Dynamics in rail infrastructure provision: maintenance and renewal costs in Sweden

Kristofer Odolinski\textsuperscript{a} and Phillip Wheat\textsuperscript{b}

\textsuperscript{a} \textit{Junior scholar.} The Swedish National Road and Transport Research Institute, Department of Transport Economics, Box 55685, 102 15 Stockholm, Sweden (kristofer.odolinski@vti.se)

\textsuperscript{b} Institute for Transport Studies, University of Leeds, 36-40, University Road, Leeds LS2 9JT, United Kingdom (p.e.wheat@its.leeds.ac.uk)

\textbf{Abstract}

In this paper, we extend to the literature on marginal wear and tear cost estimation in railways, by applying a panel vector autoregressive model to rail infrastructure renewals and maintenance costs, using an extensive dataset from Sweden. This study is significant given the inherent difficulties in modelling the substantial renewals element of infrastructure costs, as well as the need to account for the dynamics in renewals and maintenance. The dynamic model allows us to estimate equilibrium cost elasticities with respect to ton density, which are significantly larger than their static counterparts. Overall, this work highlights that dynamics in rail infrastructure costs are important to consider when setting track access charges with respect to the wear and tear caused by traffic. This is particularly important given several countries, for example France, Sweden and Switzerland, are now setting access charges based on econometric studies.

\textit{Keywords:} rail infrastructure; renewal; maintenance; panel vector autoregression; marginal cost; impulse response analysis

\textit{JEL codes:} R48, L92
1.0 Introduction

Infrastructure investments consume a large amount of resources. To reap the benefits of an investment, the infrastructure must be maintained and renewed due to the wear and tear caused by traffic and to some extent weather conditions. In other words, maintenance and renewals will affect the performance and reliability of the infrastructure. The activities performed will also determine the level of maintenance and renewals that needs to be carried out in the future. Specifically, there may be an intertemporal effect for each of the activities; the amount of maintenance (renewals) carried out in year $t$ can affect the amount of maintenance (renewals) carried out in year $t + s$. Moreover, the maintenance performed will determine when the infrastructure needs to be renewed and the size of the renewal, while a renewal can affect the level of maintenance carried out. We therefore expect interdependence between these activities.

The dynamics in maintenance and renewals implies that an infrastructure manager (IM) needs to strike a balance within and between these activities for a certain traffic level. A sudden increase in traffic may thus require an adjustment of these costs. This implies that the cost impact from traffic needs to be studied in a dynamic context, which is the objective with this paper. Specifically, the purpose of this paper is to provide empirical evidence on the interdependence between maintenance and renewals, as well as their intertemporal effects. The estimates can be used to calculate the marginal cost for traffic, which has become an important part of the track access charges that were introduced after the vertical separation between train operations and infrastructure management in Europe as of the 1990s.¹

In this study, we estimate a panel vector autoregressive (VAR) model. This is a dynamic model that considers several endogenous variables - renewals and maintenance in our case - in a multiple equation system. Our estimation approach is similar to Wheat’s (2015), with the exception that we take the panel data structure into account. Hence, we are able to model

¹ The Swedish reform took place in 1988, preceding the wider European reform.
unobserved individual heterogeneity, which in our case are unobserved effects specific for each contract area. Moreover, we have access to ton-km instead of train-km, where the former provides a better representation of wear and tear.

A central facet of the VAR model is to make structural analyses, in which identification of exogenous shocks in the system is required. In this paper, we make use of an impulse response analysis (IRA) to track how maintenance/renewal costs evolve over time due to an increase in lagged maintenance/renewal costs caused by an exogenous shock (due to for example an increase in traffic). The estimated effect a certain activity has on future activities will provide knowledge about the strategy currently used.

The paper is organized as follows. In section 2, we present the empirical context in which our study is positioned. The methodology used is described in section 3. It also includes a subsection on the equilibrium cost elasticity with respect to traffic; an elasticity that can be used in a calculation of the marginal cost for the wear and tear of the infrastructure. Section 4 comprises a description of the data. We specify our model in section 5. The estimation results are presented in section 6. Section 7 concludes.

### 2.0 Empirical context

Econometric analysis of railway maintenance and renewals costs is accepted by the European Commission as an appropriate methodology to set track access charges across EU member states (European Commission, 2015), and several countries, including Sweden, France and Switzerland have used econometric methods to inform the level of access charges within their countries. This motivates the empirical literature on marginal cost of railway infrastructure (see Link et al. 2008, Wheat et al. 2009), which has a wealth of research on the maintenance cost element (Munduch et al. 2002, Johansson and Nilsson 2004, Wheat and Smith 2008, Gaudry and Quinet 2009). However, the econometric evidence on the marginal cost associated with
renewals cost element is much less robust. Studies often add renewal costs to maintenance in the estimations (Andersson 2006, Tervonen and Pekkarinen 2007, Marti and Walker 2009, Wheat and Smith 2009), yet there are a couple of examples focusing on renewals only, although these use very disaggregate data by asset (Andersson et al. 2012 and Andersson et al. 2016). The lack of evidence on renewals cost partly reflects the lumpy nature of renewals investments, which in turn implies a long time series is required to capture the evolution of renewals expenditure to changes in traffic (Link et al. 2008 and Wheat et al. 2009).

