

Marginal cost estimation for level crossing accidents: Evidence from the Swedish railway 2000-2008

May 2, 2011

Lina Jonsson
Department of Transport Economics
Swedish National Road and Transport Research Institute (VTI)
Stockholm, Sweden
lina.jonsson@vti.se

Abstract

This study examines the relationship between train traffic and the accident risk for road users at level crossing accidents. The marginal effect of train traffic on the accident risk can be used to derive the marginal cost per train passage that is due to level crossing accidents. Based on Swedish data from 2000 to 2008 over level crossing accidents, train volume and crossing characteristics, the marginal cost per train passage is estimated to 1.13 SEK on average in 2008. The cost per train passage varies substantially depending on type of protection, road type and the traffic volume of the trains.

Keywords: Railway; Marginal Cost Estimation; Level Crossing Accidents

Acknowledgements

This study was funded by the Swedish Transport Administration (former Rail Administration) and the financial support is gratefully acknowledged. The authors are solely responsible for the results and views expressed in this paper.

1 Introduction

Rail is in general a very safe transport mode but collisions between road users and trains on level crossings are still a problem due to the often severe outcome of the accidents. During the years 2004-2008, 79 level crossing accidents occurred at the Swedish rail network, including accidents with pedestrians, leading to 42 fatalities and 42 severe injuries among the road users (SIKA, 2009b).

Marginal cost pricing is an important keystone in Swedish transport policy. The infrastructure charge the Swedish Transport Administration charges the train operators includes a part due to rail-road level crossing accidents that should be based on the marginal cost principle. This means that the train operators should be charged with the expected cost due to level crossing accidents that proceeds from driving one more train on the line. The cost of interest here is the cost that without a charge completely falls on the road users or the rest of society and therefore is external to the train operators. Charging the operators for this external marginal cost even though they don't legally bear the responsibility for the accidents is a way of internalizing the effect that train traffic has on the accident risk of the road users.

Our focus lies in estimating the marginal cost associated with rail-road level crossing accidents. I.e. how much will the expected accident cost due to collisions between trains and road vehicles at a given crossing change when one more train passes the crossing? The expected accident cost depends on both the relationship between train volume and accident risk and the expected cost per accident. The relevant accident cost is the cost that falls on the road users and is taken from the official Swedish values of fatalities and injuries used in cost benefit analysis (SIKA, 2009c).

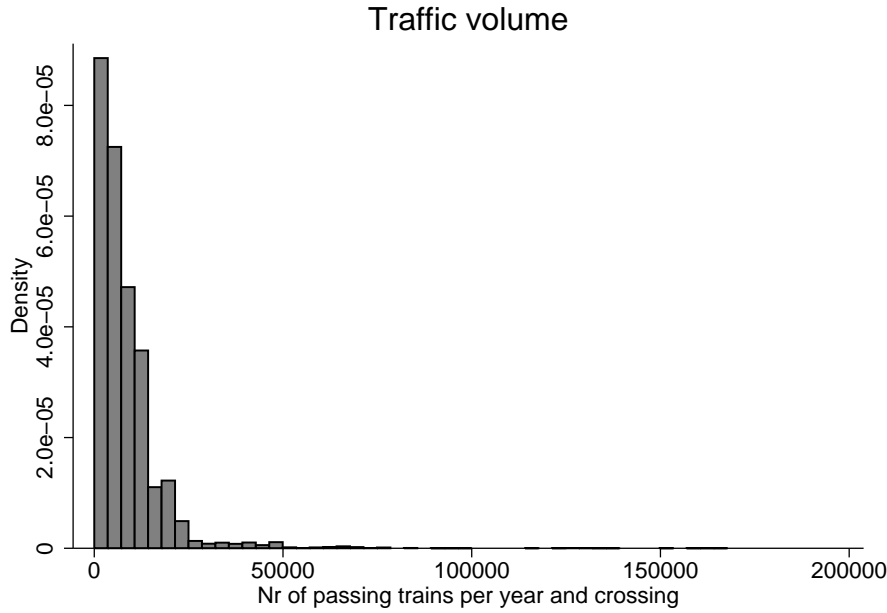
Except from Sweden, few if any other countries include the external marginal (level crossing) accident cost in the infrastructure charge for railway traffic. Studies on the relationship between train traffic and accident risk for road users on level crossings are therefore rare.

