A spline function class suitable for transport demand models

Jeppe Rich

Department of Transport, Technical University of Denmark

Bygningstorvet 116B, 2800 Kgs. Lyngby, Denmark,

jr@transport.dtu.dk

Abstract

The paper present a spline function class specifically suitable for the estimation of time and cost attributes in transport demand models with a long distance domain. The function class is designed to be monotonic decreasing in utility and have monotonic and marginal decreasing elasticities. The properties are obtained by calculating separate scaling factors and interception points at each spline knot. The resulting function class can handle a wide range of cost-damping structures, which cannot be absorbed using traditional linear-in-parameter forms or non-linear forms such as the Box-Cox function class. The function is particularly useful for models with a semi-long to long distance destination domain for which heavy damping in the tail is likely to be found. The function class is implemented in the new Danish National Model and we present simulation as well as estimation results from the National Model.

Keywords: *discrete choice models, functional form, cost damping, spline functions*

1 Introduction

The aim of this paper is to introduce a new class of parametric functions specifically aimed at modelling cost and time attributes in transport models. These functions may be relevant for all types of transport models but in particular relevant for models with a wide destination domain as such models typically require additional functional flexibility. A feature of the function class is the ability to represent a wide range of cost-damping patterns (Daly, 2010) from lightly damped cost curves to super damped curves.

In the literature piecewise linear approximations (linear splines) has been discussed in Ben-Akiva and Lerman (1985) in a logit modelling framework and more recently in Pinjari and Bhat (2006) in a mixed logit framework. Power series expansions are also discussed in Ben-Akiva and Lerman (1985) and have typically been applied in transport modelling contexts as a simple test of the base model being linear. Other work on functional forms includes early work on Box-Cox transformations in Gaudry and Wills (1978) and Hensher and Johnson (1981) and more recently in Gaudry (2010). Box-Cox function applications are found in Gaudry et al. (1989), Ben-Akiva et al. (1987), Mandel et al. (1994) and Lapparent and de Palma (2002). More recently, Rich and Mabit (2015) investigate several combinations of linear-in-parameter but non-linear-in-attribute forms which in several cases is shown to outperform the Box-Cox model.

2 Methodology

Spline functions are piecewise polynomial functions connected in knots. However, unconstrained spline functions will not apply to cost and time attributes in a utility maximization context as these functions will not in general be consistent with random utility theory.

More specifically, the class of functions we are looking for should fulfill the following properties;

i) Monotonous decreasing utility with respect to time and cost.
ii) Monotonous and differential first-order derivative of utility with respect to cost and time.
iii) Be consistent with random utility theory

None of the properties will be fulfilled by ordinary unconstrained spline functions. Although splines are monotonous they do not ensure a decreasing utility as well as marginal decreasing sensitivity. Property ii) impose structure on the functional form and cause the corresponding elasticity curve to have a smooth and monotone pattern with respect to the input attributes. Moreover, these classes facilitate a flexible cost-damping structure. Property iii) make sure that utility V is decreasing in cost and time attributes such as x, e.g. $\frac{\partial V}{\partial x} < 0$. For simplicity we consider a linear-in-parameter function $\beta F(x(d))$ of time and cost attributes $x(d)$ on choice set d where $\beta < 0$. To simplify the notation we will suppress the notation of d.

The function class we are looking for can be written as

(1)
$$F(x, c_1, \dots, c_{Q-1}) = \begin{cases} F_1(x, c_1), & c_0 \leq x < c_1 \\ F_2(x, c_1, c_2), & c_1 \leq x < c_2 \\ F_3(x, c_2, c_3), & c_2 \leq x < c_3 \\ \dots & \\ F_q\left(x, c_{q-1}, c_q\right), & c_{q-1} \leq x < c_q \\ \dots & \\ F_Q\left(x, c_{Q-1}\right), & x > c_{Q-1} \end{cases}$$

