al., 2022).

To ensure that our analysis reflects only recent derailments, we extracted from the REA database information on all derailments occurring within the past 10 years (between 2013 and 2022), about 14,000 incidents. Given that our interest was in derailments of freight trains during normal freight transport (as many derailments occur in train yards and sidings outside of freight transport), we retained in the derailment sample only derailments whose "rail equipment type" is identified as "freight train" (rather than passenger train or other types of rail equipment) and that occurred on a mainline track (as opposed to a yard, siding, or industrial track). These restrictions follow those used in prior work on derailments (e.g., Liu et al., 2017). Our final REA derailment sample, after these restrictions, included 2,906 derailments.

3.2. Highway Rail Grade Crossing Incident Data

QIE analysis requires a control type of accident that is unrelated to the independent variable of interest. Importantly, the QIE method does not require that the control accidents and the accidents of interest result from the same causes (Stamatiadis and Deacon, 1995). The primary requirement for effective control accidents is that their occurrence is independent of the explanatory variable of interest (Jiang et al., 2014). When this is the case, the incidence of control accidents in a location provides a reasonable approximation of a random sample of vehicles that transverse that location with respect to the explanatory variable being studied (Stamatiadis and Deacon, 1997). This quasi-random sample, then, may be used to represent the true exposure data (when it is not available). If the control accidents were not independent of the explanatory variable, their use to represent exposure data would bias the results of the analysis; if they were positively (negatively) correlated with the explanatory variable, their use would serve to artificially inflate (deflate) the prevenance of that variable in the exposure set. However, to the extent that the control accidents are unrelated to the explanatory variable, they can act as a quasi-random sample of exposure data in that location, whatever their causes.

As the control accident type in the analysis, we employ highway-rail grade-crossing incidents where the driver of a road vehicle attempted (unsuccessfully) to "beat the train" across the grade crossing. We chose this accident type as our control as it is unrelated to train length since road vehicle drivers at rail crossings would rarely have the sight perspective to judge train lengths greater than a few cars, making the accidents a direct result of the driver's actions and not train length (Oh, Washington, and Nam, 2006). Due to long stopping distances, freight trains of any length would not be able to slow significantly before impact in this type of accident (Bentley and Bentley, 2007). Thus, the occurrence of this type of accident should be independent of train length and should serve well as a control accident type in a QIE analysis.

The data on highway-grade accidents used to build the sample of control accidents comes from the FRA Highway-Rail Grade Crossing Accident (HRGCA) database. The HRGCA contains data on highway-rail grade-crossing incidents that are incidents involving contact between trains and road vehicles or pedestrians at locations where roads and rail tracks cross at the same grade level. Involved railroads are required to report any contact between railroad equipment and highway users at a highway-rail grade crossing, no matter the severity of the incident, to the FRA (on form FRA F 6180.57). HRGCA reports contain much of the same information about involved rail equipment as the REA reports, including all of the characteristics mentioned above (with the exception that the HRGCA data include only an indicator for whether any railcar in a train contains hazardous materials rather than the number of cars transporting hazardous materials). HRGCA reports also contain information about the grade crossing where the incident occurred, the highway user involved in the incident, and damage to highway vehicles resulting from the incident. The HRGCA database is the main source of information about grade crossing accidents in the U.S. and is commonly used in research on these events (Liu et al., 2015; Lu and Tolliver, 2016).

For the initial HRGCA sample, we again extracted only grade crossing events occurring between 2013 and 2022, about 21,000 events. We also restricted the sample to include only events that occurred on a mainline track and that involved freight trains (rather than passenger trains or other types of rail equipment). These restrictions reduced the HRGCA sample to about 14,000 events. Finally, we retained in the HRGCA sample only grade crossing events in which a highway vehicle attempted to beat the freight train across the grade crossing. We identified these events as those where the road vehicle was impacted by the front of the freight train and was moving at the time of impact (rather than having been stuck unmoving at the crossing). This restriction resulted in a final HRGCA sample of 8,092 HRGCA events. Although some events can appear both in the HRGCA database and the REA database (when a grade crossing incident results in significant damage to rail equipment), there is no overlap between our REA and HRGCA samples because FRA classifies all such events as “highway-rail” incidents and we only extract REA events classified as derailments.

3.3. Final Sample Construction

To ensure that our sample of control accidents matched our derailments sample as closely as possible (as required by the QIE method), we constructed our final analysis sample by including only events for which there was a geographic and temporal match between derailments and control accidents. Specifically, we retained in the final sample only derailments for which one or more control accidents in our control sample had occurred in the same county and year, and we retained only the control accidents that occurred in the same county and year as a derailment. This geographic and temporal matching resulted in a final analysis sample composed of 2,758 events (1,073 derailments and 1,585 control accidents). We carried out this geographic and temporal matching using the "exact" function of the Matchit package in the statistical programming language R. The distributions of train lengths for both derailments and control accidents in the sample are illustrated in Figure 1.

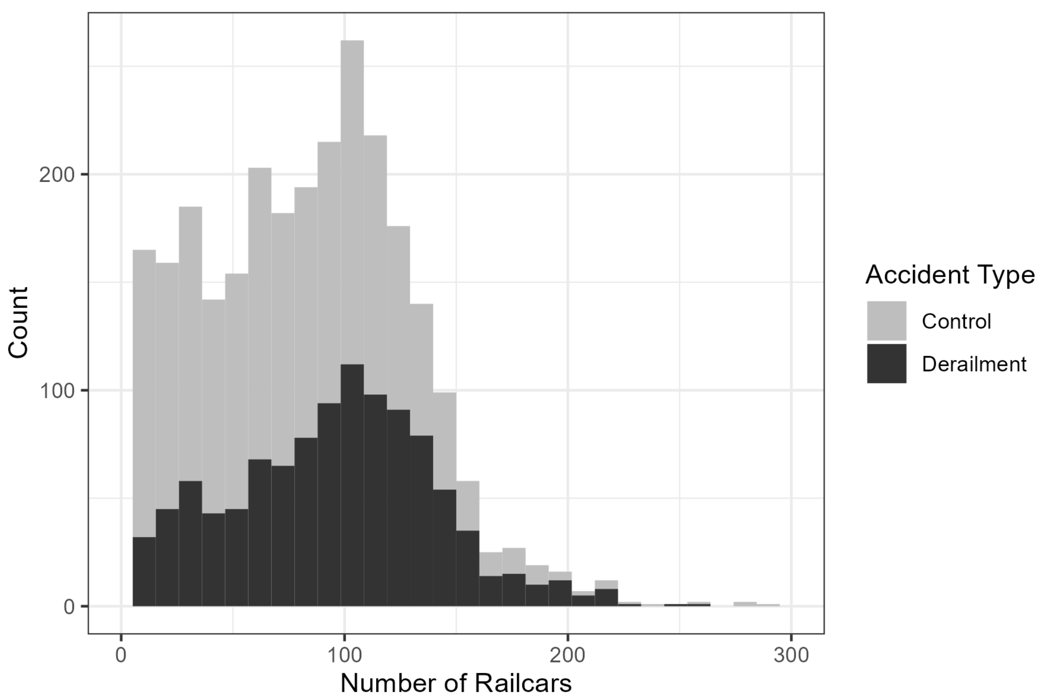

Fig. 1. Distributions of the Lengths of Trains Involved in Derailments and Control Accidents

