Estimating the impact of traffic on rail infrastructure maintenance costs; the importance of axle loads

Kristofer Odolinski

Address for correspondence. The Swedish National Road and Transport Research Institute, Department of Transport Economics, Box 55685, 102 15 Stockholm, Sweden (kristofer.odolinski@vti.se)

The author is grateful to the Swedish Transport Administration for funding this research. Special thanks to Vivianne Karlsson and Anders Nilsson for providing data. Phill Wheat, Daniel Wikström and Mattias Haraldsson have provided helpful comments on earlier version of this paper. All remaining errors are the responsibility of the author.

Abstract

In this paper, we estimate the impact of axle loads on rail infrastructure maintenance costs. The results show that cost elasticities with respect to traffic increase with axle load. Using these elasticities, we calculate marginal costs for traffic that are differentiated with respect to the trains' average tonnage per axle. The results are relevant when setting track access charges in Europe as well as for railway cost studies in general, considering that the empirical evidence in this paper gives support to the engineering perspective - that is, axle loads are important to consider when assessing the damage caused by traffic.

Date of final version: 2018-11-14

Keywords: maintenance; rail infrastructure; axle load; marginal cost; track access charge; top-down approach.

1.0 Introduction

The vertical separation between train operations and infrastructure management in Europe during the 1990s made an introduction of track access charges necessary. The charging principles, as set forth in 2012 by EU legislation (Dir. 2012/34/EU), states that these charges should be set according to the direct cost of running a vehicle on the rail infrastructure (yet, it also allows non-discriminatory mark-ups). An important part of these charges concerns the wear and tear of the infrastructure, which will differ depending on the characteristics of the vehicle running on the tracks (as well as on the characteristics of the tracks). Differentiating the charge with respect to the track damage caused by the vehicles can establish an efficient use of the infrastructure. The axle load is an important feature of the vehicles in this context.

The significance of axle load as a driver of track deterioration is for example shown in Stichel (1999), Öberg et al. (2007), and Zarembski (2015). However, studies on the direct relationship between axle loads and maintenance costs using econometric techniques cannot be found in the literature – a method generally referred to as a top-down approach, which relates traffic to actual costs. There is however an extensive literature on the relationship between tonne density (tonne-km/route-km) and costs.[1] See for example Munduch et al. (2002), Johansson and Nilsson (2004), Andersson (2008), Wheat and Smith (2008), Gaudry and Quinet (2009), Wheat et al. (2009) and Wheat et al. (2015). Still, these studies do not use (have access to) information on the axle loads (tonnes per axle) of each train and therefore use a tonne density measure; a measure that may hide information that is important for the wear and tear of the rail infrastructure and, hence, the maintenance costs. For example, a tonne density measure does not reflect the increase in axle loads when a train operator moves an additional number of tonnes with one train instead of two trains (or one wagon instead of two wagons) – that is, the measure

[1] Alternatively, tonne density can be defined as tonne-km/track-km - that is, parallel tracks are included in the denominator.

cannot distinguish between the two cases. Using this measure to estimate the cost impact of infrastructure usage may therefore produce inaccurate results.

There is thus reason to study the direct relationship between axle load and costs, which is the purpose of this paper. Specifically, we estimate the cost impact of the trains' average tonnage per axle on rail infrastructure maintenance in Sweden, using a panel dataset over the period 2011-2016. We fill a gap in the literature by estimating a direct relationship between axle loads and maintenance costs using econometric techniques. The empirical evidence in this study can also be useful when setting track access charges for the wear and tear costs caused by different vehicles.

A source of knowledge on the effect of axle loads are the studies related to the Heavy Axle Load (HAL) Research Program, initiated in 1988 in the U.S., in which predictions on the cost impacts were made using engineering models for different components and bridges. The actual costs were however much lower than expected, partly due to better technology and maintenance management (see Martland 2013 for a review of the literature related to this research program). The International Union of Railways (UIC) also made a series of studies during the 1980s on increased axle loads, generating results that are utilized in for example Öberg et al. (2007). More specifically, the calculations in Öberg et al. show that an increase in the axle load from 16 to 22 tonnes will cause a 60 per cent increase in costs per tonne-km. However, their approach relies on an assessment by experts within the Swedish Rail Administration; an assessment of the shares of maintenance and renewal costs that can be attributed to different damage mechanisms. This assessment may or may not be close to the actual cost shares of the different damages, and the calculated cost impact from an increased axle load is therefore uncertain.

Both of the above-mentioned cases are so called bottom-up approaches, which are based on engineering models. In general, this type of approach predicts the damages caused by traffic and then links it to the cost of rectifying these damages, which to some degree requires assumptions on the amount of maintenance (and renewal) activities performed and their respective costs. Other examples of when this approach has been used in the assessment of the cost impact of axle loads are Casavant and Tolliver (2001) and Bitzan and Tolliver (2001). Again, these assumptions may, or may not, generate predictions that are close to the actual costs caused by an increase in axle loads.

A combination of the bottom-up and top-down approaches was made in Smith et al. (2017), using engineering simulation approaches to generate measures on the damage incurred by different vehicles types (accounting for axle loads, among other things), and relating these damages to actual costs with econometric techniques. However, the purpose of the paper was only do demonstrate the feasibility of the approach, and the statistical significance of the estimates was limited. A swiss case study in Wheat et al. (2015) is another example of an attempt to find empirical evidence on the wear and tear costs of different vehicle types, which proved to be difficult. Trying to establish a direct relationship between axle loads and costs may on the other hand be a fruitful approach, considering its effects on track decay.

The paper is organized as follows. Section 2 describes the traffic measures considered in the study and the estimation approach. In section 3, we specify our cost model. Section 4 contains a description of our data. The estimation results are presented in section 5, while section 6 concludes.

2.0 Traffic measures and estimation approach

The purpose of using an axle load measure (and its related marginal cost) is to reflect an increase in axle loads which is not revealed by a tonne density measure; a common traffic measure in the literature on rail infrastructure wear and tear costs. For example, an axle load measure can reflect the difference between two trains with the same total weight but with different number of axles.