For this class we will have that following requirements

(2)
$$F_q(x) \leq 0 \wedge \frac{\partial F_q(x)}{\partial x} \leq 0 \; \forall x \geq 0, \forall q = 1, \dots, Q$$

$$F_1(c_1, c_1) = F_2(c_1, c_1, c_2) \quad \wedge \quad \left.\frac{\partial F_1(x)}{\partial x}\right|_{x=c_1} = \left.\frac{\partial F_2(x)}{\partial x}\right|_{x=c_1}$$

(3)
$$F_2(c_2, c_1, c_2) = F_3(c_2, c_2, c_3) \quad \wedge \quad \left.\frac{\partial F_2(x)}{\partial x}\right|_{x=c_2} = \left.\frac{\partial F_3(x)}{\partial x}\right|_{x=c_2}$$

$$\dots$$

$$F_q(c_q, c_{q-1}, c_q) = F_{q+1}(c_q, c_q, c_{q+1}) \quad \wedge \quad \left.\frac{\partial F_q(x)}{\partial x}\right|_{x=c_q} = \left.\frac{\partial F_{q+1}(x)}{\partial x}\right|_{x=c_{q+1}}$$

2.1 Function candidates

We will consider two different function candidates, which fulfills all of the requirements. The first candidate arise from a polynomial function expressed in logarithmic space. We define $c_0 = 0$ and let $\boldsymbol{c} = \{c_1, \dots, c_{Q-1}\}$ represent the vector of knot points. Hence, we are considering the function class shown in equation (4) below

(4) Log-power spline class: $\mathcal{F}(x, \boldsymbol{c}) = \sum_{q=1}^{Q} 1_q(x)\left[\theta_q(\boldsymbol{c})ln(x)^q + \alpha_q(\boldsymbol{c})\right]$

Where the indicator function$1_q(x)$ is defined such that $1_q(x) = 1 \Leftrightarrow x \in [c_{q-1}, c_q]$ and zero elsewhere.

It is worth noting that a special case of the log-power polynomial with $\alpha = \left\{\frac{1}{\gamma}, 0 \dots, 0\right\}$ and $\theta_q = \frac{\gamma^q}{q!}$ $\forall q = 1, \dots, Q$ emerges as a Taylor expansion of a Box-Cox function (γ representing the Box-Cox parameter) using L'hopital's rule. Hence, the function class on which the spline function is build have roots in the Box-Cox model but when expressed as a spline it is quite different.

To express the function as a spline we consider $Q - 1$ knot points $c_1, \dots, c_{Q-1}$. The function is defined in such a way that the $ln(x)^Q$ function operates on the first part of the curve where $c_0 \leq x < c_1$. For $c_1 \leq x < c_2$ we apply the function $\theta_{Q-1}(\boldsymbol{c})ln(x)^{Q-1} + \alpha_{Q-1}(\boldsymbol{c})$ and continue successively such that the tail of the function is modelled using a pure logarithmic form. It is clear that the class can be extended easily by relaxing the requirement that q are integers. In that case we would simply require that $q_1 > q_2 > \dots > q_Q$. However, for now we will maintain the simple formulation to ease the calculations.

The second candidate arises from a sum of power function for which the power is gradually decreasing and always less than 1. Hence, we are considering the function class shown in equation (5) below

(5) Power spline class: $\mathcal{M}(x, \boldsymbol{c}) = \sum_{q=1}^{Q-1} 1_q\theta_q(\boldsymbol{c})x^{\pi_q} + \alpha_q(\boldsymbol{c}), 1 \geq \pi_Q \geq \dots \geq \pi_1 \geq 0$

The curvature of these two classes is not very different and the one class can always be parameterized in such a way that it will mimic the other class closely.

