Marginal cost estimation for level crossing accidents: Evidence from the Swedish railways 2000-2012

Lina Jonsson & Gunilla Björklund

Department of Transport Economics

Swedish National Road and Transport Research Institute (VTI)

Stockholm, Sweden

Abstract

This study examines the relationship between train traffic and the accident risk for road users at level crossings. 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 2012 on level crossing accidents, train volume and crossing characteristics, the marginal cost per train passage is estimated at SEK 1.28 (EUR 0.13) on average in 2012. The cost per train passage varies substantially depending on type of warning device, road type and the traffic volume of the trains.

1. Introduction

Rail is in general a very safe transport mode but collisions between road users and trains at level crossings are still a problem due to the often severe outcome of the accidents. During the years 2008-2012, 59 level crossing accidents (of which 37 were collisions with motor vehicles) occurred on the Swedish rail network, leading to 34 fatalities and 29 severe injuries among the road users. Corresponding numbers for years 2003-2007 were 83 accidents (of which 65 were collisions with motor vehicles), leading to 41 fatalities and 47 severe injuries among the road users. (Trafikanalys, 2013b)

Marginal cost pricing is an important keystone in Swedish transport policy. The infrastructure charge made by the Swedish Transport Administration to the train operators includes a component for 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 results 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 is therefore external to the train operators. Charging the operators for this external marginal cost even though they do not 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 (Trafikverket, 2014).

Apart 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 at level crossings are therefore rare.

2. Marginal cost charging and level crossing accidents

Accidents between road vehicles and trains at level crossings are almost always caused by some kind of misbehaviour from the road user. Either by approaching the crossing at high speed and thereby not observing flashing lights or closing barriers or even by intentionally disregarding warning signs. It might therefore seem remarkable to put a charge on the train operators that internalizes the costs that otherwise are completely borne by the road users. A theoretical motivation for using marginal cost based charges can be found in the accident and law literature on how liabilities and costs should be split between involved parties to achieve optimal risk reduction at lowest cost presented in Shavell (2004). Accidents between road users and trains at level crossings are bilateral as the actions in the form of care taking and the activity level of both the road user and the train affect the accident risk. Even though it is almost impossible for a train to take any action to avoid a crash when approaching a crossing with a car standing on the track (due to the long stopping distance), the level of activity, i.e. the number of times a train passes a crossing, does affect the accident risk. For the road user both the amount of care taking when crossing the railway and the number of times he crosses the railway (the activity level) affect the accident risk.

There are two major rules of accident liability. Strict liability implies that the injurer is liable for the harm he causes regardless of whether he was negligent or not. Under the negligence rule on the other hand the injurer is only liable if his level of care is below some minimum standard specified by the court. As Shavell (2004) shows the rules of liability affect both the behaviour and chosen activity level of the injurer and the victim but no liability rule, neither strict liability nor negligence, will in itself lead to an optimal level of activity for both parties in bilateral accidents. A condition for an optimal choice of activity level of both parties is that they both bear the accident losses. The charges that the train operators pay in Sweden for the expected increase in accident costs for road users due to level crossing accidents is one way to make both the train operators and the road users pay for the accident losses that their use of infrastructure results in. The largest part of the losses from a level crossing accident comes from injuries of the passengers in the road vehicle and material damage to the road vehicle. These are borne by the road user and the rest of society when it comes to health care. By charging marginal cost based charges the train operators will also take into account the effect on the accident risk from train traffic. In this way, both parties, the train operator and the road user, each face the full accident consequences from level crossing accidents and will therefore both choose the optimal level of traffic.