The traffic information available in this study is the train-km, the gross weight and the number of axles of each train that has run on the Swedish railway network during years 2011 to 2016. With this information, we can calculate the number of trains and tonnes that has run on a specific part of the network, the average weight of the trains as well as their average gross tonnage per axle. Here it can be noted that we do not have access to information on the weight of each wagon in the train set. Neither do we have information on differences in characteristics of the vehicles that contribute to wear and tear of the infrastructure, such as bogie type.

Considering that maintenance costs are reported at track section level (the state-owned network comprises about 250 sections), we consider a train density measure

$$\frac{TKM_{it}}{RKM_{it}} \quad (1)$$

where $i = track\ section$ and $t = year$, TKM_{it} is train-km, and RKM_{it} is route-km.[2] A measure of the average weight of the trains on each section and year is $\frac{GTKM_{it}}{TKM_{it}}$, where $GTKM_{it}$ is gross tonne-km, while the average gross tonnage per axle is

$$\frac{\sum_{j=1}^{J} \frac{GTKM_{jit}}{AXLES_{jit}}}{TKM_{it}} \quad (2)$$

where $j = 1,2 \dots, J$ number of trains. Specifically, equation (2) implies that we calculate the average tonnage per axle for each train, which is multiplied with the number of kilometres that

[2] We include a track length variable in the model estimations to account for the fact that traffic on a section may run on a single track or multiple tracks.

the train has run on a section and year to get the gross tonne-km per axle. Taking the sum over all trains and dividing by train-km on section i in year t generates an (weighted) average tonnage per axle.

From a wear and tear perspective, the number of gross tonnes is a better traffic measure compared to the number of trains. However, high train density implies high line capacity utilisation, which can result in shorter available time slots for maintenance and/or more maintenance during night-time, which is costly (see Odolinski and Boysen 2018). Moreover, high capacity utilisation implies higher sensitivity to delays, which can result in more (preventive and corrective) maintenance. There is therefore reason to use a train density variable, where the average weight of the trains should be included in the estimations to account for variations in tonnage. Importantly, to estimate the impact of axle loads, we also include the average tonnage per axle in the estimations (equation 2) together with an interaction with train density. This interaction allows us to estimate the impact of a traffic increase that is linked to an increase in axle loads (the model is specified in section 3 below). If this effect is present, the cost elasticity with respect to trains can be used to calculate a marginal cost per train-km that varies for different axle loads.

The marginal cost per train-km is derived as (see for example Munduch et al. 2002 or Odolinski and Nilsson 2017)

$$MC_{it} = \frac{\partial C_{it}}{\partial TKM_{it}} = \frac{TKM_{it}}{C_{it}} \frac{\partial C_{it}}{\partial TKM_{it}} \frac{C_{it}}{TKM_{it}} = \frac{\partial lnC_{it}}{\partial lnT_{it}} \frac{C_{it}}{TKM_{it}} \quad (3)$$

where C_{it} is maintenance costs on track section i in year t, $\frac{\partial lnC_{it}}{\partial lnT_{it}}$ is the cost elasticity with respect to trains and $\frac{C_{it}}{TKM_{it}}$ is the average cost. The costs that we include in the estimations are maintenance costs for track substructures and track superstructures. Hence, costs for signalling and telecommunications etc. are not included, as the wear and tear of these assets does not vary with traffic and axle load.

To find the causal effect of increased traffic and axle loads on maintenance costs, we need to consider that there may be confounders – that is, variables that influence traffic and also maintenance costs.[3] Network characteristics can be confounders in our estimations. For example, prior to an increased axle load, changes in the infrastructure can be required to make it more resilient towards the deterioration caused by the higher forces from the vehicles. Indeed, investments in the rail infrastructure have been made on different parts of the railway network in Sweden, where the infrastructure manager (IM) is gradually increasing the maximum axle load allowed to 25 tonnes. Some parts of the network are even designed for 30 tonnes per axle. This implies that interventions have been made which influence the average tonnage per axle that runs on a certain section but can also influence the maintenance costs (both via tonnage per axle and directly). Hence, controlling for this network characteristic can be important in order to identify the effect axle load has on maintenance costs. In addition to the maximum axle load allowed, we use a set of control variables which we describe in the following two sections: The model we estimate is presented in the section below and a description of the available data is presented in section 4.

3.0 Model

To estimate the effect of axle loads in our econometric (top-down) approach, we use a cost function given by equation (4), with $i = 1,2 \ldots., N$ track sections observed over $t =$ years 2011 to 2016.

[3] Formally, we have a confounder if $Pr(C|do(T)) \neq Pr(C|T)$, where $Pr(C|do(T))$ is the probability of a certain maintenance cost C from the controlled intervention of a certain traffic T and $Pr(C|T)$ is the conditional probability of C when observing T.

$$C_{it} = f(\boldsymbol{Q}_{it}, \boldsymbol{X}_{it}, \boldsymbol{Z}_{it}) \quad (4)$$

C_{it} is maintenance costs, $\boldsymbol{Q}_{it}$ is a vector of traffic variables including train density (T_{it}), average weight of trains (W_{it}) and average tonnage per axle (A_{it}). $\boldsymbol{X}_{it}$ is a vector of variables for infrastructure characteristics such as track length, rail weight and maximum axle load allowed on the tracks. $\boldsymbol{Z}_{it}$ is a vector of dummy variables for the five maintenance regions in Sweden, as well as year dummy variables to capture general effects over the rail network such as variations in input prices.

We use a double-log specification – that is, the dependent and independent variables are log-transformed – which is a useful transformation of data if the estimated residuals are skewed and/or we have problems with heteroscedasticity. We consider a Translog model (proposed by Christensen et al. 1971; see Christensen and Greene 1976 for an application to cost functions), which is a second order approximation of a cost (or production) model; a flexible model that put few restrictions on the elasticities of production. The model we estimate is

$$lnC_{it} = \alpha + \beta_C lnC_{it-1} + \beta_T lnT_{it} + \frac{1}{2}\beta_{TT}(lnT_{it})^2 + \beta_W lnW_{it} + \frac{1}{2}\beta_{WW}(lnW_{it})^2 + \beta_A lnA_{it} + \frac{1}{2}\beta_{AA}(lnA_{it})^2 + \sum_{l=1}^{L}\beta_l lnX_{lit} + \frac{1}{2}\sum_{l=1}^{L}\sum_{l=1}^{L}\beta_{ll} lnX_{lit} lnX_{lit} + \beta_{TW} lnT_{it} lnW_{it} + \beta_{TA} lnT_{it} lnA_{it} + \beta_{WA} lnW_{it} lnA_{it} + \sum_{l=1}^{L}\sum_{r=1}^{R}\beta_{lr} lnX_{lit} lnX_{rit} + \sum_{l=1}^{L}\beta_{lT} lnX_{lit} lnT_{it} + \sum_{l=1}^{L}\beta_{lW} lnX_{lit} lnW_{it} + \sum_{l=1}^{L}\beta_{lA} lnX_{lit} lnA_{it} + \sum_{d=1}^{D}\vartheta_d Z_{dit} + \mu_i + v_{it} \quad (5)$$