Only $Q - 1$ spline parameters $\theta(\boldsymbol{c})$ and $\alpha(\boldsymbol{c})$ are identified, hence the first $\theta(\boldsymbol{c})$ is normalized to 1 and the first $\alpha(\boldsymbol{c})$ set to 0. In order to determine these parameters we start by making use of equation (1)-(3). The trick is to start finding all scaling parameters $\theta(\boldsymbol{c})$ and then subsequently work out the $\alpha(\boldsymbol{c})$ intercepts. We start with the log-power class for the equation system can be solved successively, hence

(6) $$\left.\frac{\partial F_1(x)}{\partial x}\right|_{x=c_1}=\left.\frac{\partial F_2(x)}{\partial x}\right|_{x=c_1}\Leftrightarrow\frac{Qln(c_1)^{Q-1}}{(Q-1)ln(c_1)^{Q-2}}=\frac{Qln(c_1)}{(Q-1)}=\theta_1(c_1,0,\dots)$$

(7) $$\left.\frac{\partial F_2(x)}{\partial x}\right|_{x=c_2}=\left.\frac{\partial F_3(x)}{\partial x}\right|_{x=c_2}\Leftrightarrow\theta_2(c_1,c_2,0,\dots)\left(\frac{Q-1}{Q-2}\right)\ln(c_2)=\frac{Qln(c_1)}{(Q-2)}\ln(c_2)=\theta_2(c_1,c_2,0,\dots)$$

…

(8) $$\left.\frac{\partial F_q(x)}{\partial x}\right|_{x=c_q}=\left.\frac{\partial F_{q+1}(x)}{\partial x}\right|_{x=c_{q+1}}\Leftrightarrow\frac{Q}{(Q-q+1)}\prod_{r=1}^{q}\ln(c_r)=\theta_q(c_1,\dots,c_q,0,\dots)$$

The equation system can be solved successively and it can be shown that for $q=1,\dots,Q-1$ then

(9) $$\theta_q(\boldsymbol{c})=\frac{Q}{Q-q+1}\prod_{r=1}^{q}\ln(c_r)\,,\forall q=1,\dots,Q-1$$

For the power-spline class using similar manipulations it can be found that

(10) $$\theta_q(\boldsymbol{c})=\frac{\pi_Q}{\pi_{Q-q}}\prod_{r=1}^{q}c_r^{\pi_{Q+1-r}-\pi_{Q-r}}\,,\forall q=1,\dots,Q-1$$

The next step is to find the interception parameters $\alpha_1(\boldsymbol{c}),\dots,\alpha_{Q-1}(\boldsymbol{c})$ in order to make sure that the utility functions are smooth and connected in all of the knots. These points can be found from the recursive equation below

(11) $$\begin{gathered}F_1(x,c_1)|_{x=c_1}=F_2(x,c_1,c_2)|_{x=c_1}\\ F_2(x,c_1,c_2)|_{x=c_2}=F_3(x,c_2,c_3)|_{x=c_2}\\ \dots\\ F_q(x,c_{q-1},c_q)\big|_{x=c_q}=F_{q+1}(x,c_q,c_{q+1})\big|_{x=c_q}\end{gathered}$$

By inserting the calculated scale parameters $\theta_1(\boldsymbol{c}),\dots,\theta_{Q-1}(\boldsymbol{c})$ these recursive equations reduces to

(12) $$\begin{gathered}\beta ln(c_1)^Q=\theta_1(\boldsymbol{c})\beta ln(c_1)^{Q-1}+\alpha_1(c_1,0,\dots)\\ \theta_1(\boldsymbol{c})\beta ln(c_2)^{Q-1}+\alpha_1(c_1,0,\dots)=\theta_2(\boldsymbol{c})\beta ln(c_2)^{Q-2}+\alpha_2(c_1,c_2,0,\dots)\\ \dots\\ \theta_{q-1}(\boldsymbol{c})\beta ln\left(c_q\right)^{Q-q+1}+\alpha_{q-1}(c_1,c_2,\dots,c_{q-1},0,\dots)=\theta_q(\boldsymbol{c})\beta ln(c_2)^{Q-q}+\alpha_q(c_1,c_2,\dots,c_q,0,\dots)\end{gathered}$$