4. ANALYSIS APPROACH

4.1. QIE Analysis

A traditional analysis of the influence of a factor on accident risk would estimate the odds of an accident given the presence of the factor of interest relative to the odds in the absence of the factor. Thus, the odds ratio would be determined by the following formula:

$$\frac{\#\,accidents\ when\ the\ factor\ is\ present \big/ \#\,non-accident\ exposures\ when\ the\ factor\ is\ present}{\#\,accidents\ when\ the\ factor\ is\ absent \big/ \#\,non-accident\ exposures\ when\ the\ factor\ is\ absent} \quad (1).$$

Thus, the odds ratio of derailment for a train of a certain length (length 1) relative to a reference length (length 2) could be estimated as:

$$\frac{derailments\ by\ trains\ of\ length\ 1 \big/ safely\ completed\ trips\ by\ trains\ of\ length\ 1}{derailments\ by\ trains\ of\ length\ 2 \big/ safely\ completed\ trips\ by\ trains\ of\ length\ 2} \quad (2).$$

The REA database contains data on train length. These data can be used to determine the number of derailments that have occurred by trains of different lengths. However, data on how many safely completed, non-accident trips have been made by trains of different lengths (exposure data) are not readily available, making analysis of the relationship between train length and derailment risk impossible using traditional techniques. We overcame the lack of available exposure data by train length by employing the QIE method and identifying a type of control accident whose occurrence is unrelated to the factor being studied. Since the control accident type is assumed to be closely correlated with exposure, but independent of the factor of interest, the control accidents provide an excellent estimate of exposure to accident risk, acting as a quasi-random sample of exposures (Jiang et al., 2014; Keall and Newstead, 2009; Leslie et al., 2021; Stamatiadis and Deacon, 1997). Using incidence of the control accident as a proxy for exposure to risk, then, the relationship between a factor of interest and the odds of a type of accident of interest may then be estimated as:

$$\frac{\text{accidents of interest when factor is present} \Big/ \text{control accidents when factor is present}}{\text{accidents of interest when factor is absent} \Big/ \text{control accidents when factor is absent}} \quad (3).$$

As discussed above, we identified "beat the train" grade crossing accidents as a good control accident type for our analysis. Studies of the causes of grade crossing accidents point to a number of factors that affect the likelihood of grade crossing accidents, including driver characteristics and road vehicle characteristics, but no study that we could find suggests train length as a possible contributing factor (Davey et al., 2008; Oh et al., 2006; McCollister and Pflaum, 2007). We summarize again the two primary reasons why this accident type can serve as a control accident type as follows: 1) because road vehicle drivers very rarely have the visibility perspective to gauge the length of an oncoming train (Oh et al., 2006; McCollister and Pflaum, 2007), the likelihood that the driver will attempt to beat the train should be independent of train length, 2) given that freight trains of any length have very long stopping distances (Bentley and Bentley, 2007), the driver of a freight train of any length would be unable to slow a train significantly in the time between observing the road vehicle attempting to cross the tracks and impact. Thus, we concluded that train length was sufficiently unrelated to "beat the train" grade crossing accidents to make this type of accident an effective control accident for our analysis of derailments, and thus serve as a quasi-random sample of exposure data.

Given that QIE has not previously been used to study rail accidents, we carried out a simulation to explore the conditions under which QIE analysis of derailments would result in unbiased vs. biased estimates of the influence of train length on derailment. To carry out this simulation analysis we created 9 different samples of simulated data representing every combination of train length being positively associated, negatively associated, or independent of both derailments and control accidents. We then analyzed the relationship between train length and derailment risk in the simulated data using both exposure-based analysis of the full samples (which is possible with the simulated data because they contain non-accident trips as well as trips that result in accidents) and QIE analysis of only the subsamples of the simulated trips that result in a derailment or a control accident. We then compared the relationship between train length and derailment risk in a sample for the exposure-based analysis (which is unbiased) compared to that for the QIE analysis (which will be biased if the QIE assumptions are not met).

Results of the simulation analysis showed that when the risk of the control accident is independent of train length, the results of the exposure-based analysis and the QIE analysis are virtually identical, indicating that the QIE results are unbiased (as discussed above). However, when the control accident risk is positively associated with train length, the QIE analysis results are biased downward, and when the control accident risk is negatively associated with train length, the QIE analysis results are biased upward. Full details of the simulation analysis and results are reported in Appendix A.

The simulation results suggest that if our assumption that the risk of the occurrence of the control, "beat the train," type of accident is independent of train length is not accurate, the QIE analysis results of the relationship between train length and derailment will be conservative (biased downward) if control accidents are more likely to occur for longer trains. In particular, if, drivers are more likely to try to beat longer trains across a grade crossing because they can see how long a train is and because waiting for a longer train would take more time then the results of the QIE analysis would be conservative. On the other hand, the size of the relationship between train length and derailment in the QIE analysis results will only be inflated in the case that the risk of the control accident is reduced for longer trains, i.e., if, drivers are less likely to try to beat longer trains across a grade crossing. It seems more likely that drivers would be more (rather than less) likely to try to beat longer trains, if indeed there is an association between train length and "beat the train" accidents. Thus, the simulation results demonstrate that even if our assumption that risk of a control accident is independent of train length does not hold, the QIE results presented below would be conservative (rather than inflated) estimates.

