Transport Policy 131 (2023) 13–21

Contents lists available at ScienceDirect

Transport Policy

journal homepage: www.elsevier.com/locate/tranpol

Damage or no damage from traffic: Re-examining marginal cost pricing for rail signalling maintenance

Kristofer Odolinski [a,b,*], Andrew Smith [a], Phill Wheat [a], Jan-Eric Nilsson [b], Clement Dheilly [c]

[a] *Institute for Transport Studies (ITS), University of Leeds, Leeds, LS2 9JT, United Kingdom*
[b] *The Swedish National Road and Transport Research Institute (VTI), Box 55685, SE-102 15, Stockholm, Sweden*
[c] *SNCF Réseau, 93418, Saint-Denis, Paris, France*

ARTICLE INFO

Keywords:
Marginal cost
Access charging
Econometric methods
Engineering methods
Rail infrastructure
Signalling maintenance

ABSTRACT

This paper re-examines the implementation of the short run marginal cost (SRMC) pricing principle with respect to rail infrastructure usage and empirically tests if there are rail infrastructure maintenance costs triggered by traffic but not caused by asset damage from traffic. This is important because current EU legislation stipulates that only costs related to infrastructure wear and tear from traffic are eligible for the direct cost-based element of track access charges. An econometric approach is applied to French panel data on signalling maintenance costs. The results show that the SRMC for infrastructure provision of these assets is not only related to asset damage costs caused by traffic, but can also be due to economic factors linked to increased line capacity utilisation: 1) higher cost per maintenance activity, and/or 2) increased preventative maintenance to curb delays. Our work offers an explanation as to why econometric and engineering approaches give different views of rail infrastructure cost variability and suggests that EU legislation on track access pricing may need to be revised.

1. Introduction

Short run marginal cost (SRMC) pricing has formed the basis for European transport pricing policy since the mid-1990s (European Commission, 1995). Considering the policy of vertically separating railway infrastructure from operations, The Single European Railway Area (SERA) Directive (2012/34/EU) and the European Commission's Implementing Regulation (EU, 2015/909) lays down the rules for how track access charges paid by train operators to infrastructure managers (IMs) are to be estimated. The application of the SRMC principle (see e.g. Nash, 2018) means that charges are to be set at the level of the direct cost emanating from trains using the railway infrastructure. According to the regulation, this implies that a direct cost-based charge cannot include costs for assets that do not deteriorate as a result of traffic on the network. Specifically, EU 2015/909 (8) states that "wear and tear of track-side signals and signal boxes does not vary with traffic and therefore should not be subject to a direct cost-based charge".

There may however be other reasons – not related to asset damage/deterioration from traffic – why traffic variations affect the extent and frequency of maintenance and renewals on these types of rail infrastructure assets. Odolinski and Boysen (2019) show that the cost of carrying out activities on the infrastructure can vary with line capacity utilisation, even when holding infrastructure deterioration from traffic constant. The reason is that the ability to obtain access to tracks for maintenance work – track possessions – is harder (more expensive) for lines that are operating closer to capacity. In addition, Nilsson and Odolinski (2018) suggest that higher line capacity utilisation may generate more maintenance and change the timing of renewals – but here the motive would be to reduce the risk of train delays. Importantly this increased maintenance or renewal activity can occur due to an increase in traffic even though the relevant assets are not actually damaged by the running of trains on the network.

The purpose of this paper is to empirically re-examine the application of SRMC pricing for rail infrastructure usage. Specifically, using a large dataset for the French railway network, we consider if there are marginal costs for infrastructure assets that are in some sense triggered by traffic, but not caused by deterioration from traffic – but rather by the economic factors noted above, namely the additional cost of track access on busy lines, and the motivation of mitigating delays. Our empirical application focuses on an asset of particular interest, namely signalling equipment. The reason is that this equipment comprises relatively few components that deteriorate with traffic, and therefore forms a highly

---

* Corresponding author. The Swedish National Road and Transport Research Institute (VTI), Box 55685, SE-102 15, Stockholm, Sweden.

E-mail addresses: kristofer.odolinski@vti.se (K. Odolinski), A.S.J.Smith@its.leeds.ac.uk (A. Smith), P.E.Wheat@its.leeds.ac.uk (P. Wheat), jan-eric.nilsson@vti.se (J.-E. Nilsson), clement.dheilly@reseau.sncf.fr (C. Dheilly).

https://doi.org/10.1016/j.tranpol.2022.11.013
Received 19 May 2022; Received in revised form 5 October 2022; Accepted 18 November 2022

relevant case for testing the effects hypothesised in this paper. This is a particularly important case from a policy perspective since, as noted above, EU legislation in respect of track access pricing specifically highlights these assets.

Further, engineering methods suggest very low estimates of SRMC (maintenance and renewals) for signalling assets (e.g. Booz Allen Hamilton and TTCI, 2005; Öberg et al., 2007; ORR, 2008; Smith and Nash, 2018). For example, as part of the process of computing SRMC for the purpose of setting track access charges in Great Britain, ORR (2008) reports that roughly 5 per cent of signalling maintenance costs can be seen to be variable with traffic, with the remaining costs fixed in the short-run. This estimate is based on engineering judgement supported by some empirical data (see also ORR, 2013 [1]). There are no econometric studies in the academic literature applying econometric techniques to rail signalling maintenance specifically, though one regulatory study indicates higher variability (in the range 24–42% in ECOPLAN/IMDM (2020)) and this is also supported by a working paper study using Swedish data (34% variability, reported in Odolinski (2018)).[2] Exploration of the non-damage related, economic factors noted above could therefore explain and bridge the gap between the estimates from econometric as opposed to engineering methods.

If the costs for signalling equipment maintenance is affected by traffic in the way(s) proposed by this paper – namely, through the non-damage related factors noted above – the EU implementation rules for track access charges need to be updated in line with the underlying logic of marginal cost pricing. Since costs for these assets comprise a significant part of IMs’ spending on infrastructure services, this can have repercussions for the revenues from pricing of infrastructure use. For example, in Great Britain and France, signalling maintenance makes up around 16 per cent of total maintenance costs (see ECOPLAN/IMDM, 2020 and Network Rail Regulatory Accounts, 2020/21). Moreover, the same logic as for signalling may be empirically relevant also for supply of other infrastructure assets, such as electrification assets and track maintenance (though in those cases – and particularly track – there is clearly also a significant deterioration component to marginal cost, in contrast to signalling maintenance).

This paper is the first in the literature to estimate impacts of traffic on signalling maintenance decisions and costs, taking account of, and also distinguishing between, damage and non-damage related (economic) factors. This builds on the work of Odolinski and Boysen (2019), who carried out a similar study in respect of an aggregated measure of maintenance costs across all rail infrastructure assets using Swedish data. A related and additional contribution of our paper is that it compares econometric and engineering approaches to determining SRMC. As noted, both methods are permitted by EU legislation for computing direct costs for the purpose of determining track access charges, and our paper highlights the different results from the two approaches. The results have substantive implications for the level of track access charges in respect of signalling assets.

