Stat Papers
DOI 10.1007/s00362-017-0940-6

REGULAR ARTICLE

Multiple mediation analysis for interval-valued data

Antonio Calcagnì[1] · Luigi Lombardi[1] · Lorenzo Avanzi[1] · Eduardo Pascali[2]

Received: 18 May 2016 / Revised: 14 July 2017

Abstract Mediation analysis is an important statistical approach to evaluate the relationships among observed variables. The most commonly used models for mediation analysis handle single-valued variables. However, there are several circumstances (e.g., dimensionality reduction of large datasets, clinical patient courses, repeated measures, masked data, uncertain data) in which the collected information can be represented more naturally by means of intervals. In these cases, standard mediation analyses can be ill-suited. Although interval-valued variables can be transformed into standard single-valued variables, such procedures may mask some relevant information provided by intervals. In this article, we present a novel and simple model (IMedA) to perform mediation analysis on interval-valued variables which is based on both the symbolic regression approach and the regression based mediation framework. We also generalize Stolzenberg's decomposition of effects to cope with interval-valued data. We further introduce a specific variance based decomposition procedure to descriptively evaluate the sizes of such effects. Finally, to better highlight the IMedA features we apply our model to a real case study from behavioral contexts.

Keywords Interval data · Mediation analysis · Path analysis · Multivariate multiple regression · Work-related burnout

Electronic supplementary material The online version of this article (doi:10.1007/s00362-017-0940-6) contains supplementary material, which is available to authorized users.

✉ Antonio Calcagnì
antonio.calcagni@unitn.it

[1] Department of Psychology and Cognitive Sciences, University of Trento, Corso Bettini 31, 38068 Rovereto, Italy

[2] Department of Mathematics and Physics 'E. De Giorgi', University of Salento, C.P. 193 Via per Arnesano, 73100 Lecce, Italy

Published online: 08 August 2017

1 Introduction

Mediation analyses are widespread in modeling underlying mechanisms of complex relationships in empirical data (MacKinnon and Fairchild 2009; Edwards and Lambert 2007). They are successfully used in several research domains, such us behavioral and social sciences, epidemiology, biology and agriculture (Claessens et al. 2004; Kristal et al. 2000; Caffo et al. 2008; Wardle et al. 2008). Conceptually, mediation models are adopted to assess situations in which the observed relation between an independent variable, x, and a dependent variable, y, is explained using a set of third variables, $m_1, m_2, \ldots, m_k$, in the relation, called mediator variables. Considering the cause-effect relation between x and y, the mediator variables are causally placed between x and y such that a change in x produces changes in m_j $(j = 1, \ldots, k)$ which, in turn, also produces changes in y. The x-m_j-y pathway explains the process through which x partially—or fully—acts on y (Baron and Kenny 1986; Yuan et al. 2013; Luo and Geng 2016). Figure 2 depicts a graphical representation for a mediation model.

The estimation of these pathways are usually performed either using a least squares regression approach (Bollen and Stine 1990; Edwards and Lambert 2007; Judd and Kenny 1981) or the so-called causal inference analysis (Imai and Van Dyk 2004; Imai et al. 2010). In the practice of research, both approaches have been successfully used in the literature (MacKinnon 2008).

Traditionally, methods to assess mediation have been developed for single-valued data only. However, in some empirical contexts the observed information may show more complex structures or patterns (e.g., interval-valued data, histogram-valued data, symbolic-valued data) than those commonly represented by single-valued data (Diday et al. 2008; Billard and Diday 2003). In this respect, *interval-valued* data are one of the simplest and most widely known types of structured data. In particular, they may arise in different cases such as, for instance, when (i) three-way datasets are reduced to two-way structures (Diday et al. 2008), (ii) clinical patient course and/or repeated measures are summarized by adopting procedures like the *response feature analysis* (Frison and Pocock 1992; Everitt 1995; Arndt et al. 2000), (iii) confidential data are masked by summarized data (Little 1993), (iv) individual sample data are incomplete (Gómez et al. 2004), (v) empirical data are modeled by interval semi-orders (Luce 1956; Fishburn 1973; Halff et al. 1976), (vi) observed measures are affected by systematic uncertainty (Augustin 2002; Parchami et al. 2012; Salicone 2007). All these sources of interval-data depend upon different assumptions about the specific data generation process (Blanco-Fernández and Winker, 2016). For instance, some experimental settings may lead to intervals as outcome of a data aggregation process (as for the above cases i–iv). By contrast, there exist other circumstances in which the data generation process produces intervals *per se* (see the above cases v–vi). Thus, choosing the most appropriate statistical approach to deal with interval-valued data (e.g., possibilistic, probabilistic, and descriptive/symbolic) should follow from substantive considerations and justifications about the data generation process underlying empirical data (Blanco-Fernández et al. 2013b; Blanco-Fernández and Winker 2016; Couso and Dubois 2014).

Bearing this in mind, in this contribution we propose a novel technique for mediation analysis, named *interval mediation analysis* (IMedA), to deal with interval-valued as well as single-valued data. Our proposal is based on the general least squares regression context (MacKinnon 2008), combining path analysis (Edwards and Lambert 2007), and symbolic data analysis (Billard and Diday 2002; Lima Neto and de Carvalho 2008).

The reminder of the article is organized as follows. Section 2 briefly recalls the basic characteristics of interval-valued data together with some applications in the behavioral sciences. Section 3 exposes the IMedA model, parameters estimation, and model evaluation. Section 4 illustrates procedures for computing and evaluating direct and indirect effects of the IMedA model. Section 5 reports a brief Monte Carlo simulation study to evaluate the performance of the IMedA algorithm whereas Sect. 6 describes a real case study showing the application of the new approach to a behavioral dataset. Finally, Sect. 7 concludes this article providing final remarks and suggestions for future extensions of the current contribution.

2 Interval-valued data

Interval data can be used in modeling empirical situations where the knowledge to be extracted is complex and/or highly structured. Unlike single-valued data, which can just represent single and point-wise information, structured-data can always take into account a set of additional information or sources at the same time. Interval-valued data may arise in many research contexts. For instance, in organizational research, studies are often conducted using the so-called within-person approach where information regarding affects, behaviors, interpersonal interactions, work events, and other workplace phenomena are collected over the time (Fisher and To 2012). Daily diary methods are the most adopted techniques to regularly collect data related to immediate or recent experiences from the same sample of people for a given interval of time. In this context, daily measurements can be naturally represented as closed and bounded intervals (Taris et al. 2010). Similarly, in other research domains like health sciences, longitudinal data can be represented according to the well-known *response feature analysis* approach where intervals are used to summarize the temporal property of the data (Arndt et al. 2000; Everitt 1995; Frison and Pocock 1992). Finally, interval-valued data might also arise when individual's measurements are collected by means of mouse-tracking instruments or other similar interfaces (e.g., trackball, joy-stick, light and laser pen, wii system, etc.). In these cases, as for reaction-times, saccadic eye movements, and brain waves, also computer-mouse movements can dynamically measure some relevant motor components of cognitive processes which are associated with individuals' responses (e.g., decisional uncertainty. See: Calcagnì et al. 2017; Calcagnì and Lombardi 2014; Johnson et al. 2012)

2.1 Basic formal definitions

The interval $\tilde{z} = [u, v]$ is the set of real numbers $\{x \in \mathbb{R} |\ u \leq x \leq v\}$ where u and v denotes the left and right endpoints of the interval. Two intervals $\tilde{z}$ and $\tilde{b}$ are *equal* if their corresponding endpoints are the same. The interval $\tilde{z}$ is said to be a *degenerated* interval if $u = v$ and in this case the interval simply reduces to the singleton $\tilde{z} = \{u\}$. The *width* of $\tilde{z}$ is defined as $wdt(\tilde{z}) = v - u$, whereas its *midpoint*