2 Data

The information on crossings, traffic and accidents are all taken from the Swedish Transport Administration. The information on traffic volume (nr of trains) is collected on a yearly basis and is an average over the whole track section. The number of track sections varies over the years as sections are divided or merged together, new opens and some are closed. In total the dataset consists of 241 different tracks sections from 1999 to 2008 while the numbers used in the analysis, sections with information on both traffic and existing crossings are only 208. The length of the track sections varies from less than one km to over 213 km and the number of crossings at each section varies from only one or two to over 200 crossings. Also the amount of traffic on each section varies substantially as shown in Figure 1. The distribution is skewed with a median value at 5696 train passages but with a few crossings with more than 100 000 passing trains per year.

The Swedish Transport Administration has a comprehensive dataset over existing crossings with information on protection devices, speed limit for the trains, and the type of road that are crossing the railway that we have been able to utilize for 2008. But to gain information back in time on crossings that have been removed or changed is harder and the comprehensive dataset has for the years 2002-2007 been supplemented with information from inspections of crossings. This data is further supplemented by information from 2000 and 2004 that come from a former analysis over accidents on road

Figure 1: Traffic volume distribution



rail level crossings presented in Lindberg (2006).

The data on crossings used in the analysis covers 9 years. During this period some crossings are being closed, others are reconstructed with a new type of protection device while also some new crossings are being built. This means that our dataset is an unbalanced panel but the variation over time within the same crossing when it comes to traffic and protection device is very small compared to the variation between crossings.

The crossings are divided into four categories based on protection device: full barriers, half barriers, light/sound and unprotected/crossings with St Andrews cross. Full barriers are barriers that close both the approach side of the crossing and also the exit side while half barriers only close the road at the approach side. The category light/sound consists of crossings without barriers but with warning devices in the form of flashing lights and/or sound. The fourth category consists of crossings with neither barriers nor lights or sounds. Some of these crossings are equipped with a St Andrews cross or other

Table 1: Nr of crossing 2000-2008

Year	Full barriers	Half barriers	Lights/sound	Unprotected	Total
2000	1.178 (3)	1.003	691	5.638 (6)	8.510 (9)
2001	1.066	943 (1)	593 (3)	4.513 (4)	7.115 (8)
2002	1.114	979	620 (2)	4.675 (9)	7.388 (11)
2003	1.120 (2)	982 (1)	606	4.653 (2)	7.361 (5)
2004	1.202 (4)	1.032 (3)	627 (3)	4.693 (8)	7.554 (18)
2005	1.213 (1)	1.046 (4)	667 (2)	4.781 (8)	7.707 (15)
2006	1.238	1.055 (4)	687	4.350 (5)	7.330 (9)
2007	1.249 (1)	1.060 (2)	682 (1)	4.281 (8)	7.272 (12)
2008	1.291	1.062	682 (1)	4.337 (2)	7.372 (4)

Number of accidents in parenthesis

simple devices while others are totally unprotected. The common category is motivated by a former study (Cedersund, 2006) on Swedish level crossings showing that crossings with and without St Andrews cross are equally risky. Due to the fact that the Swedish Transport Administration doesn't categorize accidents between pedestrians and trains as crossing accidents, crossings with a footpath are excluded from the analysis. This also means that the marginal cost estimated in the paper only covers accidents involving road vehicles, not pedestrians.

The information on accidents are collected from the Swedish Transport Administration. The accident record utilized for the analysis consists of information on level crossing accidents involving road vehicles. A description of the accident including the location is included in the record but to be able to connect all the accidents to the exact crossing some detective work has been required. For each accident also the injuries, categorized as light injuries, severe injuries and fatalities are noted. Only accidents leading to personal injuries are included in the analysis.