The remainder of the paper is structured as follows. A description of the potential links between traffic and costs for maintenance and renewal activities is provided in section 2. This forms the basis for the hypothesis in this paper, stating that traffic can trigger maintenance and renewal costs that are not caused by asset deterioration from traffic, and should be included in direct cost-based charges for infrastructure costs to contribute to an efficient use of the infrastructure. A review of the literature on rail infrastructure costs, with a focus on marginal cost estimation, is presented in section 3. Empirical tests are carried out using an econometric approach on data from the French railway infrastructure – the methodology and the model specification are presented in section 4, and the dataset is described in section 5. Section 6 presents the estimation results, whilst section 7 uses the results to calculate marginal costs. Section 8 concludes.

2. Links between traffic and costs for infrastructure provision[3]

Like all physical assets, railway infrastructure quality deteriorates over time and also when it is used by trains. This requires maintenance for preserving functionality, which can be defined as the ability of the assets to facilitate traffic according to the established timetable. To do so, the IM should seek to minimize life cycle costs (LCC) for society; assuming, as is typical in Europe, the infrastructure manager is a public body. LCC comprises costs for maintenance, renewal, and train delays. There is a trade-off between these components, where more frequent (costly) renewals and/or preventative maintenance (PM) generate lower corrective maintenance (CM) costs and fewer delays.

This paper is concerned with the way in which the IM handles this trade-off, as well as the impact on the possibility of obtaining track possessions for performing the maintenance work. The latter concerns the cost of shorter and more fragmented time slots for maintenance and/or more activities being carried out at night – which might occur on lines that are operating close to capacity. As set out in Section 3, Odolinski and Boysen (2019) also highlight these links between line capacity utilisation and costs using Swedish data, but applied to all rail infrastructure assets. One possible strategy to handle traffic increases is to keep renewal and PM constant. Over time, this would result in faster deterioration of the track quality, triggering more CM as well as increasing the frequency of train delays. To avoid this, traffic increases may induce the IM to increase the frequency of inspections, thereby identifying emerging challenges earlier and increasing the frequency of other PM activities.

Fig. 1 is used to illustrate our a priori hypotheses of the causal link between traffic and maintenance costs; building on the past literature (see section 3). Time and weather are included to show there are assets that do not significantly deteriorate with traffic, such as most parts of signalling and telecommunication equipment. Still, these assets need to be inspected to prevent failures that cause traffic delays. The frequency of inspections (classified as a maintenance activity) is typically linked to traffic; an increasing number of trains over a track section will sooner or later make train frequency pass some pre-set threshold value that is set to trigger more frequent inspections that can curtail the risk for incidents causing delays and costly CM.

Importantly, inspections and/or other preventative measures such as replacement of components are needed irrespective of whether the traffic creates damage, since a failing asset causes the same negative consequences for delays irrespective of whether it fails due to traffic-related damage or just due to age. The hypothesis is that it is therefore necessary to accelerate inspections (and/or other preventative measures) with traffic also for assets with no deterioration from traffic. This is re-emphasised with the link between line capacity utilisation and delays in Fig. 1, where more capacity being consumed will imply more delays in case of an incident on the line (grey arrows). Additionally, Fig. 1 shows a direct link from line capacity utilisation to maintenance cost (dashed arrows) to illustrate that more intensely-used lines can imply higher track possession costs because of the difficulty of gaining access for possessions activity.

In summary, the hypothesis which is empirically tested on signalling maintenance is that an increase in traffic may cause an increase in maintenance costs even without deterioration from traffic due to economic factors. This effect will come through two routes. First, when traffic increases, the cost of possessions access also increases, making

---

[1] Page 59.

[2] These are based on reported elasticities (specifically weighted average elasticities using train-km as weights); and since the elasticity can be expressed as the ratio of marginal to average costs these are typically reported as cost variability proportions in this literature – as explained further in Section 4.

[3] This section is partly based on the description in Nilsson and Odolinski (2018).

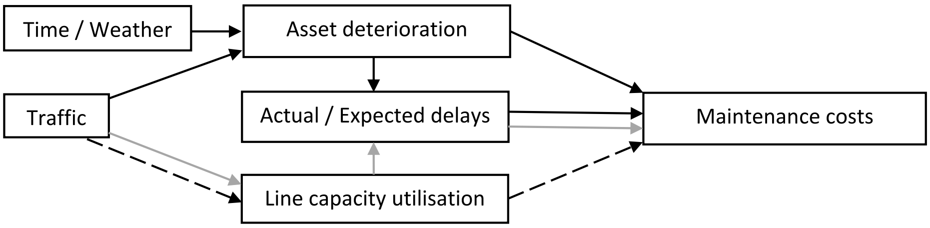

Fig. 1. Hypothesis on causal structure between traffic and costs for infrastructure provision.

maintenance more expensive (the dotted lines in Fig. 1); and second, a given asset failure will lead to higher delay costs the more trains that use the line – for this reason, an IM may increase PM to avoid these delay costs (the grey lines in Fig. 1).

3. Literature review

The vertical separation of infrastructure management and train operations in Europe has made track access pricing a topical research area. As noted in the introduction, SRMC is the basis for the transport pricing policy in the EU. This has generated an extensive literature on how costs for rail infrastructure services vary with traffic, using either cost allocation methods (see Link et al., 2008 for a review), bottom-up (engineering) approaches, top-down (econometric) methods, or a combination of approaches (Booz Allen Hamilton and TTCI, 2005; Smith et al., 2017 and 2021).

The engineering method considers physical deterioration mechanisms and models these to calculate the marginal rail infrastructure cost impact of an increase in traffic running on the network (see e.g., Öberg et al., 2007; Marschnig, 2016), whilst the econometric approach establishes a direct relationship between traffic and costs. This is illustrated in Fig. 2. The engineering method typically includes an extra step – indicated by the dashed arrows in Fig. 2 – that comprises the impact of asset deterioration/traffic on maintenance and renewal activities and then a link to costs. By considering the direct relationship between traffic and costs, the econometric approach may capture additional effects compared to the engineering approach (see Figs. 1 and 2).

The translog cost function (see Section 4) is the basis for the direct econometric approach in this area, where Münduch et al. (2002) and Johansson and Nilsson (2004) are two influential papers. This has been followed by a wealth of research using econometric methods, such as Andersson (2006), Tervonen and Pekkarinen (2007), Wheat et al. (2009), Gaudry and Quinet (2013), and more recently Walker et al. (2015), Smith and Nash (2018), ECOPLAN/IMDM (2020) and Odolinski et al. (2020). Both maintenance and renewals have been considered in this literature.