Renewals expenditure accounts for roughly one third of the sum of maintenance and renewals expenditure in Sweden (Trafikverket 2016) and so the significance of this cost category and its relationship with maintenance should not be understated.\(^2\) In general, the importance of performing renewals and maintenance activities at the right time has generated a rather extensive literature on the optimization of these activities: Gaudry et al. (2016) and Andrade and Teixeira (2011) are railway examples, while Sathaye and Madanat (2011), de la Garza et al. (2011) and Gu et al. (2012) analyse the optimization of pavement maintenance and resurfacing activities (see Sharma and Yadava (2011) for a literature review on this area). Related to this literature is Small et al. (1989), who presents an equilibrium pricing and investment model, in which optimal road durability for a certain traffic volume is calculated together with the corresponding marginal costs.

Still, there is a lack of empirical evidence on the dynamics between and within maintenance and renewal activities, especially in the literature on rail infrastructure costs. A notable exception is the study by Wheat (2015), in which a VAR model is estimated for both maintenance and renewal costs in ten zones in Britain over a 15-year period. The study finds evidence on intertemporal effects, yet not for a relationship between renewals and maintenance

\(^2\) The share can be larger (or smaller) in other countries. See for example Grimes and Barkan (2006) for renewals as a proportion of total maintenance costs in the US during 1978-2002.
costs. An intertemporal effect is also found by Odolinski and Nilsson (2016) who estimate a dynamic model (system GMM) for maintenance costs only. Similar to Wheat (2015), they find that an increase in maintenance costs in one year - due to for example a traffic increase – predicts an increase in maintenance costs in the next year. Other examples on research where the dynamics between maintenance and renewals are taken into account, is Andersson (2008) and Odolinski and Smith (2016) who both use a dummy variable approach. However, it involves an arbitrary definition of major renewals and only allows for a stepwise effect of renewals on maintenance costs.

Thus, econometric evidence on the dynamics in rail infrastructure provision is scarce, despite its relevance for track access charges. Ultimately, marginal cost estimates that take dynamic effects of renewals and maintenance into account will be closer to the actual cost of running one extra unit of traffic on the railway, compared to the cost estimates based on static models for maintenance (see for example Wheat et al. 2009) and renewals (see for example Andersson et al. 2012 and Andersson et al. 2016).

3.0 Methodology

Sims (1980) proposed the VAR model as an alternative to the simultaneous equation macroeconomic models prevalent at the time, which he criticized for its problems with arbitrary identification. The (so called) exogenous variables in the models - used for example to identify an effect on either the demand or supply - were often not strictly exogenous due to expectations in the economy that can change the behaviour of the consumer (the demand) in addition to the variable’s direct effect on the supplier and vice versa. Hence, there is a problem of simultaneity in the outcomes, which is the same type of problem we have with maintenance and renewals. The VAR framework dispenses with such arbitrary identification through the use of lagged
explanatory variables which are by definition weakly exogenous even if the values in the current time period are endogenous.

The objective in using a VAR model is to capture the effects of unexpected exogenous shocks via identification strategies which, if properly specified, can make the model useful for forecasting and policy analysis. One strategy is to make use of the temporal dependence between the variables – that is, how fast they react to a shock. Considering the endogeneity of the maintenance and renewals, where we expect the latter to react more slowly to a shock than the former, estimating a VAR model can be a fruitful approach for analyzing the dynamics in infrastructure provision.

We consider a panel VAR(p) model, where p denotes the lag length used in the model. We have two endogenous variables: renewal costs \( R_{it} \) and maintenance costs \( M_{it} \), where \( i = 1,2 \ldots, N \) contract areas and \( t = 1,2, \ldots, T \) years. \( \alpha_{1,i} \) and \( \alpha_{2,i} \) are the unobserved individual-specific effects for the renewal and maintenance equations respectively, while \( u_{1,it} \) and \( u_{2,it} \) are their respective residuals, where \( \begin{pmatrix} u'_{1,it} \ u'_{2,it} \end{pmatrix} = u_{it} \sim iid(0, \Sigma) \). \( \Sigma \) is the covariance matrix of the errors. We also include a vector of exogenous variables \( X_{it} \) with parameters \( \beta_1 \) and \( \beta_2 \) for the maintenance and renewal equations respectively. Importantly traffic is included in the vector of exogenous variables.

\[
\begin{align*}
R_{it} &= \alpha_{1,i} + \delta_{11} R_{it-1} + \theta_{11} M_{it-1} + \beta'_{1} X_{it} + u_{1,it} \\
M_{it} &= \alpha_{2,i} + \delta_{21} R_{it-1} + \theta_{21} M_{it-1} + \beta'_{2} X_{it} + u_{2,it}
\end{align*}
\]

Lagged renewal and maintenance costs are included in both equations to capture the dynamics in maintenance and renewals, as well as the interdependence between these activities.

---

\(^3\) Here we present the VAR(1) model for expositional simplicity. We consider further lags in the model estimation.
We first perform model identification by making graphs of the data to spot trends and we also choose the lag order of the model based on model selection criteria (see section 5). The lag order relates to autocorrelation in the residuals that can be removed by increasing the number of lags. For consistent estimation of the model parameters, $E(u_{it}', u_{is}) = 0$, with $t \neq s$ i.e. no autocorrelation.

