MEGI

Alternative Smoothing strategies in Smooth Partial Least Squares Path Modelling

Tiago Guia Ribeiro Lopes

Dissertation presented as partial requirement for obtaining the Master's degree in Statistics and Information Management

2020

SMOOTHING STRATEGIES IN SMOOTH PARTIAL LEAST SQUARES PATH MODELLING

by

Tiago Guia Ribeiro Lopes

Dissertation presented as partial requirement for obtaining the Master's degree in Statistics and Information Management, Specialization in Marketing Research and CRM

Advisor: Jorge Morais Mendes

ABSTRACT

The assessment of nonlinear relationships in the context of Partial Least Squares Path Modelling (PLS-PM) has received a growing interest in recent years. One important contribution to this subject has been the work of Henseler, Fassot, Dijkstra and Wilson (2012) on the analysis of four different approaches to quadratic effects. The Smooth Partial Least Squares (PLSs) estimation technique studied in this work removes any assumptions on the structure of the nonlinear relationships between latent variables, by applying smoothing spline techniques to the structural model. Performance results of the PLSs show that it is a powerful tool in the context of predictive research, for instance to support the definition of targeted policies. Building from the hybrid approach to the PLS algorithm introduced by Wold (1982), we compare the performance of alternative spline designs, including natural cubic splines, P-Splines and Thin Plate Regression Splines (TPRS). For this purpose, Monte-Carlo simulations are carried with a conceptual model drawn from a comprehensive set of nonlinear relationships, in different sample sizes. All model configurations are compared using Root Mean Squared Error (RMSE) and absolute bias results. The benchmarking exercise shows that, in most contexts, P-Splines perform slightly better than TPRS and natural cubic splines.

KEYWORDS

ACKNOWLEDGEMENTS

To Professor Doctor Jorge Mendes, for his precious suggestions and inputs and his endless patience and availability along the preparation of this dissertation.

To my family, for all the loving support throughout this journey.

INDEX

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS AND ACRONYMS

CB-SEM   Covariance Based Structural Equation Modelling

ECSI   European Consumer Satisfaction Index

GAM   Generalized Additive Models

PLS   Partial Least Squares

PLS-PM   Partial Least Squares – Path Modelling

PLSc   Consistent Partial Least Squares

PLSs   Smooth Partial Least Squares

RMSE   Root Mean Squared Error

SEM   Structural Equation Modelling

TPRS   Thin Plate Regression Splines

1. INTRODUCTION

Partial Least Squares Path Modeling (PLS-PM) was introduced by Wold (1975 and 1982) - later extended by Lohmöller (1989) - and is part of a class of multivariate techniques, defined as Structural Equation Modeling, that combine factor analysis and regression techniques to allow for the examination of relationships among measured variables and latent variables and between latent variables (Hair et al, 2014).

As summarized by Hair, Sarstedt, Ringle and Mena (2011), PLS-PM is a technique that maximizes the explained variance of endogenous latent variables by estimating partial model relationships in an iterative sequence of ordinary least squares (OLS) regressions.

PLS-PM was initially underused in detriment of the more popular covariance-based techniques (CB-SEM, e.g. Steenkamp and Baumgartner, 2000), and described as a technique with some but not all of the abilities of structural modelling (Henseler, Hubona, & Ash Ray, 2016).

This may have resulted, in part, from the original shortcomings of the model. As Dijkstra and Schermelleh-Engel (2014) describe, PLS tends to overestimate the loadings in absolute value, and to underestimate multiple and bivariate (absolute) correlations between the latent variables. This is a consequence of an inherent feature of PLS-PM: the (iterative) construction of linear compounds of indicator variables as proxies or stand-ins for the latent variables, and the use of estimated relationships between the linear compounds as estimates for the relationships between the latent variables; however, relationships between the former can never replicate those between the latter, apart from sets of measure zero in the parameter space.

Notwithstanding, PLS-PM has always been identified as a relevant alternative estimation technique given its prediction abilities, flexibility regarding the relaxation of the assumption of multivariate normality needed for maximum likelihood–based Structural Equation Models (SEM) estimations (e.g. Dijkstra, 2010) and low requirements in terms of sample size (Reinartz, Haenlein and Henseler, 2009).

Recent developments, which have provided the tools to elevate PLS-PM to a full-fledged structural equation modeling approach (Henseler, Hubona, & Ray, 2016), allowed this technique to overcome initial concerns and contributed for a considerable growth in the usage of these models in different fields.

Focus given on establishing a framework for the correct implementation of the model in empirical studies and developments in software and hardware for model estimation have also had a significant contribution to a greater implementation of this technique.

Hair et al. (2011) provide an extensive list of applications of PLS-PM in Marketing Research and Henseler et al. (2016) discuss the growing interest and usage of this modeling technique in the context of new technology adoption. Both authors support the analysis of the adoption of PLS-PM in specific fields with interesting reflections on the progress made in research.

In fact, Hair et al. (2011) cite some of the important developments implemented in PLS-PM in recent years: (1) confirmatory tetrad analysis for PLS-PM to empirically test a construct's measurement mode; (2) impact-performance matrix analysis; (3) response-based segmentation

techniques, such as finite mixture partial least squares; (4) guidelines for analyzing moderating effects; (5) non-linear effects; and (6) hierarchical component models.

More recently, Benitez, Henseler, Castillo and Schuberth (2020) incorporate the new knowledge and contrast classical views of PLS-PM modelling with a modern view, more focused in detailed performance analysis and specification testing of the model.

The concept of nonlinearity is found in PLS-PM almost since its inception. In fact, as early as the seminal work of Wold (1982), the concept of nonlinearity was already being introduced as a feature of the model when specifying relationships between variables of the structural model.

Notwithstanding, and despite evidences of the existence of nonlinear relationships between latent variables (e.g. Tuu, & Olsen, 2010), nonlinear approaches in multiple regression settings were pretty much underdeveloped for decades, even in what regards moderation effects, by means of interaction or curvilinearity, as studied by Cortina (1993).

Moreover, early examples of nonlinear approaches to SEM are mostly devoted to CB-SEM (Kenny & Judd, 1984, Klein, & Moosbrugger, 2000; Marsh, Wen, & Hau, 2004; Klein, & Muthén, 2007; Pek, Sterba, Kok, & Bauer, 2009; Kelava, Schermelleh-Engel, Moosbrugger, Zapf, Ma, Cham, Aiken, & West, (2011) or Bayesian contexts (e.g. Song, & Lu, 2010; Song, Lu, Kai, & Ip, 2013).

It wasn't until the first decade of the 2010s that these phenomena became an object of more focused and systematic investigation in PLS-PM and became a more prominent feature of PLS-PM estimating tools.

In recent years, the better understanding of non-linear effects has been an important part on the assessment improvements within PLS-PM, as significant recent research has been devoted to analyzing and incorporating these phenomena in modeling PLS-PM – despite some reservations on the actual implementation in empirical research in specific contexts, due to parsimony principles or lack of theoretical support (Hair, Ringle, & Starstedt, 2013).

Most notably, Henseler and Chin (2010) and Henseler et al. (2012), in contexts of estimating interaction and quadratic effects, identified and worked on four different approaches to these issues: (1) product indicator approach (Chin, Marcolin, & Newsted, 2003), (2) 2-stage approach (Chin et al., 2003; Henseler & Fassott, 2010), (3) hybrid approach (Wold, 1982), and (4) orthogonalizing approach (Henseler et al., 2010; Little et al., 2006).

The Smooth Partial Least Squares (PLSs) path modelling technique presented in this study addresses nonlinearity in direct relationships of the structural model, building upon the extensive work that has been developed in applying generalized additive models (GAM) to test nonlinear relationships between observed variables, based on the assumption that these relationships are not less common in the case of latent variables. Wood (2017) provides an extensive description of GAM models and associated smoothing techniques.

The proposed approach has the advantage of not limiting nonlinearity to quadratic or product-interaction terms, allowing for the assumption that the nonlinear trend is not known. In that sense, that study is groundbreaking and separated from the rest of the field, since no prior research in PLS-PM had accounted for nonlinear relationships without assuming the nature of its relationships.

In fact, even though the usage of splines in the context of the inner model in PLS-PM had been used before in practical implementations (see Jakobowicks, 2007), this is the first to address in a more general and theoretical approach.

Notwithstanding, further work lies ahead to establish PLSs as a mainstream tool to identify nonlinearities and estimate latent variable scores in the context of PLS-PM.

Particularly, for empirical purposes, it is relevant to test the model with other types of smoothers and smoother parametrizations. This dissertation draws from the groundwork laid by Mendes et al. (2018) and aims to help identifying which are the better performing smoothers in PLSs context and smoother configurations.

For this purpose, different smoothers are analyzed, based on the alternatives presented by Wood (2006). This part of the analysis should allow for the identification of the most adequate smoothers in the context of PLS-PM estimation.

To study the best smoothing choices, Monte-Carlo simulations are carried with different smoothers and parametrization strategies. The choice of smoothers and knot configuration will be based in the theoretical and practical ground laid by Wood (2006). The conclusions of the study are drawn from simulations in adequate theoretical scenarios to allow for sound generalizations and empirical application.

The software implementation and simulation plan, including the population design, will follow the methodology carried by Mendes et al. (2018) in what regards simulated population dimension, methodologies adopted for observation generation and analyzed random sample sizes

The underlying conceptual model follows the functional forms used by Bauer, Baldasaro and Gottfredson (2012), to incorporate a wide span of nonlinear relationships, with an added true linear relationship to assess to what extent the PLSs method can capture linear relationships.

Model performance is compared among all the approaches using the absolute bias and root mean square error (RMSE).

In the following paragraphs we describe the PLS model and some of its extensions (1.1). We also introduce penalized smoothing in the context of PLS (1.2), elaborate on the justification for developing the PLSs estimation methodology (1.3) and present the 3 alternative smoothing splines techniques used in our analysis. In the second part of the document (2.), the procedure to test the performance of the different smoothing splines is described (2.1) before the presentation of the results (2.2). Wrapping up and suggestions for future work are included in the last two sections (3. And 4.).

1.1. PLS AND SOME EXTENSIONS

Hair et al (2014) describe the PLS-PM model as a second-generation multivariate statistic modelling technique, being part of the family of Structural Equation Modelling (SEM) methods along with CB-SEM (Covariance-Based SEM). PLS-PM distinguishes from CB-SEM mainly by its primary focus on explaining the variance in the dependent variables when analyzing the model.

PLS-PM models are designed with two components: the structural model and the measurement models.

The structural model establishes the relationships between latent variables that are not directly measured, whereas the measurement models define the relationship between those constructs and the indicators, which are the variables that are directly measured.

The constructs may be exogenous or endogenous, whether they are, respectively, independent or dependent of other constructs in the model.

Figure 1 illustrates graphically a simple path model with both reflective and formative models. In this model, the ξ_3 and ξ_4 constructs are endogenous variables of the structural model and ξ_1 and are its ξ_2 exogenous constructs. All constructs are estimated based on reflective models except for ξ_4.
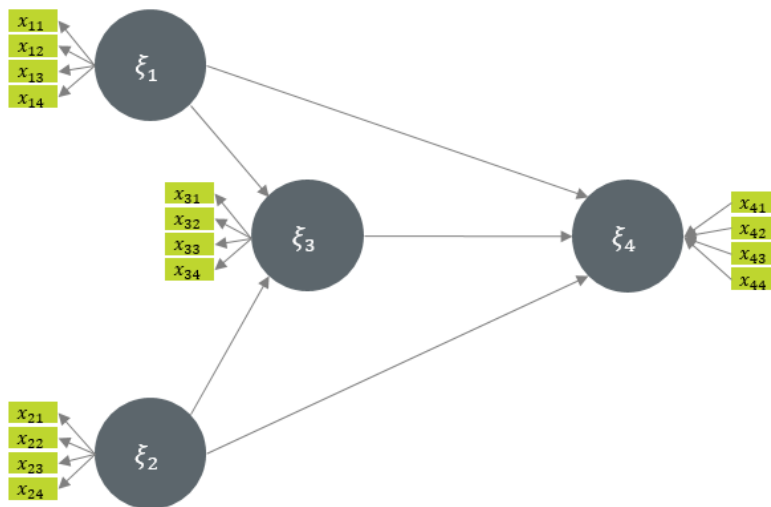
Figure 1 – Example of a simple path model

The graphical connections in the structural model define linear relationships between the constructs, which can be translated in the following equations

$$\xi_3 = \beta_{11}\xi_1 + \beta_{21}\xi_2$$

$$\xi_4 = \beta_{12}\xi_1 + \beta_{22}\xi_2 + \beta_{32}\xi_3$$

More generally, the equations of the structural model are given by:

$$\xi_i = \beta_{i0} + \sum_k \beta_{ik}\xi_k + \nu_i$$

Also, the measurement models may be formative or reflective, where formative models are linear regression relationships being the construct the dependent variable, and reflective models are models where it is assumed that the construct impacts the results of the indicator variables.

In the reflective measurement model, each indicator variable is related to the constructs by means of a simple regression:

$$x_{ij} = \varpi_{ij0} + \varpi_{ij}\xi_i + \varepsilon_{ij},$$

where ξ_i has mean m and standard deviation 1, where the only hypothesis is that ε_{ij} has zero mean and is uncorrelated with its respective construct (Tenenhaus, Vinzi, Chatelin, & Lauro, 2005), such that:

$$E(x_{ij}|\xi_i) = \varpi_{ij0} + \varpi_{ij}\xi_i$$

In the formative measurement model, the construct is generated by its indicators by means of a linear function:

$$\xi_i = \sum_j \pi_{ij}x_{ij} + \delta_i$$

Where the residual vector δ_i has a zero mean and is uncorrelated with the indicators. Hence,

$$E(\xi_i|x_{i1}, \dots, x_{ip_i}) = \sum_j \pi_{ij}x_{ij} + \delta_i$$

Tennenhaus et al. (2005) provide a detailed description of the PLS-PM methodology, including estimation techniques and statistical properties. Henseler, Ringle and Sinkovics (2009) draw from this description and sum it up has follows:

- *Step 1: Outer approximation of the latent variable scores. Outer proxies of the latent variables, $\hat{\xi}_n^{outer}$, are calculated as linear combinations of their respective indicators. These outer proxies are standardized; i.e. they have a mean of 0 and a standard deviation of 1. The weights of the linear combinations result from step 4 of the previous iteration. When the algorithm is initialized, and no weights are available yet, any arbitrary nontrivial linear combination of indicators can serve as an outer proxy of a latent variable;*
- *Step 2: Estimation of the inner weights. Inner weights are calculated for each latent variable in order to reflect how strongly the other latent variables are connected to it. There are three schemes available for determining the inner weights;*
- *Step 3: Inner approximation of the latent variable scores. Inner proxies of the latent variables, $\hat{\xi}_n^{inner}$, are calculated as linear combinations of the outer proxies of their respective adjacent latent variables, using the aforedetermined inner weights. Step 4: Estimation of the outer weights. The outer weights are calculated either as the covariances between the inner proxy of each latent variable and its indicators (in Mode A, reflective), or as the regression weights resulting from the ordinary least squares regression of the inner proxy of each latent variable on its indicators (in Mode B, formative).*
- *These four steps are repeated until the change in outer weights between two iterations drops below a predefined limit. The algorithm terminates after step 1, delivering latent variable scores for all latent variables. Loadings and inner regression coefficients are then*

calculated in a straightforward way, given the constructed indices and using the formative and reflective model equations. *In order to determine the path coefficients, for each endogenous latent variable a (multiple) linear regression is conducted.*

Aside from the linearity constraints addressed by the works of Henseler et al. (2010), Henseler et al. (2012) or Mendes et al. (2018), other problematic concerns have been raised regarding by the original PLS algorithm over time, namely:

- Inconsistency, translated in overestimation of the absolute value of loadings and underestimation of multiple and bivariate(absolute) correlations between the latent variables;
- Lack of an overall goodness of fit measure;
- Limitations addressing multicollinearity;
- Limitations addressing endogeneity;

However, recent developments in research have addressed these concerns, consolidating the validity of PLS-PM approach.