A common theme in past econometric studies is a focus on infrastructure deterioration costs caused by traffic (often referred to as 'wear and tear' costs), in line with the engineering view on how infrastructure costs vary with traffic. These estimates are obtained by modelling a relationship between costs and traffic, whilst controlling for infrastructure capability in order to derive a short-run estimate of marginal costs. Notable exceptions are Odolinski and Boysen (2019) and Odolinski and Lidén (2021) who estimate the impact of line capacity utilisation on costs for infrastructure maintenance and renewal costs, respectively, whilst controlling for infrastructure deterioration caused by traffic. Both of these recent studies consider higher costs for access to the network to explain the estimated impact of line capacity utilisation, though the former use an aggregate measure of total rail infrastructure maintenance costs. The latter study also hypothesises that the impact can be explained by more frequent renewals to prevent increased delay costs.

There is thus reason to study whether these effects also can be found for signalling maintenance, rather than the aggregate cost measures covering all rail infrastructure assets considered previously. The focus on signalling maintenance assets is particularly important as this asset class is specifically referred to in EU legislation as one that does not deteriorate with traffic and therefore should not be included in a charge for access to the rail infrastructure (see EU 2015/909 (8)). Further, as noted earlier, engineering methods suggest very low estimates of cost variability with traffic and in turn SRMC (maintenance and renewals) for signalling assets (e.g. Booz Allen Hamilton and TTCI, 2005; Öberg et al., 2007; ORR, 2008; Smith and Nash, 2018).

Overall then, the literature on marginal costs for rail infrastructure maintenance and renewals focuses on asset damage from traffic, not taking into account the economic aspects presented in section 2 that are related to increased track possession costs and the aim of preventing delays on busier lines. This background – along with the important policy relevance in respect of track access charging for signalling assets – therefore forms the entry point for our paper.

4. Methodology

The purpose of the paper is to test the hypothesis that there are marginal costs for infrastructure services triggered by traffic but not caused by asset damage from traffic. Section 2 suggested that increased line capacity utilisation implies higher marginal maintenance costs – even when holding deterioration effects fixed – because of (1) higher costs per maintenance activity, (2) more preventive maintenance carried

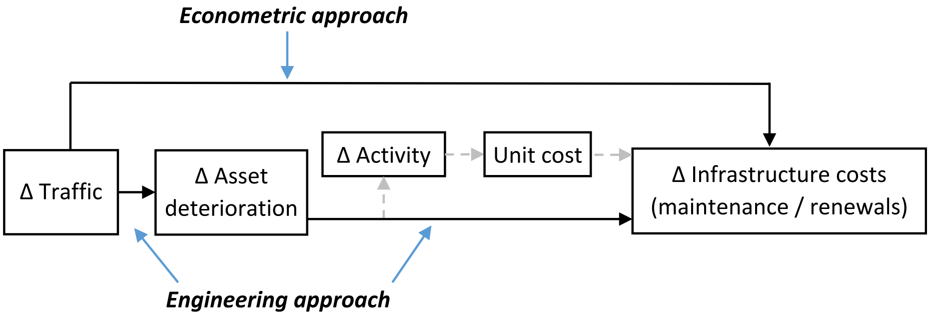

Fig. 2. Links between traffic and infrastructure costs in the engineering and econometric approaches.

out to curb an expected increase in delays, or (3) a combination of (1) and (2).

Fig. 3 explains the strategy for the empirical test. The figure includes two sections with the same number of trains (and train-km), but the line capacity utilisation is higher on the single-track section B. C_A and C_B are signalling maintenance costs on section A and B, respectively.

The assumption tested in the model estimation is that the marginal cost of traffic on section B (MC_B) is greater than the marginal cost of traffic on section A (MC_A). Since marginal cost can be written as the first derivative of costs with respect to traffic, this assumption can be also written as $\frac{\partial C_B}{\partial q} > \frac{\partial C_A}{\partial q}$, where q denotes the volume of traffic. It is important to note that the traffic levels and thus traffic-related damage to the assets are the same on both sections. The reason therefore for the difference in marginal costs between the two sections relates to the economic factors discussed above and not to any traffic-related damage. That is, resulting from a combination of higher track possession costs and/or more PM to curb delays on section B, which in turn results from the fact that section B has a higher line capacity utilisation for the same given level of traffic than section A.

A top-down (econometric) approach is used for modelling signalling maintenance costs (C) in line with the literature set out in Section 3:

$$C = f(\boldsymbol{x}) \tag{1}$$

where $\boldsymbol{x}$ includes a measure of traffic running on the network, as well as infrastructure characteristics and capability, and dummy variables, e.g., regions and years. Note that C comprises both preventative and corrective maintenance costs since – similar to most datasets in the rail cost literature – our data comprise an aggregated cost measure for these activities.

As is standard in this literature, we use a logarithmic transformation for both the dependent variable and explanatory variables, except the dummy variables and proportion variables. Specifically, the starting point for the cost model is the translog cost function, which is a second order approximation of any twice differentiable cost function (Christensen et al., 1973; Christensen and Greene, 1976) and has become the workhorse of the international cost function estimation literature (see e.g., Coelli et al., 2005).

Following Odolinski and Boysen (2019), the model is formulated with the following structure:

$$lnC_{it} = \alpha + \beta_1 lnd_{it} + \frac{1}{2}\beta_2 (lnd_{it})^2 + \beta_3 lnr_{it} + \beta_4 lnd_{it} lnr_{it} + \frac{1}{2}\beta_5 (lnr_{it})^2 + \boldsymbol{\theta z}_{\mathbf{it}} + \varepsilon_{it} \tag{2}$$

where C is signalling maintenance cost (this differs from Odolinski and Boysen, 2019, who used all rail infrastructure costs), d is train density (train-km (q) per route-km), r is number of (parallel) tracks per section, and z is a set of control variables for section i in year t. α, $\boldsymbol{\beta}$ and $\boldsymbol{\theta}$ are parameters to be estimated, and ε is an error term. Since in the data a section may comprise from one up to two parallel tracks, the variable for the average number of parallel tracks on each section is an important additional component of the model. Specifically, an interaction term between train density and number of tracks ($\beta_4 lnd_{it} lnr_{it}$) is included in the model, thus permitting the cost behaviour to vary depending on the number of (parallel) tracks per section (r), which in turn reflects the difference in line capacity utilisation for a given level of traffic on a single track as compared to a double track line.