As a model check we perform a stability test, which can reveal if a stationary process is generated by the model, where stability implies stationarity. If the process is non-stationary, first differencing would be required to avoid spurious results. The modulus of each of the eigenvalues is below one in our estimated models, which implies that our vector autoregressions are stable and therefore stationary (see Lütkepohl 2005, p.14-15). The moduli are presented in output table 2.

The model is estimated with generalized method of moments (GMM). Examples of studies that use this type of estimator for a VAR model are Love and Zicchino (2006), Tiwari (2011), Ahlfeldt et al. (2014) and Góes (2016). In doing this, we need to consider that the lagged variables are correlated with the contract area specific effects. To remove these effects, we use the transformation proposed by Arellano and Bover (1995), which is forward orthogonal deviation, or Helmert transformation (examples of studies that use this transformation is Love and Zicchino (2006), Ahlfeldt et al. (2014) and Lee and Yu (2014)). More specifically, for each year and contract area, we subtract the mean of future observations.

We need to use instruments for the lagged variables, as these are correlated with the error terms. When including a set of lags as instruments, we use the method by Holtz-Eakin et al. (1988), which basically substitutes missing values (created by increasing the lag length of the instruments) with zeros. This allows us to increase the lags of the instruments without losing the number of observations in the estimation.
### 3.1 Granger Causality and Impulse Response Analysis (IRA)

As a first test of interdependence between renewals and maintenance, we test whether lagged values of maintenance can improve the prediction of current values of renewals compared to only using lagged values of renewals (and vice versa). This approach of testing causal relations in time series is called a *Granger causality* test, proposed by Granger (1969).

A Granger causality test does however not reveal how exogenous changes in one variable affect another variable over time. To trace the effect of changes in renewals and maintenance costs, we make use of IRA, which requires identification of exogenous shocks \( \varepsilon_{it} \). We assume these shocks to be a linear function of the residuals \( u_{it} = G \varepsilon_{it} \), where \( G \) is a 2x2 matrix. Given our knowledge about the nature of renewals and maintenance, we choose *recursive identification* as the method to identify the shocks, which requires an ordering of the variables such that the \( G \) matrix can be calculated from the covariance matrix \( \Sigma \) by using the Cholesky decomposition (see for example Sims 1989 and Christiano et al. 1999. For an overview of this method, see KVA 2011, p. 15-17). Simply put, the ordering should be constructed on the basis of how fast the variables respond, from slow to fast. In our case, renewals are ordered first as we assume that the only shock that can have an effect on current renewals is a shock in renewals, while current maintenance can be influenced by both a renewal shock and a maintenance shock. Notice that we only assume that a maintenance shock will not affect current renewals, while future renewals are likely to be influenced by maintenance shocks. In other words, maintenance expenditure is more agile to shocks than renewals, - that is, an IM responds first with a change in maintenance.

When the shocks have been identified, we can use them in an impulse response function (IRF):

\[
C_t = \sum_{k=0}^{K} H_k G \varepsilon_{it-k}
\]  

(2)
where \( k = 1, 2, ..., K \) is lag length and \( \mathbf{C}_t = (R_{it}, M_{it}) \). \( H_k \) are the weights, where \( H_0 = I \) (identity matrix). We use these weights in a plot to inspect how they vary with \( k \). More specifically, we inspect how maintenance responds to a shock in either renewals or maintenance, and vice versa.

### 3.2 Equilibrium cost elasticity with respect to traffic

With lagged cost variables in our model, we are able to calculate the “equilibrium cost elasticity” with respect to traffic, both for renewals and maintenance. For maintenance costs, the logic is that an increase in traffic in year \( t - 1 \) will affect maintenance costs in year \( t - 1 \), which in turn affects maintenance costs in year \( t \). Moreover, maintenance costs that have adjusted into equilibrium imply \( \ln M_{it} = \ln M_{i,t-1} = \ln M_{i}^e \). The equation for maintenance costs in (1), with a logarithmic transformation and a traffic variable \( Q \), is then

\[
\ln M_i^e = \alpha_{2,i} + \delta_{21} \ln R_i^e + \theta_{21} \ln M_i^e + \beta_{21} \ln Q_i + \beta_2' \mathbf{X}_i + u_{2,i} \tag{3}
\]

which gives

\[
\ln M_i^e = \frac{\alpha_{2,i}}{1 - \theta_{21}} + \frac{\delta_{21}}{1 - \theta_{21}} \ln R_i^e + \frac{\beta_{21}}{1 - \theta_{21}} \ln Q_i + \frac{\beta_2'}{1 - \theta_{21}} \mathbf{X}_i + \frac{u_{2,i}}{1 - \theta_{21}} \tag{4}
\]

The equilibrium cost elasticity with respect to traffic is then

---

\(^4\) The term long-run cost is often used in the literature, which requires that there are no fixed factors in the production. However, in our analysis, the rail infrastructure is mainly fixed. We therefore prefer the term equilibrium cost.
The same logic applies for the renewal equation.

As noted by Wheat (2015), there is also a secondary effect that can be considered when estimating the cost elasticity with respect to traffic. If there is interdependence between our endogenous variables, then a change in traffic will have an impact on renewal (maintenance) costs, which in turn can have an effect on maintenance (renewal) costs, and vice versa.

