Economics of Transportation 10 (2017) 8–17

Contents lists available at ScienceDirect

Economics of Transportation

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

Estimating the marginal maintenance cost of rail infrastructure usage in Sweden; does more data make a difference?

Kristofer Odolinski *, Jan-Eric Nilsson

The Swedish National Road and Transport Research Institute, Department of Transport Economics, Box 55685, 102 15 Stockholm, Sweden

ARTICLE INFO

JEL codes:
H54
L92
R48

Keywords:
Marginal cost
Rail infrastructure
Maintenance
Dynamic model
Track access charges

ABSTRACT

This paper updates knowledge about the marginal cost of railway maintenance. Using a panel dataset comprising 16 years, we test whether more data makes a difference to conclusions. In contrast to previous estimates using a shorter panel, maintenance costs are now demonstrated to exhibit a positive dynamic effect; an increase in maintenance cost during one year indicates the need for more maintenance also the next year. Moreover, the marginal cost from the dynamic model is larger than its static counterpart. We conclude that the use of dynamic models on longer time series may have charging implications in several EU member states, considering that their track access charges are based on econometric studies that use static models and short panel datasets.

1. Introduction

The way in which railway infrastructure maintenance is affected by variations in train traffic comprises one component of the social marginal costs for using railways. The policy relevance of this relationship was formally established after the vertical separation of infrastructure management and train operations, introduced by the European Union in 1991 (Dir. 91/440). This directive required the introduction of track access charges. The charging principles of infrastructure use was further specified when Dir. 2001/14 established that track access charges should be based on the direct cost of running a vehicle on the tracks. This means that train operators (inter alia) should be charged for the impact of traffic on infrastructure maintenance.

Except for that the level of marginal cost of track use is a platform for EU's infrastructure policy, the marginal cost pricing paradigm is also one of the pillars of a policy for efficient use of societies' resources. Against this background, the purpose of this paper is to present new empirical evidence on the marginal cost for rail infrastructure maintenance, using more data than the existing literature.

Previous research has used different approaches for estimating the cost incurred by running one extra vehicle or vehicle ton on the tracks. There are examples of so-called bottom-up approaches that use engineering models to estimate track damage caused by traffic (see Booz Allen Hamilton, 2005; Öberg et al., 2007). Starting with Johansson and Nilsson (2004), most studies have, however, used econometric techniques to estimate the relationship between costs and traffic; this is referred to as a top-down approach. This line of research first estimates the cost elasticity with respect to traffic and then uses the average maintenance cost; the marginal cost is the product of these two components.

A survey of econometric rail cost studies made in Link et al. (2008) report cost elasticities in the interval 0.13–0.38. Cost elasticities between 0.2 and 0.45 are recommended by Wheat et al. (2009), based on research from several European countries. More recent evidence is provided in Wheat et al. (2015), where for example the Swiss case study have access to a panel data set comprising 10 years and reports elasticities at about 0.5. However, this value is not directly comparable to the other results since it includes a renewal cost component. Table 1 lists some results of previous studies on Swedish data which is the most direct benchmark for our analysis. It is obvious that these elasticities are within the overall range of cited observations for Europe.

Estimating a dynamic model to analyze rail infrastructure costs is rare, Andersson (2008) being one exception. He uses a difference

* Corresponding author.
E-mail address: kristofer.odolinski@vti.se (K. Odolinski).

http://dx.doi.org/10.1016/j.ecotra.2017.05.001
Received 5 November 2015; Received in revised form 5 April 2017; Accepted 13 May 2017

Table 1
Previous estimates on the marginal maintenance cost of rail infrastructure usage in Sweden.

	Model	Output variable	Cost elasticity	MC[a]	MC[a] 2014 Prices[b]
Johansson and Nilsson (2004)	Pooled OLS	Gross ton	0.17	0.0012	0.0014
Andersson (2006)	Pooled OLS	Gross ton	0.21	0.0031	0.0036
Andersson (2007)	Fixed Effects	Gross ton	0.27	0.0073	0.0084
Andersson (2008)	Fixed Effects	Gross ton	0.26	0.0070	0.0080
	Difference GMM	Gross ton	0.34^{S} 0.22^{L}	0.0092^{S} 0.0060^{L}	0.0106^{S} 0.0069^{L}
Andersson (2011)	Box-Cox	Freight gross ton	0.05	0.0014	0.0016
		Passenger gross ton	0.18	0.0108	0.0124

[a] Marginal cost.
[b] Inflation adjusted using the Swedish consumer price index, S=short-run, L=long-run.

generalized method of moments (GMM) estimator, but only has information about four years. Another exception is the study by Wheat (2015) who estimates a panel vector autoregression model on maintenance and renewal costs in ten zones in Britain over a 15-year period. However, that paper does not make use of the panel data structure, and the estimations uses train rather than ton density, where the latter is the preferred traffic variable from a wear and tear perspective.

Our paper adds to the literature on rail infrastructure costs in two ways. First and foremost, the data set covers a longer period than most previous papers. The extended dataset has motivated the title of the paper since it is relevant to consider whether longer time series makes a difference to conclusions. In particular, we contribute to the literature by addressing the presence (or not) of dynamic properties of maintenance costs – i.e. if and how spending on maintenance in one year affect costs in subsequent years – on a much longer panel compared to Andersson (2008). Moreover, as opposed to Wheat (2015), we have access to more disaggregate data by asset and we model unobserved heterogeneity. Establishing a dynamic interaction is still within the short run marginal cost paradigm, since it would only mean that the consequences of traffic in one year has implications also for maintenance over a longer period. Secondly, for the first time in this literature, our data set also includes factor prices.