3 Modelling the accident probability

Count regression models like the Poisson model or the negative binomial model are natural choices when modeling the number of events during a given time period. In situations

with a high proportion of zeroes, their zero-inflated counterparts, the ZIP and ZINB are also applicable. But looking at our dataset, at most crossings no accident at all occurs during the 9 years that our data covers and only one crossing has more than one accident during the period. A better choice is therefore to model the probability that one (or several) accident(s) will occur at a given crossing during a certain time period, in this case a year, using the logit model.

The probability that an accident occurs at a crossing during a year is a function of the number of passing trains and crossing characteristics like protection device, sight distance, number of tracks and the crossing angle. Our dataset lacks many of the variables that should be included in a complete model but we at least have access to information on protection device and train passages. A common measure of the exposure used by the U.S. Department of Transportation Accident Prediction Model among others is the product of highway and train traffic (Ogden, 2007). Our dataset lacks information on road traffic which precludes the use of this measure in the analysis. Instead, to capture the influence from road traffic, information on the type of road that crosses the railway is used as a proxy variable for road traffic flow, an approximation that has proven to work well by Lindberg (2006) in a former study on Swedish data.

For each year from 2000 to 2008 we observe whether or not an accident occurs at an existing crossing. Our dependent variable is dichotomous, accident or no accident, and we have information on the type of protection that the crossing is equipped with, the type of road that crosses the railway and the number of passing trains.

The fact that our dataset over crossings is a panel opens up for estimation methods that use the variation in accident risk, traffic and crossing characteristics within the same crossing over time to estimate the effect of traffic on the accident risk. The fixed effects

estimator uses a time-invariant individual specific constant to get unbiased and consistent estimates even in case of unobserved effects that are correlated with the regressors. The downside with the fixed effects estimator is that time-constant variables cannot be included and that the within-variation, the variation within the same crossing over time, is the only source behind the estimation of the effect of train traffic on the accident risk. In cases where the variation over time within the same crossing is very small compared to the variation between crossings the fixed effects estimator is not a suitable alternative. The random effects estimator uses both the variation within crossing and the variation between crossings and is a good choice if it can be assumed that unobserved individual specific effects are uncorrelated with the regressors. If the variation within crossing over time is very small the random effects estimator approaches the pooled estimator.

In our dataset the variation over time within the same crossing when it comes to train passages is very small. The fixed-effects estimator is therefore not an appropriate choice. The estimation of a random effects logit model shows that the within-variation is insignificant, i.e. the variation over time within the same crossing is so small that it cannot help explain the variation in accident probability. Due to this fact the models in the paper are estimated with the ordinary pooled logit with standard errors corrected for clustering on crossings. The panel character of our dataset will therefore not add any additional value to our study.

4 Results

4.1 Model specification

The focus of our study lies in estimating the effect of train traffic on the accident risk. This effect might vary depending on other crossing characteristics like type of protection

and it might also vary depending on the existing traffic volume. A hypothesis is that more frequent traffic increases the probability of an accident by increasing the occasions when a train can collide with a road vehicle. In other words will the exposure increase with the traffic volume of both trains and road vehicles. Also the speed of both the trains and the road vehicles influences the accident risk. At the same time, a crossing with more frequent train traffic will induce a safer behaviour from the road users that reduces the probability of an accident. This latter effect due to a changed behaviour among the road users could in some traffic situations override the effect from more collision occasions. In that case the accident probability would fall with the number of passing trains and the marginal cost would be negative. But a safer behaviour is not costless. This risk-reducing behaviour in the form of speed reduction or the extra anxiety that the road user feels when passing a crossing that is perceived as unsafe should be included in a full measure of the accident cost. Unfortunately, it is impossible or at least very hard to observe this risk-reducing behaviour and our measure of the accident externality from train traffic therefore only includes the estimated effect on the accident probability and not the increase in accident avoidance costs for the road users.