With the equilibrium maintenance cost in (4) and the equilibrium renewal cost

$$\ln R_i^e = \frac{\alpha_{1i}}{1-\delta_{11}} + \frac{\theta_{11}}{1-\delta_{11}} \ln M_i^e + \frac{\beta_{11}}{1-\delta_{11}} \ln Q_i + \beta'_1 X_i + \frac{u_{1i}}{1-\delta_{11}}$$  

we can derive the equilibrium maintenance cost elasticity with respect to traffic ($\gamma_M^e$) that includes the secondary effect from the renewal equation:

$$\gamma_M^e = \frac{\partial \ln M_i^e}{\partial \ln Q_i} = \frac{\beta_{21}}{1-\theta_{21}} + \frac{\delta_{21}}{1-\theta_{21}} \frac{\partial \ln R_i^e}{\partial \ln Q_i}$$  

The last term in equation (7) is the equilibrium renewal cost elasticity with respect to traffic

$$\gamma_R^e = \frac{\partial \ln R_i^e}{\partial \ln Q_i} = \frac{\beta_{11}}{1-\delta_{11}} + \frac{\theta_{11}}{1-\delta_{11}} \frac{\partial \ln M_i^e}{\partial \ln Q_i}$$  

Putting this expression into (7) gives

$$\gamma_M^e = \frac{\partial \ln M_i^e}{\partial \ln Q_i} = \frac{\beta_{21}}{1-\theta_{21}} + \frac{\delta_{21}}{1-\theta_{21}} \left( \frac{\beta_{11}}{1-\delta_{11}} + \frac{\theta_{11}}{1-\delta_{11}} \frac{\partial \ln M_i^e}{\partial \ln Q_i} \right)$$
Rearranging, we have

\[ y^{*}_{M_i} = \frac{\partial \ln M_i^e}{\partial \ln Q_i} = \frac{\beta_{21}(1-\delta_{11})+\delta_{21}\beta_{11}}{(1-\theta_{21})(1-\delta_{11})-\delta_{21}\theta_{11}} \quad (10) \]

We use equations (7) and (8) to get the corresponding renewal cost elasticity

\[ y^{*}_{R_i} = \frac{\partial \ln R_i^e}{\partial \ln Q_i} = \frac{\beta_{11}(1-\theta_{21})+\theta_{11}\beta_{21}}{(1-\delta_{11})(1-\theta_{21})-\theta_{11}\delta_{21}} \quad (11) \]

It is straightforward to include more than one lag in the above derivations.

### 4.0 Data

Data has been obtained from the Swedish Transport Administration (the Infrastructure Manager; hereafter referred to as the IM), and consists of renewal and maintenance costs, traffic, and characteristics of the railway network such as track length and rail age. We also consider an input price variable (wages) in this study, which has been obtained from the Swedish Mediation Office (via Statistics Sweden). A complete list together with descriptive statistics is provided in Table 1 below.

Maintenance is activities performed to implement railway services according to the timetable and maintain the railway assets. As of 2007, snow removal is defined as maintenance and is included in the maintenance contracts. We are however able to pinpoint the snow removal costs in the data, and we exclude these costs due to its (stochastic) weather dependence. Renewals consist of replacements or refurbishments of the railway assets.

Maintenance and renewals are procured separately by the IM. The IM used in-house production of renewals until exposure to competition was introduced in 2001, while competitive
tendering of maintenance services was introduced gradually in 2002. The effect competitive tendering had on renewal costs in Sweden has not been studied. However, in terms of maintenance costs, Odolinski and Smith (2016) find an 11 per cent reduction due to competitive tendering over the period 1999-2011. This indicates a structural change in infrastructure provision that needs to be considered when analyzing the interdependence between maintenance and renewals. Hence, Table 1 includes dummy variables for competitive tendering of railway maintenance; Mixtend which indicates the first year a contract area is tendered in cases this year is a mix between not tendered and tendered in competition; and Ctend which indicates the subsequent years an area is tendered in competition. Tendering variables for renewals are not included in this study due to missing information.

Table 1 – Descriptive statistics, 1999-2014 (480 obs.)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>St.dev.</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hourly wage, SEK*</td>
<td>156.7</td>
<td>11.7</td>
<td>128.9</td>
<td>187.4</td>
</tr>
<tr>
<td>MaintC (Maintenance costs), million SEK*</td>
<td>56.78</td>
<td>44.37</td>
<td>8.03</td>
<td>334.41</td>
</tr>
<tr>
<td>RenwC (Renewal costs), million SEK*</td>
<td>40.74</td>
<td>63.95</td>
<td>0.00</td>
<td>452.13</td>
</tr>
<tr>
<td>Route length, km</td>
<td>280</td>
<td>174</td>
<td>13</td>
<td>989</td>
</tr>
<tr>
<td>Track length, km</td>
<td>358</td>
<td>229</td>
<td>39</td>
<td>1203</td>
</tr>
<tr>
<td>Length of switches, km</td>
<td>8.68</td>
<td>6.62</td>
<td>0.58</td>
<td>37.67</td>
</tr>
<tr>
<td>Length of structures (tunnels and bridges), km</td>
<td>5.72</td>
<td>7.22</td>
<td>0.55</td>
<td>40.43</td>
</tr>
<tr>
<td>Average age of rails</td>
<td>18.83</td>
<td>5.83</td>
<td>3.76</td>
<td>38.98</td>
</tr>
<tr>
<td>Ton-density (ton-km/route-km), million</td>
<td>7.9</td>
<td>7.2</td>
<td>0.2</td>
<td>33.2</td>
</tr>
<tr>
<td>Mixtend</td>
<td>0.06</td>
<td>0.24</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Ctend</td>
<td>0.47</td>
<td>0.50</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Trend</td>
<td>8.45</td>
<td>4.50</td>
<td>1</td>
<td>16</td>
</tr>
</tbody>
</table>

* 2014 prices.