where α is a scalar, v_{it} the error term, and μ_i is the impact of unobserved track section specific effects. β_C, β_T, β_{TT}, β_W, β_{WW}, β_A, β_{AA}, β_l, β_{ll}, β_{TW}, β_{TA}, β_{WA}, β_{lr}, β_{lT}, β_{lW}, β_{lA} and ϑ_d are parameters to be estimated. The Cobb-Douglas constraint $\beta_{TT} = \beta_{WW} = \beta_{AA} = \beta_{ll} = \beta_{TW} = \beta_{TA} = \beta_{WA} = \beta_{lr} = \beta_{lT} = \beta_{lW} = \beta_{lA} = 0$ is tested using and F-test.

Lagged maintenance costs lnC_{it-1} are included in the model as there may be intertemporal effects of maintenance, where for example an increase in traffic may have an impact on costs in both the current and subsequent year(s). Such effects have been found by Andersson (2008), Wheat (2015), Odolinski and Nilsson (2017) and Odolinski and Wheat (2018). To deal with the correlation between lagged maintenance costs lnC_{it-1} and track section specific effects μ_i, we use a forward orthogonal deviation (proposed by Arellano and Bover 1995) that removes μ_i. Lagged maintenance costs are also correlated with the error terms v_{it}. We therefore use instruments for lagged maintenance costs, where further lags of maintenance are the best instruments available to us. The method by Holtz-Eakin et al. (1988) is used to not lose observations when including further lags as instruments. Specifically, missing values are substituted with zeros, generating the moment condition $\sum_{i,t} lnC_{i,t-2}\hat{v}_{it} = 0$ (see Roodman 2009 for details).

We still need to test if there is autocorrelation within the error structure, which can bias the results. The Arellano-Bond test is used and failure to reject the null hypothesis of no autocorrelation indicates that we should increase the number of lags of the dependent variable, which can remove the autocorrelation. We start with one lag and increase until we can reject the null hypothesis of no autocorrelation.

Considering that we estimate a dynamic model, we can (and should) include the intertemporal effect of maintenance in the calculation of cost elasticities. Specifically, we can calculate ‘equilibrium cost elasticities’, which are present when there is no tendency to change maintenance costs (*ceteris paribus*) – that is, we have an equilibrium cost level when $lnC_{it} = lnC_{it-1} = lnC_{it}^{e}$ (see Odolinski and Wheat 2018 for more details). Including the equilibrium cost level in equation (5), we have

$$lnC_{it}^{e} = \alpha + \beta_C lnC_{it}^{e} + \beta_T lnT_{it} + \frac{1}{2}\beta_{TT}(lnT_{it})^2 + \beta_W lnW_{it} + \frac{1}{2}\beta_{WW}(lnW_{it})^2 + \beta_A lnA_{it} + \frac{1}{2}\beta_{AA}(lnA_{it})^2 + \sum_{l=1}^{L}\beta_l lnX_{lit} + \frac{1}{2}\sum_{l=1}^{L}\sum_{l=1}^{L}\beta_{ll} lnX_{lit} lnX_{lit} + \beta_{TW} lnT_{it} lnW_{it} + \beta_{TA} lnT_{it} lnA_{it} + \beta_{WA} lnW_{it} lnA_{it} + \sum_{l=1}^{L}\sum_{r=1}^{R}\beta_{lr} lnX_{lit} lnX_{rit} + \sum_{l=1}^{L}\beta_{lT} lnX_{lit} lnT_{it} + \sum_{l=1}^{L}\beta_{lW} lnX_{lit} lnW_{it} + \sum_{l=1}^{L}\beta_{lA} lnX_{lit} lnA_{it} + \sum_{d=1}^{D}\vartheta_d Z_{dit} + \mu_i + v_{it} \quad (6)$$

which in turn can be expressed as

$$lnC_{it}^{e} = \frac{\alpha}{1-\beta_C} + \frac{\beta_T}{1-\beta_C} lnT_{it} + \frac{1}{2}\frac{\beta_{TT}}{1-\beta_C}(lnT_{it})^2 + \frac{\beta_W}{1-\beta_C} lnW_{it} + \frac{1}{2}\frac{\beta_{WW}}{1-\beta_C}(lnW_{it})^2 + \frac{\beta_A}{1-\beta_C} lnA_{it} + \frac{1}{2}\frac{\beta_{AA}}{1-\beta_C}(lnA_{it})^2 + \sum_{l=1}^{L}\frac{\beta_l}{1-\beta_C} lnX_{lit} + \frac{1}{2}\sum_{l=1}^{L}\sum_{l=1}^{L}\frac{\beta_{ll}}{1-\beta_C} lnX_{lit} lnX_{lit} + \frac{\beta_{TW}}{1-\beta_C} lnT_{it} lnW_{it} + \frac{\beta_{TA}}{1-\beta_C} lnT_{it} lnA_{it} + \frac{\beta_{WA}}{1-\beta_C} lnW_{it} lnA_{it} + \sum_{l=1}^{L}\sum_{r=1}^{R}\frac{\beta_{lr}}{1-\beta_C} lnX_{lit} lnX_{rit} + \sum_{l=1}^{L}\frac{\beta_{lT}}{1-\beta_C} lnX_{lit} lnT_{it} + \sum_{l=1}^{L}\frac{\beta_{lW}}{1-\beta_C} lnX_{lit} lnW_{it} + \sum_{l=1}^{L}\frac{\beta_{lA}}{1-\beta_C} lnX_{lit} lnA_{it} + \sum_{d=1}^{D}\frac{\vartheta_d}{1-\beta_C} Z_{dit} + \frac{\mu_i}{1-\beta_C} + \frac{v_{it}}{1-\beta_C} \quad (7)$$

Holding other interactions with train density constant, the equilibrium cost elasticity with respect to train density is

$$\frac{\partial lnC_{it}^{e}}{\partial lnT_{it}} = \hat{\gamma}_{it} = \frac{\hat{\beta}_T}{1-\hat{\beta}_C} + \frac{\hat{\beta}_{TT}}{1-\hat{\beta}_C} lnT_{it} + \frac{\hat{\beta}_{TA}}{1-\hat{\beta}_C} lnA_{it} \quad (8)$$

given that there is also a second order effect of train density. For the purpose of this paper, the parameter estimate ($\hat{\beta}_{TA}$) for the interaction between train density and the average axle load ($lnT_{it}lnA_{it}$) is of primary interest. Specifically, if $\hat{\beta}_{TA} \neq 0$, the cost elasticity with respect to train density will vary with the average tonnage per axle.