4.2. Variables

Following standard QIE methods, we carried out our analysis of the final sample using logistic regression, with an indicator variable that took a value of 1 for derailments and a value of 0 for control accidents as the dependent variable. The independent variable in the analysis was train length measured in the number of railcars composing the train. We included in the analysis as controls a set of variables that have been shown in prior work to influence derailments or that were relevant to our context. First, given prior work showing that track class influences derailment rate (Liu et al., 2017; Nayak et al., 1983; Wang et al., 2020), we included a set of indicator variables for the class of the track on which each accident occurred. Track classes in our sample included FRA track classes 1, 2, 3, 4, 5, 6, 8, 9, and X (the omitted category). Second, since prior work finds an effect of the method of operation on derailment (Liu et al., 2017), we controlled for method of operation (signaled or un-signaled) as an indicator variable. Third, to account for the possible effects of visibility on derailments and grade crossing incidents, we included a control for the visibility conditions at the time of the incident (dawn--the omitted category, day, dusk, or dark).

Fourth, because trains transporting hazardous materials may be managed with extra precautions (Zhang et al., 2022), we included an indicator variable that takes a value of one for trains transporting any hazardous materials, and a value of zero otherwise. Fifth, we included a control for the number of locomotives in the train (per 100 railcars). To control for temporal effects on derailment, we included fixed year and month effects. Finally, to account for variation in derailment risks across railroads, we included fixed railroad effects in the model. Unfortunately, we were unable to include a control for traffic density or whether a train was loaded or unloaded because, although the REA database reports these variables, the HRGCA database does not. Nonetheless, we expect that our geographic matching of derailments and control accidents partially controls for these factors.

We considered several different functional forms for modelling the relationship between train length and derailment risk, by testing linear, quadratic, cubic, logarithmic, and exponential transformations of train length. Because the logarithmic transformation produced the best model fit, we employed this functional form in the reported models. However, the reported results are robust to this modelling choice since the same pattern of results is obtained when any of the other functional forms are used instead. The logistic regression models were estimated using the glm function in R. The relationship between train length and derailment odds ratio estimated using the logistic regression model is virtually identical to that obtained estimating the basic QIE odds ratio using the following equation:

$$\frac{derailments\ among\ trains\ of\ length\ 1 \big/ control\ accidents\ among\ trains\ of\ length\ 1}{derailments\ among\ trains\ of\ length\ 2 \big/ control\ accidents\ among\ trains\ of\ length\ 2} \quad (4).$$

4. RESULTS

4.1. QIE Logistic Regression Results

The results of the logistic regression model are presented in Table I. Model 1 shows the results for the full sample, model 2 for only class 1 railroads (those with greater than $943.9M in annual revenue), and model 3 for only smaller (class 2 and class 3) railroads.

Table I. Results of the QIE Logistic Regression Analysis

	Model 1				Model 2				Model 3			
	Full Sample				Class 1 Railroads				Class 2 & 3 Railroads			
Variable	coef	se	p		coef	se	p		coef	se	p	
Train Length (logged)	1.15	0.10	0.000	***	1.19	0.11	0.000	***	1.37	0.34	0.000	***
Track Class 1	-1.01	0.63	0.108		-1.47	1.17	0.210		-0.19	0.89	0.834	
Track Class 2	-2.00	0.63	0.001	**	-2.66	1.16	0.022	*	-0.71	0.91	0.434	
Track Class 3	-2.86	0.63	0.000	***	-3.47	1.16	0.003	**	-2.55	1.03	0.013	*
Track Class 4	-2.97	0.63	0.000	***	-3.58	1.16	0.002	**	-2.28	1.31	0.082	
Track Class 5	-2.64	0.64	0.000	***	-3.27	1.17	0.005	**				
Track Class 6	-3.75	1.02	0.000	***	-4.37	1.41	0.002	**				
Track Class 8	-20.14	2400	0.993		-16.78	325	0.959					
Track Class 9	14.13	2400	0.995		9.78	325	0.976					
Signaled Operation	0.73	0.11	0.000	***	1.01	0.12	0.000	***	-1.88	0.47	0.000	***
Visibility - Day	0.01	0.15	0.947		-0.02	0.16	0.879		1.08	0.75	0.148	
Visibility - Dusk	0.03	0.20	0.877		0.02	0.20	0.919		1.34	1.12	0.233	
Visibility - Dark	0.38	0.16	0.017	*	0.37	0.16	0.023	*	0.85	0.72	0.236	
Hazmat	0.24	0.10	0.011	*	0.23	0.10	0.019	*	0.44	0.48	0.359	
Locomotives per 100 cars	0.02	0.01	0.017	*	0.03	0.01	0.002	**	0.02	0.02	0.511	
Year Fixed-Effects	included				included				included			
Month Fixed Effects	included				included				included			
Railroad Fized Effects	included				included				included			

In Model 1, the coefficient for train length for the full sample is positive and significant, indicating the risk of a train experiencing a derailment increases as train length goes up. To illustrate the magnitude of this positive relationship, it is helpful to compare the odds of derailment for a train of a given length relative to that of a baseline train length. As an example, we chose a 50-car train as the baseline train length and calculated odds ratios for the odds of derailment for trains of between 1 and 250 cars in length relative to that of a 50-car train. These odds ratios (along with 95% confidence intervals) are presented in Figure 2. Odds of derailment as well as 95% confidence intervals were computed using the "predict" function in R based on the results from Model 1 in Table I. The figure shows a positive relationship between train length and the odds of derailment. A value of 2 in the figure (for a roughly 90-car train) reflects 2 times (or double) the odds of derailment compared to a 50-car train.