Which for the general case can be stated more compactly as

(13) $$\alpha_q(\boldsymbol{c})=\alpha_{q-1}(\boldsymbol{c})+\theta_q(\boldsymbol{c})\beta\ln\left(c_q\right)^{Q-q}\left[1+\frac{\theta_{q-1}(\boldsymbol{c})}{\theta_q(\boldsymbol{c})}\ln(c_q)\right],\forall q=1,\dots,Q-1$$

For log-power splines with $Q=2$ and $Q=3$ Table 1 below summaries the spline parameterization.

Q	q	θ_q	α_q
2	2	$\frac{1}{2}\ln(c_1)$	$-\beta\ln(c_1)^2$
3	2	$\frac{3}{2}\ln(c_1)$	$-\frac{\beta}{2}\ln(c_1)^3$
3	3	$3\ln(c_1)\ln(c_2)$	$-\frac{\beta}{2}\ln(c_1)[3\ln(c_2)^2+\ln(c_1)^2]$

Table 1: Spline parametrization for the log-power spline class with $Q = 2$ and $Q = 3$.

Each of the two spline classes in (4) and (5) is formulated in such a way that the underlying log-likelihood function is in fact differential for all $x \geq 0$. This implies that the optimal knots can be found from an ordinary Newton Algorithm and in principle be estimated simultaneously with the remaining parameters. However, as it appears from Table 1 even low-dimensional cases will be non-linear and tend to be relative "ugly". Moreover, the optimization needs to be a constrained non-linear optimization in order to maintain the order of the knot points and make sure that $c_1 < c_2 < \cdots < c_Q$.

3 Results

Firstly we illustrate the curvature in a small-scale simulation context. We consider a spline of the order of $Q = 3$. The corresponding spline parameterization was shown in Table 1. Below in Figure 1 we illustrate the curvature of the utility function and the corresponding elasticity curve for a simple case where $x < 550$ and $\beta = -0.05$.