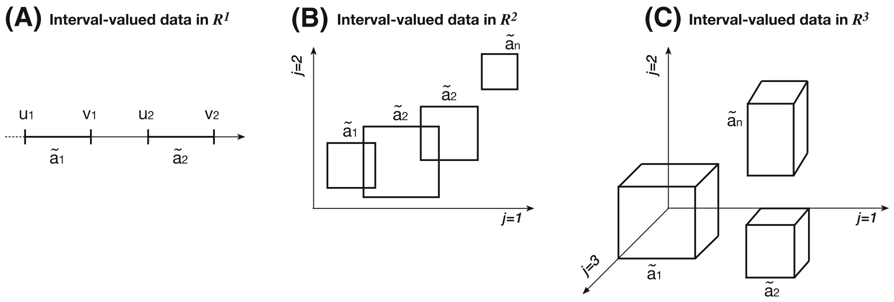

Fig. 1 Graphical representations for interval-valued data: **a** two intervals in $\mathbb{R}^1$, **b** a collection of interval in $\mathbb{R}^2$, **c** intervals in $\mathbb{R}^3$. Note that in the panel (**b**) a linear relation among the intervals is depicted

is $mid(\tilde{z}) = (u+v)/2$. The half-width of $\tilde{z}$ is called the *spread* (or *radius*) of $\tilde{z}$ and is defined as $spr(\tilde{z}) = (v-u)/2$. Interval-valued data can be easily extended to the multidimensional case. In particular, a $n \times k$ *interval matrix* $\tilde{\mathbf{Z}}$ is a matrix whose elements are interval numbers, namely $\tilde{\mathbf{z}} = (\tilde{z}_{ij}) = ([u,v]_{ij})$ with $i = 1, \ldots, n$ and $j = 1, \ldots, k$. From a geometrical point of view, the i-th row of $\tilde{\mathbf{Z}}$ can be represented as a k-dimensional hyper-rectangle. More precisely, for $k = 1$, $\tilde{z}$ is a simple interval in $\mathbb{R}$, for $k = 2$, $\tilde{\mathbf{z}}$ is a rectangle in $\mathbb{R}^2$, whereas for $k \geq 2$, $\tilde{\mathbf{z}}$ is an hyper-rectangle in $\mathbb{R}^k$ (see Fig. 1). The width of $\tilde{\mathbf{Z}}$ is the non-negative matrix of widths computed on its interval elements $\mathbf{z}^w = wdt(\tilde{Z}_{ij})$ whereas the spreads of $\tilde{\mathbf{Z}}$ is the non-negative matrix of spreads of its interval elements $\mathbf{z}^r = wdt(\tilde{Z}_{ij})/2$. Likewise, the midpoint of $\tilde{\mathbf{Z}}$ is the matrix containing the midpoints of its interval element $\mathbf{Z}^c = mid(\tilde{Z}_{ij})$. Further details about the formal properties and operations for interval-valued data can be found in Moore (1966).

2.2 Centre-range parametric representation

There are several parametric representations (e.g., centered based, min-max, centre-range) that can be adopted to describe interval-type data (Lima Neto and de Carvalho 2008). Among these, the *centre-range parametrization* allows to describe an interval $\tilde{z}$ by means of its midpoint and spread: $\tilde{z} = (c, r)_{CR}$ where $c = mid(\tilde{z})$ and $r = spr(\tilde{z})$. Unlike other parametric representations for interval-valued data, the CR-parametrization shows some nice features (Blanco-Fernández et al. 2013a). Firstly, from a computational perspective, it always ensures well-defined intervals by simply satisfying the non-negative condition $r > 0$. Secondly, the CR-representation directly works with the parameter space $\mathcal{O}_{\tilde{z}} = \{(c,r) \in \mathbb{R} \times \mathbb{R}^+\}$ of $\tilde{z}$. This would allow to extend many classical statistical approaches to interval-valued data without considering other sophisticated manipulation methods (e.g., interval algebra). Moreover, in the case of multidimensional interval data, the CR-representation can decompose the $n \times m$ *interval matrix* $\tilde{\mathbf{Z}}$ into two $n \times m$ *single-valued matrices*, $\mathbf{Z^c}$ and $\mathbf{Z^r}$, which contain all the parameters involved in $\mathcal{O}_{\tilde{z}}$. In this way, multidimensional interval-valued data may be further manipulated according to classical statistical techniques. Finally, the CR-representation may be very useful especially when intervals are used to describe empirical objects in terms of measurement precisions (by means of c) and measurement uncertainty (by means of r). Despite these gains, it should be stressed that choosing a proper parametrization for interval data should follow from a general understanding of the data generation process and the empirical conditions where intervals are used (Blanco-Fernández and Winker 2016).

3 IMedA: interval mediation analysis

In this section we illustrate the new multiple mediation model for interval-valued data (IMedA).

3.1 Model

Let $\tilde{\mathbf{x}}$ and $\tilde{\mathbf{y}}$ be n (units) $\times$ 1 interval vectors representing the independent and dependent variables, respectively. Let $\tilde{\mathbf{M}}$ be a n (units) $\times$ k (mediators) interval matrix containing the set of mediators. By adopting the CR-parametrization for interval data, the elements of $\tilde{\mathbf{x}}$, $\tilde{\mathbf{y}}$, and $\tilde{\mathbf{M}}$ can be represented by a collection of n (units) $\times$ 1 single-valued vectors $\mathbf{x^c}$, $\mathbf{x^r}$, $\mathbf{y^c}$, $\mathbf{y^r}$, and n (units) $\times$ k single-valued matrices $\mathbf{M^c}$, $\mathbf{M^r}$. The mediation model for interval-valued data can be expressed by two regression systems as follows:

$$\begin{aligned}
&\mathcal{S}_1 : \begin{cases} \mathbf{M^c} = \mathbf{1}\mathbf{A}^c + \mathbf{X}\Xi + \mathbf{E}^c \\ \mathbf{M^r} = \mathbf{1}\mathbf{A}^r + (\mathbf{1}\mathbf{A}^c + \mathbf{X}\Xi)\boldsymbol{\Pi} + \mathbf{E^r} \end{cases} \\
&\mathcal{S}_2 : \begin{cases} \mathbf{y^c} = \mathbf{1}\alpha^c + \mathbf{X}\boldsymbol{\beta} + \mathbf{M^c}\boldsymbol{\gamma}^c + \mathbf{M^r}\boldsymbol{\gamma}^r + \boldsymbol{\epsilon}^c \\ \mathbf{y^r} = \mathbf{1}\alpha^r + (\mathbf{1}\alpha^c + \mathbf{X}\boldsymbol{\beta} + \mathbf{M^c}\boldsymbol{\gamma}^c + \mathbf{M^r}\boldsymbol{\gamma}^r)\delta + \boldsymbol{\epsilon}^r \end{cases}
\end{aligned} \tag{1}$$

where in both models $\mathbf{X} = \{\mathbf{x^c}, \mathbf{x^r}\}$ is an interval valued-variable. For $\mathcal{S}_1$, the matrices $\mathbf{A}^c$, $\mathbf{A}^r$ and $\boldsymbol{\Pi}$ denote $k \times k$ diagonal matrices of intercept terms and coefficients of the ranges, $\mathbf{X}$ is a $n \times 2$ column-wise stacked matrix containing the vectors $\mathbf{x^c}$ and $\mathbf{x^r}$, whereas Ξ is a $2 \times k$ matrix of regression coefficients between the matrix of mediators $\mathbf{M^c}$ and the independent variables $\mathbf{X}$. Finally, $\mathbf{E^c}$ and $\mathbf{E^r}$ are $n \times k$ matrices of residual terms. Similarly, for the second system $\mathcal{S}_2$, the scalars α^c, α^r, δ represent the intercept terms and the range coefficient of the model. Moreover, $\boldsymbol{\beta}$ is a 2×1 vector of regression coefficients quantifying the relation between the independent variables $\mathbf{X}$ and the dependent variable $\mathbf{y^c}$ whereas $\boldsymbol{\gamma}^\mathbf{c}$ and $\boldsymbol{\gamma}^\mathbf{r}$ are $k \times 1$ vectors of regression coefficients between the matrices of mediators $\mathbf{M^c}$ and $\mathbf{M^r}$ and the dependent variable $\mathbf{y^c}$, with $\boldsymbol{\epsilon}^c$ and $\boldsymbol{\epsilon}^r$ being $n \times 1$ vectors of residual terms. Finally, $\mathbf{1}$ denotes matrices (or vectors) of appropriate orders of all ones. Figure 2a (resp. 2b) shows the compact (resp. exploded) conceptual diagram for the IMedA model.

3.2 Parameters estimation

In the IMedA model, the parameters estimates are obtained according to a standard least squares (LS) procedure which minimizes the following dissimilarity measures:

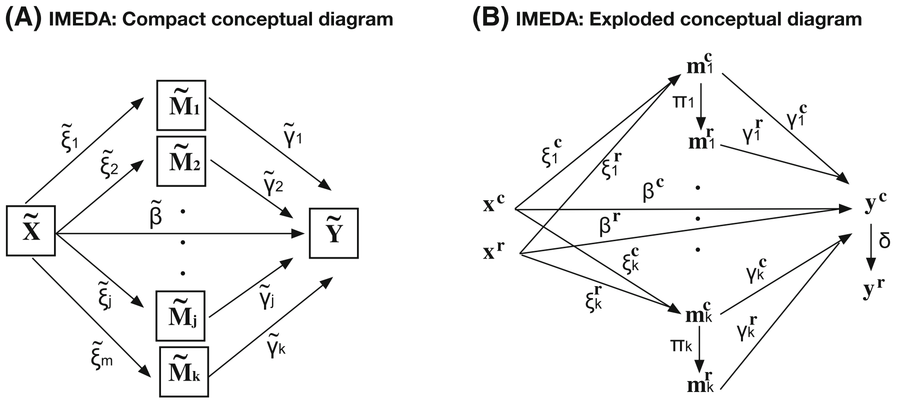

Fig. 2 Conceptual (**a**) and exploded (**b**) diagrams for the interval mediation model. Note that the *tilde symbol* denotes interval variables or interval coefficients

$$\mathcal{D}_1^2 = \|\mathbf{M^c} - \mathbf{M^{c*}}\|^2 + \|\mathbf{M^r} - \mathbf{M^{r*}}\|^2 \quad \text{and} \quad \mathcal{D}_2^2 = \|\mathbf{y^c} - \mathbf{y^{c*}}\|^2 + \|\mathbf{y^r} - \mathbf{y^{r*}}\|^2 \tag{2}$$