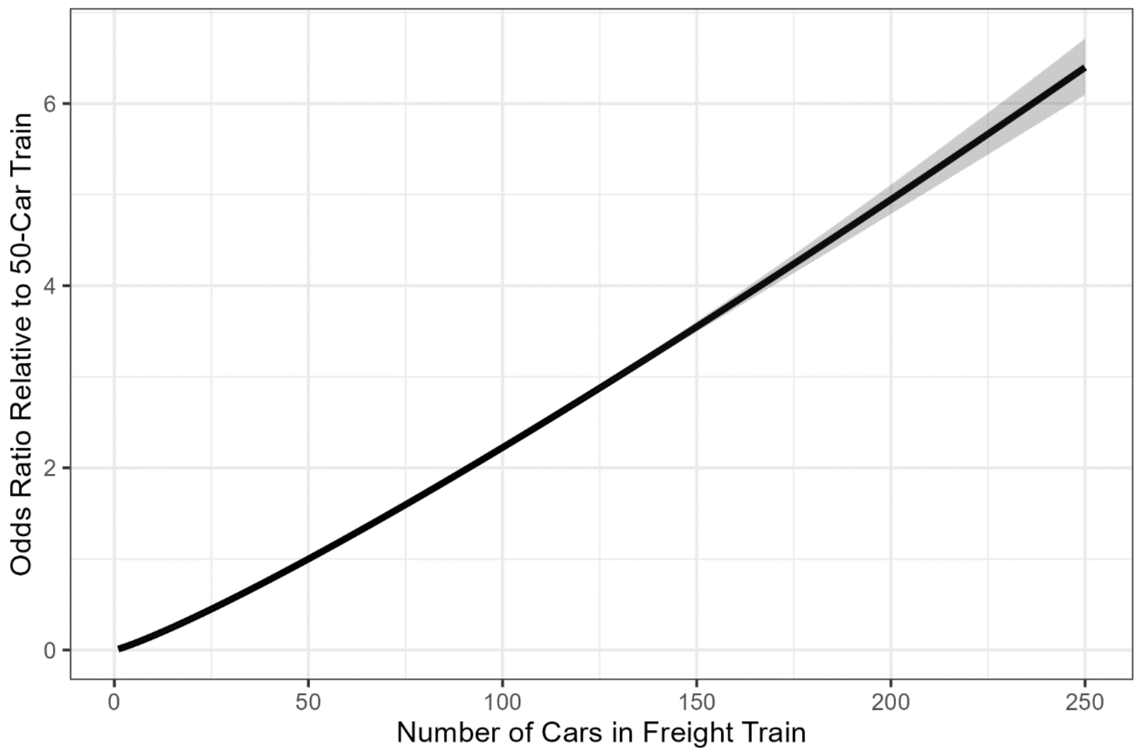

Fig. 2. The Odds Ratio of Derailment (along with 95% Confidence Interval) by Number of Cars in a Train Relative to a 50-Car Train.

Several of the control variables are also significantly related to the odds of derailment in Model 1. The odds of derailment are lower for higher track classes as found in prior work (Liu et al., 2017; Nayak et al., 1983; Wang et al., 2020). On the other hand, derailment odds are higher when the method of operation is signaled rather than non-signaled. This finding is not in line with prior work (e.g., Liu et al., 2017) but probably reflects the influence of mode of operation on "beat the train" highway-rail crossing accidents in that drivers are less likely to try to beat trains in areas where trains are signaled. The odds of derailment increase when visibility conditions are categorized as "dark", relative to the other visibility conditions (dawn, day, and dusk), and the odds of derailment also increase as hazardous material is present and as the number of locomotives per 100 railcars increases.

Model 2 reports results for logistic analysis of the subsample of our data that only includes data on class 1 (large) railroads, and Model 3 reports results for the subsample of smaller class 2 and class 3 railroads. In both models, the train length coefficient is positive and significant, indicating a meaningful positive relationship between train length and derailment risk. The magnitude of this relationship is illustrated in Figure 3, which shows the odds ratios of derailment for trains of varying lengths relative to a baseline 50-car train for both class 1 and class 2 and 3 railroads. The odds ratios are estimated using the "predict" function in R using the results from Models 2 and 3 in Table I. The relationship between train length and derailment odds for class 1 railroads is virtually identical to that for the full sample (which is not surprising given that the bulk of the sample came from class 1 railroads), while the relationship for class 2 and 3 railroads is somewhat more positive. This result indicates that the positive relationship between train length and derailment odds is somewhat stronger for class 2 and 3 railroads compared to class 1 railroads.

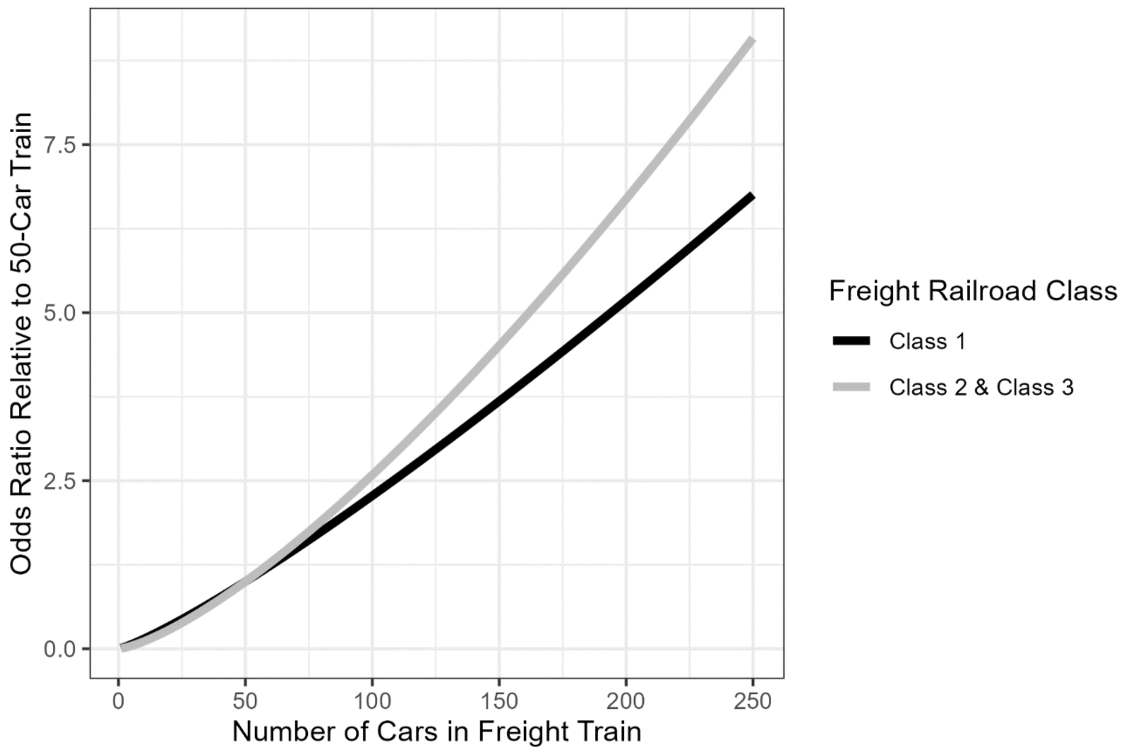

Fig. 3. The Odds Ratio of Derailment by Number of Cars in a Train Relative to a 50-Car Train for Class 1 Railroads Compared with Class 2 and 3 Railroads

4.2. Accounting for the effect of train length on rail-system-wide derailment exposure

The results displayed above suggest a strong positive relationship between train length and derailment risk. However, these results represent this relationship only for an individual freight train and, thus, may overstate the system-wide impact of train length on derailment risk for the whole freight rail system because the use of longer trains allows the same amount of freight to be transported on fewer trains (Zhang et al., 2022). The use of longer trains inherently implies fewer train trips overall, corresponding to a reduction in the aggregate exposure to derailment risk. This exposure reduction effect will operate in the opposite direction to the positive relationship between train length and derailment risk. The exposure reduction associated with freight trains of different lengths relative to a baseline 50-car train is calculated as exposure equals 50 cars divided by the number of railcars in the train length of interest. Thus 25-car trains would have twice the exposure of 50-car trains and 100-car trains would have half of the exposure. This exposure effect is illustrated in Figure 4 for trains of varying lengths.