5 Already in 1999, about 45 per cent of the reinvestment projects were produced by private companies (Trafikverket 2012).

6 We do not think this is a significant issue for the estimation results as only one year in the estimation sample (2000) would include areas not tendered when using one lag in the model.
Data on infrastructure characteristics is available at a detailed level, while costs and traffic are reported at the more aggregate track section level. Moreover, each contract area for maintenance consists of several track sections. Considering that renewals can overlap adjacent track sections, we use contract areas as the identifier in our estimations. In that way, we have less artificial splits of renewal costs.

**5.0 Model specification**

To get a first impression of the main variables of interest, we make a graph of maintenance and renewal costs during years 1999 to 2014. We use costs per ton-km as data is missing for some track sections over this period. Both maintenance and renewal costs have an upward trend, yet renewal costs have a lumpier nature with more variation during the studied period. Because in our dataset we aggregate to contract areas, the lumpy nature of renewals implies that our data have fewer observations with zero renewal costs compared to track sections (for example analyzed by Andersson et al. 2012).

![Figure 1 - Indices for maintenance (MaintC) and renewal costs (RenwC) per gross ton-km (GTkm), 1999-2014 (1999=100)](image-url)
To control for fixed (time-invariant) effects in the variables, we time-demean the log transformed variables – that is, we subtract their group means: \( \ln y_{it} = \ln y_{it} - \ln \bar{y}_i \), where \( \ln \bar{y}_i = T^{-1} \sum_{t=1}^{T} \ln y_{it} \). As previously noted, we also use a Helmert transformation to control for contract specific effects. Moreover, to impose linear homogeneity in input prices we divide maintenance and renewal costs with wages – that is, we normalize the costs with wages.

To determine the lag length of our model, we use the model fit criteria proposed by Andrews and Lu (2001), which are consistent moment and model selection criteria (MMSC) versions of the Akaike information criterion (AIC), the Bayesian information criterion (BIC), and the Hannan-Quin information criterion (HQIC).\(^7\) We include the maximum number of lags in our model checks and consider the lag orders with lowest values of the MMSC versions of AIC, BIC and HQIC. The test results show that the model with lag order 1 has the lowest values of AIC (-56.74), BIC (-262.57) and HQIC(-231.17), while the model with lag order 2 has the corresponding values -50.55, -234.62 and -208.55. However, we consider the second lag may be informative from an a priori perspective in terms of the expected behaviour of infrastructure managers, i.e. the response in maintenance/renewals from an increase in renewals/maintenance can take longer than one year. Indeed, the estimation results show that the second order lag is useful. We therefore focus on the model with lag order 2 (Model A) in the next section.\(^8\)

For comparison, we also estimate static counterparts of the models - i.e. dropping the lagged variables and the renewal/maintenance equation - with renewals, maintenance, and the sum of maintenance and renewals as dependent variables (Models B1, B2 and B3, respectively).

\(^7\) The formulations of the criteria in Abrigo and Love (2015) are used.

\(^8\) We also considered a lagged traffic variable as it could potentially add useful flexibility between traffic and costs in addition to the effect of past traffic being picked up indirectly by lagged costs. However, its coefficient was not statistically significant and other coefficients did not change.
6.0 Results

Estimation results from Model A1 and A2 are presented in Table 2, where the former only includes the endogenous variables, and the latter includes a set of exogenous variables. The static comparison models (Models B1-B3) are presented in Table 3. The models are estimated with robust standard errors, using the iterative GMM estimator. We use the maximum lag length of the instruments (up to 14), which improves the efficiency of the model estimation; reducing the lag length to 12 or 11 for the instruments generates larger standard errors. All estimations are carried out with Stata 12 (StataCorp.2011) using the package provided by Abrigo and Love (2015).

The results for lagged maintenance and renewals are similar in both model A1 and A2 with respect to the signs of the coefficients for the lagged variables, yet the estimates for lagged maintenance in the maintenance equation are significantly lower when exogenous variables are included. This indicates that we may have omitted variable bias in Model A1. We focus on the results from Model A2 which includes variables for railway characteristics, ton density, dummy variables for competitive tendering and time trends.⁹

The significance tests of the parameter estimates for lagged variables in the maintenance and renewal equations can be interpreted as Granger causality tests. The prediction of current renewals is improved by lagged values of renewals, with a coefficient at 0.3193 that is significant at the 1 per cent level. This may seem odd, but a possible explanation is that budget restrictions can make it difficult to complete renewals of the railway assets during one year, which leaves some of the required renewals for the next year. Moreover, the coefficient for renewals costs in year t-2 predicts a decrease in current renewals. More specifically, the coefficient is -0.0843, yet with p-value=0.124. The estimated intertemporal effects for renewals

---

⁹ We tried different restrictions with respect to the variables for railway characteristics. This did not have a significant effect on the results.
then suggests that renewals within a contract area are likely to overlap between two years, and seem to have the expected decreasing effect on renewal costs in the subsequent year.