where: $\mathbf{M^{c*}} = \mathbf{1A}^c + \mathbf{X}\Xi$, $\mathbf{M^{r*}} = \mathbf{1A}^r + \mathbf{M^{c*}}\boldsymbol{\Pi}$, $\mathbf{y^{c*}} = \mathbf{1}\alpha^c + \mathbf{X}\boldsymbol{\beta} + \mathbf{M^c}\boldsymbol{\beta} + \mathbf{M^r}\boldsymbol{\gamma^r}$, and $\mathbf{y^{r*}} = \mathbf{1}\alpha^r + \mathbf{y^{c*}}\delta$, respectively. To estimate the parameters contained in $\mathbf{A}^c$, $\mathbf{A}^r$, $\boldsymbol{\Pi}$, Ξ, α^c, α^r, δ, $\boldsymbol{\beta}$, and $\boldsymbol{\gamma}$, we used the alternating least squares (ALS) algorithm (Kiers 2002). By convention, this alternating gradient-descendent algorithm converges when $\|\theta - \hat{\theta}\|^2 \leq \Delta$ with θ being the array containing the model's parameters, $\hat{\theta}$ the corresponding estimated array, and Δ a small positive quantity, respectively. The detailed iterative ALS solutions of the model are reported in Appendix A.

3.3 Goodness-of-fit indices

The goodness of fit of the IMedA model can be evaluated by considering the following two normalized indices:

$$\begin{aligned} R_M^2 &= 1 - \frac{\|\mathbf{M^c} - \mathbf{M^{c*}}\|^2 + \|\mathbf{M^r} - \mathbf{M^{r*}}\|^2}{\|\mathbf{M^c} - \mathbf{1}\,\mathrm{diag}(\overline{\mathbf{M^c}})\|^2 + \|\mathbf{M^r} - \mathbf{1}\,\mathrm{diag}(\overline{\mathbf{M^r}})\|^2} \\ R_Y^2 &= 1 - \frac{\|\mathbf{y^c} - \mathbf{y^{c*}}\|^2 + \|\mathbf{y^r} - \mathbf{y^{r*}}\|^2}{\|\mathbf{y^c} - \overline{\mathbf{y^c}}\|^2 + \|\mathbf{y^r} - \overline{\mathbf{y^r}}\|^2} \end{aligned} \tag{3}$$

where $\mathrm{diag}(\overline{\mathbf{M^c}})$ and $\mathrm{diag}(\overline{\mathbf{M^r}})$ denote $k \times k$ diagonal matrices containing the column means of the matrices $\mathbf{M^c}$ and $\mathbf{M^r}$, $\mathbf{1}$ is a $n \times k$ matrix of all ones, whereas $\overline{\mathbf{y^c}}$ and $\overline{\mathbf{y^r}}$ denote $n \times 1$ vectors containing the mean values of $\mathbf{y^c}$ and $\mathbf{y^r}$, respectively. Note that R_M^2 and R_Y^2 take values in [0, 1] and compare the residual sum of squares with the observed total sum of squares. The interpretations of the goodness-of-fit indices are in line with the standard R^2 measure adopted in the regression framework.

4 Deriving direct and indirect effects

The decomposition of effects allows to quantify the amount of the total effect TE which can be ascribed to the mediators—indirect effect IE—and the residual total effect between the two variables when the mediators are held constant—direct effect DE (Alwin and Hauser, 1975; Sobel, 1982; Stolzenberg, 1980; Choi and Lee, 2016). Considering the IMedA model representation, the basic decomposition rule TE = DE + IE can be generalized as follows:

$$\underbrace{\mathrm{TE}_c + \mathrm{TE}_r}_{\text{total effect}} = \underbrace{(\mathrm{DE}_c + \mathrm{DE}_r)}_{\text{direct effect}} + \underbrace{(\mathrm{IE}_c + \mathrm{IE}_r)}_{\text{indirect effect}}$$

where TE_c, TE_r, DE_c, DE_r, IE_c and IE_r represent the interval components of TE, DE, and IE for centers and ranges. Applying the generalized Stolzenberg's decomposition procedure for interval data (see Appendix B), the final decomposed effects of the model are:

$$\begin{aligned}
\mathrm{DE}_c &= \hat{\beta}^c + \hat{\delta}\hat{\beta}^c \\
\mathrm{DE}_r &= \hat{\beta}^r + \hat{\delta}\hat{\beta}^r \\
\mathrm{IE}_c &= [(\hat{\boldsymbol{\xi}}^c \circ \hat{\boldsymbol{\gamma}}^{cT})\mathbf{1}_k + (\hat{\boldsymbol{\xi}}^c \circ \hat{\boldsymbol{\gamma}}^{rT} \circ \hat{\boldsymbol{\Pi}}^T)\mathbf{1}_k] + \hat{\delta}[(\hat{\boldsymbol{\xi}}^c \circ \hat{\boldsymbol{\gamma}}^{cT})\mathbf{1}_k + (\hat{\boldsymbol{\xi}}^c \circ \hat{\boldsymbol{\gamma}}^{rT} \circ \hat{\boldsymbol{\Pi}}^T)\mathbf{1}_k] \\
\mathrm{IE}_r &= [(\hat{\boldsymbol{\xi}}^r \circ \hat{\boldsymbol{\gamma}}^{cT})\mathbf{1}_k + (\hat{\boldsymbol{\xi}}^r \circ \hat{\boldsymbol{\gamma}}^{rT} \circ \hat{\boldsymbol{\Pi}}^T)\mathbf{1}_k] + \hat{\delta}[(\hat{\boldsymbol{\xi}}^r \circ \hat{\boldsymbol{\gamma}}^{cT})\mathbf{1}_k + (\hat{\boldsymbol{\xi}}^r \circ \hat{\boldsymbol{\gamma}}^{rT} \circ \hat{\boldsymbol{\Pi}}^T)\mathbf{1}_k]
\end{aligned} \tag{4}$$

where $\hat{\beta}^c$ and $\hat{\beta}^r$ are the estimated parameters contained in $\widehat{\boldsymbol{\beta}}$; $\hat{\boldsymbol{\xi}}^c$ and $\hat{\boldsymbol{\xi}}^r$ are $1 \times k$ row-vectors of the estimated matrix $\widehat{\Xi}$; $\widehat{\boldsymbol{\gamma}}^c$, $\widehat{\boldsymbol{\gamma}}^r$, and $\widehat{\Pi}$ are the estimated parameters previously defined; $\mathbf{1}_k$ is a $k \times 1$ vector of all ones whereas $\circ$ denotes the usual Hadamard product. Note that, the term $\mathbf{1}_k$ in the equations of effects allows to compute the total indirect effect, that is to say, the sum of the elementary indirect effects. These are the indirect effects that are associated with each of the k mediators in the model. In particular, if we omit the term $\mathbf{1}_k$ from the equations of the effects, we obtain the specific indirect effects separately for each mediator in the model.

4.1 Evaluating the size of the effects

In the context of mediation analysis, several different indices have been defined to quantify the size of the decomposed effects such as proportion mediated index, partial r^2, residual-based indices (Fairchild et al. 2009; Preacher and Kelley 2011). In this contribution, the indices λ_{DE} and λ_{IE} are used to descriptively evaluate the direct and indirect effects of the IMedA model which are obtained by decomposition of the explained variance of the model. Unlike other R^2-based decompositions, which make use of the so-called commonality analysis (e.g., see: Seibold and McPhee 1979), our indices are defined considering the so-called *reduced system* obtained by merging $\mathcal{S}_1$ with $\mathcal{S}_2$ (see Appendix C). In this new equations system, the dependent variable $\tilde{\mathbf{y}}$ is modeled as a function of all the pathways expressed by the IMedA model and its variance can be partitioned according to a dedicated regression based decomposition procedure (as described in: Mood and Graybill 1950; Fields 2003). This allows to show the contribution of the direct and indirect effects in modeling the variance of $\tilde{\mathbf{y}}$. The indices λ_{DE} and λ_{IE} are obtained as follows:

$$\lambda_{\text{DE}} = |\text{DE}_{\sigma^2_{\tilde{\mathbf{y}}}}| \cdot \Lambda^{-1} \qquad \lambda_{\text{IE}} = |\text{IE}_{\sigma^2_{\tilde{\mathbf{y}}}}| \cdot \Lambda^{-1} \qquad \text{with} \quad \lambda_{\text{DE}}, \lambda_{\text{IE}} \in [0, 1] \tag{5}$$

where $\Lambda = \text{DE}_{\sigma^2_{\tilde{\mathbf{y}}}} + \text{IE}_{\sigma^2_{\tilde{\mathbf{y}}}} + \text{RES}_{\sigma^2_{\tilde{\mathbf{y}}}}$ is equal to the observed variability explained by all the pathways in the IMedA model whereas the other components are defined as follows:

$$\text{DE}_{\sigma^2_{\tilde{\mathbf{y}}}} = \begin{bmatrix} \text{cov}\left(\mathbf{y^c}, \mathbf{x^c}\hat{\beta}^c\right) + \text{cov}(\mathbf{y^c}, \mathbf{x^r}\hat{\beta}^r) + \\ \text{cov}\left(\mathbf{y^r}, \hat{\delta}\mathbf{x^c}\hat{\beta}^c\right) + \text{cov}\left(\mathbf{y^r}, \hat{\delta}\mathbf{x^r}\hat{\beta}^r\right) \end{bmatrix} \cdot \omega^{-1} \tag{6}$$

$$\text{IE}_{\sigma^2_{\tilde{\mathbf{y}}}} = \begin{bmatrix} \text{cov}\left(\mathbf{y^c}, \mathbf{x^c}(\hat{\boldsymbol{\xi}}^c \circ \hat{\boldsymbol{\gamma}}^{c\,T})\mathbf{1}_k\right) + \text{cov}\left(\mathbf{y^c}, \mathbf{x^c}(\hat{\boldsymbol{\xi}}^c \circ \hat{\boldsymbol{\gamma}}^{r\,T} \circ \hat{\boldsymbol{\Pi}}^T)\mathbf{1}_k\right) + \\ \text{cov}\left(\mathbf{y^c}, \mathbf{x^r}(\hat{\boldsymbol{\xi}}^r \circ \hat{\boldsymbol{\gamma}}^{c\,T})\mathbf{1}_k\right) + \text{cov}\left(\mathbf{y^c}, \mathbf{x^r}(\hat{\boldsymbol{\xi}}^r \circ \hat{\boldsymbol{\gamma}}^{r\,T} \circ \hat{\boldsymbol{\Pi}}^T)\mathbf{1}_k\right) + \\ \text{cov}\left(\mathbf{y^r}, \hat{\delta}\mathbf{x^c}(\hat{\boldsymbol{\xi}}^c \circ \hat{\boldsymbol{\gamma}}^{c\,T})\mathbf{1}_k\right) + \text{cov}\left(\mathbf{y^r}, \hat{\delta}\mathbf{x^c}(\hat{\boldsymbol{\xi}}^c \circ \hat{\boldsymbol{\gamma}}^{r\,T} \circ \hat{\boldsymbol{\Pi}}^T)\mathbf{1}_k\right) + \\ \text{cov}\left(\mathbf{y^r}, \hat{\delta}\mathbf{x^r}(\hat{\boldsymbol{\xi}}^r \circ \hat{\boldsymbol{\gamma}}^{c\,T})\mathbf{1}_k\right) + \text{cov}\left(\mathbf{y^r}, \hat{\delta}\mathbf{x^r}(\hat{\boldsymbol{\xi}}^r \circ \hat{\boldsymbol{\gamma}}^{r\,T} \circ \hat{\boldsymbol{\Pi}}^T)\mathbf{1}_k\right) \end{bmatrix} \cdot \omega^{-1} \tag{7}$$

$$\text{RES}_{\sigma^2_{\tilde{\mathbf{y}}}} = \begin{bmatrix} \text{cov}\left(\mathbf{y^c}, \hat{\mathbf{E}}^c\hat{\boldsymbol{\gamma}}^c\right) + \text{cov}\left(\mathbf{y^c}, \hat{\mathbf{E}}^r\hat{\boldsymbol{\gamma}}^r\right) + \\ \text{cov}(\mathbf{y^r}, \hat{\mathbf{E}}^c\hat{\boldsymbol{\gamma}}^c\hat{\delta}) + \text{cov}\left(\mathbf{y^r}, \hat{\mathbf{E}}^r\hat{\boldsymbol{\gamma}}^r\hat{\delta}\right) \end{bmatrix} \cdot \omega^{-1} \tag{8}$$

where $\omega = [\text{var}(\mathbf{y^c}) + \text{var}(\mathbf{y^r})]$. More technical details are described in Appendix C. In the IMedA context, λ_{DE} and λ_{IE} represent the proportion of the observed variability explained by the effects of the model which is exclusively due to either the direct effect (λ_{DE}) or the indirect effect (λ_{IE}). In particular, when λ_{DE} approaches 0 (and consequently, λ_{IE} approaches 1) the representation reduces to the so-called *full mediation* case in which the variance explained by the effects is exclusively due to the mediators in the model. By contrast, if λ_{DE} approaches 1 (consequently, λ_{IE} approaches 0), then the mediation model is said to be *ill-posed* because mediators do not contribute in explaining the observed variability in the model. Note that by omitting the term $\mathbf{1}_k$ in the Eq. 7, we obtain k partial indices $\lambda^1_{\text{IE}}, \ldots, \lambda^j_{\text{IE}}, \ldots, \lambda^k_{\text{IE}}$ (with $\sum_j^k \lambda^j_{\text{IE}} = \lambda_{\text{IE}}$) where each index λ^j_{IE} represents the proportion of the variance explained by that effect which is specifically due to the corresponding mediator $\tilde{\mathbf{M}}_j$.

5 Simulation study

The aim of this simulation study is twofold. First, we will evaluate the properties of the estimators of the proposed IMedA-ALS algorithm. Although the least squares estimators for multivariate linear regression have been extensively studied in prior simulation works (e.g., see: Preacher and Hayes 2008; Zhang and Wang 2013; Nkurunziza and Ejaz Ahmed 2011; Lima Neto and de Carvalho 2008; Alkhamisi 2010; Yahya and

Olaifa 2014), in the present study we preferred to evaluate the performances of the IMedA-ALS algorithm for the sake of completeness and to further provide converging results.

Second, we will re-analyse the same simulated data by means of two alternative methods, namely a standard SEM approach and a regression-based mediation approach (2SMA). The purpose here is to evaluate whether the standard approaches can appropriately reproduce the interval-valued pathways generated by the original IMedA model representation.

The simulation study is conducted for the case of one mediator ($m = 1$) and two mediators ($m = 2$) models, respectively. However, because the results for the $m = 2$ case largely mirrored those of the simpler $m = 1$ case, in the following section we will discuss the latter case only. The results of the $m = 2$ case are provided as supplementary material of this article.

5.1 Design

Two factors were systematically varied in a complete two-factorial design:

(i) the sample size (n) at four levels: 50, 250, 500, 1000;
(ii) the amount of noise (e) at four levels: 0.10, 0.30, 0.50, 0.70. Factor e is defined as the proportion of the total variance in the data that is not accounted by the IMedA model. Technically, the proportion of error in the data is computed by modeling the variances of the error terms in the IMedA model using a predefined set of values stored in two matrices, $\mathbf{H}^E$ and $\mathbf{H}^\epsilon$, of error variances associated to the IMedA $\mathcal{S}_1$ and $\mathcal{S}_2$ regression systems, respectively. These values were defined according to a previous simulation study and reflect the condition that the proportions of explained variance accounted by the IMedA model always equal to $1 - e_k$ ($k = 1, \ldots, 4$).

5.2 Procedure

Let n_k and e_k be distinct levels of the factors n and e respectively. The following procedural steps were repeated 1000 times ($Q = 1000$) for each of the 16 combinations of levels of the simulation design:

1. Generate the interval data matrix $\tilde{\mathbf{x}}_{n_k \times 2} = ([u, v]_{ij})$ from the uniform distribution $\mathcal{U}(1, 10)$ with $[u < v]_{ij}$. Next, obtain $\mathbf{x}^\mathbf{c}$ and $\mathbf{x}^\mathbf{r}$ via the CR-parametrization on $\tilde{\mathbf{x}}_{n_k \times 2}$;
2. Generate the mediator variables $\mathbf{M^c}_{(n_k \times m)}$ and $\mathbf{M^r}_{(n_k \times m)}$ (with $m = 1$) by applying the regression system $\mathcal{S}_1$ with $\mathbf{E^c} \sim \mathcal{N}(0, \mathbf{H}^E_{k,c})$ and $\mathbf{E^r} \sim \mathcal{N}(0, \mathbf{H}^E_{k,r})$ with the following parameters: $\mathbf{A^c} = 4.8$, $\mathbf{A^r} = 3.1$, $\Xi = (2.7, 4.1)^T$, $\Pi = 2.04$;
3. Estimate the parameters $\hat{\mathbf{A}}^c_q$, $\hat{\mathbf{A}}^r_q$, $\widehat{\Xi}_q$, $\widehat{\Pi}_q$ of the system $\mathcal{S}_1$ for the q-th sample by means of the IMedA-ALS estimators (see Appendix A);
4. Generate the dependent variables $\mathbf{y^c}_{(n_k \times 1)}$ and $\mathbf{y^r}_{(n_k \times 1)}$ by applying the regression system $\mathcal{S}_2$ with $\boldsymbol{\epsilon}^c \sim \mathcal{N}(0, \mathbf{H}^\epsilon_{k,c})$, $\boldsymbol{\epsilon}^r \sim \mathcal{N}(0, \mathbf{H}^\epsilon_{k,r})$ and the following parameters: $\alpha^\mathbf{c} = 3.0$, $\alpha^\mathbf{r} = -5.3$, $\boldsymbol{\beta} = (2.3, 1.9)^T$, $\gamma^\mathbf{c} = 1.9$, $\gamma^\mathbf{r} = 0.9$, and $\delta = -3.25$.