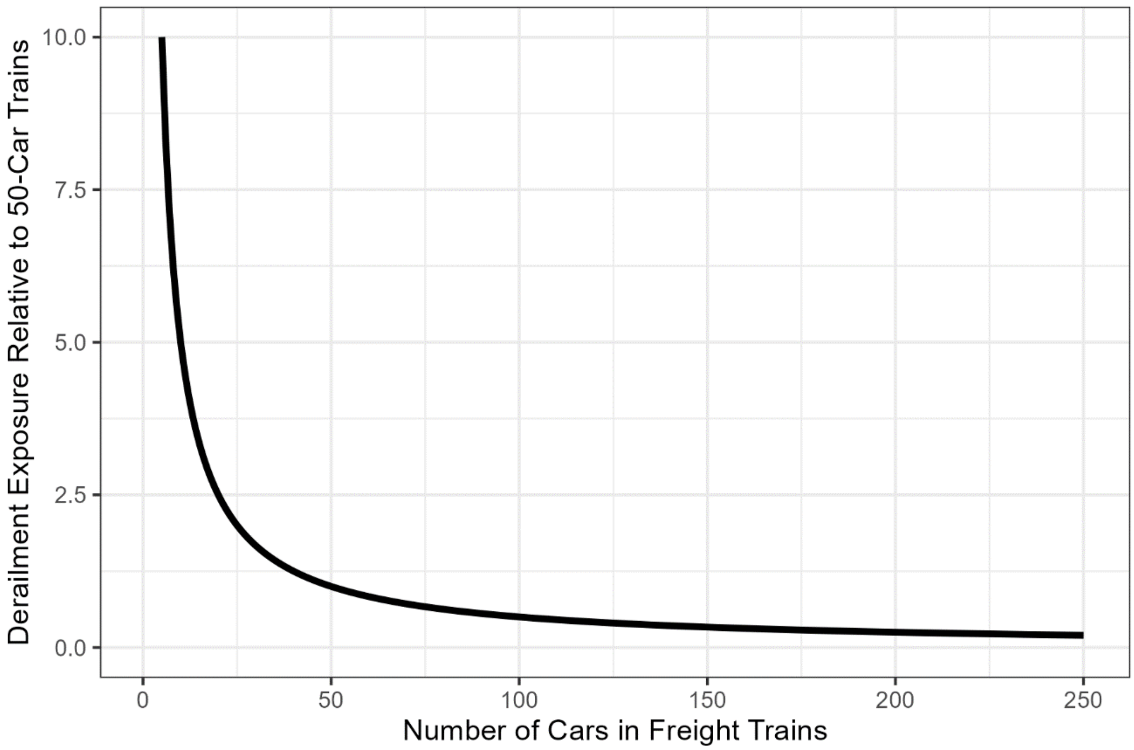

Fig. 4. The Derailment Exposure of Trains of Varying Lengths Relative to 50-Car Trains

This effect of longer trains on derailment exposure must be accounted for to understand the full relationship between train length and derailment risk for the overall U.S. rail system. An aggregate estimate of the overall relationship is estimated by multiplying the odds ratio of derailment for a given train length (relative to the baseline 50-car train) by the derailment exposure effect for that train length. This aggregate estimate is the most complete representation of the full relationship between train length and derailment risk because it accounts for both the increase in derailment risk for individual trains as they become longer and the decrease in derailment exposure that longer trains create in the overall freight rail system. This aggregate relationship (relative to a baseline 50-car train) is illustrated in Figure 5 (along with its 95% confidence interval) based on the results of the QIE logistic regression analysis in Model 1 of Table I.

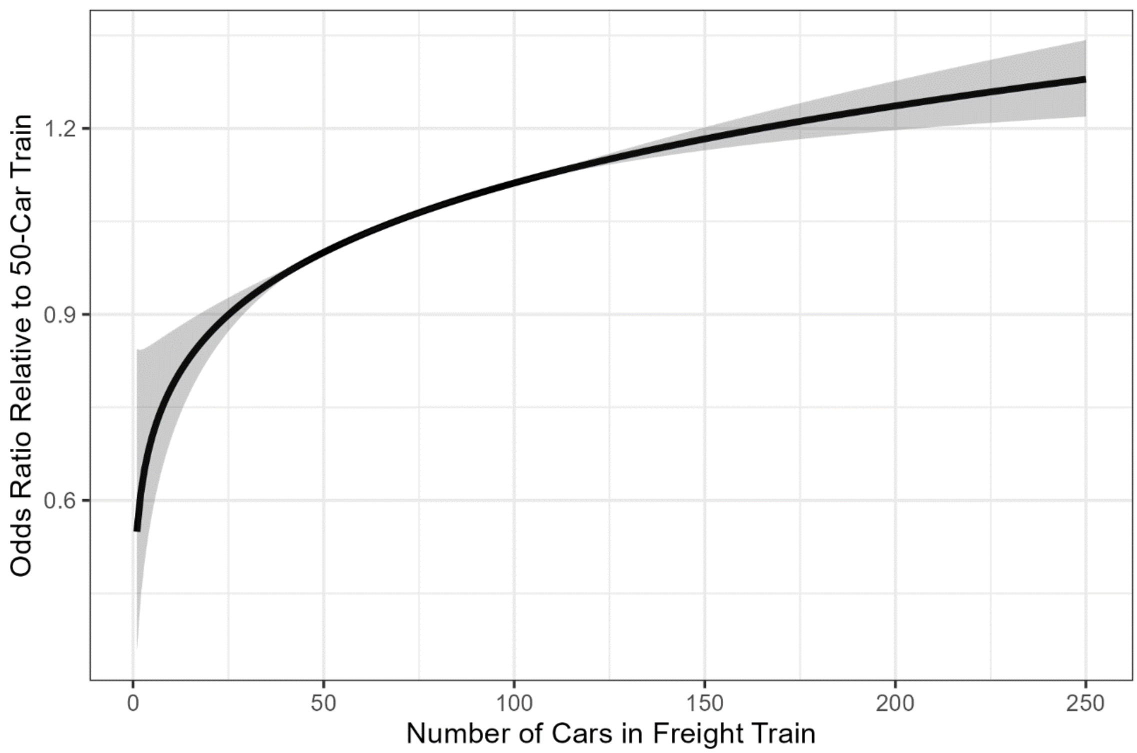

Fig. 5. The Aggregate Derailment Odds Ratio for Trains of Varying Lengths Relative to 50-Car Trains Accounting for the Reduction in Derailment Exposure for Longer Trains