The seminal contribution by Small et al. (1989) provides a generic platform for the analysis of marginal costs of infrastructure use, in that case applied to roads. The focus of that book is on how traffic affects the date of major renewals. This is obviously complementary to our focus on variations in spending on day to day activities generated by traffic variations. Using different modelling approaches, the relevance of renewal costs for appropriate levels of track user charges is addressed in Andersson et al. (2012) and Andersson et al. (2016), as well as in Yarmukhamedov et al. (2016). Except for that maintenance and renewals are complementary from a marginal-cost-pricing perspective, there may be an interaction between day-to-day maintenance and renewal activities, in so far as a change in the allocation of resources for one of the components may affect the need for resources to be spent on the other. These links are however beyond the scope of this paper.[1]

The outline of the rest of the paper is as follows. The methodology Section (2) is followed by a description of the available dataset in Section 3. We present the results in Section 4. Section 5 comprises a discussion and conclusion of the results.

[1] Several previous studies combine renewal and maintenance costs (see Andersson, 2006, Tervonen and Pekkarinen, 2007; Marti et al., 2009; Wheat and Smith, 2009; Wheat et al., 2015). In view of the lumpy nature of renewals and our focus on maintenance costs, we refrain from this in the model estimations.

2. Methodology

Several challenges must be addressed by a model that can be expected to deliver estimates of marginal costs. Section 2.1 addresses the econometric approach. The static model to be estimated is presented in Section 2.2 while Section 2.3 considers the possibility that maintenance activities in year t depend on costs in $t-1$.

2.1. Econometric approach

From an engineering perspective, the weight of the rolling stock is a driver of rail infrastructure wear and tear. Gross ton-km (GTKM, i.e. an additional ton using the tracks) has therefore become the preferred charging unit in Europe and is the output measure used in marginal cost calculations. When the impact of an additional ton on maintenance costs is estimated, there is reason to separate scale (track length) and density (tons) effects, as these dimensions of track use may have different effects. Scale effects are related to long-run expansion of the railway network. Like the literature in this field, we instead use the cost elasticity with respect to gross tons (GT) and multiply with the average cost ($\frac{C}{GTKM}$) to derive the short run marginal cost per ton-km:

$$MC = \frac{\partial C}{\partial GTKM} = \frac{GTKM}{C}\frac{\partial C}{\partial GTKM}\frac{C}{GTKM} = \frac{\partial lnC}{\partial lnGT}\frac{C}{GTKM}, \tag{1}$$

where C is maintenance costs.,[2][3] To derive the cost elasticity with respect to gross tons, we use a short run cost function given by Eq. (2) where there are $i = 1, 2, \dots, N$ track sections and $t = 1, 2, \dots, T$ years of observations.

$$C_{it} = f(\boldsymbol{P}_{it}, \boldsymbol{Q}_{it}, \boldsymbol{F}_{it}, \boldsymbol{Z}_{it}), \tag{2}$$

$\boldsymbol{P}_{it}$ are input prices, $\boldsymbol{Q}_{it}$ the volume of output (gross ton) and $\boldsymbol{F}_{it}$ is a vector of network characteristics such as track length and rail age. $\boldsymbol{Z}_{it}$ is a vector of dummy variables which includes year dummies and variables indicating whether a track section belongs to a contract area tendered in competition (a reform introduced in 2002 with a gradual transfer to competition). Since the introduction of competitive tendering in an area rarely starts at the beginning of a calendar year, we include a dummy variable for years when there is a mix between tendered and not tendered in competition. See Odolinski and Smith (2016) for more details.

A common functional form in the literature on rail infrastructure costs is the double-log specification. Indeed, agents in maintenance production are more likely to have the same reactions to relative changes than to changes in absolute levels, and a logarithmic transformation of the variables can reduce skewness and heteroscedasticity (Heij et al., 2004).

2.2. Translog model

We start with the flexible translog cost function which, for example, allows economies of scale to vary with different output levels and the production structure can be non-homothetic (input demands can vary for different output levels). See for example Christensen and Greene (1976). The translog functional form is expressed as:

[2] The result in Eq. (1) assumes that $\frac{\partial KM}{\partial GT} = 0$, i.e. an extra ton that runs on a track section will not change the length of that section. More explicitly, we consider $C = f(GT \cdot KM)$, and $\frac{\partial C}{\partial GT} = f'KM + f'GT\frac{\partial KM}{\partial GT}$. If $\frac{\partial KM}{\partial GT} = 0$, we have $\frac{\partial C}{\partial GT} = f'KM$, which implies that $\frac{\partial C}{\partial GTKM} = \frac{\partial C}{\partial GT}\frac{1}{KM}$. Note also that an interaction term between GT and track length can be added in the model estimation to allow for the cost elasticity with respect to GT to vary with track length.

[3] As indicated in the introduction, C_{it} may also include spending on renewals. Since track sections often have zero resources spent on renewal, and there is a possible interdependence between the cost categories, adding renewals to maintenance does not contribute to our understanding of the latter cost category and its own (possible) dynamics.