The model set up reflects the hypothesis illustrated in Fig. 3. That is, the expectation is that the cost response to traffic – for a given level of traffic, and thus a given level of physical damage – is lower on sections with more tracks (A); the expectation being that $\widehat{\beta}_4 < 0$ and in turn marginal costs are lower on those sections[4] (see also Odolinski and Boysen (2019)). Note that the same marginal cost estimates could have been derived using train-km per track-km, in place of train-km per route-km, as the measure of traffic density. However, and following Odolinski and Boysen (2019), we use train-km per route-km in order to more clearly highlight the hypothesis of the paper. That is, we are interested in how costs change when, for the same deterioration on a route (measured by train-km/route-km), there is a different line capacity utilisation (measured by number of tracks) which prompts additional marginal costs relating to preventing delays and/or higher cost of access to the track.[5]

A few other important points need to be made about equation (2). First of all, the density variable is based upon train-km as the underlying measure of traffic. This reflects the engineering expectation that the number of train passages is considered to be the main determinant of maintenance activities on signalling equipment (both with respect to deterioration of the assets, the benefit of preventing delays, and the cost of track possessions) rather than tonnage. Here it should be noted that the vehicle is not exerting forces on the signalling assets (see ECOPLAN/IMDM, 2020 which also uses train-km as the variable in its signalling maintenance model in work done for and in conjunction with infrastructure manager SNCF Réseau).

Secondly, it is important to control for the different types of signalling assets used on different sections to avoid omitted variable bias. Thus variables for signalling assets (e.g., number of signals, number of track circuits, block types) are included in the model. Finally, normally, a cost function would include input prices. However, in the literature these are often not included as part of within country estimations for industries with wages that follow national pay scales (with no significant differences in qualifications and age across different parts of the country), and where materials are procured using common policies; see for example Johansson and Nilsson (2004). Nevertheless, to the extent that input prices vary across the country, this can be picked up by dummy variables for regions.

In line with the literature (e.g., Johansson and Nilsson, 2004; Wheat et al., 2009; Odolinski and Boysen, 2019), we calculate marginal costs (MC_{it}) for each track section by multiplying cost elasticities with respect to traffic ($\frac{\partial lnC_{it}}{\partial lnq_{it}}$) by average costs ($AC$), where AC is cost per train-km for each observation: C_{it}/q_{it}. Specifically, and in line with the literature in this field, the marginal cost is defined as (without subscripts i and t):

$$MC = \frac{\partial C}{\partial q} = \frac{q}{C}\frac{\partial C}{\partial q}\frac{C}{q} = \frac{\partial lnC}{\partial lnq}\frac{C}{q} \tag{3}$$

Here it should be noted that, because of the logarithmic form of the model and what is held constant when evaluating the traffic elasticity (see for example, Johansson and Nilsson, 2004 and Smith, 2012) the following equality holds:

$$\frac{\partial lnC}{\partial lnq} = \frac{\partial lnC}{\partial lnd} \tag{4}$$

and thus we obtain our estimate of the traffic elasticities from the coefficients on traffic density (d) from equation (2). Finally, we calculate a traffic weighted average marginal cost for the network as follows:

$$MC^W = \sum_{it}\left(MC_{it} \bullet q_{it} \middle/ \sum_{it} q_{it}\right) \tag{5}$$

---

[4] Ultimately the hypothesis is stated in respect of marginal costs as noted earlier.

[5] The use of train-km/track-km obscures the results because this measure of traffic density measure captures both deterioration due to traffic and number of tracks in the same metric.

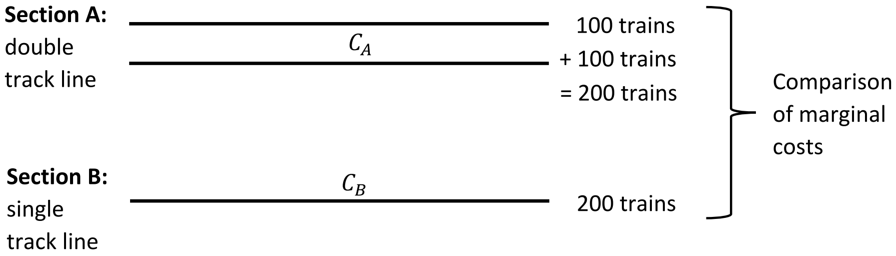

Fig. 3. Comparison in the empirical test.

5. Data

SNCF Réseau has provided data for the French railway network, comprising information about signalling maintenance costs, traffic, characteristics of the infrastructure and its capability. Examples of the latter are maximum linespeed, signals per track-km, TVM blocks ('Transmission Voie-Machine' = track-to-train transmission); these variables are important to control for since capability aspects can influence maintenance costs. If these variables are omitted in the (short-run) marginal cost estimation, there is a risk of picking up long-run effects. Overall, infrastructure characteristics and capability variables are important when estimating the short-run cost of running an extra train on various parts of the railway network as is standard in the literature (see section 3).

The French railway network is divided into around 1100 track sections. Our sample includes sections with up to two parallel tracks, comprising between 15,900 to 16,900 track-km on the French railway network. In line with Fig. 2 we focus on comparing single versus double track lines in order to isolate the relationship of interest (these comprise roughly 55% of the track sections on the network); since the relationship between capacity, number of tracks and cost can become highly complex for multiple track sections, particularly around stations, where in some cases the data indicates the presence of up to 37 parallel tracks. Results using the full sample indicates similar results to those for the restricted sample of single and double tracks, though further investigation in this area would be a useful line of future research.[6]

Information on the network's regional structure is also included in the dataset using dummy variables that indicate which region each section belongs to. Using region dummies in the model estimations can control for potential differences in the management of signalling maintenance. Moreover, as noted earlier, to the extent that input prices vary across the country, this can be picked up by the region dummies (year dummies are also included to control for general network effects over time).

There are in total 633 sections included in our sample, covering the period 2013 to 2018. The panel is unbalanced, with 533 up to 562 sections observed per year, generating 3303 observations in total. Descriptive statistics are presented in Table 1.

6. Results

Given that we have panel data we consider the following estimators: between effects, fixed (within) effects, and random effects. Note that the random effects estimator utilises both between-variation and within-variation in the data. The use of the former type of variation could generate a biased estimate depending on whether the variable in

Table 1
Descriptive statistics, observations per track section and year (3303 obs.).