The aggregate relationship between freight train length and odds of derailment is positive and meaningful in size. The results show that 100-car trains are associated with 1.11 (95% CI 1.10 to 1.12) times the derailment odds of 50-car trains (or a 11% increase), accounting for the exposure reduction given that only half as many 100-car trains would be needed to transport the same amount of freight. For 200-car trains, the derailment odds increase by about 24% (OR 1.24, 95% CI 1.20 to 1.28) net of the exposure reduction. Thus, longer trains are associated with increased derailment risk even with the reduction in exposure accounted for. This increase in derailment risk is statistically significant.

The aggregate relationship between train length and derailment risk (relative to a baseline 50-car train) for class 1 compared with class 2 and 3 railroads based on the results of the QIE logistic regression analysis in Models 2 and 3 of Table I is illustrated in Figure 6. As before, the positive relationship between train length and derailment risk is stronger for trains operated by class 2 and 3 railroads relative to those operated by class 1 railroads.

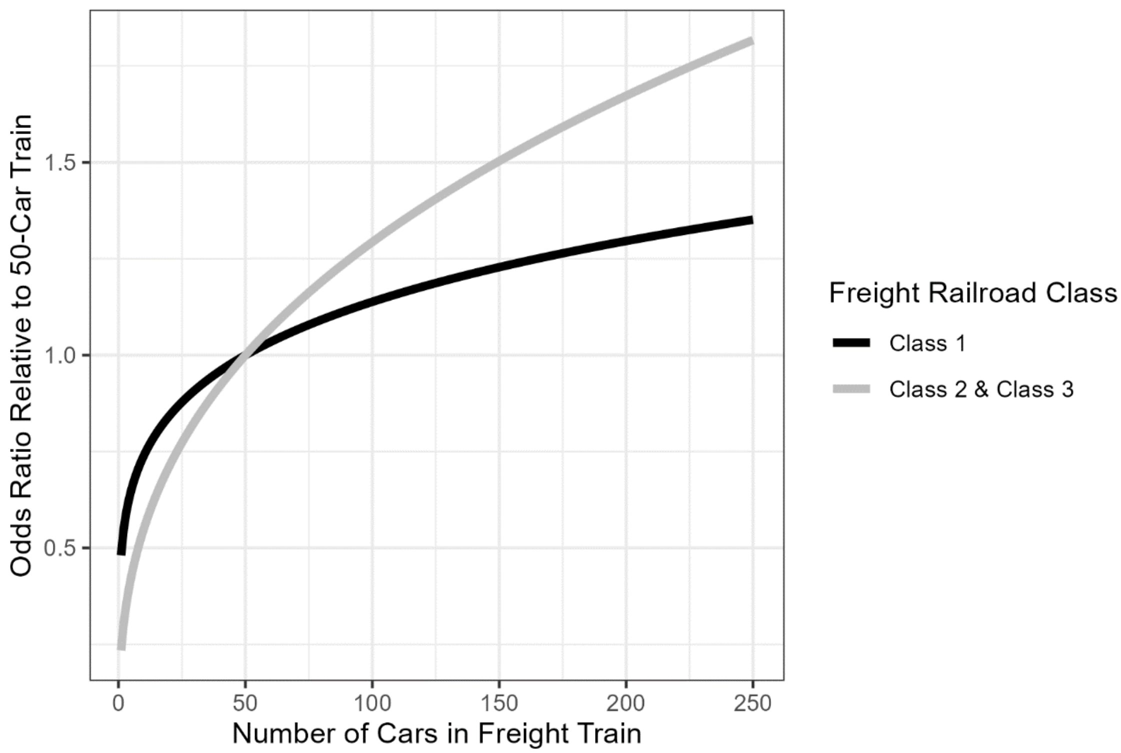

Fig. 6. The Aggregate Derailment Odds Ratio for Trains of Varying Lengths from Class 1 Compared to Class 2 and 3 Railroads Relative to 50-Car Trains Accounting for the Reduction in Derailment Exposure for Longer Trains

We should note that our analysis used the number of railcars in the train as its measure of train length. This is a common measure of train length, but other measures such as linear feet are also commonly used. Our results may be translated to train length in linear feet by multiplying the number of cars by an average car length of 62 feet (Dick et al., 2021).

5. CONCLUSIONS

This paper describes a quantitative method for analyzing the risk of derailment when considering train length. It provides a statistical procedure to examine the available accident data from the freight train industry to support policy making to address derailment risk. Derailment risk, however, is just one factor that needs to be considered by policy makers. Longer freight trains have many significant benefits for the rail system relative to shorter freight trains including greater fuel efficiency, lower emissions per ton transported, and lower operational costs than both shorter trains and many other forms of transportation (GAO, 2019; Muller et al., 2022). The operation of longer freight trains also come with costs such as increased wait times at railroad grade crossings for road vehicles in communities where freight trains frequently operate (GAO, 2019). Additionally, as recent cases like the East Palestine derailment demonstrate, derailments can have significant negative environmental and health impacts on communities, although rail remains a safer mode of transportation for hazardous chemicals than other options (Bagheri, Verma, & Verter, 2014). As the Railway Safety Act of 2023 is debated including the pros and cons of longer trains, an important consideration in these debates is the additional risk of derailment in the system that comes with longer trains. Until this time, quantifying this relationship has been elusive. However, the model described here provides a process for analyzing this relationship, and the results presented suggest a clear, monotonic, and positive relationship between freight train length and derailment risk. Even when accounting for the reduction in the number of freight trains operated when the average train is longer, longer freight trains are associated with an increase in the aggregate odds of freight train derailment.