3.1 Marginal costs

As specified in equation (3), the marginal cost per train-km can be calculated as the cost elasticity with respect to train density multiplied by the average cost. Specifically, the average cost is $\widehat{AC}_{it} = \hat{C}_{it}/TKM_{it}$, where $\hat{C}_{it}$ is the fitted maintenance cost

$$\hat{C}_{it} = \exp\left(\ln(C_{it}) - \hat{v}_{it} + 0.5\hat{\sigma}^2\right) \quad (9)$$

The specification in equation (9) is based on the double log-specification of the cost model and assumes normally distributed errors (see for example Munduch et al. 2002). The marginal cost for each track section and year is

$$MC_{it} = \widehat{AC}_{it} \cdot \hat{\gamma}_{it} \quad (10)$$

where $\hat{\gamma}_{it}$ is the cost elasticity from equation (8). The marginal cost can thus vary with axle loads, given that there is an interaction effect between train density and average axle loads. Furthermore, we calculate a weighted marginal cost

$$MC_{it}^{W} = MC_{it} \cdot \frac{TKM_{it}}{(\sum_{it} TKM_{it})/N} \quad (11)$$

where N is the number of observations that we use in the sample (which is not equal to the entire set of 1096 observations when using lagged maintenance costs). It can be noted that for example Munduch et al. (2002) and Andersson (2008) uses a slightly different expression: $MC^{W} = \sum_{it} MC_{it} \cdot \frac{TKM_{it}}{(\sum_{it} TKM_{it})}$. However, the average value of our weighted marginal cost will be equal to the overall weighted marginal cost MC^{W}. As the average tonnage per axle vary over time and track sections, we prefer the weighting procedure in equation (11), which allows use to calculate averages of the weighted marginal costs for sections with different average tonnages per axle.

4.0 Data

The railway network in Sweden is divided into five regional units, each administering several track sections for which we have data on costs, traffic and rail network characteristics. These regional units are called North, West, East, South and Central region, for which we use dummy variables to capture management effects, as well as other regional effects that otherwise might bias our estimates. Furthermore, it may be important to control for effects due to competitive tendering of maintenance production. Odolinski and Smith (2016) found an 11 per cent decrease in maintenance costs, using a dataset over the period 1999-2011. The reform was introduced in 2002 and has been gradual. As of 2015, all track sections have been tendered in competition. This implies that there are a number of track sections during the period of our dataset that were not tendered in competition, and we therefore include a set of dummy variables to capture the effect of competitive tendering.

Descriptive statistics of the data used in this study are presented in Table 1. The observations are at the track section level during years 2011 to 2016. In total, we have access to 1081 observations, where the number of track sections observed is on average 180 per year (not all sections of the state-owned railway network are included in our dataset, as some sections are marshalling yards or closed for traffic, but also due to missing information).

[Table 1 about here]

Information on maintenance costs has been retrieved from the Swedish Transport Administration and includes costs from activities performed to maintain the track substructure and track superstructure of the railway network. These costs may vary depending on the technical aspects of the rail infrastructure. For example, the maximum axle load allowed can be an important characteristic to control for in order to isolate the cost impact of axle loads. We also have information on the average rail weight, which is a proxy for track standard and to some extent the accumulated use of the tracks (heavier rails are newer). Furthermore, we have information on the average quality class on a section which indicates the maximum speed allowed and is linked to a set of requirements on track geometry. The variation in this variable is mainly between track sections rather than within sections over time. The aim with this variable is to capture differences in track standards and maintenance strategies (costs) between (and to some extent within) sections.

Some of the track sections are so called stations sections. These are track sections that primarily consists of a station, while the other sections are primarily lines between stations. A feature of stations is that they are used for shunting, changing locomotives and trains usually accelerate and brake a lot on these sections. They also have a higher share of switches (around 10 per cent of their total track length) compared to other sections (around 2 per cent). The station sections are indicated with a dummy variable in the estimations.

We also have access to the track length of structures (tunnels and bridges), where for example a transition from a stiff sub-structure in tunnels or bridges to a regular track (with a substructure that is built on softer materials such as macadam) can generate more wear and tear of the infrastructure (see for example Odolinski and Nilsson 2018).

5.0 Results

The model is estimated using the generalized method of moments (GMM). Specifically, we use the two-step System GMM to estimate our dynamic model, an approach proposed by Arellano and Bover (1995) and Blundell and Bond (1998). We use the Windmeijer (2005) correction to avoid biased standard errors when reporting the two-step results. Moreover, 42 instruments are used in the estimation.

The independent variables (except our dummy variables) have been divided by their sample median prior to the logarithmic transformation. In that way, their first order coefficients can be interpreted as elasticities at the sample median. However, the full translog model could not be retained based on F-tests: Second order effects and interactions of most of the infrastructure characteristics are included in the model, as well as the interaction between traffic and average tonnage per axle.

Estimating the model with one lag for the dependent variable does not remove autocorrelation within the error structure (the Arellano bond test indicates z=1.71, Pr>=0.087). When we include a second lag of the dependent variable, we can reject the null hypothesis of no autocorrelation (z=-0.06, Pr>z=0.954). All estimations are carried out using Stata 12 (StataCorp.2011).