Theory gives us no direct guidance when it comes to model specification. Three natural choices are to estimate the accident probability as a:

i, linear function of train passages (Q)

$$P(y = 1|X, Q) = \Lambda(X'\beta + \delta Q + \varepsilon), \quad (1)$$

ii, function including a quadratic term to capture increasing/decreasing effects

$$P(y = 1|X, Q) = \Lambda(X'\beta + \delta Q + \gamma Q^2 + \varepsilon), \quad (2)$$

Table 2: Marginal effect - different specifications

	Full dataset			Reduced dataset		
	Linear Q	Incl. Q2	Log Q	Linear Q	Incl. Q2	Log Q
dP/dQ* mean	$2.41 \cdot 10^{-8}$	$1.26 \cdot 10^{-7}$	$1.85 \cdot 10^{-7}$	$1.91 \cdot 10^{-7}$	$1.87 \cdot 10^{-7}$	$2.43 \cdot 10^{-7}$
dP/dQ* median	$1.43 \cdot 10^{-8}$	$6.13 \cdot 10^{-8}$	$9.86 \cdot 10^{-8}$	$1.31 \cdot 10^{-7}$	$1.07 \cdot 10^{-7}$	$1.35 \cdot 10^{-7}$
AIC	1240.09	1224.21	1222.39	1084.33	1083.42	1080.34
BIC	1302.77	1295.85	1285.07	1146.27	1154.21	1142.28
N	57216	57216	57216	51470	51470	51470

* Mean/median of observation specific marginal effects

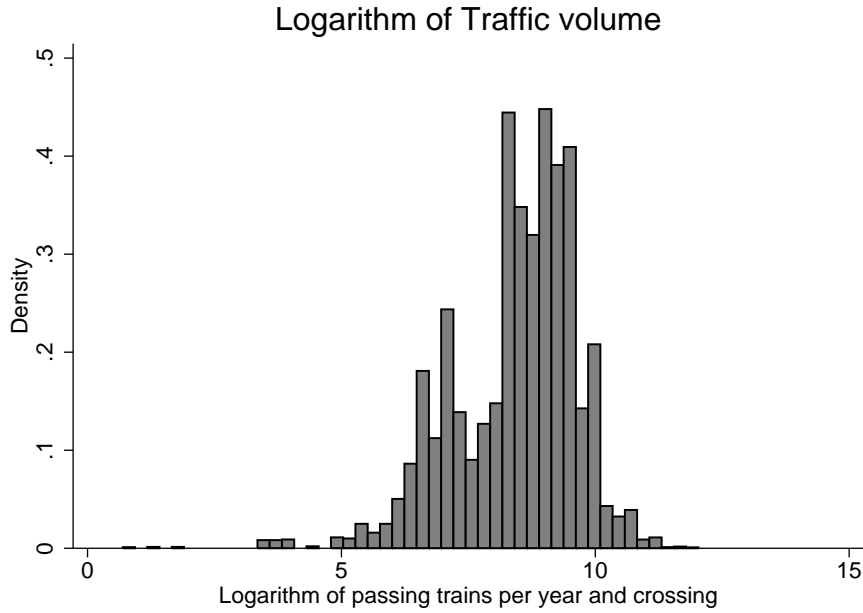
iii, function of the natural logarithm of train passages

$$P(y = 1|X, Q) = \Lambda(X'\beta + \eta \ln(Q) + \varepsilon). \quad (3)$$

The fact that the distribution of train passages is extremely skewed (see Figure 1) complicates the analysis. By taking the natural logarithm of train passages the variable becomes more symmetric as can be seen in Figure 2. Another way of reducing the problem with a few crossings with extremely high traffic volumes is to simply restrict the estimation to the crossings with more modest traffic volumes. Table 2 shows the result from three models estimated on both the full dataset and a dataset where the crossings with the 10% highest traffic volumes have been removed. In a logit model the marginal effect (dP/dQ) varies depending on the values of all independent variables. A general marginal effect has therefore been calculated by taking the mean of the crossing specific marginal effect. For comparison also the median is shown since the distribution of the marginal effect is skewed. It can be seen that the marginal effect varies substantially depending both on functional form, the sample used and also between the mean and the median.