Variable	Mean	Std. dev.	Min	Max
Cost				
Maintenance cost, EUR	91,016	126,493	15	1,114,636
Traffic				
Train density (train-km/route-km)	6456	8524	0.001	110,186
Infrastructure characteristics				
Track length, km	30.0	30.3	1.0	160.3
Route length, km	22.9	21.4	0.5	125.1
Average number of tracks	1.4	0.4	1.0	2.0
Switch density (switches per track-km)	0.6	0.7	0.0	6.4
Hot axle box detector, dummy	0.04	0.21	0	1
Capability				
Maximum speed, km/h	93.8	39.8	20	300
LGV (high-speed) line, dummy	0.03	0.17	0	1
Signals per track-km	3.41	2.89	0.3	16.8
Signalling stations per track-km	0.09	0.19	0	2.0
Track circuits per track-km	1.80	1.79	0	10.7
Proportion of signals classified as luminous, mechanic or indicator signals.	0.33	0.17	0	1
Proportion of signals classified placard signals	0.63	0.18	0	1
Manual blocks, proportion of section	0.13	0.31	0	1
TVM blocks ('Transmission Voie-Machine' = track-to-train transmission), proportion of section	0.02	0.12	0	1
Management				
Région ALCA (Alsace Lorraine Champagne-Ardenne), dummy	0.14	0.35	0	1
Région Aquitaine Poitou-Charentes, dummy	0.09	0.28	0	1
Région Bourgogne Franche-Comté, dummy	0.07	0.25	0	1
Région Bretagne Pays-de la-Loire, dummy	0.07	0.25	0	1
Région Centre Limousin, dummy	0.09	0.28	0	1
Région Haute et Basse Normandie, dummy	0.06	0.23	0	1
Région Ile-de-France, dummy	0.08	0.27	0	1
Région Languedoc-Roussillon, dummy	0.04	0.20	0	1
Région Midi-Pyrénées, dummy	0.06	0.23	0	1
Région Nord-Pas-de-Calais Picardie, dummy	0.14	0.34	0	1
Région Provence-Alpes-Côte-d'Azur, dummy	0.05	0.22	0	1
Région Rhône-Alpes Auvergne, dummy	0.13	0.33	0	1

---

[6] Whilst the elasticity relationship is similar once we include data for higher numbers of track, the marginal cost story is more nuanced, indicating a more complex data generating process when many multiple track sections are considered. We leave this for future research.

question is correlated with unobserved time-invariant characteristics – thus, the fixed (within) effects estimator guards against this type of bias. However, we focus on the random effects results, whilst also showing results from the between effects estimator and the fixed (within) effects estimator for comparison (even though the Hausman test result indicates that the fixed effects estimator is preferred[7]).

The main reason for not focusing on the fixed effects estimator is that there is (almost) no variation over time for a given track section in the key variable of interest that is used to distinguish between high and low capacity utilisation, namely the number of tracks. Further, the traffic variations required to influence preventative maintenance are less likely to be observed within track sections than between track sections. Indeed, the coefficient for number of tracks and the first and second order term for train density are not statistically significant using the fixed effects estimator (yet the interaction effect between train density and number of tracks has the expected sign). Overall, there is a trade-off between precision and bias (see e.g. Taylor 1980) and it is often the case in the literature that the choice between random effects and fixed effects is not only based on a Hausman test (see e.g. Clark and Linzer (2015)) and other factors come into play when choosing the most appropriate estimator.

Further, we have access to a rich set of control variables describing the characteristics and capability of the signalling assets (see Table 1), thus enabling us to control for heterogeneity and any potential sources of omitted variable bias that might otherwise arise, and thus isolate the effects of interest. Indeed, the interaction term between number of tracks and train density, which is the coefficient of main interest, is similar when using either the random effects or fixed effects estimator. Indeed it is for this reason – the presence of control variables in rail datasets – that the rail marginal cost literature in general tends to utilise random effects.

Estimation results are presented in Table 2. The log-transformed explanatory variables have been divided by their sample mean prior to taking logs. This means that their first order coefficients can be interpreted as elasticities at the sample mean. We started with the translog model in eq. (2) but dropped the squared term for log of number of tracks, which was not statistically significant and its exclusion generated a lower standard error for the first order coefficient. The main results are however not sensitive to this decision. The estimates for the control variables are in line with previous estimates on signalling maintenance in ECOPLAN/IMDM (2020), where for example increased switch density and signalling stations implies higher maintenance costs.

The key addition to the literature that we make here is to consider the impact of the variable of interest, average number of tracks and its interaction with train density. We find that the interaction term between number of tracks and train density is negative and statistically significant (−0.177, standard error 0.054, and p-value 0.001 when using the random effects estimator; see Table 2).

Therefore, for a given level of train density (here defined as train-km per route-km), and thus a given level of asset deterioration from that traffic, a higher number of tracks (which implies lower capacity utilisation) leads to a lower cost elasticity with respect to traffic (sensitivity of cost with respect to traffic). Put the other way round, a lower number of tracks (implying higher capacity utilisation for a given traffic level) leads to a higher cost elasticity (sensitivity to traffic). This means the traffic elasticity becomes a function of the number of tracks on a track section, and is shown to be falling with the number of tracks. By comparing lines with different levels of line capacity (number of tracks) but with the same level of train density per route-km, we are able to separate out the effects of traffic that result from factors relating to damage caused by traffic versus economic factors relating to capacity utilisation. As discussed earlier, it is argued that capacity utilisation impacts both on the cost of obtaining possessions and the desire of the infrastructure manager to avoid costly delays; this is discussed further in section 7 where we present the associated marginal cost estimates.

These findings are illustrated in Fig. 4, also including cost elasticities that are evaluated at the sample mean of number of tracks (the red line in the Figure). Fig. 4 shows the typical increasing relationship between the cost elasticity and traffic that has been found in this literature in general (see for example, Wheat et al., 2009; ECOPLAN/IMDM, 2020). However, Fig. 4 makes clear that actually there are multiple elasticity curves, depending on the average number of tracks on the section. Therefore we clearly see from Fig. 4 that the elasticity of cost with respect to traffic is lower (higher) when there are more (fewer) tracks on the route for a given level of traffic. In other words, for a given traffic level, the elasticity is lower when capacity utilisation is low and higher when capacity utilisation is high, which is the same type of relationship shown in Odolinski and Boysen (2019) for an aggregated measure of maintenance costs across all rail infrastructure assets.

One way of interpreting Fig. 4, and linking back to the literature, is that the black scatter points (double track sections) indicate the elasticity curve reflecting the very low cost impact of traffic-related deterioration, which is in line with EU legislation and the engineering view that signalling assets see little or no deterioration with traffic (as discussed earlier). However, the scatter points for the sections with less than double track (on average) see higher elasticities. This finding is in line with the hypothesis that where capacity is scarce, economic factors impact on the relationship between cost and traffic, either because of increased cost of track possessions, or because of an increased desire to limit delays. Specifically, note that the weighted average cost elasticity is 0.099 for the random effects model, and 0.137 for the between estimator, indicating that approximately 10–14% of signalling maintenance costs can be viewed as variable with traffic (as noted earlier, traffic-weighted average elasticities are typically used in this literature as a measure of cost variability). This variability is higher than the cost variability around 5% according to engineering estimates reported by ORR (2008).

7. Using results to calculate marginal costs

We use the estimated cost elasticities for traffic to calculate marginal costs. Ultimately the hypothesis in the paper is that, for a given level of traffic, the marginal cost of traffic on the infrastructure will be higher when there are fewer tracks (line closer to full capacity). The weighted average marginal cost (eq. (3)) is 0.067 euro per train-km in the model using the random effects estimator, and slightly higher (0.093 euro per train-km) when using the between effects estimator.

