Effects of Whistle-Blowing Bans on Accidents at Gated Rail-Highway Crossings: The Northeastern Illinois Experience

Author(s): Paul Metaxatos, P.S. Sriraj, Siim Sööt, and Joseph DiJohn
Source: *Journal of the Transportation Research Forum*, Vol. 43, No. 2 (Fall 2004), pp. 101-116
Published by: Transportation Research Forum
Stable URL: http://www.trforum.org/journal

Effects of Whistle-Blowing Bans on Accidents at Gated Rail-Highway Crossings: The Northeastern Illinois Experience

This paper examines the effect of whistle-blowing bans on accidents at gated rail-highway public crossings in the Chicago metropolitan region. The statistical analysis show that it is rather misleading to unconditionally associate whistle bans with accident incidence and higher collision frequencies of rail-highway crossings while ignoring other factors or combinations of factors that are probably more relevant to the operational characteristics of the crossings. A deeper one delves into the interactive effects of crossing-specific characteristics on the number of accidents, the more the impact of individual factors becomes confounded so that interaction effects may even negate the effects of individual factors.

by **Paul Metaxatos, P.S. Sriraj, Siim Sööt, Joseph DiJohn**

INTRODUCTION

Accidents involving trains and highway vehicles have been a major source of concern for the Federal Railroad Administration (FRA). The national average number of such colli-sions per year in the mid-1990s was about 4,000 at all rail crossings in the nation (USDOT/FRA, 2000a). The average number of deaths resulting from these collisions has been about 400 per year. The FRA has concluded these numbers warrant immediate action to mitigate and eliminate the reasons for the collisions.

Since the early 1990s the FRA has studied the impact that train whistles have on safety at rail crossings. Initially conducted in Florida, this study was soon conducted across the nation to evaluate the type of impact whistle bans have on accidents. In 1984, Florida authorized local governments to ban the nighttime use of whistles by trains approaching rail crossings. This resulted in a proliferation of bans against whistle blowing at rail crossings in Florida. Many local jurisdictions passed ordinances banning locomotive horns. In 1990, an FRA study on the effect of the whistle bans in Florida on the accidents at rail crossings showed that there were almost three times more collisions after the whistle bans were established. Thus, in 1991, the FRA issued an emergency order to end whistle bans. The Florida study also prompted the FRA to study the effect of whistle bans on accidents on a nation-wide basis. The results of this study and the Florida study were published in four reports (USDOT/FRA, 1995a and 1995b; USDOT/FRA, 2000a and 200b). These reports point out the significant impact that train whistles have in reducing the number of accidents at rail crossings. The FRA found that crossings with whistle bans averaged 84% more collisions than crossings that permit whistle blowing. The only exception to this finding was in the six-county northeastern Illinois Chicago region where collisions were 16% less frequent.

Moreover, the use of whistles at crossings, at the time these studies were conducted, was not governed by a national mandate and instead was dependent on local laws. Consequently, rail crossings within a state often had different stipulations regarding the blowing of whistles.

The FRA's findings gave local and municipal agencies in the Chicago area reason to keep their own operating policies regarding whistle blowing unchanged while awaiting the final ruling from the FRA. Follow-up studies were performed at the behest of many local agencies and comments have been sent to the FRA. These comments supported exemp-tions from blowing locomotive horns at crossings where accident experience was under a specified threshold.

This paper, motivated by the debate outlined above, examines whether collision rates at gated crossings in the Chicago region, which have honored (this qualification will be justified later in the discussion) whistle bans, are statistically different than the collision rates at crossings where the train whistle is routinely sounded (no-ban). The overall objective of the paper is to develop an understanding of crossing-specific factors that may have an impact on the number of collisions in Chicago gated rail-highway crossings.

The findings in this paper, applicable only to Chicago-area gated crossings, show that it is rather misleading to unconditionally associate whistle bans with accident incidence and higher collision frequencies of rail-highway crossings while ignoring other factors or combinations of factors that are probably more relevant to the operational characteristics of the crossings. The deeper one delves into the interactive effects of crossing-specific characteristics on the number of accidents, the more the impact of individual factors becomes confounded and interaction effects may even negate the effects of individual factors.

DATA ISSUES

Several types of information were needed to carry out this research: a rail-highway crossings inventory; accident data and whistle-ban status at those crossings; and vehicular and train traffic by crossing. The data sources, data reduction and data limitations are discussed in this section.

Data Sources

This research utilized available data sources from the Chicago Area Transportation Study (CATS) as follows: (a) The crossings database was developed from the Federal Inventory for Northeast Illinois. The database has 1,952 observations (crossings) and 128 crossing-specific variables with multiple missing values. The relevant set of crossings for this study is 805 gated crossings within the six-county area. (b) The 1988-1999 accident inventory for the State of Illinois has 3,318 observations (accidents) and 203 variables (also with many missing values) with collision-specific as well as crossing-specific information (e.g., vehi-cular and train traffic). The relevant data set for this study is 561 collisions at 295 crossings within the six-county area. Both fatal and non-fatal accidents are included. (c) An inventory of gated crossings with confirmed 24-hour whistle bans. The relevant set for this study is 290 crossings.

Data sources used in this research are believed to be the latest available for the study area. Note, however, the inventory update process in the individual databases from the combined data set is voluntary and rather slow. Laffey (2000) comments: "The mean age of an inventory record is 11 years while the median age is 13 years." It is regrettable that more contemporary crossing data are unavailable, but the authors do not have the resources to generate their own crossing data.

Data Reduction

This analysis is based on gated, at-grade railroad crossings in the Chicago region. If there is any increased risk of collision associated with train-whistle bans, it should be apparent on a regional level.

This research has analyzed the same types of collisions as the FRA Updated Analysis of Train Whistle Bans (USDOT/FRA, 2000b). In the FRA analysis, accidents that involved collisions beyond the fourth rail car, and collisions without a vehicle driver were not included. These types of collisions were determined to have no association with the sounding of a train whistle. Accidents that involved pedestrians were not included since there is no accident exposure data for pedestrians.

This research has included collisions during the 12-year period from 1988 through 1999, as opposed to the FRA use of only a five-year span of data from 1992 through 1996. Any association of collisions with whistle bans should be evident over this 12-year period.