Table 2 – Estimation results, Models A1 and A2, with order 2 lags (342 obs.)

<table>
<thead>
<tr>
<th>Equation</th>
<th>Variable</th>
<th>Model A1</th>
<th>Model A2</th>
</tr>
</thead>
<tbody>
<tr>
<td>RenwC</td>
<td>RenwC_t-1</td>
<td>0.3361***</td>
<td>0.0655</td>
</tr>
<tr>
<td></td>
<td>RenwC_t-2</td>
<td>-0.0718</td>
<td>0.0521</td>
</tr>
<tr>
<td></td>
<td>MaintC_t-1</td>
<td>-0.3918</td>
<td>0.2645</td>
</tr>
<tr>
<td></td>
<td>MaintC_t-2</td>
<td>0.2535</td>
<td>0.2829</td>
</tr>
<tr>
<td></td>
<td>Ton density</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Track length</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Rail age</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Switch length</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Length of struct.</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Trend</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Trend^2</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Mixtend</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Ctend</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>MaintC</td>
<td>RenwC_t-1</td>
<td>-0.0107</td>
<td>0.0099</td>
</tr>
<tr>
<td></td>
<td>RenwC_t-2</td>
<td>-0.0021</td>
<td>0.0094</td>
</tr>
<tr>
<td></td>
<td>MaintC_t-1</td>
<td>0.4665***</td>
<td>0.0570</td>
</tr>
<tr>
<td></td>
<td>MaintC_t-2</td>
<td>0.1530***</td>
<td>0.0522</td>
</tr>
<tr>
<td></td>
<td>Ton density</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Track length</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Rail age</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Switch length</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Length of struct.</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Trend</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Trend^2</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Mixtend</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Ctend</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td><strong>Moduli of eigenvalues</strong></td>
<td>0.69, 0.26, 0.26 and 0.22</td>
<td>0.42, 0.35, 0.35 and 0.11</td>
</tr>
</tbody>
</table>

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

A lagged value of maintenance improves the prediction of current values of renewals compared to only using lagged values of renewals. The estimation results show that maintenance cost in year t-2 predicts an increase in renewals (MaintC_{t-2} is 0.5776, with p-value 0.018). Hence, this
The model suggests that a shock in maintenance may increase a need for renewals in the second year, while it is unlikely to occur in the first year (coefficient is -0.1671 with p-value 0.540). The impact on renewals is rather intuitive considering that renewals should be preceded by large (corrective) maintenance costs as this is what generally motivates a renewal. Figure 2 in section 6.1 provides an illustration of this relationship.

When it comes to lagged values of renewals in the maintenance equation, we do not find a significant Granger causality, and the estimate is close to zero. However, lagged maintenance costs predict an increase in current maintenance, with a coefficient at 0.3032 (p-value= 0.000). This estimate is somewhat higher than the coefficient in Odolinski and Nilsson (2016), who estimated a system GMM on Swedish data at the track section level (more observations available compared to the contract area level), generating a coefficient for lagged maintenance costs at 0.2140.

Traffic is a key driver of cost. Therefore, the cost elasticities with respect to ton density are of particular interest which, together with coefficients for lagged costs, allows us to estimate equilibrium cost elasticities. The parameter estimate for ton density in the maintenance equation is 0.2330 (p-value=0.010), which is in line with previous results on Swedish data (see for example Odolinski and Nilsson 2016 or Andersson 2008). In the renewal equation, the coefficient for ton density is 0.2633, which is lower than previous estimates on Swedish data (however, our estimate is not significantly different from zero, p-value = 0.498); Andersson et al. (2012) find a cost elasticity with respect to ton density at 0.547, and Yarmukhamedov et al. (2016) find elasticities between 0.5258 and 0.5646.

For comparison, we estimate the static counterparts of the models, including a model with the sum of maintenance and renewal costs as the dependent variable (see Table 3). The renewal model (B1) generates non-satisfactory results due to a negative and insignificant traffic elasticity estimate, which is not surprising given the lumpy nature of renewals. The results in
models B2 and B3 are more in line with the maintenance equation results in model A2, with similar cost elasticities with respect to ton density even though renewals are included in model B3. However, adding renewals to maintenance in this model is not the preferable approach given the results in Model B1.

Table 3 – Estimation results, Models B1-B3 (342 obs.)