5.1 Econometric results

The estimation results are presented in Table 2 and are based on 675 observations comprised by 178 track sections. Note that we have 1081 observations available but lose observations when including two lags of the dependent variable. Both lags are positive and statistically significant, indicating that an increase in maintenance costs will have an impact on the maintenance cost in subsequent years, which is in line with the results found by Wheat (2015), Odolinski and Nilsson (2017) and Odolinski and Wheat (2018).

[Table 2 about here]

The first order coefficient for train density is 0.22 and significant at the 1 per cent level, whereas the estimate for the average weight of the trains is close to zero (-0.0259) and not statistically significant (p-value 0.671). What seem to matter is whether the increase in weight implies a higher axle load or not; the first order coefficient for the average tonnage per axle is 0.9879 (p-value 0.06), indicating a rather large impact of increased axle loads. Increasing the average axle loads with 10 per cent will result in an almost 10 per cent increase in maintenance costs (evaluated at the sample median – that is, we hold for example the number of trains and average train weight constant). The effect of running more trains with a higher axle load is indicated by the coefficient for the interaction between train density and average tonnage per axle (*ln(train density)ln(average tonnage per axle)*), which is 0.3705 and significant at the 5 per cent level (p-value 0.029). This result indicates that the cost elasticity with respect to train density is increasing with the average axle load (recall that we control for the average weight of the trains in this estimation).

[Figure 1 about here]

Using equation 8 (without the second order effect of train density), we calculate the equilibrium cost elasticities with respect to train density, which vary with the average tonnage per axle. The average equilibrium cost elasticity is 0.41 (standard error 0.0718 and p-value 0.000). The relationship between these elasticities and the average tonnage per axle is illustrated in Figure 1. The cost elasticities are about 0.2 to 0.4 for average axle loads between 10 and 13.9 tonnes and increase to an interval between 0.4 and 0.6 for average axle loads between 13.9 and 19 tonnes. There are few observations with average axle loads above 19 tonnes, indicating cost elasticities at 0.60 to 0.75.

The first order coefficients for the infrastructure characteristics have the expected signs, however, only track length, rail weight and average quality class are statistically significant. Newer rails are usually heavier and indicates a higher track standard, which is the reason for the negative coefficient. Including a rail age variable did not have a significant impact on the results, except for the rail weight coefficient, which is expected considering that their correlation coefficient is -0.4938. The variable for maximum axle load allowed is also used to control for track standard. Its coefficient is positive, indicating that these tracks are costlier to maintain. However, this can partly be an effect of heavier trains running on these sections (the correlation coefficient between maximum axle load allowed and average tonnage per train is 0.5725).[4] The average quality class coefficient is positive, indicating that a higher quality class – that is, lower line speed and lower requirements on track standard – increases maintenance costs. Higher requirements on track standard are to some extent captured by the rail weight variable (the correlation coefficient is -0.5659). The quality class variable then seems to capture the impact of poor track standards that generate more corrective maintenance.

The West region is the baseline in the model estimations, where the coefficients for the other regions indicate that maintenance costs are higher in this region. The coefficient for competitive tendering is negative but not statistically significant. Here we can note that as of year 2011, most track sections belonged to areas tendered in competition.

5.2 Marginal costs and discussion of results

The estimated cost elasticities are part of the marginal cost calculation. Specifically, as shown in section 3.2 and equation (10), the marginal cost is calculated by multiplying the cost

[4] The correlation coefficient between maximum axle load allowed and average tonnage per axle is 0.3962. Dropping the variable for maximum axle load allowed results in only a slight change of the estimates for average tonnage per axle.

elasticities with the (fitted) average costs. Table 3 presents the mean cost elasticities, average costs and (weighted) marginal costs.

[Table 3 about here]

The weighted marginal cost is calculated using train-km as weights (equation 11) and its mean is lower compared to the mean of the marginal costs (equation 10), indicating that track sections with high traffic volumes have lower marginal costs. This relationship is illustrated in Figure 2, where the marginal costs fall sharply with traffic, which is in line with the empirical evidence from a number of European countries (see Wheat et al. 2009).

The weighted marginal costs for different axle loads (WMC) are presented in Table 4. To show the impact axle loads has on these costs, we also present weighted marginal costs (WMC-B) based on a cost elasticity (0.41) that excludes the interaction between trains and axle loads. A charge based on this weighted marginal cost would in our case imply that trains with lower axle loads (between 8 and 14 tonnes per axle) subsidises the charges for trains with higher axle loads: See last column in Table 4, or Figure 1 where the sample median elasticity (0.41) corresponds to an average tonnage per axle at 14 tonnes (elasticities are higher for higher axle loads and vice versa).

[Figure 2 about here]

Note that WMC-B (excluding the interaction with average tonnage per axle) is still varying for different axle load intervals, however, this is due to confounding factors. For example, as shown by Figure 2, trains with the highest axle loads run on sections with a low train density, which (usually) have high marginal costs (c.f. Wheat et al. 2009) that receive a low weight in the calculation (this is also the case for the lowest axle loads). Hence, the cost impact of an increase in average tonnage per axle when running an extra train service is shown by the coefficient for the interaction between trains and tonnes per axle (generating the relationship illustrated by Figure 1) while controlling for train density and other variables (see estimation results in Table 2), and its impact on marginal costs is shown by the difference between the weighted marginal costs that includes and excludes this effect calculated for different axle load intervals.

[Table 4 about here]

An infrastructure manager might then ask how a charge (based on marginal costs) should vary for different axle loads, as the estimated costs in Table 4 also vary due to other factors than axle loads. To show this we calculate marginal costs by multiplying the cost elasticities with a weighted average cost. Specifically, we use an average cost that is weighted using train-km, and multiply by the cost elasticities that vary for different axle loads

$$[\sum_{it}(\widehat{AC}_{it} \cdot \frac{TKM_{it}}{(\sum_{it} TKM_{it})})] \cdot \hat{\gamma}_{it} \quad (12)$$

The weighted average cost is SEK 12.83 per train-km. Multiplying this average cost with the equilibrium cost elasticities ($\hat{\gamma}_{it}$) generates a marginal cost per train-km that only varies with the differences in cost elasticities caused by axle loads. These costs are presented in Table 5, indicating substantial differences in marginal costs: The lowest axle loads have a marginal cost at SEK 3.07 per train-km, while the marginal cost is SEK 8.07 per train-km for the highest axle loads (the overall mean is SEK 5.26).