Crossings that were determined to have changed the warning device type over the study period were dropped from the analysis. This includes situations such as flashing lights replaced by automatic gates. The FRA analysis also dropped crossings that changed warning device status.

The crossing file used in the study is different from the one used in the FRA's Updated Analysis of Train Whistle Bans (USDOT/FRA, 2000b). Thus the staff at CATS updated the crossing file to eliminate aban-doned crossings.

A second difference concerns how whistle-ban status is defined. Many crossings are listed as having a whistle ban, but trains routinely sound their horns anyway. Data were collected by the staff at CATS to determine which crossings in the Chicago area have whistle bans that are honored by the railroads. The honored whistle-ban crossings are the only crossings that are considered to have a whistle ban in this analysis.

Crossings that have missing or zero values for daily train volume or annual average daily (vehicular) traffic were eliminated from the analysis. The FRA accident prediction formula (APF) uses these two variables as factors, but the formula is designed to return a factor of one if information is missing. Otherwise, it returns a very small number corresponding to the very low number of expected collisions per year per crossing without train or vehicle traffic. Although not zero, the very low number may still have essentially the same meaning as a zero result, and may flag closed or abandoned crossings.

In view of the above, the previous three data sets were combined into one database with 805 records (gated crossings). The number of collisions per crossing was computed and combined with the other crossing-specific information.

Data Limitations

It is important to note that data limitations are common in this type of study. First, the lack of the date that the whistle-ban policy was enacted at a particular crossing makes it impossible to determine whether the policy was in effect at the time of the collision at that crossing. Given the resources of the study, it was impractical to investigate archival data (e.g., locomotive's event recorders) to determine whether a train's horn was sounding at the time of a collision. Moreover, it was not possible to resolve whether the driver heard the horn sounding or was distracted (e.g., music, passengers, etc.).

Secondly, the level of the annual average daily traffic (AADT) for each crossing is an average value that is not time-of-day specific. Additionally, in this study the AADT value is only available for 1999. AADT may have subsequently changed, but the authors do not have the resources to generate their own up-to-date data on a crossing-specific basis. As a result, we need to assume that vehicular traffic has not changed dramatically over the course of the study period and, furthermore, that the average value is close enough with the prevalent level of traffic at the date and time of the collision. A similar observation pertains to the number of daily trains, although this figure should be more stable over time than vehicular traffic.

A third observation is related to the fact that driving around a lowered gated crossing, probably the most frequent cause for such accidents, involves a host of factors not captured in the collision data available. Clearly, factors such as the presence of alcohol; the age and gender of the driver;

potential social pressure from the cars behind; the time of day; and the visual sight distance and angle that the road meets the track need to be controlled for in a larger-scale study.

EXPLORATORY ANALYSIS

The impact of whistle bans, daily train volume, annual average daily traffic (AADT), exposure and risk incidence is now examined. Except for the whistle-ban factor, the other factors are typically included in all methodologies used to predict collision frequencies or to prioritize rail-highway crossing safety im-provements (Elzohairy and Benekohal, 2000). In this manner, a meaningful comparison between whistle-ban and whistle-blowing crossings in terms of collision incidence and collision variability can be made. Note that in this paper, the terms 'collisions' and 'accidents' are used interchangeably. A lengthier technical discussion is available elsewhere (Metaxatos et al., 2001).

Annual Distribution of Accidents

The number of collisions in the study area has decreased almost every year in the 1988-1999 period. A chi-square analysis did not reveal any significant systematic variation in the annual distribution of the number of accidents between whistle-ban and whistle-blowing crossings.

Whistle-Ban Status and Collisions

At first, the association between whistle-ban status and collision incidence is examined without controlling for other factors of interest. In this sample of 805 crossings all the sampling assumptions for the (Pearson) chi-square statistic are met. The value of the statistic, 14.19, is clearly significant at the 0.001 level. Indeed, collisions occurred at 45% (131 out of 290) of the whistle-ban crossings, but only at 32% (164 out of 515) of the crossings without 24-hour whistle bans.

The previous analysis can be repeated to examine the association between ban status and collision frequencies. The chi-square statistic is 19.76, again significant at the 0.001 level. Indeed, 61% of the whistle-ban crossings (249 out of 408) had collisions, while only 47% (312 out of 663) of the whistle-blowing crossings had collisions.

Whistle-Ban Status and Collisions Controlling for Additional Factors

The previous analysis implies that, at a first level of analysis and if no other factors are taken into account, there appears to be a higher collision frequency at whistle-ban crossings. However, it is reasonable to assume that this association may have been affected by a number of factors. Therefore, the impact of factors such as train volume, annual average daily traffic (AADT), exposure, and a measure based on the FRA Accident Prediction Formula needs to be examined.

The Impact of Train Volume. The train volume values have been placed into ten groups with approximately the same number of crossings called deciles[1] (Table 1). The majority of crossings with high train volume in the Chicago region have whistle bans. Indeed, almost 90% of the whistle bans are enforced at crossings with 50 or more daily trains. On the contrary, almost 90% of the whistle-blowing crossings have less than 50 daily trains. Moreover, almost 90% of collision-prone crossings are associated with 50 or more daily trains and 24-hour whistle bans. In a similar fashion, more than 85% of the collisions occur at whistle-ban cross-ings with 50 or more trains. Furthermore, higher collision frequencies are associated with higher train volumes and, in general, have a lower co-efficient of variation at higher train volumes.

To avoid the effect of a potential spurious correlation between whistle-ban crossings and collisions, the effect of train volume needs to be accounted for. The Mantel-Haenszel results[2] for the stratified analysis give a test statistic[3] $Q_{CSMH} = 7.667$, which is significant ($p = 0.005$).[4] Controlling for train volume, collision occurrence is associated with whistle-ban crossings. Similarly, controlling