Dijkstra and Henseler (2015a) presented an improved estimation technique (Consistent PLS – PLSc) that addresses the shortcomings of the original PLS-PM model and leads to consistent and asymptotically normal estimators for the loadings and for the correlations between the latent variables, , by correcting for attenuation of the construct scores correlations with a new and consistent reliability coefficient for PLS. This model has had widespread usage and has become a reference in the estimation of models with reflective indicators (Henseler, Hubona, & Ray, 2016).

Also, Dijkstra and Henseler (2015b) have shown that the overall model can be assessed in two non-exclusive ways, by assessing the differences between the empirical and model implied indicator variance-covariance matrix.

Regarding multicollinearity, Jung and Park (2018) proposed the regularized PLSc, which incorporates a ridge type of regularization in the PLSc, an approach that was shown to have interesting results in contexts of strong multicollinearity.

Endogeneity as also been addressed, either by replacing OLS for 2SLS in the structural model (Benitez, Henseler, & Roldán, 2016) or alternative approaches avoiding instrumental variables, which include the Gaussian copula approach (Hult, Hair, Proksch, Sarstedt, Pinkwart, & Ringle, 2018).

Other approaches have been carried to cement the PLS-PM as a solid estimation technique, or a *silver bullet* as Hair, Ringle and Sarsted (2011) have classified it.

These works include, namely, a criterion to assess discriminant validity (Henseler, Ringle, & Sarstedt, 2015) designated Heterotrait-Monotrait ratio of correlations (HTMT), new approaches for estimating and testing second-order constructs (see Van Riel, Henseler, Kemeny, & Sasovova, 2017) and multigroup analysis to address groups with different behavior within samples (Chin & Dibbern, 2010; Sarstedt, Henseler, & Ringle, 2011).

1.2. PENALIZED SMOOTHING IN THE CONTEXT OF THE PLS

The research developed by Mendes et al. (2018) relaxes the restriction of linearity of the relationships between constructs by allowing the existence of non-linear relationships. By doing so, it proposes an alternative approach by means of *penalization*.

In this approach the Least Squares objective function is replaced by a penalized version of the traditional objective function. As a way of simplification in the context of one covariate, the objective function will be:

$$\sum_{i=1}^{n}\{y_i - f(x_i)\}^2 + \lambda \int \{f''(u)\}^2 du \ (1)$$

where

$$y_i = f(x_i) + \varepsilon_i$$

This becomes the traditional setting of a regression with smoothing spline when $f(x_i)$ is composed a set of pre-defined basis functions such that:

$$f(x_i) = \sum_{i=1}^{m} \beta_j h_j(x)$$

where the different $h_j(x)$ are piecewise polynomials joined together to compose a single smooth curve. These functions are bounded by points defined as knots of the spline. In the next chapter three alternative sets of basis functions are detailed.

In fact, this setting is based on the framework developed by Reinsch (1967) - which has served as a main reference in fitting non-linear relationships - adapted to a regression context more suited to statistical inference, where the number of functions is not the number of observations, but instead is restricted to a predefined limited number of functions.

The objective function depicted in (1) may be decomposed in two components, where $\sum_{i=1}^{n}\{y_i - f(x_i)\}^2$ minimizes the Euclidean distance, while the second term $\lambda \int \{f''(u)\}^2 du$ penalizes overfitting. This second term is 0 if f is linear. Also, λ is a smoothing parameter that controls the trade-off between the two terms, by adding a penalty to the residual sum of squares.

Applying this to a SEM context, y_i is an endogenous latent variable and x_i may be either an endogenous or an exogenous variable.

Setting matrix H, where $H_{ij} = h_j(x)$, the spline objective function comes as

$$(-H\beta)^T(-H\beta) + n\lambda\beta^T\Omega\beta$$

Where Ω is a matrix of known coefficients comprising the elements of the second derivatives of the basis functions:

$$\Omega_{ij} = \int dx h_i''(x)h_j''(x)$$

Mendes et al. (2018) show that differentiating with respect to β, the estimator of β is found through the equation:

$$\hat{\beta} = (H^T H + n\lambda\Omega)^{-1} H^T y$$

The added covariance given by $n\lambda\Omega$ depends not only of the shape of the functions, but also the desired amount of smoothing, given by the λ. The largest the λ, less is the weight of the actual data in the fit.

With this configuration, a piecewise continuous linear model is constructed, with the portioning of the range of x into K+1 intervals, by choosing K points, called knots, which may be uniformly distributed across x or corresponding to specific quantiles.

This means that each individual segment basis functions are summed to obtain a composite function $f(x)$, which is a cubic spline of order K degrees of freedom.

Wood (2017) provides detailed descriptions of how to estimate models based on natural cubic splines, including the cross-validation methodologies to calculate the optimal λ (see Wood, 2017, page 255, for a detailed description of cross validation techniques) and strategies to define the number and positioning of the knots. These authors also develop the algorithm in the context of more than one covariate.

The PLSs algorithm

The PLSs algorithm description presented below reproduces (in italic) the PLS algorithm hybrid approach developed by Wold (1982), as presented by Tennenhaus et al. (2005) with the modifications introduced by Henseler et al. (2012) and Mendes et al. (2018):

- *Step 1: Calculating outer proxies of latent variable scores: outer proxies of the latent variables $\hat{\xi}_j^{outer}$ are calculated as linear combinations of their respective indicators. The weights of the linear combinations result from step 4 of the previous iteration or are manually initialized. For each non-linear term, a new proxy is created as the element-wise transformation of the respective outer estimates.*

  The latent variable scores are estimated as a weighted sum of their respective indicators:

  $$\hat{Y}^{outer} = X\hat{W}$$

  Where W is the diagonal GxG matrix of the outer weights.

  The outer proxies of the latent variable scores are initialised by setting the weights of the previous equation to 1.

  Assuming all the MVs, $X_1, X_2, \cdots, X_K$ are scaled (mean(X_i) = 0 and Var(X_i) = 1), the latent variables are also centered (mean=0) but must be scaled to have unit variance.

  Assuming the nonlinear relationship between latent variables are approximated by smoothing splines with K-1 knots (as the constant terms vanish because we are dealing

with LVs that have been scaled to mean zero and unit variance) the $n \times G$ matrix $\hat{Y}$ is augmented to a $n \times G(K-1)$ matrix, where the ith line is given by:

$$\hat{Y}_i^{outer} = [y_{i1}, h(y_{i1}), h(y_{i1}, y_{11}^*), \cdots, h(y_{i1}, y_{1(K-1)}^*), y_{i2}, h(y_{i2}), h(y_{i2}, y_{21}^*), \cdots,$$

$$h(y_{i2}, y_{2(K-1)}^*), y_{iG}, h(y_{iG}), h(y_{iG}, y_{1G}^*), \cdots, h(y_{iG}, y_{G(K-1)}^*)]$$

where the first and second indexes of $y_{..}$ refer to the observation and to the LV, and the first and second indexes of $y_{..}^*$ refer to the LV variable and the fixed knot (for simplicity we dropped the superscript "outer"). Moreover, the columns of matrix $\hat{Y}$ should be scaled to mean zero and unit variance.

- *Step 2: Estimation of the inner weights: Inner weights are calculated for each latent variable in order to reflect how strongly the other latent variables are connected to it. There are three schemes available for determining the inner weights. Wold (1982) originally proposed the centroid scheme. Later, Lohmöller (1989) developed the factor weighting and path weighting schemes. The centroid scheme uses the sign of the correlations between a latent variable or, more precisely, the outer proxy and its adjacent latent variables; the factor weighting scheme uses the correlations. The path weighting scheme pays tribute to the arrow orientations in the path model. The weights of those latent variables that explain the focal latent variable are set to the regression coefficients stemming from a regression of the focal latent variable (regressant) on its latent regressor variables. The weights of those latent variables, which are explained by the focal latent variable, are determined in a similar manner as in the factor weighting scheme. Regardless of the weighting scheme, a weight of zero is assigned to all nonadjacent latent variables.*

  The inner weights are also determined for each nonlinear term of the splines with K -1 knots.

  An illustration of characterization the elements of the augmented matrix of the inner weights E^{Aug}, by means of the centroid method, may be found in Mendes et al. (2018).

- *Step 3: Inner approximation of the latent variable scores: Inner proxies of the latent variables, $\hat{\xi}_j^{inner}$, are calculated as linear combinations of the outer proxies of their respective adjacent latent variables, using the afore-determined inner weights. Each term of the splines with K -1 knots is used to estimate the inner proxies of the latent variables.*

$$\hat{Y}^{inner} = \hat{Y}^{inner} E^{Aug}$$

$$\hat{y}_j^{inner} = \frac{\hat{y}_j^{inner}}{\sqrt{Var(\hat{y}_j^{inner})}}, j = 1, \cdots, G,$$

  Where E^{Aug} is the augmented matrix as determined in Step 2.

- *Step 4: Estimation of the outer weights: The outer weights are calculated either as the covariances between the inner proxy of each latent variable and its indicators (in Mode A), or as the regression weights resulting from the ordinary least squares regression of the inner proxy of each latent variable on its indicators (in Mode B).*

No additional procedure is required in this step because the non-linear terms do not have any assigned indicators, as determined by the hybrid approach.

Mode A: Multivariate regression coefficient with the block of manifest variables as response and the latent variable as the regressor:

$$\begin{aligned} \widehat{w}_j^T &= \left(\widehat{y}_j^{inner^T} \widehat{y}_j^{inner} \right)^{-1} \widehat{y}_j^{inner^T} X_j \\ &= Cor(\widehat{y}_j^{inner}, X_j) \end{aligned}$$

Mode B: Multiple regression coefficient with the latent variable as response and its block of manifest variables as regressors:

$$\begin{aligned} \widehat{w}_j^T &= \left(X_j^T X_j \right)^{-1} X_j^T \widehat{y}_j^{inner} \\ &= Var(X_j)^{-1} Cor(X_j, \widehat{y}_j^{inner}) \end{aligned}$$

These steps are iterated until the change in outer weights between two consecutive iterations falls below a predefined tolerance,

$$max\left(\left| \frac{\widehat{w}_{kj}^i - \widehat{w}_{kj}^{(i+1)}}{\widehat{w}_{kj}^{(i+1)}} \right| \right) < tolerance, K = 1, \cdots, K, j = 1, \cdots, G$$

The algorithm terminates after step 1, delivering latent variable scores for all latent variables. The loadings and inner regression coefficients are then calculated in a straightforward way, given the constructed indices. The structural relationships are obtained by estimating the additive model using a smoothing spline. For each latent variable $y_j, j = 1, \cdots, G$, we have the following model:

$$y_{ij} = \sum_{l=1}^{\#pred} \beta_1^{(l)}(y_{il}) + \sum_{m=1}^{K-1} \beta_{m+1}^{(l)} h_{m+1}(y_{il}) + \epsilon_{ij}$$

Where y_{ij} is an endogenous latent variable, the summation $\sum_{l=1}^{\#pred}$ means we are regressing LV j on the set of all predecessor LVs, h is a natural cupic spline and ϵ_{ij} are i.i.d. $N(0, \sigma_\epsilon^2)$ random variables. The estimated factor scores (including the nonlinear terms) of the predecessor set of latent variables of y_j are used for this purpose.

In practice the model may be estimated using partial least squares with the *mgcv* library available through the R Project software (R Core Team, 2018).

Mendes et al. (2018) compare the performance of the PLSs with the PLS and PLSc models and shows that this model performs uniformly better, using absolute bias and Root Mean Squared Error (RMSE) measurements, in a simulated model where structural relationships have underlying real nonlinear functional forms, as proposed by Bauer et al. (2012).

1.3. JUSTIFICATION FOR PLSs

Henseler (2018) reflects on the applications PLS-PM, arguing that this estimation technique is suitable for all main lines of research, namely causal (including confirmatory and explanatory research), exploratory, descriptive and predictive research.

As what regards nonlinearity, Hair et al. (2013) state that nonlinear approaches to PLS-PM must be carried carefully, as results in one specific experiment are not easily replicable and, hence, not generalizable. Also, these approaches are usually more demanding in terms of interpretability. In most contexts, these authors recommend more parsimonious approaches, yielding, in most cases, similar and more easily interpretable results. This is especially the case for causal and experimental research.

As such, the application of PLSs for causal and experimental research is not straightforward and its usage should be taken cautiously. Nevertheless, in these contexts, it shouldn't, at all, be discarded as an important auxiliary tool, for characterization and visual analysis of results.

On the other hand, PLSs appears as a powerful tool for predictive analysis and could become one of the go to tools in specific contexts, such as strategic management, where the main objective is to accurately predict the behavior of out of sample elements and target segment specific policies.

In fact, the good performance of this method when compared with PLS and PLSc, as shown be Mendes et al. (2018), regarding absolute bias and RMSE metrics, gives indications that PLSs may be the most suitable to more accurately predict individual behaviors. Also, it shares these models' properties of transparency in how the prediction is produced, which is a plus when compared with alternative prediction tools (Henseler, 2018).

As Mendes et al. (2018) show in a practical example, visual inspection of the results of PLSs modelling is an important instrument to detect nonlinearity.

Also, the analysis of first, and even second derivatives of the results, provide valuable information about the elements of the population where specific policies may be more impactful in the presence of nonlinearity, allowing for their efficient targeting.