[Table 5 about here]

6.0 Conclusion

In this paper, we have estimated the cost impact of axle loads on maintenance costs for track substructure and track superstructure, using a top-down (econometric) approach on Swedish data during the period 2011-2016. The results show that the increase in maintenance costs from running more trains on a track will be higher when the average tonnage per axle is higher. In this estimation, we control for the average weight of the trains as well as a number of infrastructure characteristics.

The estimates in this paper can be informative for countries with a vertically separated railway system where track access charges are required, considering that our tonnes per axle measure gives a better representation of the wear and tear of the rail infrastructure, compared to only using a tonne density or train density measure. Specifically, we provide empirical evidence on the importance of axle loads with respect to maintenance costs, showing that there may be charging implications if the number of tonnes per axle is considered in the estimation of marginal costs in EU member states. Moreover, it is shown that ignoring the impact of axle loads when setting charges based on marginal costs implies that trains with a low average tonnage per axle subsidises the charge for trains with a high average tonnage per axle. Hence, the results in this paper can contribute in creating a more fair and efficient track access charge with respect to wear and tear of the tracks. Furthermore, the evidence in this paper can be relevant for research on railway costs in general. For example, analysing the effects of a policy change with respect to infrastructure costs requires that important cost drivers, such as traffic, are treated in the best possible way.

Establishing a direct relationship between axle load and maintenance costs proved to be a viable approach for the assessment of different vehicles' impact on infrastructure costs. Still, more and better data can be useful in future research. First, more observations can provide more robust results. Second, more detailed data can be rewarding, as there are for example differences in tonnes per axle between wagons on each train - information that is not available in this study. Instead, we are left with the average tonnage per axle for each train. Moreover, we do not have access to differences in characteristics between vehicles such as bogie type (which may be correlated with certain axle loads), which can also affect the results.

References

Andersson, M. (2008): 'Marginal Railway Infrastructure Costs in a Dynamic Context', *EJTIR,* 8, 268-286.

Arellano, M., and O. Bover (1995): 'Another look at the instrumental variable estimation of error-components models', *Journal of Econometrics.* 68, 29-51. DOI: https://doi.org/10.1016/0304-4076(94)01642-D

Bitzan, J. D., and D. D. Tolliver (2001): 'North Dakota Strategic Freight Analysis', Item IV. Heavier Loading Rail Cars. Upper Great Plains Transportation Institute, October 2001.

Blundell, R., and S. Bond (1998): 'Initial conditions and moment restrictions in dynamic panel data models', *Journal of Econometrics*, 87, 115-143. DOI: https://doi.org/10.1016/S0304-4076(98)00009-8

Casavant, K., and D. D. Tolliver (2001): 'Impacts of Heavy Axle Loads on Light Density Lines in the State of Washington', Washington State Department of Transportation, WA-RD 499.1, Final Report, February 2001.

Christensen, L.R, and W.H. Greene (1976): 'Economies of Scale in U.S. Electric Power Generation', *The Journal of Political Economy,* 84(4), 655-676.

Christensen, L.R, D.W., Jorgenson, and L.J. Lau (1971): 'Transcendental Logarithmic Production Frontiers', *The Review of Economics and Statistics*, 55(1), 28-45. DOI: https://doi.org/10.2307/1927992

Gaudry, M. and E. Quinet (2009): 'Track wear-and-tear cost by traffic class: Functional form, zero output levels and marginal cost pricing recovery on the French rail network', Working paper No. 2009-32, Paris-Jourdan Sciences Économiques.

Holtz-Eakin, D., W., Newey, and H. S. Rosen (1988): 'Estimating vector autoregressions with panel data', *Econometrica* 56(6), 1371-1395. DOI: https://doi.org/10.2307/1913103

Johansson, P., and J-E. Nilsson (2004): 'An economic analysis of track maintenance costs', *Transport Policy*, 11, 277-286. DOI: https://doi.org/10.1016/j.tranpol.2003.12.002

Martland, C.D. (2013): 'Introduction of Heavy Axle Loads by the North American Rail Industry', *Journal of the Transportation Research Forum*, 52(2), 103-125.

Munduch, G., A. Pfister, L. Sögner, and A. Siassny (2002): 'Estimating Marginal Costs for the Austrian Railway System', Vienna University of Economics & B.A., Working Paper No. 78, February 2002.

Odolinski, K. and H.E. Boysen (2018): 'Railway line capacity utilisation and its impact on maintenance costs', *Journal of Rail Transport Planning & Management,* (In Press).

Odolinski, K. and J-E. Nilsson (2017): 'Estimating the marginal maintenance cost of rail infrastructure usage in Sweden; does more data make a difference?', *Economics of Transportation*, 10, 8-17. DOI: https://doi.org/10.1016/j.ecotra.2017.05.001

Odolinski K. and J-E. Nilsson (2018): 'Railway Economics and Engineering; making informed decisions on railway maintenance and renewals', In NeTIRail Deliverable D1.4, Cost and User benefit report, 22 March 2018.

Odolinski, K., and P. Wheat (2018): 'Dynamics in rail infrastructure provision: Maintenance and renewal costs in Sweden', *Economics of Transportation*, 14, 21-30. DOI: https://doi.org/10.1016/j.ecotra.2018.01.001

Odolinski, K. and A.S.J. Smith (2016): 'Assessing the cost impact of competitive tendering in rail infrastructure maintenance services: evidence from the Swedish reforms (1999-2011)', *Journal of Transport Economics and Policy*, 50(1), 93-112.

Roodman, D. (2009): 'A note on the theme of too many instruments', *Oxford Bulletin of Economics and Statistics*, 71(1), 135-158. DOI: https://doi.org/10.1111/j.1468-0084.2008.00542.x

Smith, A., S. Iwnicki, A. Kaushal, K. Odolinski, and P. Wheat (2017): 'Estimation of the relative cost of track damage mechanisms: A combination of economic and engineering approaches', *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, 231(5), 620-636. DOI: https://doi.org/10.1177/0954409717698850.

StataCorp.2011. Stata Statistical Software: Release 12. College Station, TX: StataCorp LP.

Stichel, S., (1999): 'On freight wagon dynamics and track deterioration', *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, 213(4), 243-254. DOI: https://doi.org/10.1243/0954409991531182

Wheat, P., and A.S.J. Smith (2008): 'Assessing the Marginal Infrastructure Maintenance Wear and Tear Costs for Britain's Railway Network', *Journal of Transport Economics and Policy*, 42(2), 189-224.