3. Data

The information on crossings, traffic and accidents is all obtained from the Swedish Transport Administration. The information on traffic volume (no. of trains) is collected on a yearly basis and is an average over the whole track section with imputed data for the station areas. Track sections with a traffic volume of less than one train per year are excluded from the analyses. The number of track sections varies over the years as sections are divided or merged, new sections open and some are closed. The number of different track sections used in the analysis, sections with information on both traffic and existing crossings are 214. The length of the track sections varies from less than one km to over 273 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 mean traffic volume of 6 801, a median value at 4 551 and a few crossings with more than 50 000 passing trains per year.
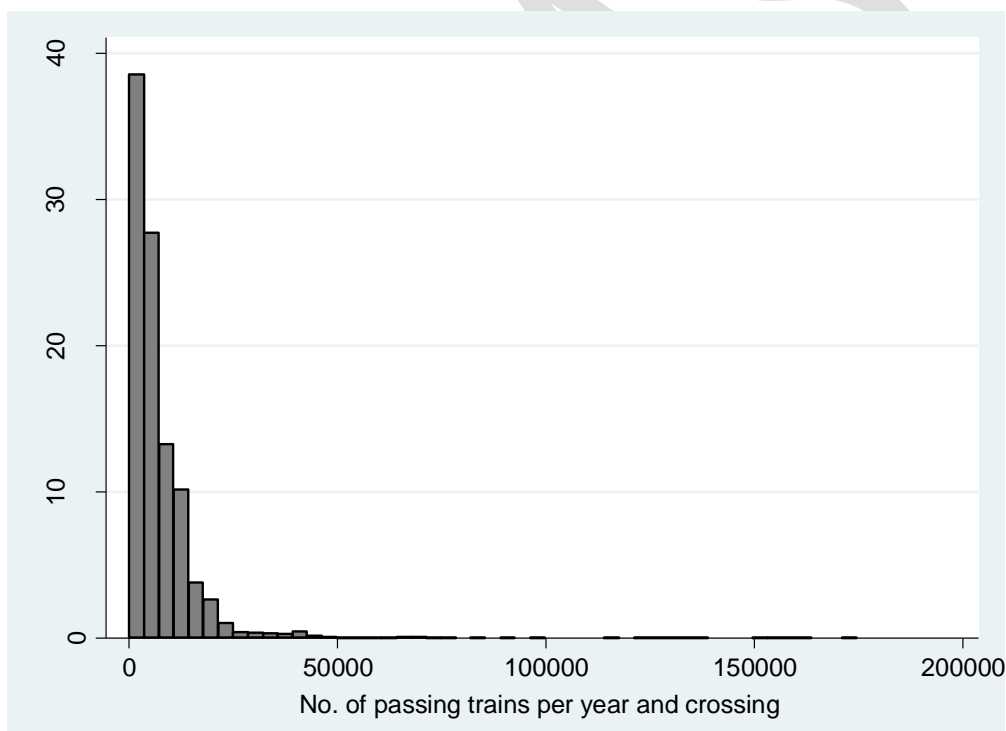

Figure 1. Traffic volume distribution

The Swedish Transport Administration has a comprehensive dataset over existing and closed crossings with information on warning devices, speed limit for the trains, and the type of road crossing the railway, among other things. From that dataset we have gained information about crossings between the years 2008 to 2012. But to gain information back in time on crossings

that have been closed 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 comes from a former analysis over accidents on road rail level crossings presented in Lindberg (2006).

The data on crossings used in the analysis covers thirteen years. During this period some crossings have been closed, others reconstructed with a new type of warning device while also some new crossings have been 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 warning devices is very small compared to the variation between crossings[1].

The crossings are divided into four categories based on warning device: full barriers, half barriers, light/sound and unprotected/crossings with crossbucks. 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 passive crossings with neither barriers nor lights or sounds. Some of these crossings are equipped with crossbucks or other 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 crossbucks are equally risky. Due to the fact that the Swedish Transport Administration earlier did not categorize accidents between pedestrians and trains as crossing accidents, these accidents are excluded from the analysis. We also do not have any information about number of pedestrian passing the crossings, i.e. the exposure. This means that the marginal cost estimated in the paper only covers accidents involving motor vehicles, not pedestrians and bicyclists.

The information on accidents has been obtained from the Swedish Transport Administration. For the years 2009 to 2012 the accident record is included in the level crossing data but for the other years the accident record has been gained from another dataset. Some detective work has been required to be able to connect all the accidents to the exact crossing. For each

---

[1] An inspection of the data shows that 94 level crossings have been reconstructed with a new warning device sometime between 2010 and 2012. However, it is difficult from the information in the data to find out when, and maybe if at all, this reconstruction has been made. For these crossings, we have set the year of the reconstruction to 2010.

accident the injuries, categorized as light injuries, severe injuries and fatalities, are also noted. Only accidents leading to personal injuries are included in the analyses.