Table 1: Number of Crossings and Collisions by Daily Train Traffic

Daily Train Traffic	Number of Crossings			Number of Crossings With Collisions (# Acc.)			Number of Crossings Without Collisions			Collisions per Crossing (totals)	
	No Ban	Ban	Total	No Ban	Ban	Total	No Ban	Ban	Total	Mean	Std. Dev.
1-5	72	2	74	7 (8)	1 (1)	8 (9)	65	1	66	0.12	0.36
6-10	85	2	87	16 (24)	2 (4)	18 (28)	69	0	69	0.32	0.70
11-18	88	1	89	29 (47)	1 (1)	30 (48)	59	0	59	0.53	0.91
20-32	72	11	83	27 (43)	5 (10)	32 (53)	45	6	51	0.63	0.98
33-41	72	2	74	25 (35)	2 (4)	27 (39)	47	0	47	0.52	0.81
42-49	56	19	75	18 (48)	6 (12)	24 (60)	38	13	51	0.80	2.20
50-63	24	34	58	12 (34)	21 (41)	33 (75)	12	13	25	1.29	2.47
64-75	20	77	97	12 (25)	32 (67)	44 (92)	8	45	53	0.94	1.43
76-88	17	71	88	10 (22)	19 (33)	29 (55)	7	52	59	0.62	1.27
89-190	9	71	80	8 (26)	42 (76)	50 (102)	1	29	30	1.27	1.36
Total	515	290	805	164 (312)	131 (249)	295 (561)	351	159	510		

Source: Chicago Area Transportation Study – Analysis by authors

for train volume, whistle-ban crossings are associated with higher number of collisions (Q_{CSMH} = 5.022, p = 0.025). The last association is significant at the 0.05 level but not at the 0.01 level.

The Impact of AADT. The AADT values have been grouped into 10 deciles (Table 2). Whistle-ban crossings are more often encountered at locations with relatively lower vehicular traffic while higher vehicular traffic is more prevalent at no-ban crossings. Moreover, 70% of collision-prone whistle-blowing crossings and a little more than 50% of whistle-ban crossings are associated with higher AADT values (more vehicular traffic). In addition, more than 80% of the collisions occur at whistle-blowing crossings associated with higher AADT values; the same is true for a little over 60% of the collisions at whistle-ban crossings. Finally, on average, more collisions with a lower coefficient of variation occur at high-traffic-volume crossings.

The Mantel-Haenszel results for the stratified analysis give a test statistic Q_{CSMH} = 22.603, which is strongly significant ($p < 0.0001$). Controlling for AADT, collision occurrence is associated with whistle-ban crossings. Similarly, controlling for AADT, whistle-ban crossings are associated with a

Table 2: Number of Crossings and Collisions by Annual Average Daily Traffic

Annual Average Daily Traffic (vehicles)	Number of Crossings			Number of Crossings with Collisions (# Acc.)			Number of Crossings Without Collisions			Collisions per Crossing (totals)	
	No Ban	Ban	Total	No Ban	Ban	Total	No Ban	Ban	Total	Mean	Std. Dev.
50-250	57	31	88	7 (11)	12 (23)	19 (34)	50	19	69	0.38	0.87
259-475	27	33	60	5 (5)	8 (11)	13 (16)	22	25	47	0.26	0.54
500-800	56	39	95	11 (15)	13 (19)	24 (34)	45	26	71	0.35	0.72
830-2050	42	37	79	7 (9)	13 (22)	20 (31)	35	24	59	0.39	0.83
2100-4200	47	33	80	15 (20)	13 (24)	28 (44)	32	20	52	0.55	0.85
4300-7000	58	25	83	16 (26)	9 (17)	25 (43)	42	16	58	0.51	0.92
7200-10370	58	19	77	27 (52)	13 (17)	40 (69)	31	6	37	0.89	1.32
10400-14300	52	31	83	24 (48)	20 (39)	44 (87)	28	11	39	1.04	1.29
14500-21900	57	23	80	33 (88)	14 (36)	47 (124)	24	9	33	1.55	2.88
22000-55000	61	19	80	19 (38)	16 (41)	35 (79)	42	3	45	0.98	1.49
Total	515	290	805	164 (312)	131 (249)	295 (561)	351	159	510		

Source: Chicago Area Transportation Study – Analysis by authors higher number of collisions (Q_{CSMH} = 41.193, p < 0.0001).

The Impact of Exposure. The previous observations imply a potential multiplicative effect of train and vehicular volumes on collision frequencies. The values for this multiplicative factor, defined as the product of AADT and the number of daily trains, have been grouped into ten deciles that comprise different levels of exposure (Table 3).

The multiplicative effect of the train traffic and AADT decile distribution for whistle-ban and whistle-blowing crossings is evident. Whistle-blowing crossings are more prevalent at low and above-average levels of exposure while whistle-ban crossings prevail at the below-average and high end of exposure levels. Moreover, collision-prone whistle-ban crossings are more prevalent at the upper two deciles of exposure, while whistle-blowing collision-prone crossings are more prevalent at the seventh and eighth deciles. A similar observation holds for the distribution of the percentage of collisions. Finally, higher collision frequencies with a lower coefficient of variation are seemingly associated with higher levels of exposure.

The Mantel-Haenszel results for the stratified analysis give a test statistic Q_{CSMH} = 14.935, which is strongly significant (p = 0.0001). Controlling for the level of exposure, collision occurrence is associated with whistle-ban crossings. Similarly, controlling for the level of exposure, whistle-ban crossings are associated with higher number of collisions (Q_{CSMH} = 20.592, p < 0.0001).

The Impact of Risk Incidence. The FRA forecasts collisions at rail-highway crossings using a regression model, the Accident Prediction Formula (APF). The model consists of three parts: an unnormalized accident prediction from the basic formula (equation 1); an adjustment to include collision histories; and a warning-specific normalizing constant that allows total collisions in the base year to equal forecast collisions. In its studies, the FRA has used the first part of its model to group gated crossings by increasing risk of collisions. Without the constants that are used to adjust total collisions, the results of the formula do not represent actual collisions, but the results serve as a tool for ranking crossings by increasing risk. In this study, the first part of the APF for gated crossings is taken to represent the level of risk and is given by equation 1.