5. Estimate the parameters $\hat{\alpha}^c_q$, $\hat{\alpha}^r_q$, $\hat{\boldsymbol{\beta}}_q$, $\hat{\gamma}^c_q$, $\hat{\gamma}^r_q$, and $\hat{\delta}_q$ of the system $\mathcal{S}_2$ for the q-th sample by means of the IMedA-ALS estimators (see Appendix A);
6. Save the estimates and proceed until $q = Q$.

This procedure was used to generate the estimated distributions of the regression parameters for each combination of levels of the simulation design. The whole procedure generated a total of $1000 \times 4 \times 4 = 16000$ new data matrices as well as an equivalent number of parameters.

5.3 Outcome measures

The sample results were evaluated considering the following *global measures* that give information about the overall performance of the IMedA-ALS estimates:

1. *Average root mean square error* (AMSE) computed as:

$$AMSE = Q^{-1} \sum_q \sqrt{J^{-1} \sum_j \left[(\widehat{\theta}_{qj} - \theta_{qj}) \cdot \theta_{qj}{}^{-1} \right]^2}$$

   with θ_q and $\widehat{\theta}_q$ being the arrays of parameters of the true and estimated model, respectively. Low values of AMSE indicate that the estimators accurately reproduce the true parameter values;
2. *Proportion of agreement* (PA) index computed as:

$$PA = Q^{-1} \sum_q \left[1 - \left(\|\widehat{\theta}_q - \theta_q\|^2 \cdot (\|\theta_q\|^2)^{-1} \right) \right] 100$$

   The index takes values in [0, 100] and assesses how much the estimated array of parameters $\hat{\theta}$ resembles the true array θ (Timmerman and Kiers 2002). When PA is closed to 100 the estimated array $\hat{\theta}$ perfectly recovers the true array θ.

5.4 Results

The first column of Table 1 reports the results of the simulation study. As expected, the AMSE index decreased almost linearly with increasing sample sizes whereas increased with increasing perturbation terms e. On the contrary, the PA index increased with increasing sample size. In particular, for $n \geq 250$, PA became unaffected by factor e. Overall, the IMedA-ALS algorithm was good and very stable also in cases of high noise terms. Clearly, these results confirmed how the ALS algorithm—upon which IMedA is based—generally shows accurate estimates. To summarize, IMedA-ALS always produced excellent estimates when the amount of noise in the data was low ($e = 0.1$) or moderate ($e = 0.3$). Moreover, also in cases of large ($e = 0.5$) or extreme amount of noise ($e = 0.7$), IMedA-ALS still showed undistorted estimates at least when $n > 50$. By contrast, for small sample sizes the performance decreased according to the amount of noise.

Table 1 Monte Carlo study: percentage of agreement (PA) index and average root mean square errors (AMSE) for the array of parameters of the single mediation model ($m = 1$)

n, e	IMedA-ALS		SEM-ML		SEM-WLS		2SMA	
	AMSE	PA	AMSE	PA	AMSE	PA	AMSE	PA
$n = 50$								
e_1	0.11	99.45	0.44	84.36	14.67	65.89	0.44	84.37
e_2	0.21	98.00	0.52	80.58	24.38	66.95	0.52	80.76
e_3	0.28	95.13	0.53	80.00	21.24	61.93	0.53	80.25
e_4	0.31	93.55	0.53	79.71	28.77	63.82	0.52	80.29
$n = 250$								
e_1	0.05	99.88	0.19	96.65	10.74	75.04	0.19	96.68
e_2	0.08	99.64	0.22	95.79	12.26	71.98	0.21	95.99
e_3	0.11	99.44	0.25	94.95	15.53	73.49	0.23	95.34
e_4	0.14	99.12	0.26	94.66	25.21	71.49	0.23	95.39
$n = 500$								
e_1	0.03	99.95	0.14	98.26	6.88	79.19	0.14	98.29
e_2	0.06	99.85	0.16	97.85	13.62	77.27	0.15	98.05
e_3	0.08	99.73	0.19	97.27	13.52	76.35	0.16	97.67
e_4	0.09	99.66	0.20	97.00	26.44	78.81	0.16	97.73
$n = 1000$								
e_1	0.02	99.98	0.10	99.17	6.79	86.15	0.10	99.20
e_2	0.04	99.92	0.12	98.78	14.16	83.31	0.11	98.98
e_3	0.05	99.87	0.15	98.37	6.97	82.51	0.12	98.77
e_4	0.07	99.84	0.16	98.27	26.89	83.50	0.11	99.00

5.5 Further analysis, results, limitations

We also evaluated the performances of a standard SEM approach for single-valued data and a least squares procedure for standard mediation analysis in reconstructing the data generated in the previous simulation design (see also Fig. 2b). In particular, SEM model fitting and estimation were implemented through the Lavaan R-package (Rosseel 2012) using the standard SEM representation for mediation analysis with both ML and WLS estimation procedures.[1] On the contrary, the standard regression approach, named 2SMA, was instead implemented through a combination of Matlab scripts that modeled a step-by-step regression procedure (the estimation algorithm is provided as supplementary material to this article). Finally, the AMSE and PA measures were computed for each condition of the simulation design and each of the three estimation procedures (SEM-ML, SEM-WLS, 2SMA).

The second, third, and fourth columns of the Table 1 report the ensuing results. An inspection of Table 1 suggests how SEM-ML and 2SMA produced comparable results

[1] In both SEM-ML and SEM-WLS procedures, interval variables were defined as standard single-valued variables in terms of centers and ranges. The R code for the Lavaan mediation model for interval variables is provided as Supplementary Material.

in estimating the IMedA parameters. As expected, both the estimation algorithms produced lower errors with increasing sample size. By contrast, SEM-WLS always showed lower performances and accuracies than the other two methods. However, it should be noted that all performances of SEM-ML, SEM-WLS, and 2SMA were not as good as those obtained using the original IMedA-ALS algorithm. Indeed, although SEM-ML and 2SMA provided acceptable results in resembling the true model structure provided by the IMedA pathways, they still showed an higher value of AMSE for each condition of the simulation design even when increasing sample sizes were considered. By and large, this can reflect relevant structural differences in the modeling and estimation procedures adopted by the other two alternative approaches. However, it should be emphasized that the simulation scenario used in this study may be a concern when more complex empirical cases are considered. But despite this limitation, we preferred to start with this simplified condition to better evaluate our IMedA-ALS method under the purest and simplest reliability scenario. The results we got largely resembled what is already known about the properties of alternating least squares estimators (e.g., see: Takane et al. 1977; Kim and Park 2007). Furthermore, other extensive studies are needed to investigate in depth the performances of IMedA-ALS algorithm with respect to other approaches (e.g., SEM) which could potentially be used with interval-data.

6 An empirical application: role and work-related burnout

Role is an important variable in many organizational research settings (Sawyer 1992; Toderi et al. 2013) and is considered a relevant dimension in predicting employee health and, more in general, organizational stress (Bliese and Castro 2000). Work-related burnout, instead, corresponds to a protracted individual response to a set of emotional and interpersonal stressors which are presented in the organization (Alarcon 2011). In this illustrative example we tested a model in which the basic linear relation between role (R) and work-related burnout (WB) was evaluated by considering job satisfaction (S) and workload (WL) as mediators.

Measures In this first application we used some recently published data (Avanzi et al. 2012). In particular, data refers to Italian teachers who participated in a psychosocial risk assessment evaluation conducted in five schools in the Trentino region (north-east of Italy). Data were collected by means of a specific questionnaire which was administered in two different occasions (T1 and T2). The variables were as follows: Job satisfaction (S) was codified by asking participants to indicate to what extent they were satisfied with their present job by responding on a single-item scale (ranging from 1: very little; 5: very much). Instead, work-related burnout (WB), which refers to the degree of fatigue and exhaustion perceived by workers as being related to their work (Kristensen et al. 2005; Avanzi et al. 2013), was measured by means of two different five-option response formats, one for intensity and the other for frequency (Kristensen et al. 2005). Finally, both workload (WL) and role (R) were measured by using an Italian version of the Indicator Tool developed by Edwards et al. (2008) and adapted by Toderi et al. (2013). In particular, workload was measured by eight items whereas role by five items. Response formats for both scales varied from 1 (never or strongly disagree, depending on the item) to 5 (often or strongly agree, depending on the item).