In Table 1, the number of crossings and accidents is presented for each of the years included in the study.

Table 1. No. of crossings in the sample 2000-2012

Year	Full barriers		Half barriers		Lights/sound		Unprotected		Total	
2000	1 016	(3)	926	-	561	-	4 446	(6)	6 949	(9)
2001	1 012	-	921	(1)	529	(1)	4 116	(4)	6 578	(6)
2002	1 075	-	947	-	634	(2)	4 221	(8)	6 877	(10)
2003	977	(2)	924	(1)	519	-	3 435	(2)	5 855	(5)
2004	1 009	(4)	947	(3)	516	(3)	3 471	(4)	5 943	(14)
2005	1 029	-	971	(4)	436	(2)	2 951	(6)	5 387	(12)
2006	1 040	-	967	(4)	428	-	2 775	(5)	5 210	(9)
2007	1 203	(1)	1 042	(1)	653	(2)	4 045	(7)	6 943	(11)
2008	1 216	(1)	1 038	-	638	(1)	3 957	(2)	6 849	(4)
2009	1 214	-	1 016	-	610	(1)	3 523	(1)	6 363	(2)
2010	1 180	(1)	958	(2)	554	(4)	3 102	(3)	5 794	(10)
2011	1 195	(2)	967	-	523	(1)	2 872	(1)	5 557	(4)
2012	1 180	(1)	963	-	498	(4)	2 718	(4)	5 359	(9)

Note. Number of accidents in parenthesis

Because several crossings have to be excluded from the analyses – due to missing data or other problems with the data – the numbers in Table 1 are smaller than the actual numbers of crossings. For example, according to official statistics (Trafikanalys, 2013a) there existed 7 380 level crossings in the Swedish rail network in 2012, including 500-600 footpath crossings.

4. Modelling the accident probability

Count regression models like the Poisson model or the negative binomial model are natural choices when modelling the number of events during a given time period. In situations with a high proportion of zeros, their zero-inflated counterparts, the ZIP and ZINB are also applicable. The theoretical motivation behind the zero-inflated models is a dual-state process which implies that, in this case crossings, exist in two states - safe and unsafe. As discussed in Lord (2005) the excess zeros in crash data often arise from low exposure or an inappropriate

selection of time/space scales and not an underlying dual-state process where some locations are totally safe. Lord (2005) therefore instead suggests a more careful selection of time/space scale for the analysis, improvements in the selection of explanatory variables, including unobserved heterogeneity effects into count regression models or applying small-area statistical methods to model motor vehicle crashes with datasets with a preponderance of zeros. Another choice of accident model is presented in Oh (2005) who models accidents at railway-highway crossings in Korea using a gamma probability count model that can deal with underdispersion as well as overdispersion.

But looking at our dataset, no accident at all occurs at most crossings during the 13 years covered by our data covers and only one crossing has more than one accident during the period. Instead of using a count model to model the number of accidents we 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.

$$P(y = 1|X) = \frac{e^{X'\beta}}{1+e^{X'\beta}} = \Lambda(X'\beta) \tag{1}$$

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. 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 been shown to work well by Lindberg (2006) in a previous study using Swedish data.

For each year from 2000 to 2012 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 warning device 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 on 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 the

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 a 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 a 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 a pooled logit with clustered robust standard errors where each cluster consists of one crossing.

5. Results

5.1. 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 number of occasions when a train can collide with a road vehicle. In other words, the exposure will increase with the traffic volume of both trains and road vehicles. The speed of both the trains and the road vehicles also influences the accident risk. At the same time, a crossing with more frequent train traffic will induce safer behaviour from the road users that reduces the probability of an accident. This latter effect due to 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 safer behaviour is not without cost. 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

8

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. A level crossing accident may also lead to costs in the form of time delays for both train users and road users. This cost is not included in our estimates.