<table>
<thead>
<tr>
<th>Dependent variable</th>
<th>Model B1</th>
<th></th>
<th>Model B2</th>
<th></th>
<th>Model B3</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Ton density</td>
<td>-0.1189</td>
<td>0.4612</td>
<td>0.2431***</td>
<td>0.0769</td>
<td>0.2506**</td>
<td>0.1197</td>
</tr>
<tr>
<td>Track length</td>
<td>3.239**</td>
<td>1.5187</td>
<td>0.5908***</td>
<td>0.2232</td>
<td>0.6333**</td>
<td>0.3236</td>
</tr>
<tr>
<td>Rail age</td>
<td>-0.2836</td>
<td>0.4919</td>
<td>0.1754***</td>
<td>0.1139</td>
<td>0.2027</td>
<td>0.1658</td>
</tr>
<tr>
<td>Switch length</td>
<td>-1.4351**</td>
<td>0.6651</td>
<td>0.4559***</td>
<td>0.1532</td>
<td>0.2406</td>
<td>0.2149</td>
</tr>
<tr>
<td>Length of structures</td>
<td>0.2130</td>
<td>0.6583</td>
<td>0.3610***</td>
<td>0.1191</td>
<td>0.3627*</td>
<td>0.1953</td>
</tr>
<tr>
<td>Trend</td>
<td>0.2501</td>
<td>0.2382</td>
<td>-0.1588***</td>
<td>0.0367</td>
<td>-0.0994</td>
<td>0.0709</td>
</tr>
<tr>
<td>Trend^2</td>
<td>-0.0323</td>
<td>0.0219</td>
<td>0.0190***</td>
<td>0.0033</td>
<td>0.0117*</td>
<td>0.0064</td>
</tr>
<tr>
<td>Mixtend</td>
<td>0.2823</td>
<td>0.4257</td>
<td>0.1019</td>
<td>0.0775</td>
<td>0.1264</td>
<td>0.1456</td>
</tr>
<tr>
<td>Ctend</td>
<td>0.1988</td>
<td>0.4526</td>
<td>-0.0051</td>
<td>0.0818</td>
<td>0.0518</td>
<td>0.1502</td>
</tr>
</tbody>
</table>

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

We calculate the equilibrium cost elasticities with respect to ton-density for both renewals and maintenance, using the results from model A2. These are presented in Table 4, where $\gamma^e$ denotes equilibrium cost elasticity and $\gamma^{es}$ equilibrium cost elasticity including secondary effects. The elasticity for renewals is not significant at the 10 per cent level, while the estimates for maintenance are significant at the 1 per cent level. Including the secondary effect does not have a large impact on the maintenance elasticity, while the elasticity becomes larger for renewals - from 0.3439 with p-value=0.506, to 0.5324 with p-value=0.350 - due to the positive coefficient for lagged maintenance costs in the renewal equation. All in all, the elasticities are larger than their static counterparts.
Table 4 – Equilibrium cost elasticities with respect to ton density, Model A2

<table>
<thead>
<tr>
<th>Cost elasticity</th>
<th>Coef.</th>
<th>Std. Err.</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\gamma^e_{\text{Maintenance}}$</td>
<td>0.3376**</td>
<td>0.1352</td>
</tr>
<tr>
<td>$\gamma^e_{\text{Renewals}}$</td>
<td>0.3439</td>
<td>0.5173</td>
</tr>
<tr>
<td>Including secondary effect</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\gamma^{es}_{\text{Maintenance}}$</td>
<td>0.3492**</td>
<td>0.1435</td>
</tr>
<tr>
<td>$\gamma^{es}_{\text{Renewals}}$</td>
<td>0.5324</td>
<td>0.5697</td>
</tr>
</tbody>
</table>

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

The dummy variable for tendering of maintenance contracts shows that maintenance costs decreased with about 11 per cent\(^{10}\), similar to the results in Odolinski and Smith (2016). In the renewal equation, the estimate for competitive tendering of maintenance is not significantly different from zero. In one way, this is not surprising considering that a decision to renew is not likely to be directly connected to the introduction of tendering of maintenance; the decision to renew ought to be more connected to the condition of the railway assets and how costly infrastructure failures are for society on a certain part of the network. However, the amount and/or type of maintenance carried out - which may have changed due to competitive tendering - is certainly connected to the condition of the railway assets, which affects the need for renewals. Still, as previously noted, the results do not indicate that competitive tendering of maintenance has affected the renewal costs. It should also be noted that a difference-in-differences approach would be a more accurate way of estimating the effect of tendering, an approach used in Odolinski and Smith (2016).

Finally, we note that the estimates for track length, average rail age, and length of structures have the expected signs in the maintenance equation. The estimate for average rail age is close to zero in the renewal equation, and not statistically significant, as is the coefficient for switch length.\(^{11}\) It is only the coefficients for switches and structures in the maintenance

\(^{10}\) $\exp(-0.1145)-1 = -0.1082$.
\(^{11}\) We also estimated the model without switch length, which did not affect the results significantly.
equation that is statistically significant. Using lag order 1 does not change these coefficients in
the maintenance equation significantly.

6.1 Impulse response analysis

We make use of IRA to trace the effect a shock in maintenance or renewals have on future costs of
these activities, which is shown in Figure 2 below. We use one standard deviation shocks
and trace its effects in each of the two equations in Model A2 (similar results are generated by
Model C2). The horizontal axis in both figures represent years. The dashed lines are 90 per cent
confidence intervals based on 200 Monte Carlo draws using Gaussian approximation.