Knowing the direction and estimated size of the relationship between train length and derailment risk is important information when considering the future development and regulation of freight rail transportation, as it allows derailment risk to be more accurately weighed against the other costs and benefits of longer freight trains for the overall system. Additionally, understanding the risk could spur additional innovations in preventive measures. As demonstrated by the differences in risk between trains operated by larger, class 1 railroads compared to small railroads, more stringent standards and better maintenance can help address the risk of longer trains. This research could encourage the FRA to collect additional data on length of trains not involved in accidents so that the exposure risk can be further studied.

Finally, this research demonstrates the applicability of QIE to safety incident analysis beyond its previous use in the study of road traffic accidents. While the procedure has broader applicability to many other industries, the limitations are that incident data for the case study of interest and a control that is independent of the factor of interest are needed, and the results are relative ratios not explicit rates. Therefore, the technique can be useful in many situations but not for all applications.

Acknowledgements: The research was supported by the National Science Foundation Grant: 2051685, however, any opinions, findings, and conclusions or recommendations in this document are those of the authors.

APPENDIX A.

As noted above, quasi-induced exposure (QIE) is an analytic technique designed to study how factors influence the likelihood of accident occurrence in contexts where exposure data are limited or unavailable. QIE is an accepted methodology in the study of road accidents, but does not appear to be well known in the broader risk analysis community. For example, we can find no examples of QIE being used to study rail accidents.

Here, we use QIE to study the influence of train length on the likelihood of derailment. Our QIE results suggest a significant positive relationship between freight train length and the risk of a derailment. However, the validity of this finding is contingent on the key assumption of the QIE method being met—namely that the "control" accident type is genuinely independent of train length. Given the data limitations that require us to use QIE analysis in the first place (i.e., the lack of exposure data by train length), we cannot obtain the data necessary to fully test this assumption empirically. To explore the ramifications for our QIE analysis if this assumption is not met, we conducted an analysis of simulated data allowing for the observation of exposure by train length. In the analysis of the simulated data, we explored how results of a QIE analysis on a subset of a data set would differ from those of a traditional, exposure-based analysis of the whole data set.

We first simulated 9 different samples of simulated data representing every combination of train length being positively associated with, negatively associated with, or independent of both the risk of derailments and control accidents. Each of the 9 simulated samples contained 500,000 observations of simulated train trips. In all 9 samples, the length of the freight train making a given trip was randomly drawn from a normal distribution with a mean of 100 cars and a standard deviation of 25 cars. Indicators of whether a given trip ended in a derailment or a control accident were randomly drawn from binomial distributions with different levels of risk. The risk of derailment was set as 0.001 (or 1 in 1000 trips) for the conditions in which derailment was assumed to be independent of train length, 0.001 * (train length / 100 cars) for the conditions in which derailment was assumed to be positively associated with train length, and 0.001 * (100 cars / train length) for the conditions in which derailment was assumed to be negatively associated with train length. Because control accidents were more common in our data than derailments, the risk of a control accident was set at 0.005 when control accidents were independent of train length, 0.005 * (train length / 100 cars) when control accidents were positively associated with train length, and 0.005 * (100 cars / train length) for conditions where control accidents were negatively associated with train length. For each simulated observation, a train length was randomly determined and then given the train length, whether that train observation was a derailment accident or a control (crossing) accident was determined.

We then analyzed the relationship between train length and derailment risk in each of the simulated data samples using logit regression with the derailment indicator as the dependent variable and the train length as the independent variable. First, exposure-based analysis was carried out via logit regression of the full sample. Second, QIE analysis was carried out via logit regression of the subsample of trips that ended in either derailment or a control accident. This second model mimics the real-world conditions under which QIE is used, when data on accidents is available but data on non-accident exposure is not. Because the exposure-based analyses return unbiased estimates of the relationships between train length and derailment risk, comparison between the results of the exposure-based analysis and the QIE analysis shows the conditions under which the QIE analysis results are biased and unbiased.

Table AI presents the results of the analysis of the 9 different simulated samples representing different assumptions about the relationships between train length and derailments and control accidents.

Table AI. Results of Analysis of Simulated Data. Exposure Coef. stands for the coefficient of train length in the exposure-based analysis. QIE Coef. stands for the coefficient of train length in the QIE analysis.

Condition	Exposure Coef.	p-value	QIW Coef.	p-value	QIE Sample Size
Derail = Positive; Control = Positive	0.006	0.000	-0.004	0.045	2999
Derail = Positive; Control = Independent	0.012	0.000	0.012	0.000	2993
Derail = Positive; Control = Negative	0.011	0.000	0.022	0.000	3110
Derail = Independent; Control = Positive	-0.001	0.483	-0.012	0.000	3027
Derail = Independent; Control = Independent	0.002	0.390	0.002	0.400	2965
Derail = Independent; Control = Negative	0.000	0.988	0.011	0.000	3226
Derail = Negative; Control = Positive	-0.011	0.000	-0.020	0.000	3043
Derail = Negative; Control = Independent	-0.012	0.000	-0.011	0.000	3064
Derail = Negative; Control = Negative	-0.013	0.000	0.001	0.770	3284

As can be seen in the table, when train length is independent of control accident risk, the QIE analysis results are virtually identical to the exposure-based analysis results, showing that in that case, the QIE results are unbiased. This finding is fully in line with our assumptions about the need for the control accident to be independent of the independent variable in QIE analysis. However, when train length is positively related to control accident risk, the QIE analysis results are biased downward, such that the relationship between train length and derailment risk appears more negative in the QIE analysis than in the exposure-based analysis. And when train length is negatively related to control accident risk, the QIE analysis results are biased upward, such that the relationship between train length and derailment risk appears more positive in the QIE analysis than in the exposure-based analysis.

Given that we observe a significant positive relationship between train length and derailment risk in the QIE analysis of our real-world data, if the risk of control accidents was actually positively related to train length, our reported results would be conservative and the true relationship between train length and derailment risk would be even more strongly positive than what we report. In this case, the magnitude of the relationship we report in the main analysis may be conservative, but the direction of this relationship would remain correct. On the other hand, if the risk of control accidents was negatively related to train length, our reported results could be completely spurious in that a false positive relationship between train length and derailment risk could be created by the QIE methodology in this case. Thus, the greatest threat to the validity of the reported results is the possibility of a negative effect of train length on the risk of the control, "beat the train" grade crossing, accidents.