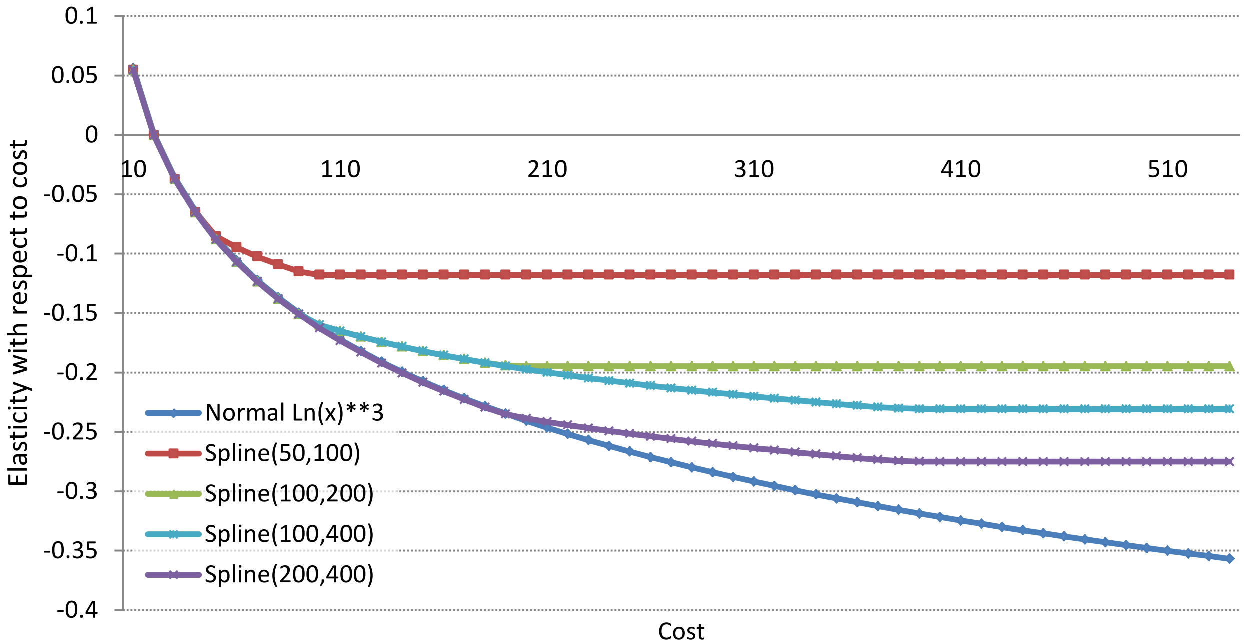

Figure 1: Comparison between a base $\beta\ln(x)^3$ and spline-curves with different degrees of tail damping.

The real benefit of applying the spline function is that we are able to control more effectively what happens in the tail. In particularly we are able to impose a heavy damping characteristic from a certain threshold. This is not accommodated by any other simple functional form such as the Box-Cox or alternative linear approximations of the Box-Cox as presented in Rich and Mabit (2015).

What happens on the first part of the curve is less interesting and less problematic as we can easily blend other function classes with the spline function. A good blending candidate is a pure logarithmic form as it will be active on the first part of the curve and will not violate the properties of the spline function as the log-terms with a power higher than 1 will out weight the logarithmic curve.

3.1 Results from the Danish national model

Until now we have assumed the underlying discrete choice model to be multinomial and operate on a destination choice domain. However, it may be more relevant to consider a nested-logit model by adding a mode-choice dimension. It is quite easy to see, that if the logsum parameters does not change with the choice of destination or mode, all of the calculations carries over and the spline parameters are similar to those found above. If this is not the case it will generally affect the scaling and thereby the interception points. The calculations are relative straightforward and the scaling parameters are essentially identical to the ones found above except for additional scaling from the logsum parameters.

In the Danish national model we apply nested logit models with logsum parameters that does not change with the choice of destination and mode. The models are characterized by having a relative wide destination domain and it has been difficult to find appropriate function candidates applicable for the entire destination domain. As a result, we have turn to a log-power spline formulation which has turned out to be particularly well functioning for leisure travel and shopping travel as these trip purposes does have heavy damping in the tail. This has not been the case for business and commuting, which is far less damped. The core function implemented in the Danish National model (only for weekday travel and only for primary trips) is shown in (14) below

$$F_n(GTT_{m,d|n}, \beta, c_1, c_2) = \sum_{q=1}^{3} 1_q \left[\beta\theta_q(\mathrm{c}_1, \mathrm{c}_2)\ln\left(GTT_{m,d|n}\right)^{Q+1-q} + \alpha_q(\mathrm{c}_1, \mathrm{c}_2)\right] \quad (14)$$

With $GTT_{m,d|n} = \frac{cost_{m,d}}{\mu_n} + time_{m,d}$ where μ_n is the value-of-time for person n. The value-of-time is income dependent but is fixed prior to the estimation. The function in (14) is then further “blended” with simple log-functions to introduce a more flexible description for the shorter trips. Below we illustrate the immediate goodness-of-fit impact of using the log-power spline class compared to the previous model for which the base model was actually identical to the parametrization with $c_1 = \infty$ and $c_2 = \infty$ which is shown to the right.