In Fig. 5 we plot marginal costs for sections with different intervals of number of tracks. In order to clearly isolate the impact of train density and its interaction with number of tracks, we multiply the cost elasticities by average costs that are evaluated (predicted) at the sample mean of all variables except train density, number of tracks, and track length. Before turning to discuss the main point of interest, firstly it should be noted that, as with previous studies on rail infrastructure maintenance and renewals (e.g., Wheat et al., 2009; Odolinski and Boysen, 2019; ECOPLAN/IMDM, 2020), the estimated marginal costs fall with traffic (marginal cost is below average cost, indicating significant economies of density, and both marginal and average costs fall with traffic).[8]

Secondly and more importantly, Fig. 5 shows that marginal costs are higher for sections with few tracks when comparing estimates at a certain point on the x-axis (i.e. for a certain train density level), which is in line with the hypothesis. Specifically, the higher marginal costs on

---

[7] We use a robust version of the test by including group means of the time-varying variables and estimate the model with random effects. The parameters for the group means are jointly significant, indicating that fixed effects is preferred (see e.g. Wooldridge, 2019).

[8] Note, as is common in the literature, we observe a small spike in marginal costs for very low traffic sections.

Table 2
Estimation results.

Dependent variable:	*Between effects*		*Fixed effects*		*Random effects*	
ln(signalling maintenance cost)	Coef.	Std. Err.	Coef.	Rob. Std. Err.	Coef.	Rob. Std. Err.
Constant	13.131***	0.454	−3.886	3.778	12.704***	0.389
ln(train density)	0.140***	0.040	0.062	0.050	0.102***	0.032
1/2ln(train density)^2	0.020**	0.008	0.007	0.015	0.014*	0.008
ln(train density)ln(no. of tracks)	−0.237***	0.061	−0.178	0.116	−0.177***	0.054
ln(no. of tracks)	−0.293*	0.155	0.744	1.046	−0.137	0.145
ln(track length)	0.864***	0.052	−1.102	1.631	0.898***	0.048
ln(max speed)	0.801***	0.130	0.308	0.310	0.691***	0.108
ln(switch den)	0.201***	0.065	−0.046	0.108	0.063	0.067
ln(signal per km)	0.094	0.075	1.775**	0.794	0.206**	0.083
ln(signal stat. per km)	0.109*	0.058	−0.464	0.384	0.145***	0.050
ln(track circ. per km)	0.206***	0.041	−1.674	1.813	0.229***	0.043
lumin. & mech signal prop.	−0.660	0.433	12.042	8.126	−0.604	0.463
placard signal prop.	−1.350***	0.387	14.870***	4.650	−1.497***	0.410
manual blocks prop.	0.153	0.120	0.240	0.154	0.191**	0.089
TVM blocks prop.	1.274***	0.449	0.259	0.414	0.864***	0.278
D. hot axle detector	0.333*	0.176	–	–	0.334**	0.135
D. LGV	−2.197***	0.307	–	–	−2.007***	0.328
D. zero signal stations per km	−0.275***	0.091	−0.082	0.365	−0.322***	0.090
D. zero track circuits per km	−0.483***	0.161	–	–	−0.565***	0.190
D. zero switch density	−0.369***	0.133	0.031	0.182	−0.099	0.134
Region dummies	Yes	–	–	–	Yes	–
Year dummies 2014–2018	Yes	–	Yes	–	Yes	–
R-squared	0.781		0.020		0.694	

***, **, *: Significance at the 1%, 5%, and 10% level, respectively.

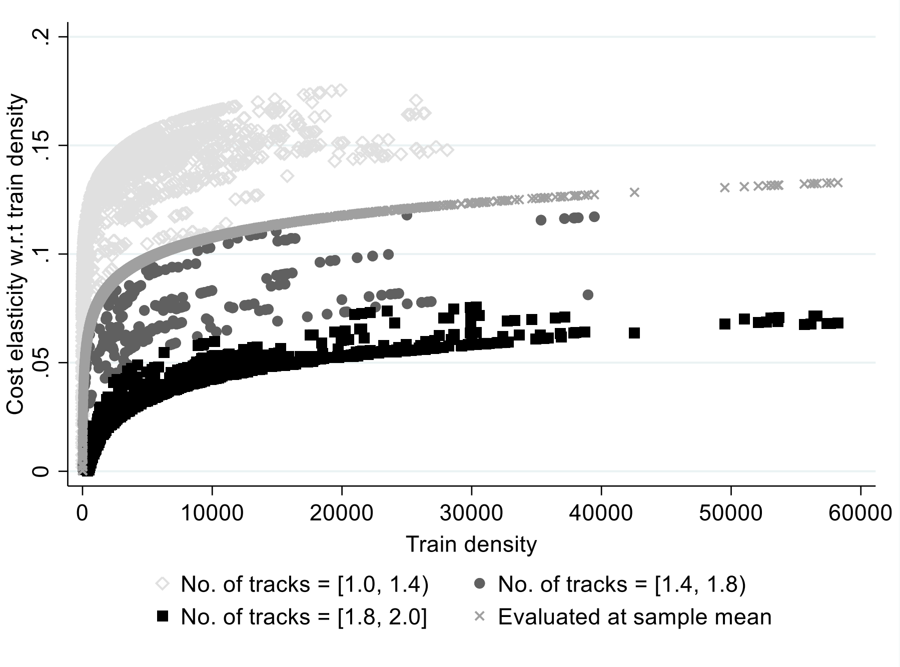

Fig. 4. Cost elasticities (capped at 60K train density): Model using random effects.

sections with fewer tracks is not driven by deterioration of the assets caused by traffic damage because the level of traffic running on the different sections, and hence direct damage to the assets, is the same. Rather, therefore, the variation in marginal costs depending on the number of tracks can be argued to be driven by other, economic factors as discussed earlier, namely: 1) higher cost of carrying out maintenance work on more intensely utilised routes; and 2) the desire to reduce the number of delay minutes on highly utilised routes through increased preventative maintenance.

Overall, the presence of line capacity-related differences in marginal costs illustrated in Fig. 5 show that there are additional cost components than just damage caused by traffic that can be included in track access charges based on direct costs/SRMC.[9] A differentiated charge using the variation in Fig. 5 could also be considered. However, the main point is that including these cost components – e.g. using an econometric approach similar to the method in this paper which picks up the economic effects described earlier – will generate higher direct cost-based charges overall, *ceteris paribus*. For example, using the engineering cost variability at 5% in ORR (2008) – which only considers damage from traffic – and the average costs in our sample gives a weighted

---

[9] As Nash (2018) notes, the concept of direct cost is intended to capture short-run marginal cost (SRMC).