Table 2 Case study: estimated model parameters and effects

	First mediator (WL)		Second mediator (S)	
	Values	95% CIs	Values	95% CIs
Model parameters (M)				
$\mathbf{A}^c$	3.92	(2.88, 4.88)	0.81	(−0.45, 1.94)
$\mathbf{A}^r$	0.36	(−1.59, 2.46)	0.17	(−0.99, 1.42)
Π	−0.08	(−0.73, 0.69)	0.08	(−0.27, 0.38)
$\boldsymbol{\xi}^c$	−0.26	(−0.50, −0.09)	0.68	(0.41, 0.96)
$\boldsymbol{\xi}^r$	0.17	(0.09, 0.55)	0.18	(−0.14, 0.57)
Model parameters (Y)				
α^c	1.50	(−0.22, 2.37)		
α^l	0.18	(0.03, 0.38)		
δ	0.03	(0.05, 0.25)		
β^c	0.06	(−0.06, 0.39)		
β^r	0.21	(0.04, 0.46)		
$\boldsymbol{\gamma}^c$	0.68	(0.54, 0.81)	−0.38	(−0.48, −0.24)
$\boldsymbol{\gamma}^r$	0.14	(0.07, 0.39)	−0.02	(−0.16, 0.07)
Effects				
DE_c	0.06	(−0.06, 0.38)		
DE_r	0.19	(0.05, 0.46)		
IE_c	−0.16	(−0.33, −0.02)	−0.27	(−0.40, −0.12)
IE_r	0.12	(0.02, 0.39)	−0.10	(−0.24, −0.01)

$R_M^2 = 0.39$, $R_Y^2 = 0.48$, $\lambda_{\mathrm{DE}} = 0.22$, $\lambda_{\mathrm{IE}}^1 = 0.36$, $\lambda_{\mathrm{IE}}^2 = 0.42$, CIs indicate the 95% confidence intervals obtained by bias-corrected and accelerating (BCa) bootstrap with 5000 bootstrap samples (see Appendix D for further details about the BCa procedure)

Data The sample was composed of $n = 140$ teachers (83% females), with a mean age of 41.2 years (ranging from 23 to 62 with SD = 10.7) and a mean workload of 19.2 years (ranging from 2 to 37 with SD = 10.3). Because in this case the variables are collected longitudinally, we decided to pre-process the T1–T2 variables according to the *response feature analysis* (Senn et al. 2000; Everitt 1995). In particular, considering $\min(T1, T2)$ and $\max(T1, T2)$ as the lower and the upper bounds of intervals, centers and ranges were defined according to the CR-parametrization. Note that, in this case, centers express the averaged point-valued evaluation on a 5-point scale whereas ranges directly refer to a *change score* over the time.

Data analysis and results The IMedA algorithm required 781 iterations to estimate the parameters in the $\mathcal{S}_1$ system and 1084 iterations to estimate those in the $\mathcal{S}_2$ system. The performance of the IMedA model were acceptable ($R_M^2 = 0.39$ and $R_Y^2 = 0.48$) and in line with the literature (Alarcon 2011; Sutton 1998). Table 2 shows the final estimated parameters together with the corresponding direct and indirect effects for the model. A quick inspection of Table 2 reveals that the variance explained by all the pathways of the model was $\Lambda = 0.60$ where the direct effect, the first mediator (WL), and the second mediator (S) contributed for 22, 36, and 42% of the variance, respectively. The results indicate that the centers of R did not have a linear impact on WB ($\beta^c = 0.06$) whereas the ranges of R showed a weak but significant impact on WB ($\beta^r = 0.21$). The centers of R were negatively related to WL ($\xi_1^c = -0.26$) but positively associated with S ($\xi_2^c = 0.68$). Considering the second regression system $\mathcal{S}_2$, the centers and ranges of WL were positively associated with WB ($\gamma_1^c = 0.68$ and $\gamma_1^r = 0.14$), whereas the centers and ranges of S were negatively related to WB ($\gamma_2^c = -0.38$, $\gamma_2^r = -0.02$). Interestingly, in this example the direct effect of R was mainly produced by the ranges ($\text{DE}_r = 0.19$). By contrast, the centers did not show a significant effect ($\text{DE}_c = 0.06$). The results also suggest that the indirect effects for the centers and ranges were both significant. In particular, the indirect effect through the centers of WL ($\text{IE}_c = -0.16$) was negative and weaker than the corresponding effect through S ($\text{IE}_c = -0.27$). Instead, the indirect effect through the ranges of WL ($\text{IE}_r = 0.12$) and S ($\text{IE}_r = -0.10$) showed opposite directions and both the effects were significant. Overall, considering the centers component, the results indicated that R has a protective impact on WB by reducing the positive relation between WL and WB. In a similar way, R improves the negative relation between S and WB, that is to say, the more the role is perceived as clear by the teacher, the higher is the perceived work satisfaction and the lower is the perceived work-related burnout. By contrast, considering the ranges component, R improves the positive impact of WL on WB whereas it makes the relation between S and WB to vanish. Because in this context ranges are interpreted as change scores, this result would possibly highlight how the more the teachers experience unstable clarity of role, the more they experience workload and work-related burnout.

7 Conclusions

In this article, we developed a novel and simple model (IMedA) to perform mediation analysis on interval-valued variables. As far as we know, IMedA is the first proposal that is devoted to mediation analyses of interval data. Globally, the main characteristic of this model is its use of two linear equations systems for modeling the interval pathways among the independent, mediators, and dependent observed variables. This involved the extension of the well-known Stolzenberg's decomposition to handle with interval-valued causal effects. Relatedly, a set of variance-based indices was also defined to quantify the sizes of such effects in the interval context. Finally, we used a simulation study and a real application to highlight some characteristics of the proposed model. In particular, the simulation study revealed that the IMedA model is sufficiently accurate to reproduce the observed relationships among the interval variables. Moreover, our findings also showed how IMedA outperforms existing mediation approaches for single-valued variables that might be eventually used in modeling interval pathways.

7.1 Model's advantages

One nice property of the IMedA representation is that in estimating the model's parameters, the IMedA algorithm works similarly to an alternating recursive two-steps procedure where the reconstruction of ranges proceeds conditionally on the reconstruction of centers. This centre-range dependence assumption is straightforward in an interval framework, where both the centers and ranges are key components in determining the observed interval data structures. Moreover, IMedA-ALS allows a simple generalization of the single-valued case as it subsumes the mediation model for single-valued variables as a special case. Indeed, when data are expressed in terms of degenerated intervals, the CR-parametrization always boils down to single-valued variables (with $\mathbf{x}^{\mathbf{r}} = \mathbf{y}^{\mathbf{r}} = \mathbf{0}_{n\times 1}$ and $\mathbf{M}^{\mathbf{r}} = \mathbf{0}_{n\times k}$) and the regression systems $\mathcal{S}_1$ and $\mathcal{S}_2$ simply reduce to the regressions for the ordinary multiple mediation analysis (MedA), namely $\mathbf{M} = \mathbf{1A} + \mathbf{x}\xi + \mathbf{E}$ and $\mathbf{y} = \mathbf{1}\alpha + \mathbf{x}\beta + \mathbf{M}\boldsymbol{\gamma} + \boldsymbol{\epsilon}$ (note that in this special case also the IMedA estimators as well as the effects decomposition reduce accordingly).

7.2 Model's limitations

However, as for other statistical procedures, also the proposed method can potentially suffer from some limitations. First, if the IMedA model is fitted to empirical data which are largely corrupted by noise, the corresponding estimates may violate the natural constraints of interval-valued data (namely: $\mathbf{M}^{\mathbf{r}} > \mathbf{0}_{n\times k}$ and $\mathbf{y}^{\mathbf{r}} > \mathbf{0}_{n\times 1}$) thus possibly yielding unfeasible solutions. In these situations, a constrained version of the algorithm based on specific optimization techniques should instead be used (Lima Neto and de Carvalho 2010; Carpita and Ciavolino 2017). However, for the standard unconstrained algorithm, a simple way out might consist in setting to zero all the negative range estimates so that their natural constraints are numerically satisfied. Second, some empirical contexts may require more complex models to better evaluate the relationships among the observed variables. For instance, we may think to moderated mediation models in which the indirect paths are partially or completely moderated by other intervening variables (e.g., age, gender, income. See: Edwards and Lambert 2007). Lastly, the standard assumption that mediators work in parallel could be contrived in some particular contexts where models with serial mediators would be instead preferred (Taylor et al. 2008). In this latter case, the IMedA representation appears clearly inadequate to achieve such an advantage.

7.3 Future extensions

Various possible extensions of our approach can be considered in future works. For example, moderated mediation models for interval-valued data would allow the modeling of more complex situations in which the mediated paths vary as a function of third moderator variables. Likewise, interval mediation models with serial mediators may also extend our proposal to represent situations in which researchers evaluate three-path mediated effects. Further, also modeling correlational paths among the mediator directly, by means of parametric covariance matrix estimators, can constitute a future target. Several empirical situations may also require the use of mediation models handling with non-linear decomposition of effects (e.g., see: Hayes and Preacher 2010). The extension of the IMedA representation to deal with non-linear pathways among variables can surely be considered an interesting future extension of the present article. Finally, further simulation and benchmarking studies could be considered to extensively assess IMedA-ALS properties with regards to statistical methods that can be used in the case of interval-valued data.