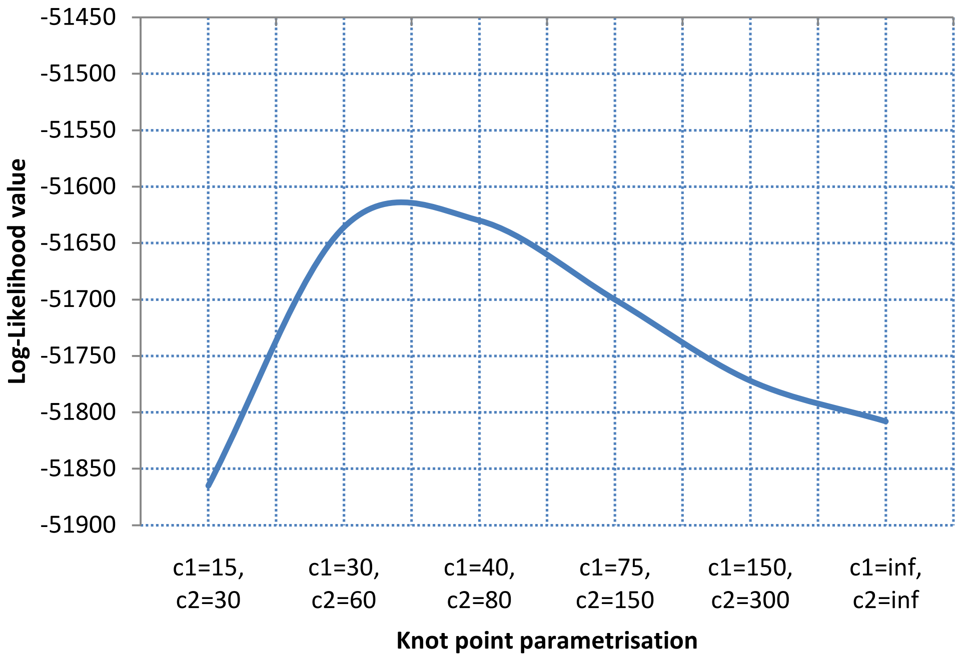

Figure 2: Log-Likelihood performance of leisure trip segment as a function of different knot point parameters.

As can be seen, there is a relative massive improvement in likelihood corresponding to close to 200 likelihood points. It should be said that the previous model with $c_1 = \infty$ and $c_2 = \infty$ was in fact a well-tested model, which replicated the distance profile of the segments closely.

The implication for the model sensitivity of choosing a near-optimal combination of c_1 and c_2 in contrast to choosing $c_1 = \infty$ and $c_2 = \infty$ is significant. The overall mean of the elasticities will not be affected to a great deal as only a small share of the trips exceed c_1 and c_2. However, for the longer trips the spline-function prevents the elasticities to “sky-rock” as it introduces a maximum elasticity.

This have several implications, firstly, for large-scale infrastructure projects for which the share of long-distance trips are high, using a proper spline-form will significantly reduce the elasticities for these trips. Also, the elasticity will tend to move from being sensitive to distance to being sensitive to the choice of mode. Hence, whereas a form which is too aggressive in the tails (such as the $c_1 = \infty$ and $c_2 = \infty$ model) will be too active in the destination choice (and thereby the implicit choice of distance) a spline-form will lower the distance sensitivity and have more sensitivity in the choice of mode. Figure 3 and Figure 4 illustrate the impact on the elasticity curve.