To briefly illustrate this potential, we draw from the application presented by Mendes et at. (2018).

These authors use a real data set produced in the context of the European Customer Satisfaction Index to investigate nonlinear relationships between satisfaction and loyalty and between perceived quality and satisfaction latent variables.

The structural equation model is the reduced ECSI model presented below.
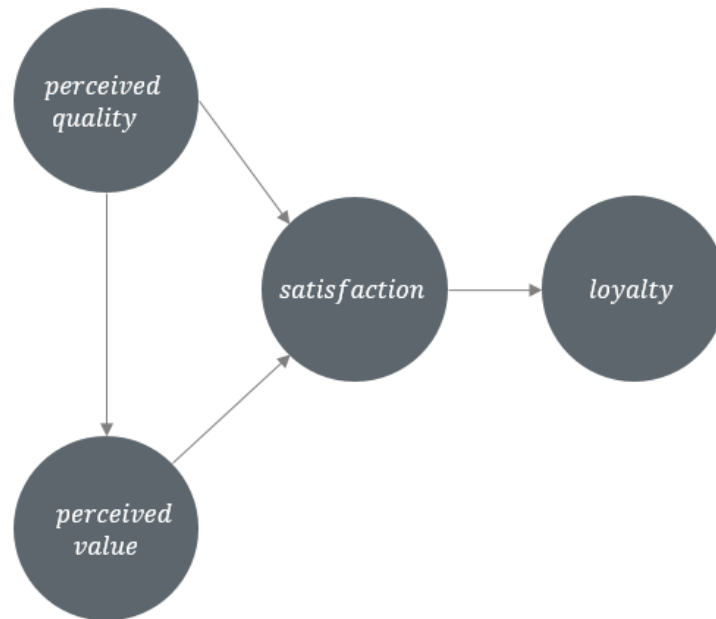
Figure 2 – Structural Equation Model of the reduced ECSI

In the table below we also reproduce the indicators of the measurement model of the analysis. All the included indicators are measured on a 10-point scale, from 1 to 10, where 1 traduces a very negative opinion and 10 a very good opinion. Cronbach's alpha (α) and Dillon-Goldstein's rho (ρ) unidimensionality measures are presented for each group of indicators.

Latent variables	Manifest variables
Perceived quality $\alpha = 0.93$; $\rho = 0.94$	(a) Overall perceived quality (b) Quality of products and services (c) Customer service and personal advice (d) Availability of contact channels (e) Reliability of products and services (f) Diversity of products and services (g) Clarity and transparency of information provided (h) Accessibility (i) Quality of physical facilities
Perceived value $\alpha = 0.91$; $\rho = 0.95$	(a) Evaluation of prices given quality (b) Evaluation of quality given prices
Customer satisfaction $\alpha = 0.84$; $\rho = 0.91$	(a) Overall satisfaction (b) Fulfillment of expectations (c) Distance to ideal company
Customer loyalty $\alpha = 0.89$; $\rho = 0.95$	(a) Intention to remain customer (b) Recommendation to friends and colleagues

Table 1 – Measurement model of the reduced ECSI.

The authors show that the PLSs method behaves slightly better than the PLSc and significantly better than the traditional PLS approach, in that regards R-squared values.

The study presents the estimated relationships of the latent variables graphically. In particular, the visual inspection of the data and estimation results shows the ability of the PLSs to capture

apparent non-negligible nonlinearities in the case of the direct relationship between satisfaction and loyalty. Figure 3 illustrates these conclusions, whereby the PLSs fitting is represented by the continuous line.

The results presented by Mendes et al. (2018) are interesting and show the potential of the model for further graphical analysis.
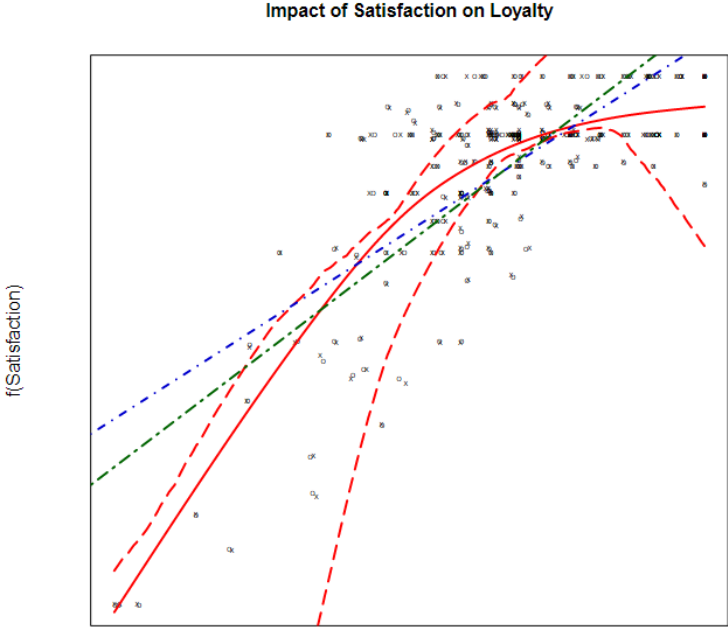
Figure 3 – Estimated direct relationships between Satisfaction and Loyalty in the reduced ECSI model. Circles (○) represent the traditional PLS estimated factor scores. Crosses (x) represent Smooth PLS estimated factor scores. The (non)-linear relationship estimated using Smooth PLS is represented by the red solid line. PLS and PLSc linear relationships are represented by the blue dashed line and green long-dashed lines.

Going somewhat further, below we present a depiction of the derivatives of the estimated direct relationships between satisfaction and loyalty, alongside the estimated curves produced by Mendes et al. (2018).

The concave form of the direct relationship is evidently translated in a downward line of the derivatives. In fact, this illustrates that targeted satisfaction-based policies applied toward loyalty increase are likely to have a much stronger impact in elements of the population where satisfaction is lower than in those clients that already show high levels of satisfaction.
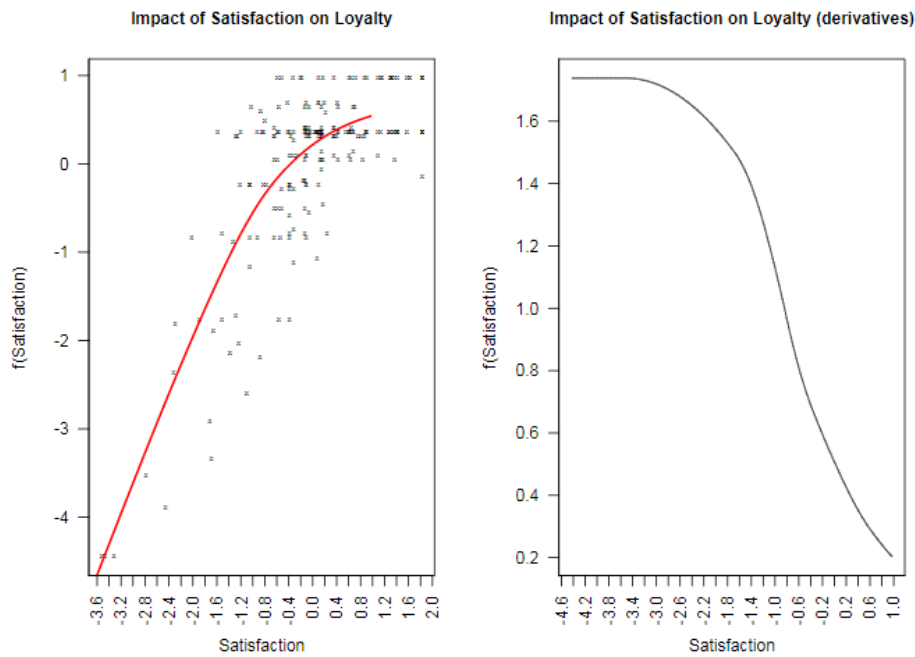
Figure 4 – Estimated direct relationships between Satisfaction and Loyalty and their derivatives in the reduced ECSI model

In instances where nonlinear direct relationships between variables show inflections in the rate of variation (from concave to convex or vice-versa), inspection of second derivatives should also prove useful to detect the respective inflection points.

Surely, consumer behavior analysis for targeted policies definition shouldn't forego other sources of information regarding satisfaction and loyalty (e.g. regarding client value), but in marketing research and strategic management perspectives the PLSs produces valuable auxiliary and complementary insights towards efficiently addressing and predicting consumer behavior.

1.4. ALTERNATIVE SMOOTHING SPLINES

As stated by Wood (2017), penalised regression splines are low rank smoothing splines, that are designed to provide efficient compromise between retaining the good properties of splines and computational efficiency.

In this section, we present a brief description of the three alternative smoothing methods based on penalized regression splines used in our analysis.

Natural cubic splines

First, we start by presenting the natural cubic regression splines used by Mendes et al. (2018) in their PLSs study, which is drawn from Wahba (1990).

The basis functions in this case are:

$$h_1(x) = 1$$

$$h_2(x) = x$$

$$h_{i+2}(x, x_1) = \frac{(x - x_i)_+^3 - (x - x_i)_+^3}{x_n - x_i} - \frac{(x - x_i)_+^3 - (x - x_i)_+^3}{x_n - x_{n-1}}$$

Where $(\cdot)_+$ is the positive portion of its argument:

$$(x - x_i)_+ = \begin{cases} x - x_i, & x - x_i \geq 0 \\ 0, & x - x_i < 0 \end{cases}$$

In a comparison analysis with other penalized regression splines, Wood (2006) states that the main advantages of natural cubic splines are the fact these are computationally cheap and have directly interpretable parameters. Its main disadvantage is the dependence on knots, which adds a degree of subjectivity in the model design.

P-Splines

P-Splines were introduced by Eillers, & Marx (1996) and are a low rank version of the B-Splines developed by Boor (2001).

Eillers et al (1996) propose the following objective function with basis splines of any order q:

$$\sum_{i=1}^{n} \left\{ y_i - \sum_{j=1}^{m} a_j B_j(x_i) \right\}^2 + \lambda \int \sum_{j=k+1}^{m} (\Delta^k a_j)^2$$

Where the B-splines basis functions of order q are defined recursively:

$$B_j^q(x) = \frac{x - x_j}{x_{j+m-+1} - x_j} B_j^{m-1}(x) + \frac{x_{j+m+2} - x}{x_{j+m+2} - x_{j+1}} B_{j+1}^{m-1}(x) \; j = 1, \ldots m$$

And

$$B_j^{-1}(x) = \begin{cases} 1 & x_i \leq x \leq x_{i+1} \\ 0 & otherwise \end{cases}$$

That is, the authors propose to base the penalty on higher order finite differences of the coefficients of adjacent B-splines. By doing this, the dimensionality of the problem is reduced from the number of observations to the number of B-splines.

The authors describe P-splines as also not being computationally expensive, including for cross-validation (indispensable for calculating λ). These splines have no boundary effects, conserve moments (means, variances) of the data, and have polynomial curve fits as limits. The main limitations of P-splines are the fact that they require equally spaced knots and the interpretation of penalties is far from being straightforward (Wood, 2006).

Thin-plate regression splines

The third type of penalized splines tested in this analysis is Thin Plate Regression Splines (TPRS). TPRS were first introduced by Wood (2002) and built from the Thin Plate Splines (Duchon, 1977).

The following paragraphs follow very closely the description and notation carried by Wood (2016)

In thin plate splines smoothing, the objective function to be minimized is:

$$\|y - f\|^2 + \lambda J_{md}(f) \ (2)$$

Where f is a function of the covariates and J is a penalty functional measuring the wiggliness of f and λ *is a smoothing parameter. The wiggliness* penalty may be written as:

$$J_{md} = \int \cdots \int_{\mathbb{R}^d} \sum_{\nu_1 + \ldots + \nu_d = m} \frac{m!}{\nu_1! \ldots \nu_d!} \left(\frac{\partial^m f}{\partial x_1^{\nu_1} \ldots \partial x_d^{\nu_d}} \right)^2 dx_1 \ldots dx_d$$

where d is the number of covariates.

Defining m such that $2m > d + 1$, function (2) may be written as:

$$\hat{f}(x) = \sum_{i=1}^{n} \delta_i \eta_{md}(\|x - x_i\|) + \sum_{j=1}^{m} \alpha_j \phi_j(x)$$

where δ and α are vectors of the coefficients to be estimated, with δ being subject to the constrains $T^T \delta = 0$, where $T_{ij} = \phi_j(x_i)$ and

$$\eta_{md}(r) = \begin{cases} \frac{(-1)^{m+1+d/2}}{2^{2m-1} \pi^{d/2}(m-1)!\,(m-d/2)!} r^{2m-d} \log r & d \ even \\ \frac{\Gamma(d/2 - m)}{2^{2m} \pi^{d/2}(m-1)!} r^{2m-d} & d \ odd \end{cases}$$

At last, by defining the matrix E by $E_{ij} \equiv \eta_{md}(\|x_i - x_j\|)$, the objective function may be written as:

$$\|y - E\delta - T\alpha\|^2 + \lambda \delta^T E \delta, \text{ subject to } T^T \delta = 0$$

As Wood (2016) states and describes, thin plate regression splines are based on the idea of truncating the space of wiggly components of the thin plate spline (the components with parameters δ), while leaving the components of zero wiggliness unchanged (the $\alpha\ components$).

The author also defends that the advantages of this approach are the absence of knots and some optimality properties, which regularly translate in better performances than the other spline methods described above. Regarding disadvantages, computational costs in large data sets is the one that may have more impact in the context of our analysis.

2. APPLICATION

2.1. DATA DESCRIPTION AND PROCEDURE

To study the best smoothing choices, Monte-Carlo simulations are carried with different smoothers and knot choice placement strategies.

The software implementation and simulation plan, including the population design, will follow the methodology carried by Mendes et al. (2018) in what regards simulated population dimension, methodologies adopted for observation generation and analyzed random sample sizes.

Hence, the underlying conceptual model will follow the functional forms used by Bauer et al. (2012), to incorporate a wide span of nonlinear relationships, with an added true linear relationship to assess to what extent the PLSs method can capture linear relationships.

In the next paragraphs we describe the steps carried to generate the simulation data, following the method implemented by Mendes et al. (2018).

The Monte Carlo procedure follows three steps:

1. Define an underlying true model.
2. Generate random data emerging from the defined model.
3. Use PLSs to estimate the models with different smoothing splines and compare the different approaches using the absolute bias and RMSE.