Wheat, P., A.S.J. Smith, and C. Nash (2009): 'CATRIN (Cost Allocation of TRansport INfrastructure cost)', Deliverable 8 – Rail Cost Allocation for Europe, VTI, Stockholm

Wheat, P., A. Smith, and B. Matthews (2015): 'The sustainable freight railway: Designing the freight vehicle–track system for higher delivered tonnage with improved availability at reduced cost SUSTRAIL', Deliverable 5.3: Access Charge Final Report.

Wheat, P., (2015): The Sustainable Freight Railway: Designing the Freight Vehicle-track System for Higher Delivered Tonnage with Improved Availability at Reduced Cost SUSTRAIL, Deliverable 5.3: Access Charge Final Report Annex 4, British Case Study.

Windmeijer, F. (2005): ‘A finite sample correction for the variance of the linear efficient two-step GMM estimators’, *Journal of Econometrics*, 126(1), 25-51. DOI: https://doi.org/10.1016/j.jeconom.2004.02.005

Zarembski, A. (2015): ‘Determination of the Impact of Heavy Axle Loads on Short Lines, Delaware Center for Transportation, DCT 251.

Öberg, J., E. Andersson, and J. Gunnarsson (2007): ‘Track Access Charging with Respect to Vehicle Characteristics’, second edition, Rapport LA-BAN 2007/31, Banverket.

Tables and Figures

Table 1

Costs, traffic, infrastructure characteristics and dummy variables, track sections during years 2011-2016 (1081 obs.)

Variable	*Median*	*Mean*	*Std. dev.*	*Min.*	*Max.*
Maintenance costs for track substructure and track superstructure, million SEK*	4.64	8.26	12.24	0.01	153.27
Traffic variables					
Train-km**	471	780	929	0	4 738
Gross tonne-km**	156 760	375 655	534 964	1	4 219 003
Train density (train-km/route-km)** Eq. 1	11	18	23	0	191
Gross tonne density (gross tonne-km/route-km)** Eq. 2	4 517	7 784	8 909	0	65 855
Average weight of trains (gross tonne-km/train-km)	369	541	633	51	6 152
Average tonnage per axle (gross tonnes/number of axles) Eq. 3	14.00	14.12	1.31	8.01	23.48
Infrastructure characteristics					
Route length, km	40.9	52.4	40.7	1.8	219.4
Track length, km	57.4	71.8	53.1	3.2	305.5
Average number of tracks (Track length/Route length)	1.2	1.7	1.1	1.0	8.5
Average rail weight, kg per meter of rail	50.6	52.4	4.8	41.9	60.0
Average quality class (1 to 6)*** (Qual_ave)	3.0	3.0	1.2	1.0	6.0
Switches, km	1.3	1.7	1.7	0.1	13.7
Maximum axle load allowed, tonnes	0.4	1.6	3.7	0.0	22.1
Track length structures (tunnels and bridges), km	22.5	23.1	1.7	16.0	30.0
Station section, dummy variable	0.0	0.1	0.3	0.0	1.0
Organisational dummy variables					
West region	0.0	0.2	0.4	0.0	1.0
North region	0.0	0.1	0.3	0.0	1.0
Central region	0.0	0.2	0.4	0.0	1.0
South region	0.0	0.3	0.4	0.0	1.0
East region	0.0	0.2	0.4	0.0	1.0
Dummy when mix between tendered and not tendered	0.0	0.0	0.1	0.0	1.0
Dummy when tendered in competition	1.0	1.0	0.2	0.0	1.0

* 2016 prices (deflated using consumer price index), ** Thousands, *** Track quality class ranges from 0-5 on parts of a track section (from low to high line speed), but 1 has been added to avoid observations with value 0.

Table 2

Estimation results (675 obs.)

	Coef.	Corr. Std. Err.	[95% Conf. Interval]	
Cons.	8.7138***	1.2469	6.2531	11.1746
ln(maintCt-1)	0.3378***	0.0657	0.2081	0.4675
ln(maintCt-2)	0.1161*	0.0627	-0.0077	0.2400
ln(train density)	0.2212***	0.0505	0.1217	0.3208
ln(average train weight)	-0.0259	0.0609	-0.1461	0.0943
ln(average tonnage per axle)	0.9879*	0.5210	-0.0402	2.0160
ln(max. axle load)	0.2315	0.5009	-0.7569	1.2200
ln(track length)	0.5110***	0.0895	0.3344	0.6875
ln(number of tracks)	0.0802	0.1305	-0.1774	0.3378
ln(lenght of structures)	0.0154	0.0308	-0.0453	0.0761
ln(rail weight)	-1.3131**	0.5419	-2.3825	-0.2437
ln(quality class)	0.6640***	0.1586	0.3510	0.9771
ln(length of switches)	0.1240	0.0755	-0.0250	0.2731
ln(train density)ln(average tonnage per axle)	0.3705**	0.1682	0.0386	0.7025
ln(track length)^2	0.0033	0.1488	-0.2903	0.2969
ln(track length)ln(quality class)	0.3464***	0.1337	0.0825	0.6104
ln(track length)ln(length of switches)	-0.0320	0.0908	-0.2113	0.1472
ln(quality class)^2	1.1876***	0.3469	0.5031	1.8721
ln(quality class)ln(length of switches)	-0.0754	0.1113	-0.2951	0.1443
ln(length of switches)^2	0.1271	0.0839	-0.0385	0.2926
D.year14	0.1220*	0.0695	-0.0152	0.2591
D.year15	0.0044	0.0723	-0.1383	0.1471
D.year16	-0.2055**	0.0923	-0.3877	-0.0234
D.North region	-0.1016	0.1583	-0.4140	0.2109
D.Central region	-0.2861***	0.0962	-0.4760	-0.0962
D.South region	-0.3413***	0.0979	-0.5345	-0.1481
D.East region	-0.1483*	0.0881	-0.3222	0.0257
D.mix between tend. and not tend. in competition	-0.2739	0.3043	-0.8745	0.3267
D. tendered in competition	-0.0289	0.0854	-0.1975	0.1397
D.station section	0.1182	0.1694	-0.2162	0.4526

Notes: ***, **, *: Significance at the 1%, 5%, 10% level. The variables were divided with the sample median prior to taking logs. The first order coefficients can therefore be interpreted as cost elasticities at the sample median.