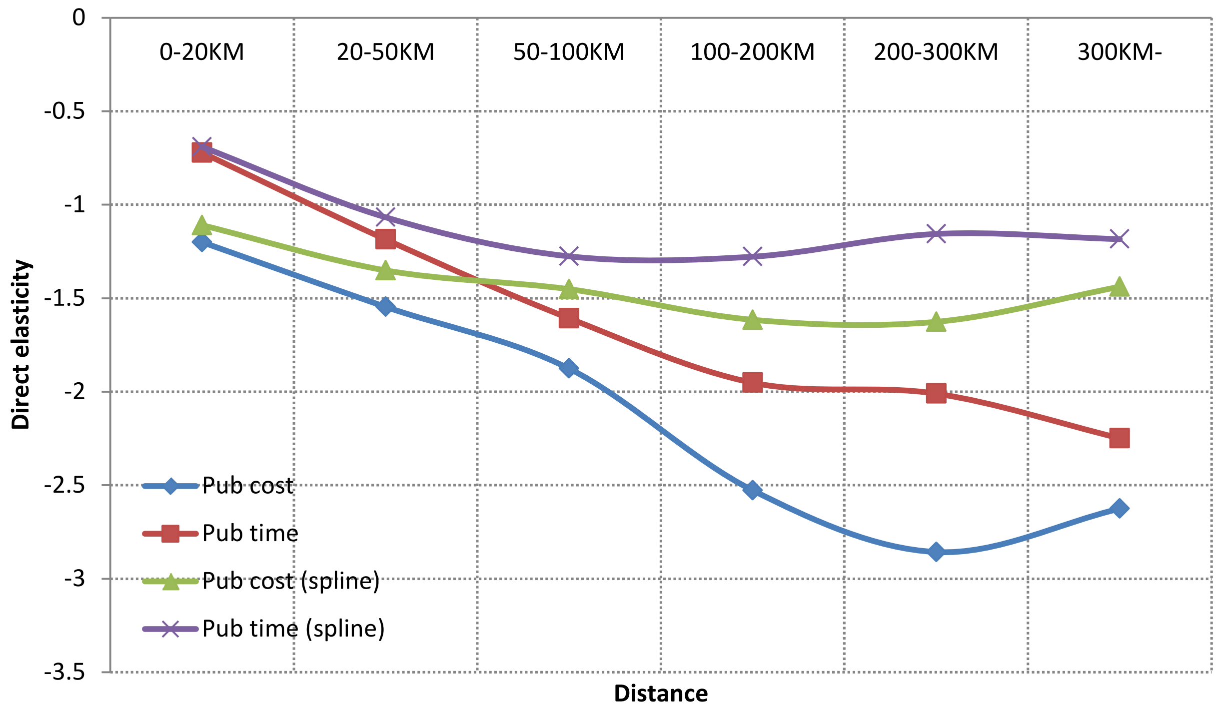

Figure 3: Elasticity curve for public cost and time for leisure travel for spline model versus normal model with $c_1 = \infty$ and $c_2 = \infty$.