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)

$$P(y = 1|X, Q) = \Lambda(X'\beta + \delta Q), \tag{2}$$

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

$$P(y = 1|X, Q) = \Lambda(X'\beta + \delta Q + \gamma Q^2) \tag{3}$$

iii, function of the natural logarithm of train passages

$$P(y = 1|X, Q) = \Lambda(X'\beta + \eta \ln(Q)) \tag{4}$$

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.
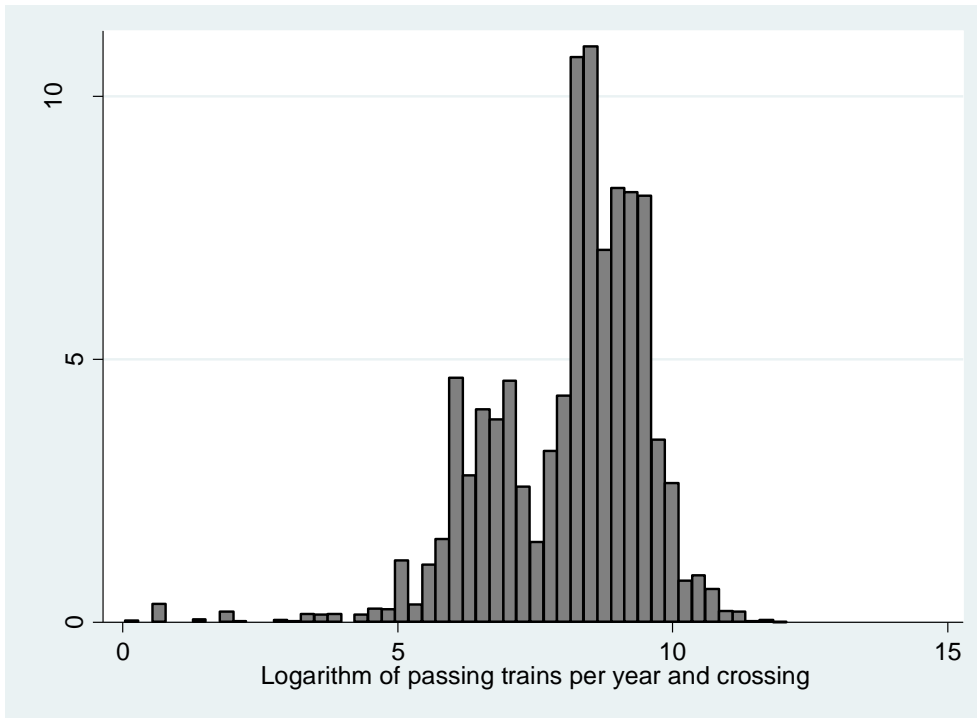
Figure 2. Logarithm of traffic volume distribution

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. In Table 2 it can be seen that the marginal effect varies substantially depending both on functional form and between the mean and the median.

Table 2. Marginal effect – different specifications

	Linear Q	Incl. Q^2	Log Q
dP/dQ×mean	$2.04 \cdot 10^{-8}$	$1.38 \cdot 10^{-7}$	$2.79 \cdot 10^{-7}$
dP/dQ×median	$1.27 \cdot 10^{-8}$	$6.47 \cdot 10^{-8}$	$9.63 \cdot 10^{-8}$
AIC	1547.38	1523.66	1519.11
BIC	1612.38	1597.94	1584.11
N	79 664	79 664	79 664

The Akaike Information Criteria (AIC) and the Bayes Information Criteria (BIC) point towards using the model with the logarithm of train passages compared to the model with train passages directly.

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 are shown in Figure 3 and 4, respectively. 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.