The defined model consists of one exogenous latent variable and four endogenous latent variables, all depicting different nonlinear relationships with the exogenous latent variable. The chosen functional forms were those used by Bauer et al. (2012). A true linear relationship is added to assess to what extent the PLSs method can capture linear relationships. Figure 5 presents all the designed relationships.
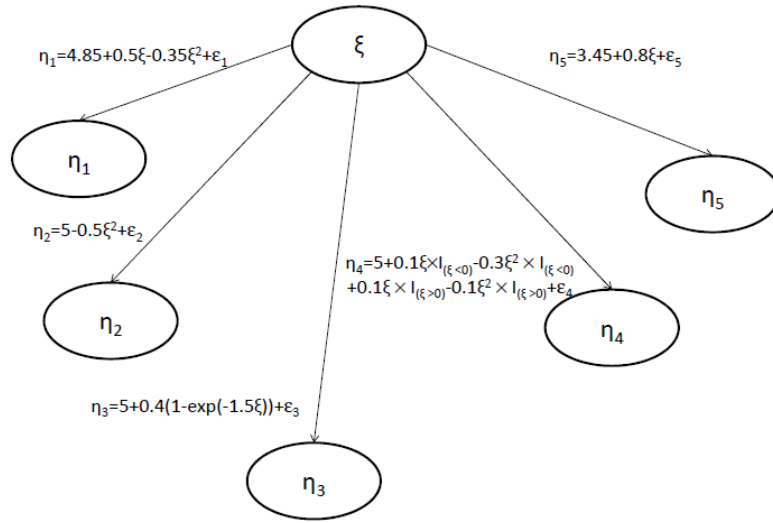
Figure 5 – Structural Equation Model implemented in the Monte Carlo simulation

where $I_{\xi<0} < 0$ is the indicator function, taking the value 1 if $\xi < 0$ and 0 otherwise, and the disturbances are sampled as follows:

$$\epsilon_j \sim N(0; 0,5); \; j \; = \; 1; \; 2; \; 3; \; 4; \; 5$$

To ensure consistency in the PLS-PM, the measurement model has a fixed number of five indicators and latent variables with unit loadings:

$$x_l^{\xi} \; = \; \xi + \delta_0, \qquad l \; = \; 1; \; 2; \; 3; \; 4; \; 5$$

$$y_l^{\eta_k} \; = \; \eta_k \; + \; \delta_k; \; k \; = \; 1; \; 2; \; 3; \; 4; \; 5; \; l \; = \; 1; \; 2; \; 3; \; 4; \; 5$$

where residual variance $\delta_{()}$, follows a normal distribution $N(\sigma_\delta^2)$ with 3 different levels of variance (3, 1 and 0,33), corresponding to communalities of 0.25, 0.50 and 0.75, respectively.

The experiment is implemented with different samples sizes, ranging from 50 to 900 observations (namely, 50, 100, 150, 250, 300, 500, 600, 750 and 900).

Following this methodology, three populations of 10 000 units, will be generated (corresponding to the three levels of residual variance).

Then, K = 1000 random samples of size n = 50; 100; 150; 250; 300; 500; 600; 750; 900, are randomly drawn from the three levels of communality.

2.2. RESULTS

This section presents the results drawn from our simulation plan.

After generating the samples, SEM models where estimated through PLSs with the three different types of splines using the "gam" function of the *mgcv* R Package (Wood, 2018).

To analyze the performance of the models, we measured the absolute bias (B) and the RMSE. The calculations of these indicators are given by:

$$B_l \approx \frac{1}{300} \sum_{i=1}^{300} |\bar{g}_i(\xi) - g_i(\xi)|$$

$$RMSE_l \approx \frac{1}{K} \sum_{K=1}^{K} \left(\frac{1}{300} \sum_{i=1}^{300} (\hat{g}_{ki}(\xi) - g_i(\xi))^2 \right)$$

Where the index l=1,2,3 represents the different sample sizes, and $\bar{g}_k(\xi) = \frac{1}{1000} \sum_{k=1}^{1000} \hat{g}_{ki}(\xi)$ is the mean estimated functional relationship between ξ and $\eta(.)$ evaluated at a fixed, even grid of $i = 1, \ldots, 300$ points in the range of ξ, and $g(\xi)$ is the known functional relationship (see Mendes et al, 2018).

Regarding the parametrization of the splines, the following configurations were implemented:

1. Natural cubic splines with 10 equidistant knots (GAM function with bs="cr" and k=10);
2. P-Splines with a 2nd order spline basis and a third order difference penalty (GAM function with bs="ps" and m=2).
3. P-Splines with a 3rd order spline basis and a third order difference penalty (GAM function with bs="ps" and m=3).
4. P-Splines with a 4rd order spline basis and a fourth order difference penalty (GAM function with bs="ps" and m=4).
5. Thin Plate Regression splines with 2nd order derivatives (GAM function with bs="tp" and m=2).
6. Thin Plate Regression splines with 3rd order derivatives (GAM function with bs="tp" and m=3).
7. Thin Plate Regression splines with 4th order derivatives (GAM function with bs="tp" and m=4).

The natural cubic splines configuration is the configuration used by Mendes et al. (2018).

Firstly, we compare the performance of splines with 2nd order derivatives. Figures 3 and 4 below compare the performances of these spline configurations across all sample sizes, regarding, respectively, the RMSE and absolute bias.

η_1		n=	50	100	150	250	300	500	600	750	900
h=0,25	Cs		0,2684	0,1982	0,1800	0,1694	0,1644	0,1577	0,1550	0,1530	0,1525
	Ps2		0,2691	**0,1954**	**0,1773**	**0,1675**	**0,1619**	**0,1553**	**0,1530**	**0,1507**	**0,1503**
	TPRs2		**0,2630**	0,1966	0,1792	0,1693	0,1643	0,1575	0,1550	0,1528	0,1524
h=0,5	Cs		0,2666	0,2047	0,1844	0,1702	0,1669	0,1610	0,1578	0,1575	0,1549
	Ps2		0,2643	**0,2019**	**0,1816**	**0,1680**	**0,1646**	**0,1587**	**0,1552**	**0,1551**	**0,1523**
	TPRs2		**0,2616**	0,2036	0,1830	0,1701	0,1665	0,1609	0,1574	0,1573	0,1546
h=0,75	Cs		0,2756	0,1979	0,1849	0,1716	0,1673	0,1599	0,1581	0,1561	0,1549
	Ps2		0,2735	**0,1955**	**0,1835**	**0,1699**	**0,1657**	**0,1577**	**0,1562**	**0,1542**	**0,1530**
	TPRs2		**0,2714**	0,1973	0,1848	0,1714	0,1676	0,1597	0,1582	0,1564	0,1550

η_2		n=	50	100	150	250	300	500	600	750	900
h=0,25	Cs		0,2842	0,2375	0,2206	0,2052	0,1997	0,1922	0,1890	0,1865	0,1844
	Ps2		**0,2825**	**0,2356**	**0,2174**	**0,2021**	**0,1966**	**0,1887**	**0,1858**	**0,1833**	**0,1814**
	TPRs2		0,2827	0,2377	0,2201	0,2047	0,1995	0,1920	0,1889	0,1861	0,1840
h=0,5	Cs		0,2939	0,2455	0,2269	0,2128	0,2086	0,1982	0,1959	0,1933	0,1911
	Ps2		**0,2899**	**0,2427**	**0,2246**	**0,2103**	**0,2063**	**0,1960**	**0,1936**	**0,1912**	**0,1891**
	TPRs2		0,2909	0,2450	0,2269	0,2129	0,2092	0,1988	0,1965	0,1940	0,1917
h=0,75	Cs		0,2858	0,2412	0,2224	0,2071	0,2024	0,1923	0,1912	0,1870	0,1862
	Ps2		0,2833	**0,2392**	**0,2198**	**0,2042**	**0,1997**	**0,1890**	**0,1881**	**0,1842**	**0,1835**
	TPRs2		**0,2832**	0,2407	0,2226	0,2071	0,2026	0,1923	0,1910	0,1870	0,1863

η_3		n=	50	100	150	250	300	500	600	750	900
h=0,25	Cs		0,6156	0,5037	0,4855	0,4717	0,4649	0,4555	0,4518	0,4480	0,4479
	Ps2		0,6129	**0,5002**	**0,4820**	**0,4687**	**0,4617**	**0,4511**	**0,4476**	**0,4439**	**0,4439**
	TPRs2		**0,6081**	0,5020	0,4842	0,4715	0,4648	0,4539	0,4506	0,4466	0,4469
h=0,5	Cs		0,5867	0,5047	0,4790	0,4600	0,4580	0,4474	0,4437	0,4427	0,4381
	Ps2		0,5856	**0,4987**	**0,4751**	**0,4563**	**0,4534**	**0,4421**	**0,4389**	**0,4365**	**0,4331**
	TPRs2		**0,5821**	0,5017	0,4770	0,4587	0,4577	0,4462	0,4424	0,4410	0,4367
h=0,75	Cs		0,6111	0,4914	0,4788	0,4604	0,4550	0,4426	0,4434	0,4395	0,4365
	Ps2		0,6066	**0,4887**	**0,4762**	**0,4557**	**0,4508**	**0,4372**	**0,4373**	**0,4335**	**0,4299**
	TPRs2		**0,6050**	0,4915	0,4784	0,4589	0,4537	0,4407	0,4411	0,4375	0,4344

η_4		n=	50	100	150	250	300	500	600	750	900
h=0,25	Cs		0,8080	0,6929	0,6761	0,6644	0,6607	0,6479	0,6437	0,6376	0,6333
	Ps2		0,8064	0,6914	**0,6752**	**0,6621**	**0,6580**	**0,6462**	**0,6415**	**0,6356**	**0,6311**
	TPRs2		**0,8001**	**0,6913**	0,6760	0,6658	0,6607	0,6496	0,6448	0,6388	0,6339
h=0,5	Cs		0,8026	0,6956	0,6809	0,6610	0,6579	0,6432	0,6395	0,6328	0,6276
	Ps2		0,8003	0,6944	**0,6792**	**0,6569**	**0,6554**	**0,6391**	**0,6364**	**0,6289**	**0,6240**
	TPRs2		**0,7914**	**0,6928**	0,6812	0,6611	0,6592	0,6443	0,6406	0,6340	0,6292
h=0,75	Cs		0,8012	0,6874	0,6702	0,6508	0,6482	0,6299	0,6215	0,6158	0,6092
	Ps2		0,7936	**0,6861**	**0,6691**	**0,6499**	**0,6457**	**0,6264**	**0,6195**	**0,6129**	**0,6072**
	TPRs2		**0,7888**	0,6866	0,6708	0,6534	0,6501	0,6320	0,6239	0,6175	0,6114

η_5		n=	50	100	150	250	300	500	600	750	900
h=0,25	Cs		0,2495	0,1605	0,1484	0,1400	0,1373	0,1354	0,1341	0,1334	0,1330
	Ps2		0,2457	0,1572	**0,1456**	**0,1394**	**0,1361**	**0,1343**	**0,1335**	**0,1324**	**0,1322**
	TPRs2		**0,2362**	**0,1553**	0,1457	0,1400	0,1364	0,1349	0,1340	0,1332	0,1329
h=0,5	Cs		0,2313	0,1622	0,1467	0,1384	0,1357	0,1327	0,1314	0,1308	0,1306
	Ps2		0,2232	0,1590	0,1449	**0,1360**	**0,1337**	**0,1315**	**0,1302**	**0,1298**	**0,1296**
	TPRs2		**0,2186**	**0,1575**	**0,1440**	0,1366	0,1348	0,1324	0,1309	0,1308	0,1304
h=0,75	Cs		0,2321	0,1573	0,1443	0,1344	0,1325	0,1288	0,1276	0,1261	0,1255
	Ps2		0,2245	0,1541	0,1416	**0,1330**	**0,1312**	**0,1275**	**0,1267**	**0,1252**	**0,1249**
	TPRs2		**0,2188**	**0,1527**	**0,1412**	0,1332	0,1318	0,1282	0,1274	0,1261	0,1256

Table 2 - RMSE for Natural Cubic Splines and P-Splines and TPRS with 2nd order derivatives

η_1		n=	50	100	150	250	300	500	600	750	900
h=0,25	Cs		0,1427	0,1265	0,1244	0,1223	0,1199	0,1173	0,1162	0,1156	0,1157
	Ps2		0,1432	0,1253	0,1233	0,1209	0,1183	0,1154	0,1142	0,1136	0,1136
	TPRs2		0,1440	0,1275	0,1248	0,1226	0,1200	0,1173	0,1161	0,1154	0,1154
h=0,5	Cs		0,1397	0,1319	0,1270	0,1231	0,1219	0,1198	0,1181	0,1187	0,1169
	Ps2		0,1392	0,1309	0,1258	0,1217	0,1204	0,1180	0,1163	0,1169	0,1151
	TPRs2		0,1412	0,1332	0,1280	0,1235	0,1222	0,1199	0,1183	0,1188	0,1170
h=0,75	Cs		0,1429	0,1271	0,1267	0,1234	0,1211	0,1186	0,1175	0,1172	0,1162
	Ps2		0,1426	0,1262	0,1260	0,1218	0,1200	0,1169	0,1158	0,1154	0,1143
	TPRs2		0,1451	0,1280	0,1273	0,1239	0,1217	0,1186	0,1177	0,1173	0,1163

η_2		n=	50	100	150	250	300	500	600	750	900
h=0,25	Cs		0,1747	0,1637	0,1617	0,1568	0,1542	0,1517	0,1501	0,1491	0,1480
	Ps2		0,1728	0,1625	0,1606	0,1547	0,1523	0,1493	0,1478	0,1472	0,1463
	TPRs2		0,1763	0,1645	0,1626	0,1567	0,1545	0,1518	0,1500	0,1489	0,1479
h=0,5	Cs		0,1780	0,1731	0,1680	0,1636	0,1626	0,1579	0,1568	0,1556	0,1544
	Ps2		0,1782	0,1717	0,1665	0,1621	0,1611	0,1563	0,1554	0,1544	0,1534
	TPRs2		0,1807	0,1742	0,1685	0,1643	0,1632	0,1583	0,1572	0,1560	0,1547
h=0,75	Cs		0,1763	0,1679	0,1637	0,1580	0,1562	0,1518	0,1520	0,1493	0,1496
	Ps2		0,1753	0,1660	0,1616	0,1555	0,1543	0,1497	0,1501	0,1477	0,1481
	TPRs2		0,1785	0,1685	0,1643	0,1582	0,1562	0,1521	0,1519	0,1494	0,1497