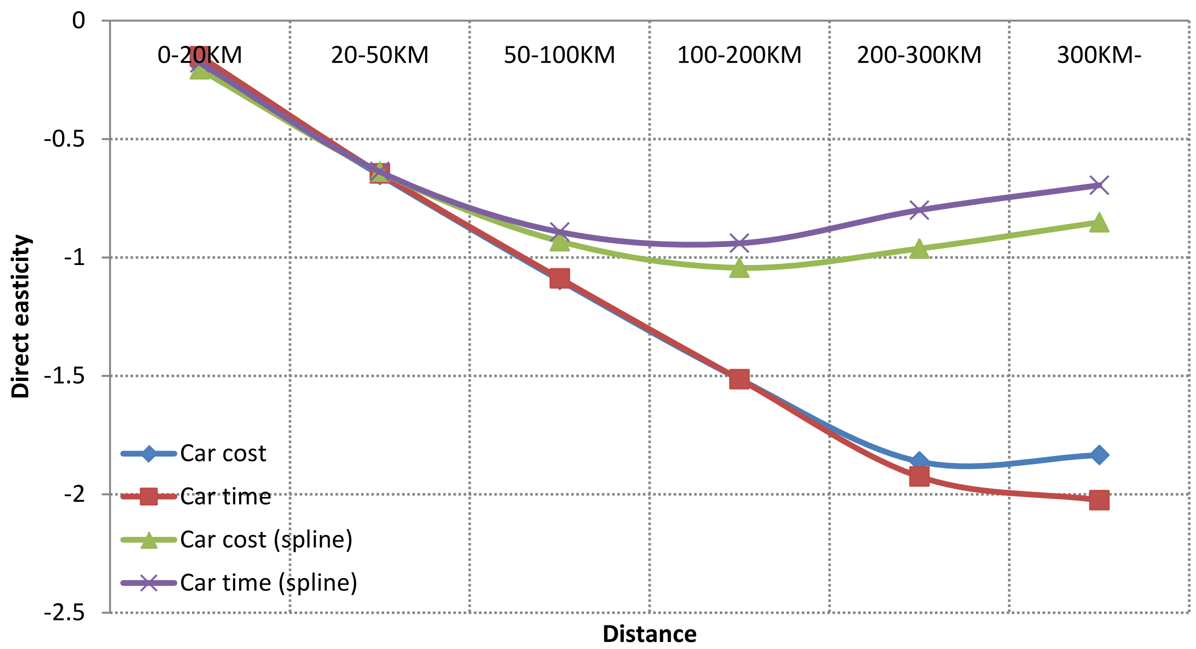

Figure 4: Elasticity curve for car cost and time for leisure travel for spline model versus normal model with $c_1 = \infty$ and $c_2 = \infty$.

As seen there is a significant difference, not only in the tail of the distribution but also on semi-long distances. Even the mean is relative different.

For the most distant destinations above 300 KM the elasticity curve is actually declining. This can happen as a result of mode-shifts but it can also happen because of shifts in the destination choice. The above simulations are based on a cost increase of 10% and this will generally cause the average distance to decline. Hence, there will be a move from right to left. If there are very few trips for some of the longest distance intervals the share of these will tend to go down on the expense of shorter trips and this movement will result in a backward bending curve.

4 Conclusion

The paper present a new function class based on a restricted non-linear spline functions. The function class is particularly suited for transport models with a long to semi-long destination domain as it is RUM consistent and allow for flexible cost-damping patterns. The damping can range from lightly damped on certain parts of the curve to super-damped on other parts. This feature is not supported by know functional forms such as Box-Cox.

We illustrate the potential of the function class by applying it to the New Danish National Model for which it has been implemented. The new function class result in massive goodness-of-fit improvement for certain trip segments (leisure and shopping) and the elasticity curve are shown to be very different.

References

Ben-Akiva M, Lerman S (1985) Discrete Choice Analysis: Theory and Application to Travel Demand. MIT Press, Cambridge

Ben-Akiva M, Daly AJ, Gunn HF (1987) Destination Choice Models: Design and Appraisal. Proceedings of the 10th PTRC Summer Annual Meeting, University of Warwick

Daly A (2010) Cost damping in travel demand models. Technical report, Rand Europe, see http://www.rand.org/pubs/technical_reports/TR717.html (accessed 15-12-12)

Gaudry M (2010) Quebec-Windsor Corridor High Speed Rail Market Forecast Profiles in Context: Level-of-Service Response Curvature Sensitivity and Attitude to Risk or to Distance in Forty Logit Core Model Applications of the Law of Demand, Agora Jules Dupuit, AJD-127, www.e-ajd.org

Gaudry M, Jara-Diaz S, Ortuzar JD (1989) Value of time sensitivity to model specification. Transportation Research B 23(2):151-158

Gaudry M, Wills M (1978) Estimating the Functional Form of Travel Demand Models. Trans. Research 12, pp. 257-289

Hensher D, Johnson L (1981) Applied Discrete Choice Modelling. Croom Helm, London

Lapparent M, de Palma A (2002) Nonlinearities in the Valuations of Time Estimates. Proceedings of the 30th PTRC Annual Meeting, PTRC, London

Mandel B, Gaudry M, Rothengatter W (1994) Linear or Nonlinear Utility Function in Logit Models? The Impact on German High-Speed Rail Demand Forecast. Transportation Research B, 28(2):91-101

Rich, J., Mabit, S.L. (2015). Cost damping and functional form in transport models. *Transportation*. forthcoming.