The top left graph in Figure 2 shows how a shock in maintenance costs affects future
maintenance costs, while the top right graph shows how a shock in maintenance costs affects
future renewal costs. The lower graphs in Figure 2 show how a shock in renewal cost affects
the different cost categories. As noted in section 3, we use recursive identification of the shocks,
where we assume that current renewals can be affected by a shock in renewals but not by a
contemporaneous shock in maintenance. Still, past shocks in maintenance may influence
renewals according to this assumption. For comparison, we estimate the impulse response
function (IRF) with the opposite (wrong) ordering in the recursive identification of the shocks.
These functions are presented in Figure 3 in appendix, showing that the IRFs for maintenance
vs. renewals and renewals vs. maintenance are different with respect to levels and shapes,
compared to the corresponding graphs in Figure 2. As expected, the IRFs for maintenance vs.
maintenance and renewals vs. renewals do not change with the ordering in the identification
method.
As noted earlier, we found Granger Causality between lagged renewals and current renewals. The IRF indicates that a shock in renewals seems to be accompanied by more renewal costs within one year (and to some extent two years). This is not very surprising given budget restrictions and the lumpy nature of renewals - that is, these costs will probably stretch over more than one year in the accounting system.

From the upper left graph in Figure 2, we can see that a shock in maintenance costs has a direct impact on maintenance cost within a year, with a recovery within two or three years. This suggests that the IM adjusts rather quickly to a sudden increase in traffic, without making an over-investment in preventive maintenance. An over-investment would manifest itself as a decrease in maintenance cost for at least one year, and then adjust back to zero in the IRF.

**Figure 2 – Impulse response functions**
An impulse response function for the response in renewals from a maintenance shock is illustrated in the upper right graph in Figure 2. A shock in maintenance predicts a decrease in renewals (yet, it is not significantly different from zero). However, in the second year, the maintenance shock results in an increase in renewal costs. This suggests that an increase in maintenance costs may be a signal that it can be costly to continue with maintenance activities, and that a renewal is warranted. A renewal activity can be difficult to perform within a year due to the planning procedures required (procurement of the project and getting access to the tracks), making it more probable that the renewal response to a shock in maintenance mainly occurs in the second year.

7.0 Conclusions

This paper is the first paper in the literature to have estimated a panel vector autoregressive model for rail infrastructure costs, using data for Sweden. We provide estimates on dynamic effects in infrastructure provision that can be important to consider when pricing infrastructure use. Specifically, we provide empirical evidence on the relationship between maintenance and renewals, as well as evidence on intertemporal effects for each of these activities.

As well as insights into the interplay between maintenance and renewals cost over time, our results highlight that analysing costs within such a holistic dynamic model can result in different elasticities of cost and thus marginal cost, with respect to traffic. In particular, we find that once we take into account both the contemporaneous change in cost and future changes in cost from increasing traffic, the ‘equilibrium’ maintenance cost elasticity, and thus marginal cost of traffic, is substantially higher than that found for static models (both the estimate in this dataset and in comparison to those in the literature on Swedish data). We do acknowledge the uncertainty around our estimates (particularly with respect to renewals cost) and so do not advocate this approach to set charges in this specific case. However, our finding could have
important implications subject to further research, since econometric analysis of railway maintenance and renewals costs is accepted by the European Commission as an appropriate methodology to set track access charges across EU member states (European Commission, 2015). Thus, there could be substantial charging implications for adopting this methodology in the future. Moreover, we have contributed to the literature on renewal costs, showing that modelling this cost category in a dynamic setting can be a fruitful approach for future research.

Our results show that past maintenance costs can improve the prediction of current values of renewals compared to only using past values of renewals. We also found intertemporal effects for both renewal and maintenance costs; an increase in renewals (maintenance) during one year predicts an increase in renewals (maintenance) during the next. Our IRA shows how the intertemporal effects evolve over time, where the IM seems to adjust the maintenance costs quickly to a sudden increase in for example ton-density. The IRF for renewals has a similar shape, indicating that a renewal during one year is followed by additional renewal costs in the next year. It is probably the lumpy nature of renewals together with budget restrictions that makes it difficult to completely serve a need to renew the railway assets during one year in a contract area, leaving some of the required renewals to be made in the next year.

The type of results provided in this paper can also be a useful demonstration of the maintenance and renewal strategy currently used. Stripping out the effects from the current strategy is essential for making improvements, where a proper balance between maintenance and renewal activities can generate higher benefits at a lower cost. For example, the estimate for the second order lag of maintenance cost in the renewal equation gives us a hint on how sensitive renewal costs are to prior increases in maintenance. Moreover, the intertemporal effect for maintenance reveals how quickly this cost adjusts to equilibrium.

There are opportunities for more methodological research in this area over and above collecting more data and undertaking more empirical applications with a view of providing
more robust estimates. For example, the analysis in this paper is not able to answer if the quick adjustment in maintenance costs is avoiding an over-investment - that is, doing more than is necessary to uphold the performance of the infrastructure. In fact, the IM may well be over- or under-investing in maintenance after a sudden increase in traffic. User costs (values of train delays for passengers and freight companies) must be considered in this type of analysis. That is, with access to data on train delaying failures and delay costs for passengers and freight companies, it could be a step towards a cost-benefit analysis of maintenance and renewals which in turn can generate economically efficient levels of these activities. This is an area for future research.

Acknowledgements

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

References


European Commission, 2015. 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.


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


Appendix

Figure 3 – IRFs with wrong ordering in the recursive identification, Model A2