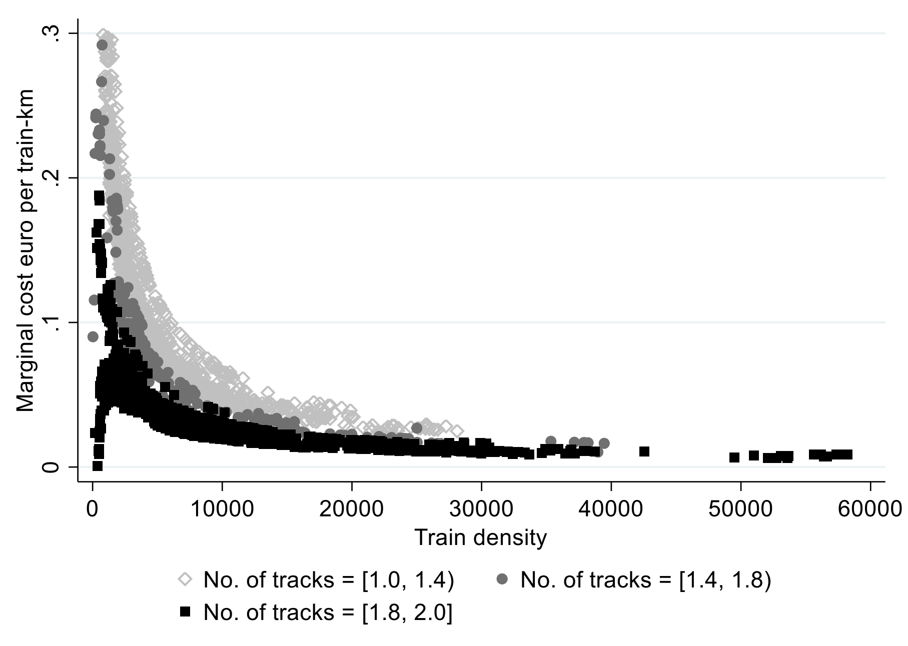

Fig. 5. Marginal costs (capped at 60K train density).

average marginal cost at 0.034 euro per train-km (see eq. (3) and eq. (4)), which is substantially lower than the weighted average marginal costs (0.067–0.093) calculated using the cost elasticities from our econometric approach.

8. Conclusions

The starting point for this paper is the hypothesis that there may be consequences of traffic variations for railway signalling maintenance costs that do not take the form of infrastructure damage caused by traffic. This is important from a policy perspective since engineering evidence regards deterioration of signalling assets with respect to traffic to be very low or even zero; and the current EU legislation on track access charges specifically notes that such costs should not form part of direct costs for charging purposes.

Our results lend strong support to this hypothesis, suggesting that the direct cost/SRMC for signalling assets is not only related to infrastructure deterioration costs caused by usage. Our results indicate that SRMC for these assets could also include costs relating to the higher cost of track possessions on lines with higher line capacity utilisation, and costs related to the associated desire to curb delays on such routes through increased preventative maintenance activity. These additional, economic factors form part of the objective function of an IM that tries to minimize LCC, including costs for train delays, where maintenance and/or renewals is implemented to prevent delays. When traffic increases, *ceteris paribus*, it is necessary to inspect the asset and possibly replace it slightly earlier than originally planned to avoid extra delays. The IM should increase these preventative measures until its extra production cost is equal to the marginal benefit of reducing costs for corrective maintenance and delays. This strategy may be optimal even in the presence of zero asset deterioration from traffic.

The importance of this paper is that it highlights additional components of SRMC that need to be included within direct costs and track access charges in order to contribute to an efficient use of the infrastructure. Indeed our econometric model indicates that the marginal signalling maintenance cost would increase from around 0.03 Euros per train-km (based on the engineering approach of capturing damage-related marginal cost) to 0.09 Euros per train-km if the non-damage related, economic factors are included within track access charges, based on the results of our econometric model. Thus our work would imply substantially higher track access charges in respect of signalling maintenance costs than implied by engineering approaches and when compared to the statement within EU legislation that signalling maintenance not caused by wear and tear from traffic should be excluded from a direct cost charge.

Further, our work also helps explain the differences in results from econometric and engineering methods reported in the literature – both of which are permitted under EU legislation as a basis for computing direct costs – with the former producing higher estimates of cost variability for signalling assets than the latter. Our work highlights that the reason for this finding can be found in the fact that econometric methods can capture economic factors that would not be picked-up in the engineering approach, which only reflects damage related elements.

The results in this paper therefore have important implications for EU track access charging policy. They suggest that EU legislation may need to be revised to be in line with the underlying logic of marginal cost pricing – recognising that some infrastructure assets will have costs that vary with traffic, even though there is no associated damage, and that these elements are a valid component of direct costs. Our results also have important implications for the choice of methodology to use for computing direct costs, given that engineering methods do not capture some of the economic factors highlighted.

Finally, as noted in Nash (2018), direct costs are only the starting point for computing track access charges, and non-discriminatory mark-ups are also permitted, as well as charges for scarcity/congestion and environmental costs. The design of the track access charging system may therefore comprise a number of different elements. The element of marginal cost related to the impact of line capacity on the ease of possessions access would be passed straight through to an increased direct cost computation, though potentially this may then lead to a situation where mark-ups are lower, based on an analysis of what the market can bear. Congestion charges, where they have been introduced within Europe, will reflect the cost of delays imposed by higher traffic over and above those reflected in the performance regime. However, provided the congestion charge and the direct cost charge are calibrated through the same dataset there should be no double counting of congestion charges (where they apply) and the element of direct costs linked to additional preventative maintenance required to mitigate delays discussed above.

Overall, a case-by-case approach to setting track access charges is needed, depending on the specific situation in each country to set charges appropriately – but starting with an accurate measure of direct costs is an important first step in the calculations and we argue that EU legislation may need to be re-considered in the light of the evidence presented in this paper.

CRediT authorship contribution statement

Kristofer Odolinski: Conceptualization, Formal analysis, Methodology, Writing – original draft, Writing – review & editing. **Andrew Smith:** Conceptualization, Formal analysis, Methodology, Project administration, Supervision, Writing – original draft, Writing – review & editing. **Phill Wheat:** Methodology, Writing – review & editing. **Jan-Eric Nilsson:** Conceptualization, Supervision, Writing – review & editing. **Clement Dheilly:** Data curation, Writing – review & editing.

Declaration of competing interest

None.

Data availability

The authors do not have permission to share data.

Acknowledgements

We are grateful to Prof. Chris Nash for comments on an earlier version of this paper, which has also been presented on the ITEA Annual Conference and School on Transportation Economics, Rome, 26 June 2021. We acknowledge the receipt of funding and provision of data from SNCF Réseau in support of this research, and also comments received on earlier versions.