The model including a quadratic term gives a decreasing accident probability for high train volumes and thereby a negative marginal effect for crossings with high train volumes, something that is problematic from the view of charging the marginal cost to the train operators. For the model with logarithmic traffic the marginal effect as a function of train traffic is continuously decreasing but positive, as seen in Figure 3 and 4, which is reassuring given that the train volume influences the behaviour of the road users.
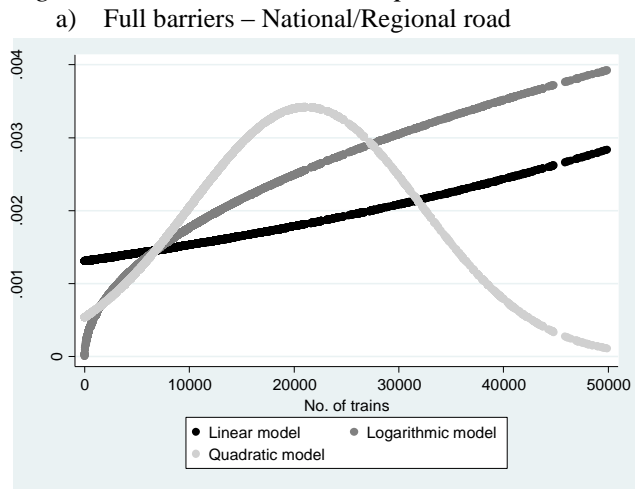
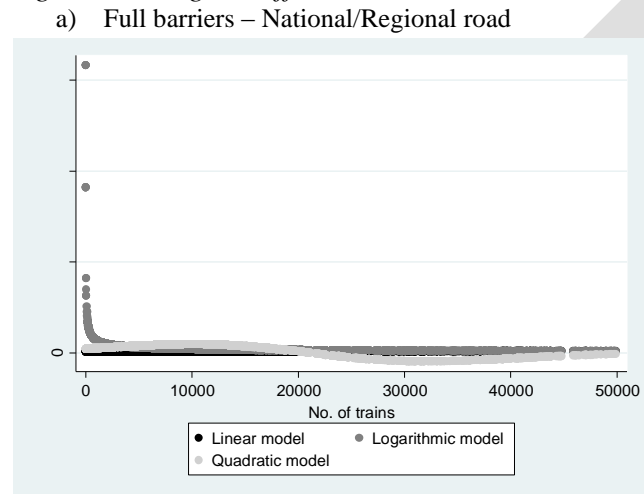*Figure 3. Predicted accidents probabilities*

a) Full barriers – National/Regional road

b) Unprotected crossing – Private road

Figure 4. Marginal effects

a) Full barriers – National/Regional road

c) Unprotected crossing – Private road

Based on both the AIC/BIC results and the shape of the marginal effect the model with logarithmic traffic volume 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 have a significant lower accident probability than the reference category national/regional roads. Crossings with full and half barriers have a significant lower accident probability than the reference category crossings with lights/sound while the unprotected crossings does not differ from the reference category. Train speed probably also influences the accident probability and one way of capturing train speed is to distinguish between freight trains and passenger trains where freight trains in general are slower than passenger trains. However, we did not found any separate effect from different train types in the estimation.

Table 3. Regression results from the logarithmic model

	b	S.E.
Constant	-9.03***	.70
Ln(Q)	.50***	.07
Street/other road	-1.02***	.28
Private road	-2.98***	.47
Full barrier	-1.90***	.37
Half barrier	-1.66***	.32
Unprotected	-.06	.28
AIC	1519.11	
BIC	1584.11	
N	79 664	

Note. Standard errors are corrected for clustering on crossing.

*p < .05, **p < .01, ***p < .001

5.2. Marginal effects and crossing characteristics

The marginal effect varies depending on crossing characteristics as well as the traffic volume. Table 5 shows calculated marginal effects from the model using the logarithm of train traffic for crossing with different warning devices and road types. The marginal effects are calculated as a weighted average marginal effect over all train passages year 2012 in each crossing type. In that way, crossings with larger traffic volume have a larger impact on the

marginal effect than crossings with smaller traffic volume. Because of different number of train passages comparisons between warning devices or road types can not be done.

Table 5. Marginal effect for different crossings – weighted average traffic

	Full barrier	Half barrier	Light/sound	Unprotected
National/regional	6.57×10^{-8}	9.98×10^{-8}	1.14×10^{-6}	-
Street/other road	2.63×10^{-8}	3.71×10^{-8}	2.58×10^{-7}	2.39×10^{-7}
Private road	3.09×10^{-9}	4.20×10^{-9}	2.56×10^{-8}	3.91×10^{-8}

Some crossing types are more common than others as can be seen in Table 6. There is a clear tendency that barriers are more common on crossings with road types with larger traffic volumes, which also is supported by a Chi-square test ($\chi^2(6)=2300$, $p<.001$).