η_3		n=	50	100	150	250	300	500	600	750	900
h=0,25	Cs		0,3275	0,2853	0,2796	0,2780	0,2745	0,2716	0,2694	0,2676	0,2686
	Ps2		0,3277	0,2830	0,2781	0,2764	0,2730	0,2695	0,2674	0,2659	0,2670
	TPRs2		0,3287	0,2852	0,2796	0,2781	0,2747	0,2711	0,2689	0,2671	0,2683
h=0,5	Cs		0,3175	0,2879	0,2793	0,2723	0,2722	0,2677	0,2660	0,2665	0,2645
	Ps2		0,3153	0,2851	0,2769	0,2702	0,2697	0,2651	0,2635	0,2633	0,2621
	TPRs2		0,3189	0,2874	0,2787	0,2717	0,2719	0,2672	0,2655	0,2657	0,2638
h=0,75	Cs		0,3261	0,2823	0,2800	0,2733	0,2706	0,2663	0,2676	0,2652	0,2637
	Ps2		0,3254	0,2804	0,2783	0,2709	0,2689	0,2638	0,2646	0,2624	0,2607
	TPRs2		0,3272	0,2829	0,2799	0,2729	0,2700	0,2655	0,2665	0,2644	0,2629

η_4		n=	50	100	150	250	300	500	600	750	900
h=0,25	Cs		0,5641	0,5284	0,5209	0,5151	0,5124	0,5007	0,4975	0,4913	0,4877
	Ps2		0,5642	0,5269	0,5210	0,5131	0,5097	0,4998	0,4956	0,4899	0,4859
	TPRs2		0,5654	0,5304	0,5230	0,5176	0,5135	0,5036	0,4993	0,4932	0,4888
h=0,5	Cs		0,5655	0,5282	0,5209	0,5061	0,5052	0,4917	0,4885	0,4814	0,4766
	Ps2		0,5645	0,5266	0,5195	0,5035	0,5035	0,4881	0,4866	0,4791	0,4741
	TPRs2		0,5662	0,5285	0,5244	0,5082	0,5082	0,4935	0,4904	0,4836	0,4786
h=0,75	Cs		0,5604	0,5238	0,5132	0,4979	0,4978	0,4809	0,4734	0,4676	0,4610
	Ps2		0,5591	0,5229	0,5125	0,4977	0,4954	0,4781	0,4719	0,4656	0,4602
	TPRs2		0,5605	0,5264	0,5162	0,5021	0,5003	0,4837	0,4762	0,4697	0,4639

η_5		n=	50	100	150	250	300	500	600	750	900
h=0,25	Cs		0,1113	0,0865	0,0900	0,0903	0,0891	0,0893	0,0877	0,0887	0,0880
	Ps2		0,1111	0,0861	0,0901	0,0900	0,0886	0,0884	0,0868	0,0880	0,0870
	TPRs2		0,1117	0,0872	0,0908	0,0904	0,0895	0,0894	0,0876	0,0889	0,0882
h=0,5	Cs		0,1038	0,0870	0,0862	0,0881	0,0854	0,0860	0,0846	0,0844	0,0845
	Ps2		0,1033	0,0865	0,0863	0,0880	0,0852	0,0857	0,0843	0,0840	0,0842
	TPRs2		0,1041	0,0874	0,0871	0,0886	0,0860	0,0863	0,0848	0,0846	0,0845
h=0,75	Cs		0,1013	0,0852	0,0851	0,0845	0,0847	0,0841	0,0831	0,0824	0,0811
	Ps2		0,1011	0,0855	0,0849	0,0842	0,0841	0,0831	0,0821	0,0812	0,0798
	TPRs2		0,1013	0,0863	0,0857	0,0850	0,0849	0,0841	0,0833	0,0826	0,0812

Table 3 - RMSE for Natural Cubic Splines and P-Splines and TPRS with 2nd order derivatives

In the context of very small samples (n=50), the results vary depending on the performance indicator. As PLSs with TPRS have less RMSE, the absolute bias (Bias) indicator shows better performances of Cubic Splines and P-Splines depending on the type of nonlinear relationship.

However, from n=100 all models with P-Splines perform better in both Bias and RMSE when estimating the relationships η_1 to η_4.

In the specific case of η_5, TPRS have lower RMSE in samples n=100 and 150 and Natural Cubic Splines show the best Bias results with n=100. Hence, P-Splines tend to be a less obvious choice when the underlying relationship between the latent variables is linear.

The results presented in the Annex confirm and deepen these conclusions when splines of higher order are used. In fact, P-Splines fare better in all contexts when the underlying relationships are nonlinear.

This advantage disappears when dealing with a linear relationship. In this case, higher order TPRS present better results than the other cases in what regards RMSE. Regarding Bias performance, the outcome is more mixed and inconclusive. Nevertheless, both TPRS and P-Splines tend to perform better than that natural cubic splines originally employed by Mendes et al. (2018).

Hence, with these results, P-splines become the obvious choice for estimating PLSs, since they fare better than the other methodologies in most contexts.

This conclusion is further justified as P-Splines are also faster producing results, as stated by Wood (2016).

Hence, the recommendation drawn from this analysis is that P-Splines should be preferred when using PLSs when nonlinear relationships are present.

3. CONCLUSIONS

In the last decade and a half, extensive research in PLS-PM has endowed this estimation technique with an extensive set of tools to overcome or at least mitigate its original shortcomings.

Evenso, nonlinear approaches to PLS-PM have been limited and mostly circumscribed to nonlinear relationships of known functional form.

Building from the work carried by Mendes et al. (2018), which introduces a method that embeds natural cubic splines with fixed knots in the PLS algorithm, aptly named smooth PLS (PLSs), we studied the performance of PLSs with different penalized regression splines.

For that, we embedded the PLS algorithm with P-Splines and Thin Plate Regression Splines in structural model direct relationships, testing different sets of parameters.

The different configurations were tested using a simulated dataset. The simulation framework included four different nonlinear structural relationships and a linear relationship for control purposes. The studied scenarios included six different sample sizes (from n=50 to n= 900), and 3 different levels of communality between latent variables and their indicators.

For comparison of the performance of these configurations we used the usual metrics: Root Mean Squared Error (RMSE) and absolute bias (Bias).

Our results show that P-Splines and TPRS are valid alternatives to natural cubic splines when studying non-linear relationships of unknown form in the context of PLS-SEM. More so, P-Splines fare better than the rest of the configurations in almost all the studied scenarios, except for very small samples contexts (n=50) and linear relationships, where TPRS, in some cases, presented better results.

We also illustrate how PLSs can be a powerful for inspection of nonlinear relationships and be a powerful predictive in corporate management and marketing research work frameworks, namely by assessing first and second derivatives of the direct relationships.

4. LIMITATIONS AND FUTURE WORK

Despite the observed trend of better results from P-Splines and TPRS, further simulation work should be developed to test in a more robust fashion whether these splines are significantly better suited to fit nonlinear relationships.

Besides RMSE and Bias, these different parametrizations could also be analyzed in an efficiency perspective. That is, it should be studied whether residual gains in RMSE and Bias with higher order parameters are sufficiently important to compensate in higher estimation times, since the identified differences are residual.

More generally, the full impact of PLSs in predictive research could be object of a structured and detailed approach, namely in what regards the analysis of derivatives of the direct relationships' functions in the structural model.

Also, the performance of the method in models with formative constructs should be studied to evaluate if the conclusions presented here may be generalized.

As the model behavior hasn't fared consistently better than the classic PLS approach in contexts where relationships between latent variables are in fact linear, the development of tools to automatically opt for the most adequate model configuration would be a valuable development for practical contexts, since it could provide substantial gains regarding parameter interpretability.

Finally, as the classic PLS has evolved to PLSc to address its original consistency issues, future work on PLSs could pursue a similar path, since the nonlinear approximations carried in this methodology approximation relies on linear combinations of functions.

5. BIBLIOGRAPHY

Bauer, D. J., Baldasaro, R. E. & Gottfredson, N. C. (2012). Diagnostic procedures for detecting nonlinear relationships between latent variables *Structural Equation Modeling. A Multidisciplinary Journal, 19*(2), 157-177.

Benitez, J., Henseler, J., Roldan, J. (2016). How to address endogeneity in partial least squares path modeling, *Proceedings of the 22th Americas Conference on information Systems*, 1–10.

Benitez, J., Henseler, J., Castillo, A., & Schuberth, F. (2020). How to perform and report an impactful analysis using partial least squares: Guidelines for confirmatory and explanatory IS research. *Information & Management,* 57(2), 103168.

Boor, C. (2001). *A Practical Guide to Splines 2nd edition*. New York: Springer.

Chin, W. W., Marcolin, B. L., & Newsted, P. R. (2003). A partial least squares latent variable modeling approach for measuring interaction effects. Results from a Monte Carlo simulation study and an electronic-mail emotion/adoption study. *Information Systems Research, 14*, 189 217.

Cortina, J. M. (1993). Interaction, Nonlinearity, and Multicollinearity: implications for Multiple Regression. *Journal of Management, 19*(4), 915–922.

Dijkstra, T. K. (2010). Latent variables and indices. *Handbook of Partial Least Squares: Concepts, methods and applications*. Berlin Heidelberg: Springer, 23–46.

Dijkstra, T.K., & Schermelleh-Engel, K. (2014). Consistent Partial Least Squares for Nonlinear Structural Equation Models. *Psychometrika, Volume 79, Issue 4*, 585-604.

Dijkstra, T. K., & Henseler, J. (2015). Consistent and asymptotically normal PLS estimators for linear structural equations. *Computational Statistics & Data Analysis, 81*, 10–23.

Dijkstra, T.K., & Henseler, J. (2015). Consistent partial least squares path modeling. *MIS Quarterly ,* 39(2), 297-316.

Duchon, J. (1977). Splines minimizing rotation-invariant semi-norms in Sobolev spaces. *Constructive Theory of Functions of Several Variables Lecture Notes in Mathematics*, 85–100.

Eilers, P. H. C., & B. D. Marx (1996). Flexible smoothing with B-splines and penalties, *Statistical Science, 11(*2), 89–121.

Green, P.J., & Silverman, B.W. (1994). *Nonparametric regression and generalized linear models – A roughness penalty approach*. London: Chapman & Hall

Hair, J.F., Ringle, C.M., & Sarstedt, M. (2011). PLS-SEM: indeed a silver bullet. *J. Market. Theory Pract. 18*(2), 139–152 (2011).

Hair, J. F., Ringle, C. M., & Sarstedt, M. (2013). Partial Least Squares Structural Equation Modeling: Rigorous Applications, Better Results and Higher Acceptance. *Long Range Planning, 46*(1-2), 1–12.

Hair, J. F., Sarstedt, M., Ringle, C. M., & Mena, J. A. (2011). An assessment of the use of partial least squares structural equation modeling in marketing research. *Journal of the Academic Marketing Science, 40*(3), 414–433.

Hair, J. F., Hult, G. T., Ringle, C. M., & Sarstedt, M. (2014). *A primer on partial least squares structural equation modeling (PLS-PM)*. Thousand Oaks, CA: Sage.

Henseler, J. (2018). Partial least squares path modeling: Quo vadis? *Quality & Quantity, 52*(1), 1–8.

Henseler, J., & Chin, W. W. (2010). A Comparison of Approaches for the Analysis of Interaction Effects Between Latent Variables Using Partial Least Squares Path Modeling. *Structural Equation Modeling: A Multidisciplinary Journal, 17*(1), 82–109.

Henseler, J., & Fassott, G. (2010). Testing moderating effects in PLS path models: An illustration of available procedures. *Handbook of Partial Least Squares: Concepts, methods and applications,* 713–735.

Henseler, J., Hubona, G., & Ray, P. A. (2016). Using PLS path modeling in new technology research: updated guidelines, *Industrial Management & Data Systems, 116*(1), 2-20.

Henseler, J., Ringle, C. M., & Sinkovics, R. R. (2009). The use of partial least squares path modeling in international marketing, *Advances in International Marketing - New Challenges to International Marketing*, 277-319.

Hult, G. T. M., J. F. Hair, D. Proksch, M. Sarstedt, A. Pinkwart, & C. M. Ringle (2018). Addressing Endogeneity in International Marketing Applications of Partial Least Squares Structural Equation Modeling. *Journal of International Marketing,* forthcoming.

Jung, S., & Park, J. (2018). Consistent Partial Least Squares Path Modeling via Regularization. *Frontiers in Psychology, 9*.

Kelava, A., Werner, C. S., Schermelleh-Engel, K., Moosbrugger, H., Zapf, D., Ma, Y., Cham H., Aiken, L. S. & West, S.G. (2011). Advanced nonlinear latent variable modeling: Distribution analytic LMS and QML estimators of interaction and quadratic effects, *Structural Equation Modeling: A Multidisciplinary Journal, 18*(3), 465-491.

Klein, A. G. & Moosbrugger, H. (2000). Maximum likelihood estimation of latent interaction effects with the LMS method, *Psychometrika, 65*(4), 457-474.

Klein, A. G., & Muthén, B. O. (2007). Quasi-maximum likelihood estimation of structural equation models with multiple interaction and quadratic effects, *Multivariate Behavioral Research, 42*(4), 647-673.

Kline, R. B. (2005). *Principles and practice of structural equation modeling, (2nd ed.)*. New York, NY: Guilford.

Lohmöller, J.-B. (1989). *Latent variable path modeling with partial least squares*. Heidelberg: Physica.

Marsh, H. W., Wen, Z., & Hau, K (2004). Structural equation models of latent interactions: evaluation of alternative estimation strategies and indicator construction. *Psychological methods, 9*(3), 275-300.

Mendes, J., Coelho, P. S., & Hanseler J., (2018). Smooth Partial Least Squares Path Modelling, *forthcoming*.

R Core Team (2018). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. Vienna: Austria.

Pek, J., Sterba, S. K., Kok, B. E. & Bauer, D. J. (2009). Estimating and visualizing nonlinear relations among latent variables: a semiparametric approach. *Multivariate Behavioral Research*, *44*(4), 407-436.

Reinartz, W. J., Haenlein, M., & Henseler, J. (2009). An Empirical Comparison of the Efficacy of Covariance-Based and Variance-Based SEM. *International Journal of Research in Marketing*, *26*(4), 332-344.

Reinsch, C. H. (1967). Smoothing by spline functions. *Numerische Mathematik 10*, 177–183.

Ruppert, D. (2002). Selecting the Number of Knots for Penalized Splines, *Journal of Computational and Graphical Statistics, 11*(4), 735-757.

Sarstedt, M., Henseler, J., Ringle, C.M. (2011). Multi-group analysis in Partial Least Squares (PLS) path modeling: alternative methods and empirical results. *Advances in International Marketing* 22, 195-218.

Song, X.-Y. & Lu, Z.-H. (2010). Semiparametric latent variable models with Bayesian P-splines, Journal of Computational and Graphical Statistics. 19(3), 590-608.

Song, X.-Y., Lu, Z.-H., Cai, J.-H., & Ip, E. H.-S. (2013). A Bayesian modeling approach for generalized semiparametric structural equation models, *Psychometrika, 78*(4), 624-647.