References

Andersson, M., 2006. Marginal cost pricing of railway infrastructure operation, maintenance, and renewal in Sweden – from policy to practice through existing data. Transport. Res. Rec. 1–11. https://doi.org/10.1177/0361198106194300101, 1943.

Booz, Allen, Hamilton, TTCI, 2005. Review of Variable Usage and Electrification Asset Usage Charges: Final Report. Report R00736a15, London, June.

Christensen, L.R., Greene, W.H., 1976. Economies of scale in U.S. electric power generation. J. Polit. Econ. 84 (1), 655–676.

Christensen, L.R., Jorgenson, D.W., Lau, L.J., 1973. Transcendental logarithmic production frontiers. Rev. Econ. Stat. 55 (1), 28–45.

Clark, T.S., Linzer, D.A., 2015. Should I use fixed or random effects? Politic. Sci. Res. Method. 3 (2), 399–408. https://doi.org/10.1017/psrm.2014.32.

Coelli, T.J., Rao, D.S.P., O'Donnell, C.J., Battese, G.E., 2005. An Introduction to Efficiency and Productivity Analysis, second ed. Springer.

ECOPLAN/IMDM, 2020. Modelling railway infrastructure maintenance and renewal costs in France. Overview of estimates, Final report. Mimeo.

European Commission, 1995. Towards fair and efficient pricing in transport. Policy options for internalising the external costs of transport in the European Union. COM 95, 691.

EU 2015/909, 'Commission implementing regulation (EU) 2015/909 of 12 June 2015 on the modalities for the calculation of the cost that is directly incurred as a result of operating the train service' Off. J. Eur. Union L 148.

Gaudry, M., Quinet, É., 2013. 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, version 35, 8th January 2013.

Johansson, P., Nilsson, J.-E., 2004. An economic analysis of track maintenance costs. Transport Pol. 11, 277–286. https://doi.org/10.1016/j.tranpol.2003.12.002.

Link, H., Stuhlehemmer, A., Haraldsson, M., Abrantes, P., Wheat, P., Iwnicki, S., Nash, C., Smith, A.S.J., 2008. CATRIN (Cost Allocation of TRansport INfrastructure Cost), Deliverable D 1, Cost Allocation Practices in European Transport Sector. VTI, Stockholm. March 2008.

Marschnig, S., 2016. Innovative track access charges. Transport. Res. Procedia 14, 1884–1893.

Münduch, G., Pfister, A., Sögner, L., Siassny, A., 2002. Estimating marginal costs for the Austrian railway system. In: Vienna University of Economics & B.A., Working Paper No. 78, February 2002.

Nash, C., 2018. Track access charges: reconciling conflicting objectives. Project Report. CERRE, Centre on Regulation in Europe, 9 May. https://cerre.eu/publications/track-access-charges-reconciling-conflicting-objectives/.

Nilsson, J.-E., Odolinski, K., 2018. 'Marginalkostnader För Reinvesteringar I Järnvägsanläggningar: En Delrapport Inom SAMKOST 3', CTS Working Paper 2018:22. Centre for Transport Studies, Stockholm (In Swedish).

Öberg, J., Andersson, E., Gunnarsson, J., 2007. Track Access Charging with Respect to Vehicle Characteristics, second ed. Rapport LA-BAN 2007/31.

Odolinski, K., 2018. Marginalkostnader för järnvägsunderhåll: trafikens påverkan på olika anläggningar. In: CTS Working Paper 2018:24. Centre for Transport Studies, Stockholm (In Swedish).

Odolinski, K., Boysen, H.E., 2019. Railway line capacity utilisation and its impact on maintenance costs. Journal of Rail Transport Planning & Management 9, 22–33. https://doi.org/10.1016/j.jrtpm.2018.12.001.

Odolinski, K., Nilsson, J.-E., Yarmukhamedov, S., Haraldsson, M., 2020. The marginal cost of track renewals in the Swedish railway network: using data to compare methods. Economics of Transportation 22, 100170. https://doi.org/10.1016/j.ecotra.2020.100170.

Odolinski, K., Lidén, T., 2021. Railway Line Capacity and its Impact on Marginal Cost of Renewals. VTI Working paper 2021.

ORR, 2008. Periodic review 2008. Determination of Network Rail's outputs and funding for 2009-14. Offc. Rail Regulation. https://www.orr.gov.uk/media/10686.

Smith, A.S.J., Nash, C., 2018. Track Access Charges: Reconciling Conflicting Objectives. Case study – Great Britain'. CERRE Centre on Regulation in Europe, 9 May. https://cerre.eu/publications/track-access-charges-reconciling-conflicting-objectives/.

ORR, 2013. Periodic review 2013: final determination of network rail's outputs and funding for 2014-19. Offc. Rail Regulation. https://www.orr.gov.uk/sites/default/files/om/pr13-final-determination.pdf.

Smith, A.S.J., 2012. The application of stochastic frontier panel models in economic regulation: experience from the European rail sector. Transport. Res. E Logist. Transport. Rev. 48 (2), 503–515. https://doi.org/10.1016/j.tre.2011.10.003.

Smith, A.S.J., Iwnicki, S., Kaushal, A., Odolinski, K., Wheat, P., 2017. Estimating the relative cost of track damage mechanisms: combining economic and engineering approaches. Proc. Inst. Mech. Eng. - Part F J. Rail Rapid Transit 231 (5), 620–636. https://doi.org/10.1177/0954409717698850.

Smith, A.S.J., Odolinski, K., Hossein-Nia, S., Jönsson, P.-A., Stichel, S., Iwnicki, S., Wheat, P., 2021. Estimating the marginal maintenance cost of different vehicle types on rail infrastructure. Proc. Inst. Mech. Eng. - Part F J. Rail Rapid Transit 235 (10), 1191–1202. https://doi.org/10.1177/09544097211991309.

Taylor, W.E., 1980. Small sample considerations in estimation from panel data. J. Econom. 13, 203–223. https://doi.org/10.1016/0304-4076(80)90015-9.

Tervonen, J., Pekkarinen, S., 2007. 'Marginal Rail Infrastructure Costs in Finland 1997-2005', Finnish Rail Administration. Publications of the Finnish Rail Administration A3, 2007.

Walker, P., Wheat, P., Marti, M., Smith, A.S.J., 2015. Swiss case study, annex 3 to deliverable 5.3 access charge. In: The Sustainable Freight Railway: Designing the Freight Vehicle – Track System for Higher Delivered Tonnage with Improved Availability at Reduced Cost, SUSTRAIL.

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

Wooldridge, J.M., 2019. Correlated random effects models with unbalanced panels. J. Econom. 211 (1), 137–150. https://doi.org/10.1016/j.jeconom.2018.12.010.