Table 3

Cost elasticity, average cost and marginal cost (675 obs.)

Variable	Mean	Std. Err.	[95% Conf.	Interval]
Equilibrium cost elasticity w.r.t. train density	0.41	0.00	0.41	0.42
Average cost (per train-km)	75.62	22.86	30.72	120.51
Marginal cost per train-km	21.87	3.26	15.47	28.27
Weighted marginal cost per train-km	5.29	0.32	4.67	5.92

Table 4

Weighted marginal costs, including (WMC) and excluding (WMC-B) the interaction effect between train density and average tonnage per axle (SEK)

Average tonnage per axle, intervals	Obs.	WMC, including interaction between trains and axle load[a]	WMC-B, excluding interaction between trains and axle load[a]	Difference, WMC and WMC-B
[8, 12)	12	**1.44** (0.36) [0.65, 2.23]	**2.33** (0.51) [1.20, 3.46]	0.90
[12, 13)	66	**3.87** (0.45) [2.98, 4.76]	**4.66** (0.54) [3.57, 5.74]	0.79
[13, 14)	240	**4.75** (0.44) [3.88, 5.61]	**5.04** (0.45) [4.15, 5.94]	0.29
[14, 15)	220	**6.34** (0.78) [4.80, 7.89]	**6.00** (0.75) [4.53, 7.47]	-0.34
[15, 16)	105	**5.80** (0.55) [4.71, 6.89]	**5.01** (0.47) [4.08, 5.95]	-0.79
[16, 17)	22	**4.58** (1.19) [2.10, 7.07]	**3.60** (0.93) [1.66, 5.54]	-0.98
[17, 23.5]	10	**5.61** (2.02) [1.04, 10.18]	**3.77** (1.52) [0.33, 7.21]	-1.84
Overall mean	*675*	*5.29 (0.32) [4.67, 5.92]*	*5.20 (0.31) [4.59, 5.80]*	*0.10*

[a] Standard errors are in parentheses and 95 per cent confidence intervals are in brackets.

Table 5

Marginal costs per train-km for different axle load intervals

Average tonnage per axle, intervals	Obs.	Marginal cost per train-km (SEK)[a]
[8, 12)	12	**3.07** (0.29) [2.43, 3.71]
[12, 13)	66	**4.32** (0.02) [4.27, 4.37]
[13, 14)	240	**4.86** (0.01) [4.83, 4.88]
[14, 15)	220	**5.50** (0.01) [5.48, 5.53]
[15, 16)	105	**6.01** (0.02) [5.98, 6.04]
[16, 17)	22	**6.59** (0.04) [6.51, 6.67]
[17, 23.5]	10	**8.07** (0.28) [7.45, 8.69]
Overall mean	*675*	*5.26 (0.03) [5.21, 5.32]*

[a] Standard errors are in parentheses and 95 per cent confidence intervals are in brackets.

Figure 1

Cost elasticities with respect to train density for different levels of average tonnage per axle

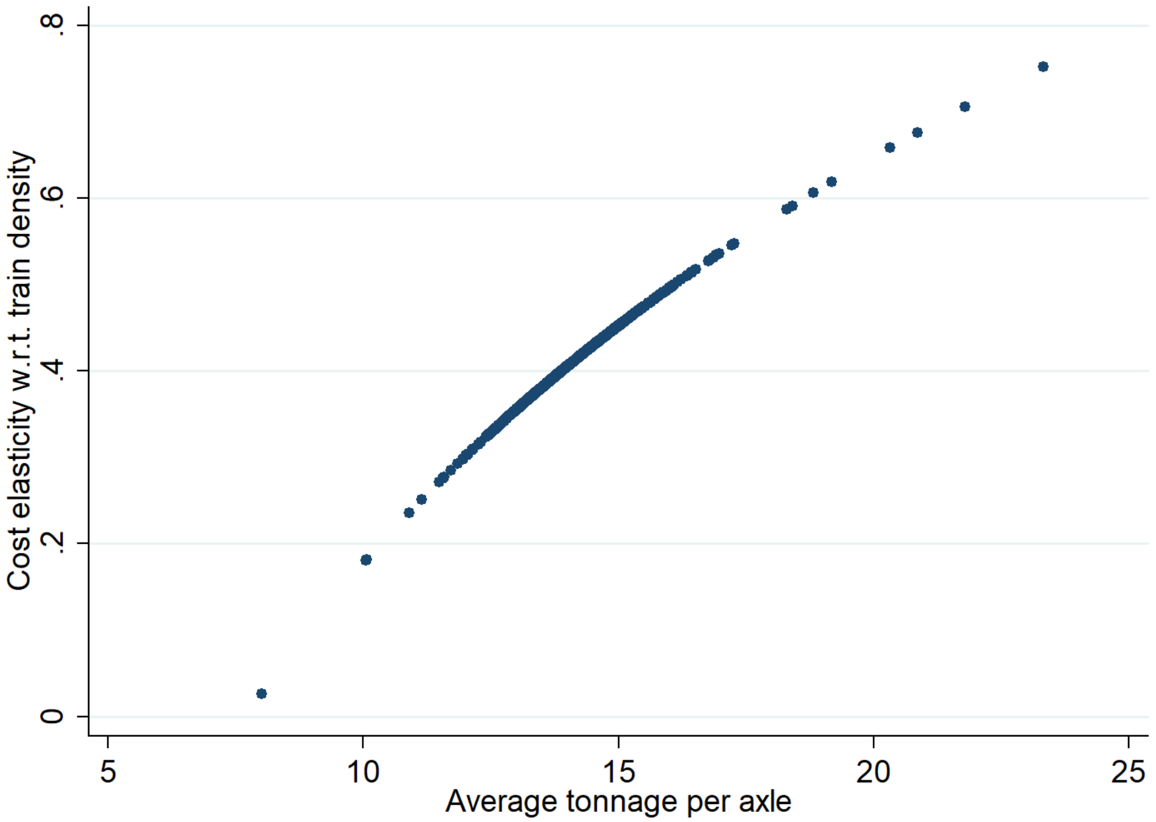

Figure 2

Marginal costs per train-km for different axle load intervals