We remain convinced that it is most likely that the occurrence of "beat the train" grade crossings is indeed independent of train length for the reasons argued in the paper. However, even if this logic is inaccurate and drivers' decisions to attempt to beat trains across grade crossings are related to train length, it seems more likely that attempts to beat the train (and thus the likelihood of being hit by the train) would be more common for longer trains than for shorter trains because the amount of time that drivers could save by beating a train would be greater for longer trains. It appears unlikely that if drivers could perceive the length of an approaching freight train, they would preferentially choose to attempt to beat shorter trains relative to longer trains. Thus, the simulation results suggest that even if the assumptions we made in conducting the QIE analysis do not fully hold, the most likely result would be that the reported QIE results are somewhat conservative in magnitude, but correct in direction, rather than spurious.

REFERENCES

AAR. (2023). Freight Train Length. Retrieved from https://www.aar.org/wp-content/uploads/2023/03/AAR-Train-Length-Fact-Sheet.pdf

Anderson, R. T., & Barkan, C. P. (2004). Railroad accident rates for use in transportation risk analysis. Transportation Research Record, 1863(1), 88-98.

Bagheri, M., Verma, M., & Verter, V. (2014). Transport mode selection for toxic gases: rail or road?. Risk Analysis, 34(1), 168-186.

Bentley, J. M., Bentley, Jr. J. (2007). The freight train emergency brake system and a method to calculate stop distance and time. SAE Transactions, 74-89.

Bernton, H. (2023). Lawmakers consider limits to length of freight trains moving through WA. The Seattle Times. https://www.seattletimes.com/seattle-news/lawmakers-consider-limits-to-length-of-freight-trains-moving-through-wa/

Cao, T., Mu, W., Gou, J., & Peng, L. (2020). A study of risk relevance reasoning based on a context ontology of railway accidents. Risk analysis, 40(8), 1589-1611.

CBS. (2023). States clamp down on freight trains fearing derailments and federal gridlocks. CBS News. https://www.cbsnews.com/pittsburgh/news/states-clamp-down-on-freight-trains-fearing-derailments-and-federal-gridlock/

Congress.gov. (2023). Text - S.576. Washington D.C.: 118th Congress (2023-2024); Retrieved May 10, 2023, from https://www.congress.gov/bill/118th-congress/senate-bill/576/text

Davey, J., Wallace, A., Stenson, N., & Freeman, J. (2008). The experiences and perceptions of heavy vehicle drivers and train drivers of dangers at railway level crossings. Accident Analysis & Prevention. 40(3), 1217-1222.

Dick, C. T., Zhao, J., Liu, X., & Kirkpatrick, S. W. (2021). Quantifying recent trends in class 1 freight railroad train length and weight by train type. Transportation research record, 2675(12), 890-903.

Ghofrani, F., Sun, H., & He, Q. (2022). Analyzing risk of service failures in heavy haul rail lines: A hybrid approach for imbalanced data. Risk Analysis, 42(8), 1852-1871.

Jiang, X., Lyles, R. W., & Guo, R. (2014). A comprehensive review on the quasi-induced exposure technique. Accident Analysis & Prevention, 65, 36-46.

Kaeeni, S., Khalilian, M., & Mohammadzadeh, J. (2018). Derailment accident risk assessment based on ensemble classification method. Safety science, 110, 3-10.

Keall, M., & Newstead, S. (2009). Selection of comparison crash types for quasi-induced exposure risk estimation. Traffic Injury Prevention, 10(1), 23–29.

Leslie, A. J., Kiefer, R. J., Meitzner, M. R., & Flannagan, C. A. (2021). Field effectiveness of General Motors advanced driver assistance and headlighting systems. Accident Analysis & Prevention, 159, 106275.

Li, W., Roscoe, G. S., Zhang, Z., Saat, M. R., & Barkan, C. P. (2018). Quantitative analysis of the derailment characteristics of loaded and empty unit trains. Transportation Research Record, 2672(10), 156-165.

Liu, J., Khattak, A. J., Richards, S. H., & Nambisan, S. (2015). What are the differences in driver injury outcomes at highway-rail grade crossings? Untangling the role of pre-crash behaviors. Accident Analysis & Prevention, 85, 157-169.

Liu, X. (2016). Analysis of collision risk for freight trains in the United States. Transportation Research Record, 2546(1), 121-128.

Liu, X., Saat, M. R., & Barkan, C. P. (2012). Analysis of causes of major train derailment and their effect on accident rates. Transportation Research Record, 2289(1), 154-163.

Liu, X., Saat, M. R., & Barkan, C. P. L. (2017). Freight-train derailment rates for railroad safety and risk analysis. Accident Analysis & Prevention, 98, 1–9.

Lu, P., & Tolliver, D. (2016). Accident prediction model for public highway-rail grade crossings. Accident Analysis & Prevention, 90, 73-81.

McCollister, G. M., & Pflaum, C. C. (2007). A model to predict the probability of highway rail crossing accidents. Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit, 221(3), 321-329.

Multer, J., Melnik, G., Safar, H., & King, S. (2022). Stakeholder Perceptions of Longer Trains (No. DOT/FRA/ORD-22/43, DOT-VNTSC-FRA-22-01). United States Department of Transportation, Federal Railroad Administration Technical Reports.

Nayak, P. R., Rosenfield, D. B., & Hagopian, J. H. (1983). Event probabilities and impact zones for hazardous materials accidents on railroads (Report No. DOT-TSC-FRA-83-5). United States. Federal Railroad Administration.

Oh, J., Washington, S. P., & Nam, D. (2006). Accident prediction model for railway-highway interfaces. Accident Analysis & Prevention, 38(2), 346-356.

Stamatiadis, N., & Deacon, J. A. (1995). Trends in highway safety: Effects of an aging population on accident propensity. Accident Analysis & Prevention 27, 443-459.

Stamatiadis, N., & Deacon, J. A. (1997). Quasi-induced exposure: methodology and insight. Accident Analysis & Prevention, 29(1), 37-52.

United States Government Accountability Office (US GAO). (2019). Rail Safety: Freight Trains Are Getting Longer, and Additional Information Is Needed to Assess Their Impact (Report No. GAO-19-443). United States Government Accountability Office.

Wang, B. Z., Barkan, C. P., & Rapik Saat, M. (2020). Quantitative analysis of changes in freight train derailment causes and rates. Journal of transportation engineering, Part A: Systems, 146(11), 04020127.

Zhang, Z., Lin, C. Y., Liu, X., Bian, Z., Dick, C. T., Zhao, J., & Kirkpatrick, S. (2022). An Empirical analysis of freight train derailment rates for unit trains and manifest trains. Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit, 236(10), 1168-1178.