Table 3: Number of Crossings and Collisions by Range of Exposure

Range of Exposure (in 000s)	Number of Crossings			Number of Crossings with Collisions (# Acc.)			Number of Crossings Without Collisions			Collisions per Crossing (totals)	
	No Ban	Ban	Total	No Ban	Ban	Total	No Ban	Ban	Total	Mean	Std. Dev.
0.177-6.45	73	7	80	4 (4)	3 (8)	7 (12)	69	4	73	0.15	0.57
6.8-15.6	56	20	76	11 (16)	9 (16)	20 (32)	45	11	56	0.42	0.83
16-26.468	51	34	85	13 (17)	9 (13)	22 (30)	38	25	63	0.35	0.71
26.6-50	44	38	82	9 (12)	12 (16)	21 (28)	35	26	61	0.34	0.65
50.4-102.4	51	29	80	10 (15)	6 (7)	16 (22)	41	23	64	0.27	0.61
104.4-172.5	54	25	79	15 (26)	13 (27)	28 (53)	39	12	51	0.67	1.14
172.8-306.9	63	19	82	27 (40)	6 (12)	33 (52)	36	13	49	0.63	0.90
308-500	59	22	81	33 (60)	10 (18)	43 (78)	26	12	38	0.96	1.18
500.5-924	38	42	80	21 (43)	23 (35)	44 (78)	17	19	36	0.97	1.29
959.4-3822	26	54	80	21 (79)	40 (97)	61 (176)	5	14	19	2.20	2.95
Total	515	290	805	164 (312)	131 (249)	295 (561)	351	159	510		

Source: Chicago Area Transportation Study – Analysis by authors. Exposure is defined as the product between AADT and number of trains.

$$(1)\ RISK = \{[(AADT \times TRAINS) + 0.2]/0.2\}^{0.2942} \times [(DAY\_THRU\_TRAINS + 0.2)/0.2]^{0.1781} \times e^{MAIN\_TRACKS \times 0.1512} \times e^{(LANES - 1) \times 0.1420}$$

where, RISK is the risk score, AADT is the annual average daily vehicular traffic, TRAINS is the number of trains per day, DAY_THRU_TRAINS is the number of through trains per day during daylight, MAIN_TRACKS is the number of tracks, and LANES is the number of highway lanes.

Whistle-ban crossings are, in general, more prevalent at high-risk crossings compared to whistle-blowing crossings that are more prevalent at low-risk crossings. A similar observation holds for the distribution of collision-prone crossings and the distribution of collision frequencies in those crossings. Moreover, higher collision frequencies with a lower spread around the mean occur at higher risk crossings.

The values for the level of risk as computed above have been placed into ten deciles (Table 4) and comprise the ten risk groups. The Mantel-Haenszel results for the stratified analysis give a test statistic Q_{CSMH} = 5.888, which is strongly significant (p = 0.01). Controlling for the level of risk, collision occurrence is associated with whistle-ban crossings. Similarly, controlling for the level of risk, whistle-ban crossings are associated with higher number of collisions (Q_{CSMH} = 8.582, p = 0.003).

The Impact of Combinations of Factors. The previous sections explored the association between whistle bans and collisions controlling for individual factors, such as train traffic, AADT, level of exposure, and level of risk. This section presents the results, summarized in Table 5, from a Mantel-Haenszel stratified analysis for various combinations of these factors.

The Mantel-Haenszel procedure potentially removes the confounding influence of explanatory variables that comprise the stratification and provides greater power to detect an association by comparing like objects. The procedure requires minimal assumptions and thus the conclusions of the

Table 4: Number of Crossings and Collisions by Risk Group

Risk Group	Number of Crossings			Number of Crossings with Collisions (# Acc.)			Number of Crossings Without Collisions			Collisions per Crossing (totals)	
	No Ban	Ban	Total	No Ban	Ban	Total	No Ban	Ban	Total	Mean	Std. Dev.
13.3-52.3	76	4	80	6 (7)	3 (5)	9 (12)	70	1	71	0.15	0.45
52.5-80.5	75	6	81	11 (15)	1 (4)	12 (19)	64	5	69	0.23	0.67
80.8-107.6	63	17	80	11 (16)	5 (7)	16 (23)	52	12	64	0.28	0.67
108.5-135.0	35	46	81	7 (7)	12 (17)	19 (24)	28	34	62	0.29	0.62
135.1-175.2	55	25	80	14 (17)	11 (18)	25 (35)	41	14	55	0.43	0.76
175.6-214.3	52	29	81	18 (30)	9 (13)	27 (43)	34	20	54	0.53	0.90
214.5-272.7	52	29	81	29 (44)	6 (11)	35 (55)	23	23	46	0.67	1.01
272.9-342.1	45	35	80	21 (33)	21 (35)	42 (68)	24	14	38	0.85	0.98
342.2-458.2	37	44	81	25 (54)	27 (62)	52 (116)	12	17	29	1.43	1.59
460.3-1036.7	25	55	80	22 (89)	36 (77)	58 (166)	3	19	22	2.07	2.93
Total	515	290	805	164 (312)	131 (249)	295 (561)	351	159	510		

Source: Chicago Area Transportation Study – Analysis by authors. Risk is defined by the value obtained for each crossing based on the APF (equation 1).

Table 5: Whistle-Ban Impacts on Collisions Controlling for Different Factor Combinations

Controlling for Factor Combination	Collision Incidence		Higher Number of Collisions	
	Q_{CSMH}	Probability[a]	Q_{CSMH}	Probability
Train traffic and AADT	8.072	0.004*	13.199	0.0003*
Train traffic and level of exposure	7.009	0.008*	7.936	0.004*
Train traffic and level of risk	1.226	0.268	2.113	0.146
AADT and level of exposure	3.483	0.062	9.583	0.002*
AADT and level of risk	4.771	0.028**	9.534	0.002*
Level of exposure and level of risk	3.808	0.051	2.255	0.133
Train traffic, AADT and level of exposure	0.021	0.884	0.0006	0.981
Train traffic, AADT and level of risk	0.412	0.520	0.527	0.467
AADT, level of exposure and level of risk	3.252	0.071	6.290	0.012**
Train traffic, AADT, level of exposure and level of risk	0.014	0.903	0.001	0.974

[a]Probability of obtaining a higher test-statistic value under the null hypothesis of no association; *significant at the 0.01 level; **significant at the 0.05 level. Source: Chicago Area Transportation Study – Analysis by authors

analysis may be restricted to the study population at hand, versus inference to a larger population.

In Table 5, the first column displays the factor combination being controlled for. The second and third columns present, respectively, the chi-square statistics (Q_{CSMH} values) and associated probabilities for testing the impact of whistle bans on collision incidence, controlling for the particular factor combination. Finally, the fourth and fifth columns present, respectively, the test statistic value and associated probability for testing the impact of whistle bans on collision frequencies, again, controlling for that particular factor combination.