Appendix A: Solutions for IMedA model

$$\mathrm{vec}(\widehat{\mathbf{A}}^r) = (\mathbf{I}_k \otimes \mathbf{1}^T\mathbf{1})^{-1} \cdot (\mathbf{I}_k \otimes \mathbf{1})^T\, \mathrm{vec}(\mathbf{M}^\mathbf{r} - \mathbf{X}\Xi\mathbf{\Pi} - \mathbf{1}\mathbf{A}^c\mathbf{\Pi}); \tag{A1}$$

$$\mathrm{vec}(\widehat{\mathbf{\Pi}}) = \begin{bmatrix} (\mathbf{I}_k \otimes \Xi^T\mathbf{X}^T\mathbf{X}\Xi)+ \\ (\mathbf{I}_k \otimes \Xi^T\mathbf{X}^T\mathbf{1}\mathbf{A}^c)+ \\ (\mathbf{I}_k \otimes \mathbf{A}^{cT}\mathbf{1}^T\mathbf{1}\mathbf{A}^c) \end{bmatrix}^{-1} \cdot \begin{bmatrix} (\mathbf{I}_k \otimes \mathbf{X}\Xi)^T\, \mathrm{vec}(\mathbf{M}^\mathbf{r} - \mathbf{1}\mathbf{A}^\mathbf{r})+ \\ (\mathbf{I}_k \otimes \mathbf{1}\mathbf{A}^c)^T\, \mathrm{vec}(\mathbf{M}^\mathbf{r} - \mathbf{1}\mathbf{A}^\mathbf{r}) \end{bmatrix}; \tag{A2}$$

$$\mathrm{vec}(\widehat{\Xi}) = \begin{bmatrix} (\mathbf{I}_k \otimes \mathbf{X}^T\mathbf{X})+ \\ (\mathbf{\Pi} \otimes \mathbf{X}^T\mathbf{X}) \end{bmatrix}^{-1} \cdot \begin{bmatrix} (\mathbf{I}_k \otimes \mathbf{X})^T\, \mathrm{vec}(\mathbf{M}^\mathbf{c} - \mathbf{1}\mathbf{A}^c)+ \\ (\mathbf{\Pi} \otimes \mathbf{X})^T\, \mathrm{vec}(\mathbf{M}^\mathbf{r} - \mathbf{1}\mathbf{A}^r - \mathbf{1}\mathbf{A}^c\mathbf{\Pi}) \end{bmatrix}; \tag{A3}$$

$$\mathrm{vec}(\widehat{\mathbf{A}}^c) = \begin{bmatrix} (\mathbf{I}_k \otimes \mathbf{1}^T\mathbf{1})+ \\ (\mathbf{\Pi} \otimes \mathbf{1}^T\mathbf{1}) \end{bmatrix}^{-1} \cdot \begin{bmatrix} (\mathbf{I}_k \otimes \mathbf{1})^T\, \mathrm{vec}(\mathbf{M}^\mathbf{c} - \mathbf{X}\Xi)+ \\ (\mathbf{\Pi} \otimes \mathbf{1})^T\, \mathrm{vec}(\mathbf{M}^\mathbf{r} - \mathbf{1}\mathbf{A}^r - \mathbf{X}\Xi\mathbf{\Pi}) \end{bmatrix}; \tag{A4}$$

$$\widehat{\delta} = \begin{bmatrix} \alpha^c\mathbf{1}^T\mathbf{1}\alpha^c + 2\alpha^c\mathbf{1}^T\mathbf{X}\boldsymbol{\beta} + 2\alpha^c\mathbf{1}^T\mathbf{M}\boldsymbol{\gamma}+ \\ \boldsymbol{\beta}^T\mathbf{X}^T\mathbf{1}\alpha^c + 2\boldsymbol{\beta}^T\mathbf{X}^T\mathbf{X}\boldsymbol{\beta} + 2\boldsymbol{\beta}^T\mathbf{X}^T\mathbf{M}\boldsymbol{\gamma}+ \\ \boldsymbol{\gamma}^T\mathbf{M}^T\mathbf{1}\alpha^c + 2\boldsymbol{\gamma}^T\mathbf{M}^T\mathbf{X}\boldsymbol{\beta} + 2\boldsymbol{\gamma}^T\mathbf{M}^T\mathbf{M}\boldsymbol{\gamma} \end{bmatrix}^{-1} \cdot \begin{bmatrix} \alpha^c\mathbf{1}^T(\mathbf{y}^r - \mathbf{1}\alpha^r)+ \\ \boldsymbol{\beta}^T\mathbf{X}^T(\mathbf{y}^r - \mathbf{1}\alpha^r)+ \\ \boldsymbol{\gamma}^T\mathbf{M}^T(\mathbf{y}^r - \mathbf{1}\alpha^r) \end{bmatrix}; \tag{A5}$$

$$\widehat{\boldsymbol{\beta}} = \begin{bmatrix} \mathbf{X}^T\mathbf{X}+ \\ \delta\mathbf{X}^T\mathbf{X}\delta \end{bmatrix}^{-1} \cdot \begin{bmatrix} \mathbf{X}^T(\mathbf{y}^c - \mathbf{1}\alpha^c - \mathbf{M}\boldsymbol{\gamma})+ \\ \mathbf{X}^T[\,\mathbf{y}^r - \mathbf{1}\alpha^r + (-\mathbf{1}\alpha^c - \mathbf{M}\boldsymbol{\gamma})\delta\,]\delta \end{bmatrix}; \tag{A6}$$

$$\widehat{\boldsymbol{\gamma}^c} = \begin{bmatrix} \mathbf{M}^{\mathbf{c}T}\mathbf{M}^\mathbf{c}+ \\ \delta\mathbf{M}^{\mathbf{c}T}\mathbf{M}^\mathbf{c}\delta \end{bmatrix}^{-1} \cdot \begin{bmatrix} \mathbf{M}^{\mathbf{c}T}(\mathbf{y}^c - \mathbf{1}\alpha^c - \mathbf{X}\boldsymbol{\beta} - \mathbf{M}^\mathbf{r}\boldsymbol{\gamma}^\mathbf{r})+ \\ \mathbf{M}^{\mathbf{c}T}[\,\mathbf{y}^r - \mathbf{1}\alpha^r + (-\mathbf{1}\alpha^c - \mathbf{X}\boldsymbol{\beta} - \mathbf{M}^\mathbf{r}\boldsymbol{\gamma}^\mathbf{r})\delta\,]\delta \end{bmatrix}; \tag{A7}$$

$$\widehat{\boldsymbol{\gamma}^r} = \begin{bmatrix} \mathbf{M}^{\mathbf{r}T}\mathbf{M}^\mathbf{r}+ \\ \delta\mathbf{M}^{\mathbf{r}T}\mathbf{M}^\mathbf{r}\delta \end{bmatrix}^{-1} \cdot \begin{bmatrix} \mathbf{M}^{\mathbf{r}T}(\mathbf{y}^r - \mathbf{1}\alpha^r - \mathbf{X}\boldsymbol{\beta} - \mathbf{M}^\mathbf{c}\boldsymbol{\gamma}^\mathbf{c})+ \\ \mathbf{M}^{\mathbf{r}T}[\,\mathbf{y}^r - \mathbf{1}\alpha^r + (-\mathbf{1}\alpha^r - \mathbf{X}\boldsymbol{\beta} - \mathbf{M}^\mathbf{c}\boldsymbol{\gamma}^\mathbf{c})\delta\,]\delta \end{bmatrix}; \tag{A8}$$

$$\widehat{\alpha}^c = \frac{1}{n(1+\delta^2)} \cdot \begin{bmatrix} \mathbf{1}^T(\mathbf{y}^c - \mathbf{X}\boldsymbol{\beta} - \mathbf{M}\boldsymbol{\gamma})+ \\ \mathbf{1}^T[\,\mathbf{y}^r - \mathbf{1}\alpha^r + (-\mathbf{X}\boldsymbol{\beta} - \mathbf{M}\boldsymbol{\gamma})\delta\,]\delta \end{bmatrix}; \tag{A9}$$

$$\widehat{\alpha}^r = \frac{1}{n} \cdot \mathbf{1}^T[\,\mathbf{y}^\mathbf{r} - (\mathbf{1}\alpha^c - \mathbf{X}\boldsymbol{\beta} - \mathbf{M}\boldsymbol{\gamma})\delta]; \tag{A10}$$

where vec(.) is the linear operator that converts a $n \times k$ matrix into a $kn \times 1$ vector, $\otimes$ denotes the Kronecker product, $\mathbf{I}_k$ is a $k \times k$ identity matrix whereas $\mathbf{1}$ is a $n \times k$ matrix of all ones.