Steenkamp, J.-B., & Baumgartner, H., 2000. On the use of structural equation models for marketing modeling, *International Journal of Research in Marketing 17*(2-3), 195-202.

Tenenhaus, M., Vinzi, V. E., Chatelin, Y-M., & Lauro, C. (2005). PLS path modeling. *Computational Statistics & Data Analysis, 48*(1), 159 – 205.

Tuu, H. H., & Olsen, S. O. (2010). Nonlinear effects between satisfaction and loyalty: An empirical study of different conceptual relationships, *Journal of Targeting, Measurement and Analysis for Marketing*, *18(*3-4), 238–251.

van Riel, A.C., Henseler, J., Kemény, I., Sasovova, Z. (2017). Estimating hierarchical constructs using consistent partial least squares: the case of second-order composites of common factors. *Industrial Management & Data Systems 117*(3), 459–477.

Wold, H. O. A. (1975). Path models with latent variables: The NIPALS approach. *IH. M. Blalock (Ed.), Quantitative sociology*. New York, NY: Academic Press., 307–359.

Wold, H. O. A. (1982). Soft modelling: The basic design and some extensions. *K. G. Joreskog & H. O. A. Wold (Eds.), Systems under indirect observation, part II.* Amsterdam, The Netherlands: North-Holland, 1–55.

Wood, S. (2006). *Generalized Additive Models: an introduction with R.* Boca Raton: Chapman & Hall/CRC Press.

Wood, S. (2017). *Generalized Additive Models: an introduction with R. 2nd Edition.* Boca Raton: Chapman & Hall/CRC Press.

Wood, S. N. (2018). Package 'mgcv' - R.

Wahba, G. (1990). *Spline models for observational data.* Philadelphia: SIAM.

RMSE η_1

n=	50	100	150	250	300	500	600	750	900
h1									
Cs	0,2684 (3)	0,1982 (7)	0,1800 (7)	0,1694 (7)	0,1644 (7)	0,1577 (7)	0,1550 (6)	0,1530 (7)	0,1525 (7)
Ps2	0,2691 (5)	0,1954 (5)	0,1773 (5)	0,1675 (5)	0,1619 (5)	0,1553 (5)	0,1530 (5)	0,1507 (5)	0,1503 (5)
Ps3	0,2684 (4)	0,1852 (3)	0,1682 (4)	0,1609 (4)	0,1555 (3)	0,1497 (3)	0,1480 (3)	0,1458 (3)	0,1456 (3)
Ps4	0,2805 (7)	0,1876 (4)	0,1672 (2)	0,1588 (2)	0,1531 (2)	**0,1471 (1)**	**0,1454 (1)**	**0,1432 (1)**	**0,1429 (1)**
TPRs2	0,2630 (2)	0,1966 (6)	0,1792 (6)	0,1693 (6)	0,1643 (6)	0,1575 (6)	0,1550 (6)	0,1528 (6)	0,1524 (6)
TPRs3	**0,2559 (1)**	0,1817 (2)	0,1672 (3)	0,1604 (3)	0,1563 (4)	0,1508 (4)	0,1490 (4)	0,1470 (4)	0,1470 (4)
TPRs4	0,2705 (6)	**0,1810 (1)**	**0,1648 (1)**	**0,1575 (1)**	**0,1528 (1)**	0,1476 (2)	0,1458 (2)	0,1437 (2)	0,1436 (2)
h2									
Cs	0,2666 (6)	0,2047 (7)	0,1844 (7)	0,1702 (7)	0,1669 (7)	0,1610 (7)	0,1578 (7)	0,1575 (7)	0,1549 (7)
Ps2	0,2643 (5)	0,2019 (5)	0,1816 (5)	0,1680 (5)	0,1646 (5)	0,1587 (5)	0,1552 (5)	0,1551 (5)	0,1523 (5)
Ps3	0,2636 (4)	0,1918 (3)	0,1728 (4)	0,1614 (3)	0,1583 (3)	0,1532 (3)	0,1506 (3)	0,1502 (3)	0,1479 (3)
Ps4	0,2728 (7)	0,1928 (4)	0,1702 (2)	0,1591 (2)	0,1551 (2)	**0,1501 (1)**	**0,1480 (1)**	**0,1474 (1)**	**0,1455 (1)**
TPRs2	0,2616 (3)	0,2036 (6)	0,1830 (6)	0,1701 (6)	0,1665 (6)	0,1609 (6)	0,1574 (6)	0,1573 (6)	0,1546 (6)
TPRs3	**0,2527 (1)**	0,1892 (2)	0,1719 (3)	0,1620 (4)	0,1591 (4)	0,1545 (4)	0,1518 (4)	0,1515 (4)	0,1494 (4)
TPRs4	0,2616 (2)	**0,1876 (1)**	**0,1686 (1)**	**0,1585 (1)**	**0,1553 (2)**	0,1506 (2)	0,1486 (2)	0,1482 (2)	0,1462 (2)
h3									
Cs	0,2756 (6)	0,1979 (7)	0,1849 (7)	0,1716 (7)	0,1673 (6)	0,1599 (7)	0,1581 (6)	0,1561 (6)	0,1549 (6)
Ps2	0,2735 (5)	0,1955 (5)	0,1835 (5)	0,1699 (5)	0,1657 (5)	0,1577 (5)	0,1562 (5)	0,1542 (5)	0,1530 (5)
Ps3	0,2693 (3)	0,1858 (3)	0,1747 (4)	0,1635 (4)	0,1598 (3)	0,1525 (3)	0,1516 (3)	0,1498 (3)	0,1487 (3)
Ps4	0,2797 (7)	0,1867 (4)	0,1725 (2)	0,1617 (2)	0,1573 (2)	**0,1500 (1)**	**0,1494 (1)**	**0,1475 (1)**	**0,1465 (1)**
TPRs2	0,2714 (4)	0,1973 (6)	0,1848 (6)	0,1714 (6)	0,1676 (7)	0,1597 (6)	0,1582 (7)	0,1564 (7)	0,1550 (7)
TPRs3	**0,2613 (1)**	0,1833 (2)	0,1737 (3)	0,1638 (4)	0,1603 (4)	0,1537 (4)	0,1525 (4)	0,1510 (4)	0,1499 (4)
TPRs4	0,2670 (2)	**0,1815 (1)**	**0,1696 (1)**	**0,1604 (1)**	**0,1566 (1)**	0,1503 (2)	0,1497 (2)	0,1481 (2)	0,1472 (2)

RMSE η_2

n=	50	100	150	250	300	500	600	750	900
h1									
Cs	0,2842 (7)	0,2375 (6)	0,2206 (7)	0,2052 (7)	0,1997 (7)	0,1922 (7)	0,1890 (7)	0,1865 (7)	0,1844 (7)
Ps2	0,2825 (5)	0,2356 (5)	0,2174 (5)	0,2021 (5)	0,1966 (5)	0,1887 (5)	0,1858 (5)	0,1833 (5)	0,1814 (5)
Ps3	0,2687 (3)	0,2229 (3)	0,2062 (3)	0,1932 (3)	0,1882 (3)	0,1820 (3)	0,1797 (3)	0,1779 (3)	0,1763 (3)
Ps4	0,2777 (4)	0,2208 (2)	**0,2025 (1)**	**0,1886 (1)**	**0,1840 (1)**	**0,1785 (1)**	**0,1764 (1)**	**0,1745 (1)**	**0,1731 (1)**
TPRs2	0,2827 (6)	0,2377 (7)	0,2201 (6)	0,2047 (6)	0,1995 (6)	0,1920 (6)	0,1889 (6)	0,1861 (6)	0,1840 (6)
TPRs3	**0,2634 (1)**	0,2231 (4)	0,2075 (4)	0,1946 (4)	0,1897 (4)	0,1836 (4)	0,1816 (4)	0,1794 (4)	0,1779 (4)
TPRs4	0,2675 (2)	**0,2185 (1)**	0,2027 (2)	0,1893 (2)	0,1845 (2)	0,1792 (2)	0,1770 (2)	0,1753 (2)	0,1739 (2)
h2									
Cs	0,2939 (7)	0,2455 (7)	0,2269 (6)	0,2128 (6)	0,2086 (6)	0,1982 (6)	0,1959 (6)	0,1933 (6)	0,1911 (6)
Ps2	0,2899 (5)	0,2427 (5)	0,2246 (5)	0,2103 (5)	0,2063 (5)	0,1960 (5)	0,1936 (5)	0,1912 (5)	0,1891 (5)
Ps3	0,2750 (3)	0,2301 (4)	0,2143 (3)	0,2021 (3)	0,1988 (3)	0,1894 (3)	0,1872 (3)	0,1857 (3)	0,1840 (3)
Ps4	0,2825 (4)	0,2289 (2)	0,2107 (2)	0,1981 (2)	**0,1949 (1)**	**0,1860 (1)**	**0,1840 (1)**	**0,1827 (1)**	**0,1810 (1)**
TPRs2	0,2909 (6)	0,2450 (6)	0,2269 (7)	0,2129 (7)	0,2092 (7)	0,1988 (7)	0,1965 (7)	0,1940 (7)	0,1917 (7)
TPRs3	**0,2706 (1)**	0,2296 (3)	0,2155 (4)	0,2031 (4)	0,2003 (4)	0,1910 (4)	0,1893 (4)	0,1874 (4)	0,1855 (4)
TPRs4	0,2743 (2)	**0,2261 (1)**	**0,2092 (1)**	**0,1980 (1)**	0,1954 (2)	0,1866 (2)	0,1850 (2)	0,1836 (2)	0,1819 (2)
h3									
Cs	0,2858 (7)	0,2412 (7)	0,2224 (6)	0,2071 (6)	0,2024 (6)	0,1923 (6)	0,1912 (7)	0,1870 (7)	0,1862 (6)
Ps2	0,2833 (6)	0,2392 (5)	0,2198 (5)	0,2042 (5)	0,1997 (5)	0,1890 (5)	0,1881 (5)	0,1842 (5)	0,1835 (5)
Ps3	0,2661 (3)	0,2271 (4)	0,2099 (3)	0,1959 (3)	0,1921 (3)	0,1829 (3)	0,1823 (3)	0,1795 (3)	0,1792 (3)
Ps4	0,2738 (4)	0,2258 (2)	0,2066 (2)	**0,1915 (1)**	**0,1880 (1)**	**0,1794 (1)**	**0,1788 (1)**	**0,1759 (1)**	**0,1756 (1)**
TPRs2	0,2832 (5)	0,2407 (6)	0,2226 (7)	0,2071 (7)	0,2026 (7)	0,1923 (7)	0,1910 (6)	0,1870 (6)	0,1863 (7)
TPRs3	**0,2624 (1)**	0,2265 (3)	0,2114 (4)	0,1977 (4)	0,1938 (4)	0,1847 (4)	0,1840 (4)	0,1809 (4)	0,1807 (4)
TPRs4	0,2642 (2)	**0,2228 (1)**	**0,2061 (1)**	0,1922 (2)	0,1888 (2)	0,1803 (2)	0,1797 (2)	0,1768 (2)	0,1766 (2)

RMSE η_3

n=	50	100	150	250	300	500	600	750	900
h1									
Cs	0,6156 (7)	0,5037 (7)	0,4855 (7)	0,4717 (7)	0,4649 (7)	0,4555 (7)	0,4518 (7)	0,4480 (7)	0,4479 (7)
Ps2	0,6129 (6)	0,5002 (5)	0,4820 (5)	0,4687 (5)	0,4617 (5)	0,4511 (5)	0,4476 (5)	0,4439 (5)	0,4439 (5)
Ps3	0,5933 (2)	0,4822 (3)	0,4649 (4)	0,4561 (3)	0,4519 (3)	0,4435 (3)	0,4404 (3)	0,4373 (3)	0,4378 (3)
Ps4	0,6108 (5)	0,4832 (4)	0,4625 (2)	**0,4525 (1)**	**0,4471 (1)**	**0,4400 (1)**	**0,4370 (1)**	**0,4337 (1)**	**0,4345 (1)**
TPRs2	0,6081 (4)	0,5020 (6)	0,4842 (6)	0,4715 (6)	0,4648 (6)	0,4539 (6)	0,4506 (6)	0,4466 (6)	0,4469 (6)
TPRs3	**0,5834 (1)**	0,4789 (2)	0,4644 (3)	0,4586 (4)	0,4530 (4)	0,4448 (4)	0,4421 (4)	0,4382 (4)	0,4384 (4)
TPRs4	0,5975 (3)	**0,4766 (1)**	**0,4614 (1)**	0,4537 (2)	0,4479 (2)	0,4407 (2)	0,4380 (2)	0,4348 (2)	0,4353 (2)
h2									
Cs	0,5867 (6)	0,5047 (7)	0,4790 (7)	0,4600 (7)	0,4580 (7)	0,4474 (7)	0,4437 (7)	0,4427 (7)	0,4381 (7)
Ps2	0,5856 (5)	0,4987 (5)	0,4751 (5)	0,4563 (5)	0,4534 (5)	0,4421 (5)	0,4389 (5)	0,4365 (5)	0,4331 (5)
Ps3	0,5724 (2)	0,4792 (2)	0,4613 (3)	0,4473 (3)	0,4443 (3)	0,4349 (3)	0,4331 (3)	0,4313 (3)	0,4280 (3)
Ps4	0,5902 (7)	0,4821 (4)	0,4598 (2)	0,4431 (2)	**0,4410 (1)**	**0,4320 (1)**	**0,4295 (1)**	**0,4284 (1)**	**0,4251 (1)**
TPRs2	0,5821 (4)	0,5017 (6)	0,4770 (6)	0,4587 (6)	0,4577 (6)	0,4462 (6)	0,4424 (6)	0,4410 (6)	0,4367 (6)
TPRs3	**0,5581 (1)**	0,4798 (3)	0,4630 (4)	0,4483 (4)	0,4458 (4)	0,4365 (4)	0,4345 (4)	0,4327 (4)	0,4293 (4)
TPRs4	0,5737 (3)	**0,4783 (1)**	**0,4586 (1)**	**0,4431 (1)**	0,4415 (2)	0,4322 (2)	0,4304 (2)	0,4292 (2)	0,4261 (2)
h3									
Cs	0,6111 (6)	0,4914 (6)	0,4788 (7)	0,4604 (7)	0,4550 (7)	0,4426 (7)	0,4434 (7)	0,4395 (7)	0,4365 (7)
Ps2	0,6066 (5)	0,4887 (5)	0,4762 (5)	0,4557 (5)	0,4508 (5)	0,4372 (5)	0,4373 (5)	0,4335 (5)	0,4299 (5)
Ps3	0,5928 (2)	0,4718 (4)	0,4620 (3)	0,4459 (3)	0,4401 (3)	0,4288 (3)	0,4293 (3)	0,4261 (3)	0,4238 (3)
Ps4	0,6120 (7)	0,4711 (3)	**0,4564 (1)**	**0,4408 (1)**	**0,4353 (1)**	**0,4241 (1)**	**0,4252 (1)**	**0,4215 (1)**	**0,4193 (1)**
TPRs2	0,6050 (4)	0,4915 (7)	0,4784 (6)	0,4589 (6)	0,4537 (6)	0,4407 (6)	0,4411 (6)	0,4375 (6)	0,4344 (6)
TPRs3	**0,5840 (1)**	0,4708 (2)	0,4632 (4)	0,4482 (4)	0,4439 (4)	0,4313 (4)	0,4319 (4)	0,4284 (4)	0,4261 (4)
TPRs4	0,5943 (3)	**0,4679 (1)**	0,4565 (2)	0,4421 (2)	0,4380 (2)	0,4262 (2)	0,4274 (2)	0,4236 (2)	0,4211 (2)