The results in Table 5 show a statistically significant association between whistle bans and collision incidence after controlling for three combinations of factors: train traffic and AADT; train traffic and level of exposure; and AADT and level of risk. There is also a statistically significant association between whistle bans and higher number of collisions (last two columns in Table 5) after controlling for the same factors. Three additional combinations of factors that, when controlled for, render the association significant are: AADT and level of exposure; AADT and level of risk, and AADT, level of exposure and level of risk.

Interestingly enough, controlling for two factor combinations, namely, train traffic, AADT and level of exposure, and train traffic, AADT, level of exposure, and level of risk, render the impact of whistle bans on collision incidence or collision frequencies non significant. This is true with almost all combinations of three or more factors. The fact that AADT and level of exposure appear almost always in all significant factor combinations should not be considered a factual observation. After all, AADT, a one-time statistical average value over the 12-year period, serves this analysis only as a proxy of the prevailing vehicular traffic at the time of the collision.

In summary, the previous analysis implies that it may be rather misleading to unconditionally associate whistle bans with collision incidence and higher collision frequencies of gated rail-highway crossings, ignoring the synergy of a number of factors or combination of factors that are probably more relevant to the operational characteristics of the crossings.

REGRESSION ANALYSIS

While the association between whistle bans and collisions along with the synergistic role of other factors has been investigated in the previous section, the strength and direction of the association has yet to be determined. This section shifts the focus to statistical models, methods aimed at describing the nature of the association in terms of a potentially parsimonious number of parameters. It is important to note at this point that this section is not describing a new accident prediction model. Clearly, such an effort would require many more resources as well as an entirely different formal approach.

Model Formulation

The customary assumption for accident data is that they are observations of a Poisson variable. This is well-supported by the theory of the Poisson distribution and indeed the Poisson distribution was originally developed precisely for accident statistics. The Poisson distribution is the nominal distribution for counted data in much the same way that the Normal distribution is the benchmark for continuous data. The frequency distributions (Table 6) seem to corroborate the conjecture that the Poisson distribution is appropriate as a first approximation.

Poisson regression has the advantage of being precisely tailored to the discrete, highly-skewed distributions of accident data. The Poisson regression model gets its name from the assumption that the dependent variable has a Poisson distribution.

The regression model will estimate the number of collisions at a crossing given a number of crossing-specific explanatory variables that are typically found in the literature concerning rail-highway crossing safety. These include the whistle-blowing status, the number of all types of daily trains, the AADT value, an interaction term for the multiplicative effect of train and vehicular volume (level of exposure), and the level of risk as estimated by the first part of the APF. Recall that the risk scores from the formula are a function of the level of exposure, the number of daily through trains, the number of main tracks and the number of highway lanes. Because the level of exposure is included in the risk score, we decided to partition the scores intro three groups, high, medium, and low, and treat the scores as a discrete rather than a continuous variable. A technical discussion on the functional form of the model and its estimation is provided elsewhere (Metaxatos et al., 2001).

Thus, the only concern with this method lies with the credibility of the FRA's APF. However, since the FRA has been working with and improving the APF for a long time, one would expect that it is of adequate quality for the purpose of the models in this paper. Also, there is no other practical alternative to the FRA risk measure.

Model Estimation

Assuming a linear relationship among the (untransformed) independent variables, we obtained initial parameter estimates. There is undoubtedly some inter-crossing variability in the number of collisions that cannot be

Table 6: Frequency Distributions of Collisions

	All Crossings		Whistle-Blowing Crossings		Whistle-Ban Crossings	
Number of Collisions	Number of Crossings	Percent of Crossings	Number of Crossings	Percent of Crossings	Number of Crossings	Percent of Crossings
0	510	63.35	351	68.16	159	54.83
1	161	20.00	98	19.03	63	21.72
2	76	9.44	36	6.99	40	13.79
3	33	4.10	17	3.30	16	5.52
4	8	0.99	3	0.58	5	1.72
5	9	1.12	5	0.97	4	1.38
6	5	0.62	2	0.39	3	1.03
8	1	0.12	1	0.19	0	0.00
17	2	0.25	2	0.39	0	0.00

Source: Chicago Area Transportation Study – Analysis by authors accounted for by the Poisson model. This calls for adjusting for overdispersion.

Two Unusual Cases

Interestingly, only the intercept and the variable representing the risk incidence of the crossing appear to be significant at the 0.01 level after adjusting for overdispersion. Before the adjustment, the AADT variable and the variable representing the exposure factor were also significant at the 0.05 level. Before we say more, however, we need to look for unusual cases that do not belong in the model. This is an important part of the analysis because such points might suggest de-ficiencies in the model (e.g., a missing independent variable) or that the algebraic form of the model is incorrect (e.g., need for data transformations).

In view of the above, we examined the residuals for potential outliers and influential observations. It was expected that the two observations (crossings) with 17 collisions each (Table 6) would be problematic. Indeed, upon examining the plot of standardized deviance residuals[5] against the predicted values, those two cases stand out with residual values of more than 5. The expected average number of collisions under the model are 1.48 and 1.52, respectively for those two cases.

The two cases involve no-ban crossings. The first crossing (ID# 372177T) is in Wood Dale on Irving Park Road, and the second one (ID# 478713F) in Chicago on East 130th Street. The first one has 46 daily trains and the second one has 52. Recall that the average number of collisions for the category is 0.8 for the first crossing and 1.29 for the second one (Table 1). Both crossings experience AADT values within the 9th decile of the distribution (Table 2) with an 1.55 average number of collisions for the category. The high level of train and vehicular activity puts both crossings in the 10th decile on the exposure scale (Table 3) with 2.2 average number of collisions for the category. Finally, both crossings score in the 10th decile of the risk distribution (Table 4) with 2.07 average number of collisions for the category.

Clearly both cases require further investigation. We intend to bring both cases back into the model later during the variable selection stage. At this point, however, we simply choose to eliminate these obviously unusual points because we decided that they should not be considered in an attempt to find a general relationship between the average number of collisions and the factors of interest.

Upon deletion of the two cases, the normal probability plot (Sen and Srivastava, 1990) has been mostly straightened out. We now need to focus our attention to the possibility that the algebraic form of the model is not correct.

Data Transformations, Additional Terms, Multicollinearity and Variable Search

In fitting the Poisson regression model, we have already transformed the dependent variable to a logarithmic scale. Here, we will examine whether there is a need to transform the independent variables or if additional terms involving the same independent variables would be helpful.