According to the Akaike Information Criteria (AIC) and the Bayes Information Criteria (BIC) the model using the logarithm of train passages is the preferred model for both datasets. The AIC and BIC cannot be used to compare models between samples. The

Figure 2: Logarithm of Traffic volume distribution



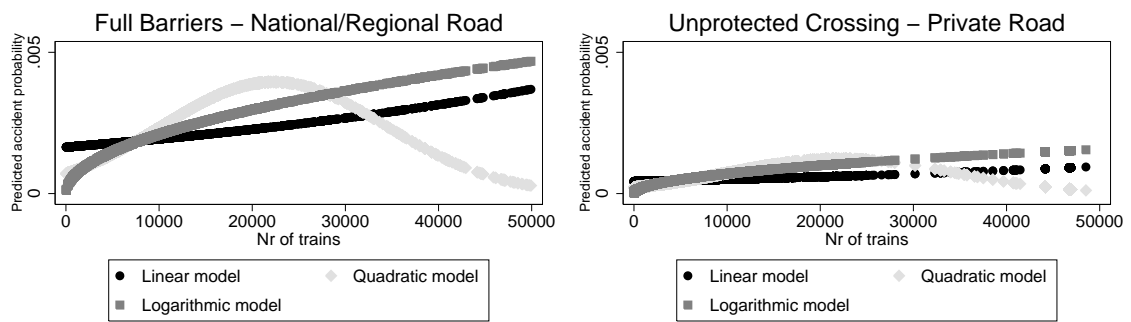
fact that the model with the logarithm of train passages is preferred makes the reduction of the sample unnecessary.

The choice of functional form influences how the predicted accident probabilities as well as the marginal effect vary over the traffic interval. Predicted accident probabilities and marginal effects for crossings with full barriers crossing a national/regional road and unprotected crossings crossing a private road for all three models using the full sample are shown in figure 3. To make the graphs easier to read only predicted probabilities and marginal effects for traffic up to 50 000 passages/year are shown, thereby reducing the dataset by less than 1%.

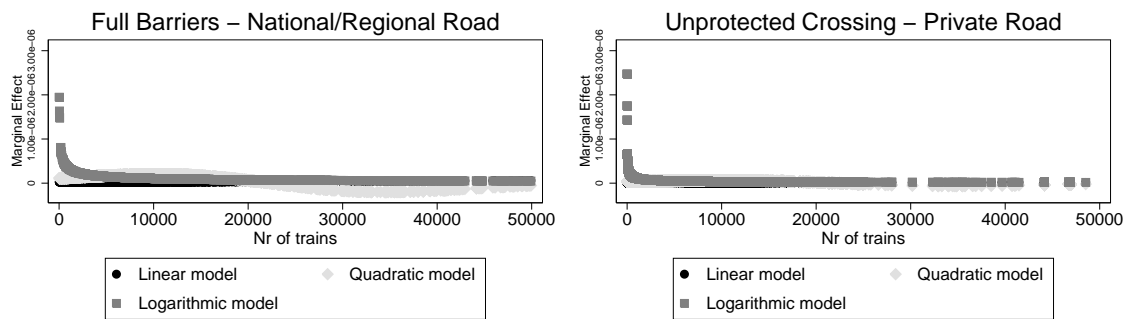
The marginal effect of train passages on the accident probability varies in different ways over the traffic interval depending on functional form. Since the marginal cost is a direct function of the marginal effect this will have a large impact on the accident charge if the charge should vary depending on traffic volume. For the model with logarithmic

Figure 3: Predicted Accident Probabilities and Marginal Effects

Predicted Accident Probabilities



Marginal Effect



traffic the marginal effect as a function of train traffic is continuously decreasing, as seen in figure 3, which is reassuring given that the train volume influences the behaviour of the road users. Based on both the AIC/BIC results and the shape of the marginal effect the model with logarithmic traffic volumes and the full dataset is used in the rest of the analysis. Regression results from this model are shown in table 3. The logarithm of train passages ($\ln(Q)$) increases the accident probability and is highly significant. The road type variables are significant and with the expected signs where crossings with streets/other roads and private roads has a lower accident probability than the reference category national/regional roads. Crossings with full and half barriers have a lower accident probability than the reference category crossings with lights/sound while the unprotected crossings has a (insignificantly) higher accident probability. Train speed probably also influences the accident probability and one way to capture train speed is to distinguish between freight trains and passenger trains where freight trains in general are slower than the passenger trains. Unfortunately we have not been able to separate the effect from different train types in the estimation.