Table 6. No. of crossings in the sample year 2012

	Full barrier	Half barrier	Light/sound	Unprotected	Total
National/regional	428	410	109	0	947
Street/other road	733	542	323	1 256	2 854
Private road	19	11	66	1 462	1 558
Total	1 180	963	498	2 718	5 359

6. Marginal cost

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

$$MC = dP/dQ \times E(Cost) \tag{5}$$

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 7.

14

Table 7. Accident costs

	Fatality	Severe	Light injury
Valuation (SEK)	23 739 000	4 412 000	217 000

Note. Official values taken from Trafikverket (2014). Price level of 2010. SEK 1 = EUR 0.1

The 106 accidents included in the study resulted in 54 fatalities, 28 severe injuries, and 51 light injuries. The average accident cost for the accidents is SEK 14 026 453. Table 8 shows weighted average marginal cost estimates per passage for each combination of road type and warning device where crossings with many passages have a higher weight than crossings with few passages. Because the marginal effect decreases with the number of passages, the differences between the crossings increases when weighting by the number of passages than when taking an unweighted average across the crossings. The difference between crossings reflect both differences in warning device, road type, and number of train passages.

Table 8. Marginal cost per train passage for different crossings – weighted average traffic year 2012 (SEK)

	Full barrier	Half barrier	Light/sound	Unprotected
National/regional	0.92	1.40	15.97	-
Street/other road	0.37	0.52	3.62	3.35
Private road	0.04	0.06	0.36	0.55

Note. SEK 1 = EUR 0.1

7. 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 SEK 1.28 in 2012.

According to official statistics (Trafikanalys, 2013a) the Swedish state-owned rail network with traffic consisted of 11 136 route km and 7 380 level crossings, including footpath crossings, in 2012. There is no official statistics of how many the footpath crossings are, but according to the dataset used in this paper, there were 576 such crossings 2012. Using this as an approximation of actual number of foothpath crossings and deduct these from the official numbers gives 0.61 level crossings per km and an accident charge per km at SEK 0.78, for motor vehicle accidents.

To model the accident probability for foothpath accidents one must have a proxy variable for exposure. In this paper we used road type as proxy variable for volume of road traffic, but for pedestrians/bicyclists another proxy variable must be found. One suggestion is to use the amount of persons living in a close area to the footpath crossing as a proxy for exposure.

The accident charge today in Sweden due to level crossing accidents is set to 0.34 – 0.80 SEK/train km (Trafikanalys, 2014) based on a similar study using accident records for 2000-2008 (Jonsson, 2011) In that study, the average marginal cost per train passage was estimated to a value of SEK 1.13 in 2008, compared to SEK 1.28 in 2012 in the present study. In the present paper we have enlarged the data set, which now also includes the years 2009 to 2012. We also base the estimation on other data regarding traffic volume, which now includes station areas. The values for the injuries are also updated according to the official Swedish values which can explain the, relatively small, difference in average marginal cost between the two studies.

References

Cedersund, H.-Å.: 2006, 'Trafiksäkerhet i plankorsningar mellan väg och järnväg 1999-2004'. Report 540, VTI. In Swedish.

Jonsson, L., 2011. 'Marginal cost estimation for level crossing accidents: Evidence from the Swedish railways 2000-2008'. Scandinavian working papers in Economics, Nr. 2011:15.

Lindberg, G., 2006, 'Valuation and Pricing of Traffic Safety'. Örebro studies in economics 13, Örebro University.

Lord, D., Washington, S., Ivan, J., 2005. 'Poisson, poisson-gamma and zero-inflated regression models of motor vehicle crashes: Balancing statistical fit and theory'. Accident Analysis and Prevention 37 (1), 35-46.

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

Shavell, S., 2004. 'Foundations of economic analysis of law'. Belknap Press of Harvard University Press.

Trafikanalys, 2013a. 'Bantrafik 2012. Statistik 2013:28'. In Swedish.

Trafikanalys, 2013b. 'Bantrafikskador 2012. Statistik 2013:15'. In Swedish.

Trafikanalys, 2014. 'Transportsektorns samhällsekonomiska kostnader. Bilagor 2014'. PM 2014:3. In Swedish.

Trafikverket, 2014. 'Samhällsekonomiska principer och kalkylvärden för transportsektorn: ASEK 5.1. Kapitel 9 Trafiksäkerhet'. Version 2014-04-01. In Swedish.