Component plus residual plot analysis (Sen and Srivastava, 1990) indicated the need for transformations of the independent variables. We tested a number of transformations and found that a transformation between the log and the square root was most appropriate. We chose to work with the log transformation because the coefficients then have a much simpler interpretation.

As we have already observed, not only individual factors but also combinations of factors may have an impact on the number of collisions. This conjecture can be further tested in the regression model by means of additional terms representing factor combinations. This course of action raises the issue of multicollinearity. The basic point is that, if two or more independent variables are closely related to each other the quality of the estimates, as measured by their variances, can be seriously and adversely affected making it harder to get good estimates of the distinct effect on the dependent variable. Although multicollinearity does not bias the coefficients, it does make them more unstable. Standard errors may become large, and variables that appear to have weak effects, individually, may actually have quite strong effects as a group. We could also obtain counter-intuitive results, especially in the signs of the coefficients (positives may become negatives and vice versa).

We used a number of multicollinearity diagnostics, such as tolerances, variance inflation factors, eigenvalues and condition numbers (Sen and Srivastava, 1990) and tested a fairly long list of independent variables (individual factors and factor combinations). Subsequently, we culled the independent variable list to obtain a more parsimonious model that is easier to work with, reduce the multicollinearity in the model, and reduce the ratio of the number of variables to the number of observations, which is statistically beneficial. The practice of variable search is often a matter of making the best compromise between keeping the standard errors and bias low and achieving parsimony and reducing multicollinearity (Sen and Srivastava, 1990).

Estimation Results

During variable selection, one frequently finds, clustered around the chosen model, other models which are nearly as good and not statistically distinguishable. The six different models shown in this section were found to be comparably good candidates in terms of measure of fit, moderate bias, and little multicollinearity. The variable combinations in the six models are because of transformations of the independent variables to address multicollinearity and heteroscedasticity issues, as well as introduction of second order effects. Considering that the models estimated in this study will not be used as tools for forecasting collision frequencies or for prioritizing safety upgrades in crossings, we will not recommend a 'best' model among the six candidates.

In Table 7 we show only the significant parameter estimates for each of the models presented by increasing level of fit, as measured by their deviance in column two. Models 1 to 6 have been adjusted for overdispersion. An estimate of the overdispersion is given by the ratio of the deviance value divided by the degrees of freedom (third column in Table 7). Parameter estimates for the factors and factor combinations in the fifth column are shown in the sixth column while their standard errors are in column seven. The 95% confidence intervals for the estimated parameters are shown in columns eight and nine.

The scale parameter (the square root of the dispersion parameter in Table 7) is computed as the square root of the Pearson chi-square divided by the degrees of freedom. Its function is similar to the scale parameter in linear regression, known as root mean square error. The adjusted chi-square for each coefficient (in column 10) is the unadjusted coefficient divided by the scale parameter. As with linear regression, it tests for the null hypothesis that the parameter estimated is zero. The probability of obtaining a higher chi-square value is reported in column 11. Note that the adjusted standard error of each coefficient is the unadjusted standard error multiplied by the square root of the scale parameter.

The estimated parameters give the log-odds increase/decrease (recall that the dependent variable is logged) for every one-unit increase/decrease in the explanatory variable. Whenever the independent variable has multiple levels (e.g., the 'Ban' variable has two levels, yes and no; the 'Rgroup' variable has three levels, high, medium and low) the comparison is made with respect to the reference level (here, the lowest classification level is always the reference level).

Note that models 1 to 6 were also estimated as negative binomial models with almost no qualitative difference. Almost the entire set of factors and factor combinations appearing significant under the Poisson assumption appear to be significant under the

Table 7: Analysis of Significant Parameter Estimates

Model	Deviance	Value/DF	Log Likelihood	Factors and Factor Combinations*	Estimate	Standard Error	Wald 95% Confidence Limits[6] Lower	Upper	Chi-Square	Pr>Chi
1	627.56	0.79	-396.38	Intercept	-2.6099	0.3435	-3.2831	-1.9367	57.74	<0.0001
				Ban	2.5801	0.5263	1.5487	3.6116	24.04	<0.0001
				Trains	0.0096	0.0038	0.0022	0.0169	6.46	0.0110
				Lexpo	0.3104	0.124	0.0661	0.5546	6.20	0.0128
				Ban * AADT	0.0622	0.0198	0.0235	0.1009	9.93	0.0016
				Ban * Lexpo	-0.6517	0.1917	-1.0275	-0.2759	11.55	0.0007
				Dispersion Parameter	1.5217	0.3250	1.0011	2.3128		
2	626.92	0.79	-395.27	Intercept	-1.7381	0.2314	-2.1917	-1.2845	56.40	<0.0001
				Ban	2.5801	0.5263	1.5487	3.6116	24.04	<0.0001
				Trains * AADT	0.0006	0.0003	0.0001	0.0011	6.17	0.0130
				Rgroup(3)	1.5347	0.3061	0.9348	2.1346	25.14	<0.0001
				Rgroup(2)	0.9325	0.2630	0.4170	1.4480	12.57	0.0004
				Ban * AADT	0.0507	0.0253	0.0010	0.1004	4.00	0.0454
				Ban * Trains * AADT	-0.0007	0.0004	-0.0015	-0.0001	4.03	0.0447
				Ban * Rgroup(2)	-0.9891	0.4204	-1.8131	-0.1651	5.53	0.0186
				Dispersion Parameter	1.5366	0.3370	0.9997	2.3622		
3	626.57	0.79	-395.12	Intercept	-1.9645	0.2180	-2.3918	-1.5372	81.19	<0.0001
				Ban	1.3747	0.3940	0.6024	2.1469	12.17	0.0005
				Trains	0.0137	0.0047	0.0044	0.0229	8.42	0.0037
				Rgroup(3)	0.9526	0.3666	0.2340	1.6711	6.75	0.0094
				Ban * Trains	-0.0224	0.0661	-0.0343	-0.0105	13.62	0.0002
				Ban * LAADT	-0.4074	0.1979	-0.7951	-0.0196	4.24	0.0395
				Dispersion Parameter	1.5361	0.3472	0.9862	2.3924		
4	625.38	0.79	-394.24	Intercept	-1.7085	0.4434	-2.5775	-0.8395	14.85	0.0001
				Trains * AADT	0.0008	0.0002	0.0003	0.0012	12.23	0.0005
				Rgroup(3)	1.6409	0.3596	0.9362	2.3457	20.83	<0.0001
				Rgroup(2)	0.9959	0.3000	0.4080	1.5838	11.02	0.0009
				Ban * AADT	0.0531	0.0240	0.0059	0.1002	4.87	0.0273
				Ban * Trains*AADT	-0.0008	0.0003	-0.0015	-0.0002	5.83	0.0158
				Ban * Rgroup(2)	-0.9976	0.4530	-1.8855	-0.1096	4.85	0.0277
				Dispersion Parameter	1.5413	0.3454	0.9934	2.3911		
5	623.96	0.78	-393.19	Intercept	-2.5896	0.4781	-3.5267	-1.6526	29.34	<0.0001
				Ban	3.4968	0.8109	1.9074	5.0862	18.59	<0.0001
				Ltrains	0.3664	0.1748	0.0238	0.7089	4.39	0.0360
				Ban * Ltrains	-0.9012	0.2683	-1.4271	-0.3753	11.28	0.0008
				Ban * LAADT	-0.3855	0.1962	-0.7699	-0.0010	3.86	0.0494
				Dispersion Parameter	1.5477	0.3497	0.9937	2.4102		
6	620.88	0.78	-390.18	Intercept	-3.0619	0.4222	-3.8993	-2.2345	52.61	<0.0001