4.2 Marginal effects and crossing characteristics

The marginal effect varies depending on crossing characteristics as well as the traffic volume. Table 4 shows calculated marginal effects from the model using the logarithm of train traffic on the full sample for crossing with different protection devices and road types. The marginal effects are calculated at mean traffic (7982 train passages/year) for the sample. The safer the protection device the lower is the estimated marginal effect where the unprotected crossings have a marginal effect that is almost 7 times higher than the safest crossings with full barriers, given the same number of train passages. The

Table 3: Regression Results

	b/t
ln(Q)	0.501*** (5.41)
Street/other road	-1.246*** (-3.98)
Private road	-3.003*** (-6.16)
Full barrier	-1.277** (-2.93)
Half barrier	-0.869* (-2.20)
Unprotected	0.631 (1.62)
Constant	-9.504*** (-11.05)
N	57216
AIC	1222.385
BIC	1285.067

Standard errors corrected for clustering on crossing
Significance levels: * 5%, ** 1%, *** 0.1%

Table 4: Marginal effect for different crossings - mean traffic

	Full barrier	Half barrier	Light/sound	Unprotected
National/Regional	$1.17 \cdot 10^{-7}$	$1.76 \cdot 10^{-7}$	$4.17 \cdot 10^{-7}$	$7.75 \cdot 10^{-7}$
Street/other road	$3.39 \cdot 10^{-8}$	$5.09 \cdot 10^{-8}$	$1.21 \cdot 10^{-7}$	$2.27 \cdot 10^{-7}$
Private road	$5.85 \cdot 10^{-9}$	$8.80 \cdot 10^{-9}$	$2.10 \cdot 10^{-8}$	$3.94 \cdot 10^{-8}$

road types seems to work good as proxies for road traffic volume where the national and regional roads have a marginal effect that is around 20 times as high as the smallest roads (private roads).

Some crossing types are more common than others as can be seen in Table 5. There is a clear tendency that the barriers are more common on crossings with road types with larger traffic volumes.

Table 5: Nr of crossings 2008 in estimation sample

	Full barrier	Half barrier	Light/sound	Unprotected	Total
National/Regional	423	391	85	3	902
Street/other road	687	562	378	1281	2908
Private road	20	11	2	1659	1692
Total	1130	964	465	2943	5502

Table 6: Accident cost

	Fatality	Severe Injury	Light Injury
Valuation (SEK)	22 321 000	4 147 000	199 000

Official Swedish values taken from SIKÅ (2009c)

1 SEK \approx 0.1 Euro

5 Marginal cost

The marginal cost per train passage can be calculated as the marginal effect multiplied with the expected accident cost. Since the marginal effect is crossing specific also the marginal cost will vary depending on traffic volume, protection device and type of road.

$$MC = dP/dQ * E(Cost), \quad (4)$$

The accident cost relevant for the accident charge is the cost that without a charge will be external to the train operators. We have taken this cost to equal the cost that is due to injuries and fatalities among the road users involved in the accidents. For each crossing we have information on the number of fatalities, severe injuries and light injuries among the road users involved. The values for the injuries come from the official Swedish values used in cost benefit analysis and cover both material costs in the form of lost income and health care and the risk valuation, see table 6.

The average accident cost for the accidents used in the analysis is 12 084 635 SEK. No relation can be seen between the accident cost and crossing characteristics. Table 7 shows marginal cost estimates per passage for different crossing types at mean traffic volumes (7982 train passages/year).

Table 7: Marginal cost per train passage for different crossings - mean traffic (SEK)

	Full barrier	Half barrier	Light/sound	Unprotected
National/Regional	1.41	2.12	5.03	9.36
Street/other road	0.41	0.615	1.46	2.74
Private road	0.0707	0.106	0.254	0.476

1 SEK \approx 0.1 Euro

6 Discussion

The part of the access charge that relates to level crossing accidents can be based on the marginal cost per train passage estimated in this study. An extremely differentiated charge can be set where the train operators are charged for every crossing passage depending on the characteristics of the crossing including the traffic volume. A more realistic approach is probably to instead calculate a charge per km that varies depending on track section.