Table A.1 - RMSE for Natural Cubic Splines and P-Splines and TPRS

RMSE

η_4

	n= 50	100	150	250	300	500	600	750	900
h1									
Cs	0,8080 (3)	0,6929 (3)	0,6761 (5)	0,6644 (4)	0,6607 (5)	0,6479 (4)	0,6437 (4)	0,6376 (4)	0,6333 (4)
Ps2	0,8064 (2)	0,6914 (2)	**0,6752** (2)	0,6621 (3)	0,6580 (3)	0,6462 (3)	0,6415 (3)	0,6356 (3)	0,6311 (3)
Ps3	0,8317 (5)	0,6978 (4)	0,6757 (3)	0,6587 (2)	0,6534 (2)	0,6396 (2)	0,6344 (2)	0,6298 (2)	0,6255 (2)
Ps4	0,8646 (7)	0,7044 (6)	0,6750 (1)	0,6546 (1)	0,6498 (1)	0,6343 (1)	0,6310 (1)	**0,6256** (1)	**0,6205** (1)
TPRs2	**0,8001** (1)	**0,6913** (1)	0,6760 (4)	0,6658 (5)	0,6607 (4)	0,6496 (5)	0,6448 (5)	0,6388 (5)	0,6339 (5)
TPRs3	0,8201 (4)	0,6994 (5)	0,6793 (6)	0,6697 (6)	0,6630 (6)	0,6535 (6)	0,6501 (6)	0,6423 (6)	0,6375 (6)
TPRs4	0,8515 (6)	0,7098 (7)	0,6858 (7)	0,6704 (7)	0,6655 (7)	0,6536 (7)	0,6517 (7)	0,6447 (7)	0,6393 (7)
h2									
Cs	0,8026 (3)	0,6956 (3)	0,6809 (4)	0,6610 (4)	0,6579 (4)	0,6432 (4)	0,6395 (4)	0,6328 (4)	0,6276 (4)
Ps2	0,8003 (2)	0,6944 (2)	**0,6792** (1)	0,6569 (3)	0,6554 (3)	0,6391 (3)	0,6364 (3)	0,6289 (3)	0,6240 (3)
Ps3	0,8270 (5)	0,6982 (4)	0,6794 (3)	0,6541 (2)	0,6496 (2)	0,6322 (2)	0,6296 (2)	0,6213 (2)	0,6177 (2)
Ps4	0,8646 (7)	0,7049 (6)	**0,6793** (2)	0,6489 (1)	0,6450 (1)	0,6273 (1)	**0,6251** (1)	**0,6176** (1)	**0,6142** (1)
TPRs2	**0,7914** (1)	**0,6928** (1)	0,6812 (5)	0,6611 (5)	0,6592 (5)	0,6443 (5)	0,6406 (5)	0,6340 (5)	0,6292 (5)
TPRs3	0,8107 (4)	0,7001 (5)	0,6836 (6)	0,6643 (6)	0,6624 (6)	0,6464 (6)	0,6444 (6)	0,6350 (6)	0,6303 (6)
TPRs4	0,8448 (6)	0,7107 (7)	0,6905 (7)	0,6665 (7)	0,6638 (7)	0,6478 (7)	0,6461 (7)	0,6364 (7)	0,6317 (7)
h3									
Cs	0,8012 (3)	0,6874 (3)	0,6702 (4)	0,6508 (4)	0,6482 (4)	0,6299 (4)	0,6215 (4)	0,6158 (4)	0,6092 (4)
Ps2	0,7936 (2)	**0,6861** (1)	0,6691 (3)	0,6499 (3)	0,6457 (3)	0,6264 (3)	0,6195 (3)	0,6129 (3)	0,6072 (3)
Ps3	0,8199 (5)	0,6924 (4)	0,6690 (2)	0,6444 (2)	0,6394 (2)	0,6191 (2)	0,6124 (2)	0,6058 (2)	0,6009 (2)
Ps4	0,8531 (7)	0,6949 (6)	**0,6656** (1)	0,6384 (1)	0,6331 (1)	0,6131 (1)	0,6067 (1)	0,5999 (1)	0,5957 (1)
TPRs2	**0,7888** (1)	0,6866 (2)	0,6708 (5)	0,6534 (5)	0,6501 (5)	0,6320 (5)	0,6239 (5)	0,6175 (5)	0,6114 (5)
TPRs3	0,8119 (4)	0,6936 (5)	0,6757 (6)	0,6594 (6)	0,6549 (6)	0,6361 (6)	0,6291 (6)	0,6210 (6)	0,6135 (6)
TPRs4	0,8414 (6)	0,7027 (7)	0,6797 (7)	0,6600 (7)	0,6557 (7)	0,6385 (7)	0,6308 (7)	0,6246 (7)	0,6181 (7)

RMSE

η_5

	n= 50	100	150	250	300	500	600	750	900
h1									
Cs	0,2495 (3)	0,1605 (4)	0,1484 (5)	0,1400 (6)	0,1373 (7)	0,1354 (7)	0,1341 (7)	0,1334 (7)	0,1330 (7)
Ps2	0,2457 (2)	0,1572 (2)	0,1456 (2)	0,1394 (3)	0,1361 (5)	0,1343 (5)	0,1335 (5)	0,1324 (5)	0,1322 (5)
Ps3	0,2674 (6)	0,1644 (6)	0,1493 (6)	0,1395 (4)	0,1357 (4)	0,1318 (4)	0,1310 (4)	0,1297 (4)	0,1291 (4)
Ps4	0,2839 (7)	0,1718 (7)	0,1512 (7)	0,1395 (5)	0,1355 (3)	0,1312 (3)	0,1304 (3)	0,1288 (2)	0,1283 (2)
TPRs2	**0,2362** (1)	**0,1553** (1)	0,1457 (3)	0,1400 (7)	0,1364 (6)	0,1349 (6)	0,1340 (6)	0,1332 (6)	0,1329 (6)
TPRs3	0,2503 (4)	0,1595 (3)	**0,1452** (1)	0,1375 (2)	0,1340 (2)	0,1312 (2)	0,1303 (2)	0,1295 (3)	0,1290 (3)
TPRs4	0,2610 (5)	0,1611 (5)	0,1457 (4)	**0,1344** (1)	**0,1317** (1)	**0,1284** (1)	**0,1276** (1)	**0,1264** (1)	**0,1255** (1)
h2									
Cs	0,2313 (4)	0,1622 (5)	0,1467 (5)	0,1384 (7)	0,1357 (7)	0,1327 (7)	0,1314 (7)	0,1308 (6)	0,1306 (7)
Ps2	0,2232 (2)	0,1590 (2)	0,1449 (4)	0,1360 (5)	0,1337 (5)	0,1315 (5)	0,1302 (5)	0,1298 (5)	0,1296 (5)
Ps3	0,2401 (6)	0,1654 (6)	0,1478 (6)	0,1357 (4)	0,1321 (4)	0,1288 (4)	0,1270 (4)	0,1265 (3)	0,1262 (3)
Ps4	0,2566 (7)	0,1719 (7)	0,1498 (7)	0,1356 (3)	0,1318 (3)	0,1278 (2)	0,1257 (2)	0,1252 (2)	0,1249 (2)
TPRs2	**0,2186** (1)	**0,1575** (1)	0,1440 (3)	0,1366 (6)	0,1348 (6)	0,1324 (6)	0,1309 (6)	0,1308 (7)	0,1304 (6)
TPRs3	0,2306 (3)	0,1603 (3)	0,1435 (2)	0,1338 (2)	0,1310 (2)	0,1287 (3)	0,1266 (3)	0,1266 (4)	0,1262 (4)
TPRs4	0,2399 (5)	0,1613 (4)	**0,1424** (1)	**0,1314** (1)	**0,1278** (1)	**0,1256** (1)	**0,1235** (1)	**0,1233** (1)	**0,1230** (1)
h3									
Cs	0,2321 (4)	0,1573 (5)	0,1443 (5)	0,1344 (7)	0,1325 (7)	0,1288 (7)	0,1276 (7)	0,1261 (7)	0,1255 (6)
Ps2	0,2245 (2)	0,1541 (2)	0,1416 (3)	0,1330 (5)	0,1312 (5)	0,1275 (5)	0,1267 (5)	0,1252 (5)	0,1249 (5)
Ps3	0,2407 (6)	0,1602 (6)	0,1447 (6)	0,1327 (4)	0,1296 (4)	0,1245 (4)	0,1238 (4)	0,1222 (3)	0,1221 (4)
Ps4	0,2598 (7)	0,1643 (6)	0,1464 (7)	0,1315 (3)	0,1288 (3)	0,1226 (2)	0,1219 (2)	0,1206 (2)	0,1206 (2)
TPRs2	**0,2188** (1)	**0,1527** (1)	0,1412 (2)	0,1332 (6)	0,1318 (6)	0,1282 (6)	0,1274 (6)	0,1261 (6)	0,1256 (7)
TPRs3	0,2310 (3)	0,1558 (3)	0,1417 (4)	0,1309 (2)	0,1290 (3)	0,1244 (3)	0,1235 (3)	0,1223 (4)	0,1221 (3)
TPRs4	0,2424 (6)	0,1568 (4)	**0,1398** (1)	**0,1283** (1)	**0,1266** (1)	**0,1221** (1)	**0,1210** (1)	**0,1200** (1)	**0,1198** (1)

Table A.1 **(continued) - RMSE for Natural Cubic Splines and P-Splines and TPRS**

BIAS

η_1

	n=	50		100		150		250		300		500		600		750		900	
h1																			
Cs		0,1427	(5)	0,1265	(6)	0,1244	(6)	0,1223	(6)	0,1199	(6)	0,1173	(6)	0,1162	(7)	0,1156	(7)	0,1157	(7)
Ps2		0,1432	(6)	0,1253	(5)	0,1233	(5)	0,1209	(5)	0,1183	(5)	0,1154	(5)	0,1142	(5)	0,1136	(5)	0,1136	(5)
Ps3		0,1247	(4)	0,1103	(3)	0,1106	(3)	0,1125	(3)	0,1105	(3)	0,1095	(3)	0,1089	(3)	0,1087	(3)	0,1092	(3)
Ps4		0,1206	(2)	0,1058	(2)	0,1065	(1)	0,1093	(2)	**0,1075**	(1)	**0,1070**	(1)	**0,1066**	(1)	**0,1065**	(1)	**0,1070**	(1)
TPRs2		0,1440	(7)	0,1275	(7)	0,1248	(7)	0,1226	(7)	0,1200	(7)	0,1173	(7)	0,1161	(6)	0,1154	(6)	0,1154	(6)
TPRs3		0,1242	(3)	0,1111	(4)	0,1121	(4)	0,1130	(4)	0,1114	(4)	0,1105	(4)	0,1098	(4)	0,1097	(4)	0,1102	(4)
TPRs4		**0,1180**	(1)	**0,1051**	(1)	0,1068	(2)	**0,1090**	(1)	0,1076	(2)	0,1075	(2)	0,1070	(2)	0,1071	(2)	0,1076	(2)
h2																			
Cs		0,1397	(6)	0,1319	(6)	0,1270	(6)	0,1231	(6)	0,1219	(6)	0,1198	(6)	0,1181	(6)	0,1187	(6)	0,1169	(6)
Ps2		0,1392	(5)	0,1309	(5)	0,1258	(5)	0,1217	(5)	0,1204	(5)	0,1180	(5)	0,1163	(5)	0,1169	(5)	0,1151	(5)
Ps3		0,1219	(3)	0,1160	(3)	0,1147	(3)	0,1135	(3)	0,1132	(3)	0,1125	(3)	0,1118	(3)	0,1125	(3)	0,1113	(3)
Ps4		**0,1184**	(1)	0,1116	(2)	**0,1105**	(1)	**0,1102**	(1)	**0,1100**	(1)	**0,1102**	(1)	**0,1099**	(1)	**0,1106**	(1)	**0,1096**	(1)
TPRs2		0,1412	(7)	0,1332	(7)	0,1280	(7)	0,1235	(7)	0,1222	(7)	0,1199	(7)	0,1183	(7)	0,1188	(7)	0,1170	(7)
TPRs3		0,1227	(4)	0,1172	(4)	0,1159	(4)	0,1146	(4)	0,1140	(4)	0,1136	(4)	0,1127	(4)	0,1134	(4)	0,1124	(4)
TPRs4		0,1187	(2)	**0,1114**	(1)	0,1107	(2)	0,1106	(2)	0,1106	(2)	0,1108	(2)	0,1104	(2)	0,1110	(2)	0,1101	(2)
h3																			
Cs		0,1429	(6)	0,1271	(6)	0,1267	(6)	0,1234	(6)	0,1211	(6)	0,1186	(6)	0,1175	(6)	0,1172	(6)	0,1162	(6)
Ps2		0,1426	(5)	0,1262	(5)	0,1260	(5)	0,1218	(5)	0,1200	(5)	0,1169	(5)	0,1158	(5)	0,1154	(5)	0,1143	(5)
Ps3		0,1256	(3)	0,1114	(3)	0,1141	(3)	0,1139	(3)	0,1127	(3)	0,1111	(3)	0,1109	(3)	0,1109	(3)	0,1101	(3)
Ps4		0,1230	(2)	**0,1062**	(1)	**0,1094**	(1)	**0,1106**	(1)	**0,1096**	(1)	**0,1089**	(1)	**0,1088**	(1)	**0,1090**	(1)	**0,1086**	(1)
TPRs2		0,1451	(7)	0,1280	(7)	0,1273	(7)	0,1239	(7)	0,1217	(7)	0,1186	(7)	0,1177	(7)	0,1173	(7)	0,1163	(7)
TPRs3		0,1263	(4)	0,1129	(4)	0,1157	(4)	0,1154	(4)	0,1138	(4)	0,1124	(4)	0,1118	(4)	0,1119	(4)	0,1112	(4)
TPRs4		**0,1217**	(1)	0,1068	(2)	0,1098	(2)	0,1108	(2)	0,1099	(2)	0,1093	(2)	0,1093	(2)	0,1096	(2)	0,1091	(2)