*<u>Variables</u>: Ban=Whistle-ban status; Trains=number of daily trains; Ltrains=log(Trains); LAADT=log(AADT); Lexpo=log(Trains*AADT); Rgroup(1)=low level of risk; Rgroup(2)=medium level of risk; Rgroup(3)=high level of risk.

Source: Chicago Area Transportation Study – Analysis by authors. Dependent variable is number of collisions at crossings.

negative binomial assumption. This is the reason separate parameter estimates are not shown for the negative binomial models. The very few qualitative differences (borderline significance becoming borderline non-significance and vice versa) shown in Table 8 do not affect the overall interpretation of the results. This should not come as a surprise because for modest amounts of overdispersion (which is the case here given the low ratio of deviance and degrees of freedom values), it may be shown that the difference between two sets of parameter estimates, one based on the negative binomial likelihood and the other on the Poisson likelihood, may be neglected (see McCullagh and Nelder, 1989, p. 199).

Note that the estimated dispersion parameter for the six negative binomial models varied between 0.5986 and 0.6659 showing a level of overdispersion that is not dramatically different than the estimate for the respective Poisson model adjusted for overdispersion (column 3 in Table 7). As already noted, there seems to exist inter-crossing variability with respect to the number of collisions, perhaps because crossings with no collisions or one collision are qualitatively different from crossings with two or more collisions.

Interpretation of Model 3

In Model 3 in Table 7, for example, the coefficient for the 'Ban' variable (a yes/no variable) is 1.3747. This means that the expected number of collisions in whistle-ban crossings is 100(exp(1.3747)-1)=295% higher than those for whistle-blowing crossings. Similarly, the number of collisions per crossing increases, on average, by 100(exp(0.0137)-1)=1.37% with each additional train. Moreover, the number of collisions per crossing is, on average, 100(exp(0.9526)-1)=159% higher in high-risk vs. low-risk crossings.

The interpretation of the coefficients when interaction terms are involved follows that in multiple linear regression. Notice, for example, that the percent change in the expected number of accidents with each unit increase in the number of trains for ban crossings is 100*(exp(1.3747-0.0224*trains)-1) that of no-ban crossings. That percentage becomes increasingly negative when train volume increases above 61 trains/per day. Note that only about 9% of the no-ban crossings, but almost 75% of the ban crossings have train traffic above that level (Table 1). Therefore, the expected number of accidents in almost three out of four ban crossings (those with the heavier train traffic) is lower than in no-ban crossings with the same level of traffic, a difference that increases with the number of trains (above the 61 trains per day threshold). For example, at 64 trains per day, the expected number of accidents in ban crossings is already 5.7% lower than in no-ban crossings. More importantly, however, this difference is significant at any reasonable level.

Similar observations can be made for the other interaction term in Model 3. Notice, for example, that the percent change in the expected number of accidents with each unit increase in AADT for ban crossings is 100*(exp(1.3747-0.4074*LAADT)-1) that of no-ban crossings. That percentage becomes increasingly negative when AADT is higher than 29.2. Note, however, that all 805 crossings have an AADT value of at least 50 (Table 2). Therefore, the expected number of accidents in all ban crossings is lower than in no-ban crossings with the same level of traffic, a difference that increases with AADT. For example, at AADT=50, 100, 500, 1000 and 10,000, the expected number of accidents in ban crossings is, respectively, 19.6%, 39.4%, 68.5%, 76.2% and 90.7% lower than in no-ban crossings. More importantly, this difference is significant at any reasonable level.

CONCLUSIONS

The analysis implies that the study of collisions at gated crossings is anything but obvious. A first-level approach reveals positive effects between whistle bans and individual operational characteristics, collision incidence, and number of collisions. For example, collisions occurred at 45% of the whistle ban crossings, but only 32% of the no-ban crossings when other factors are not controlled for. This type of study, however, is generally insufficient in revealing the contributing factors in a complex problem, such as collisions at highway-rail crossings. Individual factors are never at work in isolation. The deeper one delves into the interactive effects of crossing-specific characteristics on the number of collisions, the more confounded the impact of individual factors becomes so that interaction effects may even negate the effects of individual factors.

Additional research is needed to investigate the effects of collision-specific environmental and human factors on the number of collisions. There are still unaccounted for factors in rail-highway collisions that warrant further study. If nothing more, however, this research has not found evidence to support the rather simplistic view that whistle bans are responsible for an increase in the number of collisions at gated crossings. The presence or absence of whistle blowing at gated crossings does not occur in a vacuum, but rather along with other phenomena that are related to the operation of the crossings. It is rather misleading, therefore, to attribute to a single factor, such as whistle bans, responsibility for the variation in collision occurrences and frequencies without assess-ing the compound effects of all other contributing factors.