The accident charge set by the Swedish Transport Administration is now a uniform charge per km independent on section of the rail network. A uniform charge per km can be calculated using the crossing specific calculated marginal cost weighted by the train traffic, i.e. crossings with a lot of train traffic will be given a heavier weight than the crossings on the part of the network that is sparsely used. Such a calculation gives an average marginal cost per train passage at 1.13 SEK in 2008.

According to official statistics (SIKA, 2009a) the Swedish state owned rail network with traffic consisted of 9 830 km of rail and 8 054 level crossings including foot path crossings in 2008. Using these official numbers gives 0.82 level crossings per km and an accident charge per km at 0.92 SEK. The official numbers differ quite substantially from the numbers given by our data. A part of the difference in the number of crossings is a consequence of the fact that our dataset excludes crossings with footpaths due to the

fact that collisions involving pedestrians is excluded from the accident record utilized for the analysis. In 2008, 531 crossings with footpaths existed on the State owned rail network according to our data giving a total number of crossings including footpaths of 7900. The discrepancy in track length is much larger and can be explained by the fact that our dataset over track length excludes many station areas, marshalling yards and also the part of the state owned network that is managed by Inlandsbanan AB. Using the track length and number of crossings according to our dataset for the track sections where we have information on traffic instead gives 0.66 crossings per km and a marginal cost of 0.74 SEK/train km. This charge should be used for track sections **excluding** Inlandsbanan and not for station areas or marshalling yards. This charge also excludes crossings with footpaths. Instead calculating a charge per km based on the official track length and number of crossings less the number of footpaths according to our dataset gives a charge of 0,86 SEK/train km.

The accident charge today in Sweden due to level crossing accidents is set to 0.24 SEK/train km (Swedish Rail Administration, 2009) based on a similar study using accident records for 1995-2004 (Lindberg, 2006). The values presented in this paper would imply a substantial increase in the part of the accident charge that is due to level crossing accidents.

The disparity between our results and the results in Lindberg (2006) is mostly due to the choice of functional form. The marginal cost in Lindberg (2006) is based on an estimation of the accident probability using the number of train passages per se, not the logarithm of train passages. Estimating the linear model (eq. 1) instead of the loglinear model (eq. 3) on our dataset will result in a lower marginal effect and thereby a lower (weighted) marginal cost at 0.42 SEK/train passage. This gives a marginal cost

per km at 0,28-0,34 SEK/train km, only slightly higher than the results in Lindberg (2006). The choice of functional form has accordingly a substantial influence on the calculated marginal cost but is somewhat arbitrary and our choice is mainly based on test diagnostics, i.e. the AIC-test.

References

- Cedersund, H.-Å.: 2006, 'Trafiksäkerhet i plankorsningar mellan väg och järnväg 1999-2004'. Report 540, VTI. In Swedish.
- Lindberg, G.: 2006, 'Valuation and Pricing of Traffic Safety'. Örebro studies in economics 13, Örebro University.
- Ogden, B. D.: 2007, 'Railroad-Highway Grade Crossing Handbook - Revised Second Edition 2007'. Technical Report FHWA-SA-07-010, Office of Safety Design Federal Highway Administration.
- SIKA: 2009a, 'Bantrafik 2008'. Statistik 22, Statens institut för kommunikationsanalys. In Swedish.
- SIKA: 2009b, 'Bantrafikskador 2008'. Statistik 15, Statens institut för kommunikationsanalys. In Swedish.
- SIKA: 2009c, 'Värden och metoder för transportsektorns samhällsekonomiska analyser ASEK 4'. Rapport 3, Statens institut för kommunikationsanalys. In Swedish.
- Swedish Rail Administration: 2009, 'Underlagsrapport avgifter i Banverkets järnvägsnätsbeskrivning 2011'. Rapport, Banverket. In Swedish.