BIAS

η_2

	n=	50		100		150		250		300		500		600		750		900	
h1																			
Cs		0,1747	(6)	0,1637	(6)	0,1617	(6)	0,1568	(7)	0,1542	(6)	0,1517	(6)	0,1501	(7)	0,1491	(7)	0,1480	(7)
Ps2		0,1728	(5)	0,1625	(5)	0,1606	(5)	0,1547	(5)	0,1523	(5)	0,1493	(5)	0,1478	(5)	0,1472	(5)	0,1463	(5)
Ps3		0,1502	(3)	0,1467	(3)	0,1477	(3)	0,1457	(3)	0,1440	(3)	0,1435	(3)	0,1427	(3)	0,1427	(3)	0,1422	(3)
Ps4		0,1455	(2)	0,1409	(1)	0,1428	(1)	0,1417	(1)	0,1401	(1)	0,1405	(1)	0,1399	(1)	0,1399	(1)	0,1394	(1)
TPRs2		0,1763	(7)	0,1645	(7)	0,1626	(7)	0,1567	(6)	0,1545	(7)	0,1518	(7)	0,1500	(6)	0,1489	(6)	0,1479	(6)
TPRs3		0,1514	(4)	0,1488	(4)	0,1497	(4)	0,1469	(4)	0,1451	(4)	0,1445	(4)	0,1438	(4)	0,1436	(4)	0,1431	(4)
TPRs4		0,1454	(1)	0,1421	(2)	0,1443	(2)	0,1427	(2)	0,1409	(2)	0,1413	(2)	0,1405	(2)	0,1406	(2)	0,1401	(2)
h2																			
Cs		0,1780	(5)	0,1731	(6)	0,1680	(6)	0,1636	(6)	0,1626	(6)	0,1579	(6)	0,1568	(6)	0,1556	(6)	0,1544	(6)
Ps2		0,1782	(6)	0,1717	(5)	0,1665	(5)	0,1621	(5)	0,1611	(5)	0,1563	(5)	0,1554	(5)	0,1544	(5)	0,1534	(5)
Ps3		0,1540	(3)	0,1563	(3)	0,1549	(3)	0,1536	(3)	0,1535	(3)	0,1503	(3)	0,1496	(3)	0,1496	(3)	0,1490	(3)
Ps4		**0,1499**	(1)	**0,1509**	(1)	**0,1505**	(1)	**0,1498**	(1)	**0,1499**	(1)	**0,1474**	(1)	**0,1468**	(1)	**0,1468**	(1)	**0,1464**	(1)
TPRs2		0,1807	(7)	0,1742	(7)	0,1685	(7)	0,1643	(7)	0,1632	(7)	0,1583	(7)	0,1572	(7)	0,1560	(7)	0,1547	(7)
TPRs3		0,1564	(4)	0,1580	(4)	0,1565	(4)	0,1545	(4)	0,1543	(4)	0,1512	(4)	0,1508	(4)	0,1505	(4)	0,1497	(4)
TPRs4		0,1512	(2)	0,1519	(2)	0,1508	(2)	0,1504	(2)	0,1505	(2)	0,1480	(2)	0,1477	(2)	0,1477	(2)	0,1471	(2)
h3																			
Cs		0,1763	(6)	0,1679	(6)	0,1637	(6)	0,1580	(6)	0,1562	(6)	0,1518	(6)	0,1520	(7)	0,1493	(6)	0,1496	(6)
Ps2		0,1753	(5)	0,1660	(5)	0,1616	(5)	0,1555	(5)	0,1543	(5)	0,1497	(5)	0,1501	(5)	0,1477	(5)	0,1481	(5)
Ps3		0,1516	(3)	0,1517	(3)	0,1503	(3)	0,1472	(3)	0,1470	(3)	0,1447	(3)	0,1454	(3)	0,1441	(3)	0,1447	(3)
Ps4		**0,1461**	(1)	**0,1468**	(1)	**0,1462**	(1)	**0,1436**	(1)	**0,1434**	(1)	**0,1417**	(1)	**0,1426**	(1)	**0,1412**	(1)	**0,1419**	(1)
TPRs2		0,1785	(7)	0,1685	(7)	0,1643	(7)	0,1582	(7)	0,1562	(7)	0,1521	(7)	0,1519	(6)	0,1494	(7)	0,1497	(7)
TPRs3		0,1532	(4)	0,1531	(4)	0,1527	(4)	0,1487	(4)	0,1478	(4)	0,1455	(4)	0,1462	(4)	0,1446	(4)	0,1454	(4)
TPRs4		0,1470	(2)	0,1474	(2)	0,1468	(2)	0,1444	(2)	0,1441	(2)	0,1425	(2)	0,1432	(2)	0,1418	(2)	0,1425	(2)

BIAS

η_3

	n=	50		100		150		250		300		500		600		750		900	
h1																			
Cs		0,3275	(5)	0,2853	(7)	0,2796	(6)	0,2780	(6)	0,2745	(6)	0,2716	(7)	0,2694	(7)	0,2676	(7)	0,2686	(7)
Ps2		0,3277	(6)	0,2830	(5)	0,2781	(5)	0,2764	(5)	0,2730	(5)	0,2695	(5)	0,2674	(5)	0,2659	(5)	0,2670	(5)
Ps3		0,3069	(2)	0,2682	(2)	0,2653	(2)	0,2669	(2)	0,2656	(2)	0,2639	(3)	0,2625	(3)	0,2617	(3)	0,2632	(4)
Ps4		**0,3069**	(1)	**0,2665**	(1)	**0,2637**	(1)	**0,2653**	(1)	**0,2638**	(1)	**0,2628**	(1)	**0,2614**	(1)	**0,2604**	(1)	**0,2620**	(1)
TPRs2		0,3287	(7)	0,2852	(6)	0,2796	(7)	0,2781	(7)	0,2747	(7)	0,2711	(6)	0,2689	(6)	0,2671	(6)	0,2683	(6)
TPRs3		0,3086	(4)	0,2693	(4)	0,2663	(4)	0,2687	(4)	0,2661	(4)	0,2644	(4)	0,2628	(4)	0,2616	(4)	0,2629	(3)
TPRs4		0,3073	(3)	0,2669	(2)	0,2646	(2)	0,2666	(2)	0,2647	(2)	0,2631	(2)	0,2618	(2)	0,2609	(2)	0,2623	(2)
h2																			
Cs		0,3175	(6)	0,2879	(7)	0,2793	(7)	0,2723	(7)	0,2722	(7)	0,2677	(7)	0,2660	(7)	0,2665	(7)	0,2645	(7)
Ps2		0,3153	(5)	0,2851	(5)	0,2769	(5)	0,2702	(5)	0,2697	(5)	0,2651	(5)	0,2635	(5)	0,2633	(5)	0,2621	(5)
Ps3		0,2979	(4)	0,2695	(2)	0,2657	(3)	0,2625	(3)	0,2625	(3)	0,2602	(3)	0,2597	(3)	0,2602	(3)	0,2590	(3)
Ps4		0,2962	(2)	**0,2688**	(1)	**0,2641**	(1)	**0,2611**	(1)	**0,2612**	(1)	**0,2592**	(1)	**0,2585**	(1)	**0,2594**	(1)	**0,2581**	(1)
TPRs2		0,3189	(7)	0,2874	(6)	0,2787	(6)	0,2717	(6)	0,2719	(6)	0,2672	(6)	0,2655	(6)	0,2657	(6)	0,2638	(6)
TPRs3		0,2973	(3)	0,2719	(4)	0,2674	(4)	0,2635	(4)	0,2630	(4)	0,2602	(3)	0,2598	(4)	0,2605	(4)	0,2591	(4)
TPRs4		**0,2952**	(1)	0,2702	(3)	0,2654	(2)	0,2616	(2)	0,2618	(2)	0,2594	(2)	0,2588	(2)	0,2595	(2)	0,2583	(2)
h3																			
Cs		0,3261	(6)	0,2823	(6)	0,2800	(7)	0,2733	(7)	0,2706	(7)	0,2663	(7)	0,2676	(7)	0,2652	(7)	0,2637	(7)
Ps2		0,3254	(5)	0,2804	(5)	0,2783	(5)	0,2709	(5)	0,2689	(5)	0,2638	(5)	0,2646	(5)	0,2624	(5)	0,2607	(5)
Ps3		**0,3060**	(1)	0,2672	(3)	0,2674	(3)	0,2633	(3)	0,2610	(3)	0,2584	(3)	0,2596	(3)	0,2581	(3)	0,2574	(3)
Ps4		0,3061	(2)	**0,2649**	(1)	**0,2645**	(1)	**0,2612**	(1)	**0,2591**	(1)	**0,2565**	(1)	**0,2579**	(1)	**0,2560**	(1)	**0,2554**	(1)
TPRs2		0,3272	(7)	0,2829	(7)	0,2799	(6)	0,2729	(6)	0,2700	(6)	0,2655	(6)	0,2665	(6)	0,2644	(6)	0,2629	(6)
TPRs3		0,3088	(4)	0,2688	(4)	0,2691	(4)	0,2645	(4)	0,2627	(4)	0,2590	(4)	0,2605	(4)	0,2587	(4)	0,2580	(4)
TPRs4		0,3073	(3)	0,2660	(2)	0,2667	(2)	0,2624	(2)	0,2607	(2)	0,2575	(2)	0,2590	(2)	0,2568	(2)	0,2561	(2)

Table A.2 – Absolute Bias for Natural Cubic Splines and P-Splines and TPRS

BIAS
η_s

	n= 50		100		150		250		300		500		600		750		900	
h1																		
Cs	0,1113	(6)	0,0865	(6)	0,0900	(5)	0,0903	(6)	0,0891	(6)	0,0893	(6)	0,0877	(7)	0,0887	(6)	0,0880	(6)
Ps2	0,1111	(5)	0,0861	(5)	0,0901	(6)	0,0900	(5)	0,0886	(5)	0,0884	(5)	0,0868	(5)	0,0880	(5)	0,0870	(5)
Ps3	0,1076	(3)	0,0804	(3)	0,0839	(3)	0,0851	(3)	0,0842	(3)	0,0850	(3)	0,0837	(3)	0,0852	(3)	0,0845	(3)
Ps4	0,1013	(2)	0,0749	(2)	0,0759	(2)	0,0781	(2)	0,0781	(2)	0,0804	(2)	0,0798	(2)	0,0814	(2)	0,0812	(2)
TPRs2	0,1117	(7)	0,0872	(7)	0,0908	(7)	0,0904	(7)	0,0895	(7)	0,0894	(7)	0,0876	(6)	0,0889	(7)	0,0882	(7)
TPRs3	0,1081	(4)	0,0824	(4)	0,0855	(4)	0,0863	(4)	0,0850	(4)	0,0858	(4)	0,0844	(4)	0,0860	(4)	0,0853	(4)
TPRs4	**0,0993**	(1)	**0,0741**	(1)	**0,0757**	(1)	**0,0774**	(1)	**0,0775**	(1)	**0,0799**	(1)	**0,0793**	(1)	**0,0810**	(1)	**0,0807**	(1)
h2																		
Cs	0,1038	(6)	0,0870	(6)	0,0862	(5)	0,0881	(6)	0,0854	(6)	0,0860	(6)	0,0846	(6)	0,0844	(6)	0,0845	(6)
Ps2	0,1033	(5)	0,0865	(5)	0,0863	(6)	0,0880	(5)	0,0852	(5)	0,0857	(5)	0,0843	(5)	0,0840	(5)	0,0842	(5)
Ps3	0,0984	(3)	0,0809	(3)	0,0799	(3)	0,0828	(3)	0,0803	(3)	0,0822	(3)	0,0811	(3)	0,0812	(3)	0,0814	(3)
Ps4	0,0912	(2)	0,0772	(2)	0,0746	(2)	0,0757	(2)	0,0751	(2)	0,0775	(2)	0,0770	(2)	0,0775	(2)	0,0782	(2)
TPRs2	0,1041	(7)	0,0874	(7)	0,0871	(7)	0,0886	(7)	0,0860	(7)	0,0863	(7)	0,0848	(7)	0,0846	(7)	0,0845	(7)
TPRs3	0,0998	(4)	0,0820	(4)	0,0814	(4)	0,0840	(4)	0,0817	(4)	0,0828	(4)	0,0816	(4)	0,0818	(4)	0,0820	(4)
TPRs4	**0,0899**	(1)	**0,0746**	(1)	**0,0728**	(1)	**0,0756**	(1)	**0,0741**	(1)	**0,0772**	(1)	**0,0764**	(1)	**0,0773**	(1)	**0,0778**	(1)
h3																		
Cs	0,1013	(6)	0,0852	(5)	0,0851	(6)	0,0845	(6)	0,0847	(6)	0,0841	(7)	0,0831	(6)	0,0824	(6)	0,0811	(6)
Ps2	0,1011	(5)	0,0855	(6)	0,0849	(5)	0,0842	(5)	0,0841	(5)	0,0831	(5)	0,0821	(5)	0,0812	(5)	0,0798	(5)
Ps3	0,0984	(3)	0,0799	(3)	0,0793	(3)	0,0789	(3)	0,0793	(3)	0,0791	(3)	0,0788	(3)	0,0784	(3)	0,0774	(3)
Ps4	0,0930	(2)	0,0719	(1)	0,0717	(2)	**0,0714**	(1)	**0,0719**	(1)	**0,0734**	(1)	0,0741	(2)	0,0743	(2)	**0,0739**	(7)
TPRs2	0,1013	(7)	0,0863	(7)	0,0857	(7)	0,0850	(7)	0,0849	(7)	0,0841	(6)	0,0833	(7)	0,0826	(7)	0,0812	(7)
TPRs3	0,0994	(4)	0,0816	(4)	0,0808	(4)	0,0803	(4)	0,0805	(4)	0,0802	(4)	0,0796	(4)	0,0794	(4)	0,0783	(4)
TPRs4	**0,0923**	(1)	**0,0721**	(2)	**0,0716**	(1)	0,0716	(2)	0,0723	(2)	0,0735	(2)	**0,0739**	(1)	**0,0742**	(1)	0,0740	(2)

Table A.2 **(continued) – Absolute Bias for Natural Cubic Splines and P-Splines and TPRS**