Endnotes

1. For a set of measurements arranged in order of magnitude, the *d*-th decile is the value that has *d*% of the measurements below it and (100-*d*)% above it. Deciles, as other quantiles, cannot be calculated algebraically.

2. Kuritz, Landis and Koch (1988) present a useful overview of the Mantel-Haenszel strategy. The Mantel-Haenszel procedure provides statistics that detect general association, mean score differences, and linear correlation as alternatives to the null hypothesis of no association. The procedure potentially removes the confounding influence of the explanatory variables.

3. Q_{CSMH} is called the extended Mantel-Haenszel correlation statistic (Mantel, 1963). It approximately follows the chi-square distribution with one degree of freedom when the combined strata sizes are sufficiently large, that is 40 or more.

4. The p-value is the probability that a normal random variable has an absolute value larger than the z-score (estimated coefficient minus zero divided by the estimated standard error) obtained. If the p-value is small, we have good evidence that the corresponding variable is significant and that the difference between the coefficient estimate and zero arises not from chance but from a systematic effect.

5. The adjusted Pearson, deviance, and likelihood residuals are defined by Agresti (1990). These residuals are useful for outlier detection and for assessing the influence of single observations on the fitted model.

6. The confidence limits shown are the two-sided Wald confidence intervals for all model parameters based on the asymptotic normality of the parameter estimates, a reasonable assumption given the sample size. A small sample would require the computation of the likelihood-ratio confidence intervals.

References

Agresti, A. *Categorical Data Analysis*. John Wiley & Sons, New York, 1990.

Elzohairy, Y.M. and R.F. Benekohal. *Evaluation of Expected Accident Frequency Formulas for Rail-Highway Crossings*. Report No. ITRC FR 98-2. Illinois Department of Transportation, Springfield, Ill., 2000.

Kuritz, S.J., J.R. Landis, and G.G. Koch. "A General Overview of Mantel-Haenszel Methods: Applications and Recent Developments." *Annual Review of Public Health*, (1988): 123-160.

Laffey, S. "Grade Crossing Safety in the Chicago Area: An Environmental Analysis of the Potential Noise Impacts from the Swift Rail Development Act's Locomotive Horn Sounding Requirement." *Transportation Quarterly* 54, (2000): 1, 69-82.

Mantel, N. "Chi-square Tests with One Degree of Freedom; Extensions of the Mantel-Haenszel Procedure." *Journal of the American Statistical Association* 58, (1963): 690-700.

McCullagh, P. and J.A. Nelder. *Generalized Linear Models*. Second Edition. Chapman and Hall, London, England, 1989.

Metaxatos, P., P.S. Sriraj, S. Sööt, and J. DiJohn. *Factors Affecting Collisions at Gated Rail-Highway Crossings: A Northeastern Illinois Case Study.* Final Report. Illinois Department of Transportation, October 2001.

Sen, A. and M. Srivastava. *Regression Analysis: Theory, Methods and Applications*. Springer-Verlag, New York, 1990.

U.S. Department of Transportation, Federal Railroad Administration, 2000a. *Use of Locomotive Horns at Highway-Rail Grade Crossings; Proposed Rule.* 49 CFR Parts 222 and 229. Federal Register, Vol. 65, No. 9, Jan. 13, 2000.

U.S. Department of Transportation, Federal Railroad Administration, Office of Safety, 2000b. *Updated Analysis of Train Whistle Bans.* Washington, D.C., January 2000.

U.S. Department of Transportation, Federal Railroad Administration, Office of Safety, 1995a. *Florida's Train Whistle Ban.* Washington, D.C., September 1995.

U.S. Department of Transportation, Federal Railroad Administration, Office of Safety, 1995b. *Nationwide Study of Train Whistle Bans.* Washington, D.C., April 1995.

U.S. Department of Transportation, Federal Railroad Administration, Office of Safety. *Rail-Highway Crossing Accident/Incident and Inventory Bulletin.* Washington, D.C. Published annually.

Acknowlegements

The paper presents results from a study undertaken under the sponsorship of the Illinois Department of Transportation (IDOT). The contents of the paper strictly reflect the views of the authors who gratefully acknowledge the support of Don Kopec and Parry Frank, CATS. The comments from two anonymous referees, as well as those from the JTRF General Editor are greatly appreciated.

Paul Metaxatos *is a Research Assistant Professor at the University of Illinois-Chicago (UIC) Urban Transportation Center (UTC). He holds a Ph.D. in Public Policy Analysis from UIC. Dr. Metaxatos has had almost 15 years of experience in transportation research. He has conducted research in the following areas: (a) transit planning, (b) transportation and land use, (c) intelligent transportation systems, (d) transportation modeling and database development, and (e) transportation impact analysis.*

P.S. Sriraj *is a Research Assistant Professor in the UTC/UIC. He holds a Ph.D. in Civil Engineering from the Illinois Institute of Technology. He has worked on research projects ranging from traffic flow simulation, land-use modeling and pedestrian analysis. Dr. Sriraj has worked on applying systems intervention techniques to urban problems, as well as other socioeconomic problems such as cost benefit analysis of impacts of transit, and environmental justice issues associated with transportation planning.*

Siim Sööt *is a Research Professor in the UTC/UIC. He holds a Ph.D. in Geography from the University of Washington. He was an associate professor of Geography at UIC. He has been associated with the UTC since its inception in the late 1970s and was its Director from 1998-2000. He has conducted research in transportation and geographic information systems and has published extensively. Since 1996 he has also been president of the Illinois Universities Transportation Consortium. The consortium includes UIC, University of Illinois at Urbana Champaign, Illinois Institute of Technology and Northwestern University.*

Joseph DiJohn *is a Research Professor at the University of Illinois-Chicago Urban Transportation Center. Mr. DiJohn received a bachelor's degree from Marquette University and a master's in Business Administration from DePaul University. He is director of the Metropolitan Transportation Support Initiative at the Urban Transportation Center. He is an Advisory Board member of the National Transportation Institute, a Project Panel member of the Transportation Cooperative Research Program, a member of Business Leaders for Transportation, a member of the Greater North Michigan Avenue Association's Traffic and Transportation Committee, and a director of the George Krambles Scholarship Foundation. Prior to joining the faculty in 1998, he was executive director of Pace, the Suburban Bus Division of the Regional Transportation Authority. Prior to that, he held various management positions in the transit, railroad and airline industries.*