Appendix B: Decomposition of effects for IMedA model

In order to derive direct and indirect effects for the IMedA model, we proceed as follows. Consider the regression systems $\mathcal{S}_1$ and $\mathcal{S}_2$ shown in Eq. 1:

$$\mathcal{S}_1 : \begin{cases} \mathbf{M^c} = \mathbf{1A}^c + \mathbf{X}\Xi + \mathbf{E}^c \\ \mathbf{M^r} = \mathbf{1A}^r + (\mathbf{1A}^c + \mathbf{X}\Xi)\mathbf{\Pi} + \mathbf{E^r} \end{cases}$$

$$\mathcal{S}_2 : \begin{cases} \mathbf{y^c} = \mathbf{1}\alpha^c + \mathbf{X}\boldsymbol{\beta} + \mathbf{M^c}\boldsymbol{\gamma^c} + \mathbf{M^r}\boldsymbol{\gamma^r} + \boldsymbol{\epsilon^c} \\ \mathbf{y^r} = \mathbf{1}\alpha^r + (\mathbf{1}\alpha^c + \mathbf{X}\boldsymbol{\beta} + \mathbf{M^c}\boldsymbol{\gamma^c} + \mathbf{M^r}\boldsymbol{\gamma^r})\delta + \boldsymbol{\epsilon^r} \end{cases}$$

Firstly, substitute the equations of $\mathbf{M^c}$ and $\mathbf{M^r}$ into $\mathbf{y^c}$ and $\mathbf{y^r}$, as follows:

$$\mathcal{S}'_2 : \begin{cases} \mathbf{y^c} = \mathbf{1}\alpha^c + \mathbf{X}\boldsymbol{\beta} + [\mathbf{1A}^c + \mathbf{X}\Xi + \mathbf{E}^c]\boldsymbol{\gamma}^c + \\ \quad + [\mathbf{1A}^r + (\mathbf{1A}^c + \mathbf{X}\Xi)\mathbf{\Pi} + \mathbf{E^r}]\gamma^r + \boldsymbol{\epsilon}^c \\ \mathbf{y^r} = \mathbf{1}\alpha^r + (\mathbf{1}\alpha^c + \mathbf{X}\boldsymbol{\beta} + [\mathbf{1A}^c + \mathbf{X}\Xi + \mathbf{E}^c]\boldsymbol{\gamma}^c + \\ \quad + [\mathbf{1A}^r + (\mathbf{1A}^c + \mathbf{X}\Xi)\mathbf{\Pi} + \mathbf{E^r}]\gamma^r)\delta^r_y + \boldsymbol{\epsilon^r} \end{cases} \tag{B1}$$

Multiplying through and expanding terms, using a little algebra, we obtain the following reduced form system $\mathcal{S}'_2$:

$$\mathcal{S}'_2 : \begin{cases} \mathbf{y^c} = \mathbf{1}\alpha^c + \mathbf{1A}^c\boldsymbol{\gamma}^c + \mathbf{1A}^r\boldsymbol{\gamma}^r + \mathbf{1A}^c\mathbf{\Pi}\boldsymbol{\gamma}^r + \mathbf{x^c}[\boldsymbol{\beta}^c + \boldsymbol{\xi}^c(\boldsymbol{\gamma}^c + \mathbf{\Pi}\boldsymbol{\gamma}^r)] \\ \quad + \mathbf{x^r}[\boldsymbol{\beta}^r + \boldsymbol{\xi}^r(\boldsymbol{\gamma}^c + \mathbf{\Pi}\boldsymbol{\gamma}^r)] + \mathbf{E}^c\boldsymbol{\gamma}^c + \mathbf{E}^r\boldsymbol{\gamma}^r + \boldsymbol{\epsilon}^c \\ \mathbf{y^r} = \mathbf{1}\alpha^r + \mathbf{1}\alpha^c\delta + \mathbf{1A}^c\boldsymbol{\gamma}^c\delta + \mathbf{1A}^r\boldsymbol{\gamma}^r\delta + \mathbf{1A}^c\mathbf{\Pi}\boldsymbol{\gamma}^r\delta + \\ \quad + \mathbf{x^c}[\boldsymbol{\beta}^c + \boldsymbol{\xi}^c(\boldsymbol{\gamma}^c + \mathbf{\Pi}\boldsymbol{\gamma}^r)]\delta + \mathbf{x^r}[\boldsymbol{\beta}^r + \boldsymbol{\xi}^r(\boldsymbol{\gamma}^c + \mathbf{\Pi}\boldsymbol{\gamma}^r)]\delta \\ \quad + \mathbf{E}^c\boldsymbol{\gamma}^c\delta + \mathbf{E}^r\boldsymbol{\gamma}^r\delta + \boldsymbol{\epsilon}^r \end{cases} \tag{B2}$$

Next, taking the partial derivatives of $\mathbf{y^c}$ and $\mathbf{y^r}$ with respect to $\mathbf{x^c}$ and $\mathbf{x^r}$ we have the equations for the total effect (TE) of the model, as follows:

$$\begin{cases} \frac{\partial\, \mathbf{y^c}}{\partial\, \mathbf{x^c}} = \delta\beta^c + \boldsymbol{\xi}^c(\boldsymbol{\gamma}^c + \mathbf{\Pi}\boldsymbol{\gamma}^r) & \frac{\partial\, \mathbf{y^c}}{\partial\, \mathbf{x^r}} = \delta\beta^r + \boldsymbol{\xi}^r(\boldsymbol{\gamma}^c + \mathbf{\Pi}\boldsymbol{\gamma}^r) \\ \frac{\partial\, \mathbf{y^r}}{\partial\, \mathbf{x^c}} = \delta\beta^c + \boldsymbol{\xi}^c(\boldsymbol{\gamma}^c + \mathbf{\Pi}\boldsymbol{\gamma}^r)\delta & \frac{\partial\, \mathbf{y^r}}{\partial\, \mathbf{x^r}} = \delta\beta^r + \boldsymbol{\xi}^r(\boldsymbol{\gamma}^c + \mathbf{\Pi}\boldsymbol{\gamma}^r)\delta \end{cases}$$

Finally, collecting and simplifying the ensuing terms, we obtain the following equations for TE:

$$\begin{aligned} \mathrm{TE}^{y^c} &= \beta^c_y + \beta^r_y + (\boldsymbol{\beta}^c_m \circ \boldsymbol{\gamma}^{c\,T}_m)\mathbf{1}_m + (\boldsymbol{\beta}^c_m \circ \boldsymbol{\gamma}^{r\,T}_m \circ \mathbf{\Pi})\mathbf{1}_m + (\boldsymbol{\beta}^r_m \circ \boldsymbol{\gamma}^{c\,T}_m)\mathbf{1}_m + \\ &\quad + (\boldsymbol{\beta}^r_m \circ \boldsymbol{\gamma}^{r\,T}_m \circ \mathbf{\Pi})\mathbf{1}_m \\ \mathrm{TE}^{y^r} &= \delta^r_y[\beta^c_y + \beta^r_y + (\boldsymbol{\beta}^c_m \circ \boldsymbol{\gamma}^{c\,T}_m)\mathbf{1}_m + (\boldsymbol{\beta}^c_m \circ \boldsymbol{\gamma}^{r\,T}_m \circ \mathbf{\Pi})\mathbf{1}_m + \\ &\quad + (\boldsymbol{\beta}^r_m \circ \boldsymbol{\gamma}^{c\,T}_m)\mathbf{1}_m + (\boldsymbol{\beta}^r_m \circ \boldsymbol{\gamma}^{r\,T}_m \circ \mathbf{\Pi})\mathbf{1}_m] \end{aligned} \tag{B3}$$

which are in the general form of $\mathrm{TE} = \mathrm{DE}_c + \mathrm{DE}_r + \mathrm{IE}_{c/c} + \mathrm{IE}_{c/r} + \mathrm{IE}_{r/c} + \mathrm{IE}_{r/r}$. Note that the equation TE^{y^r} for $\mathbf{y^r}$ is obtained as linear combination of TE^{y^c} through the parameter δ.

Appendix C: Decomposition of variance for IMedA model

Considering the reduced form system $\mathcal{S}'_2$:

$$\mathcal{S}'_2 : \begin{cases} \mathbf{y^c} = \mathbf{1}\alpha^c + \mathbf{1A}^c\boldsymbol{\gamma}^c + \mathbf{1A}^r\boldsymbol{\gamma}^r + \mathbf{1A}^c\boldsymbol{\Pi}\boldsymbol{\gamma}^r + \mathbf{x^c}[\boldsymbol{\beta}^c + \boldsymbol{\xi}^c(\boldsymbol{\gamma}^c + \boldsymbol{\Pi}\boldsymbol{\gamma}^r)] \\ \qquad + \mathbf{x^r}[\boldsymbol{\beta}^r + \boldsymbol{\xi}^r(\boldsymbol{\gamma}^c + \boldsymbol{\Pi}\boldsymbol{\gamma}^r)] + \mathbf{E}^c\boldsymbol{\gamma}^c + \mathbf{E}^r\boldsymbol{\gamma}^r + \boldsymbol{\epsilon}^c \\ \mathbf{y^r} = \mathbf{1}\alpha^r + \mathbf{1}\alpha^c\delta + \mathbf{1A}^c\boldsymbol{\gamma}^c\delta + \mathbf{1A}^r\boldsymbol{\gamma}^r\delta + \mathbf{1A}^c\boldsymbol{\Pi}\boldsymbol{\gamma}^r\delta + \\ \qquad + \mathbf{x^c}[\boldsymbol{\beta}^c + \boldsymbol{\xi}^c(\boldsymbol{\gamma}^c + \boldsymbol{\Pi}\boldsymbol{\gamma}^r)]\delta + \mathbf{x^r}[\boldsymbol{\beta}^r + \boldsymbol{\xi}^r(\boldsymbol{\gamma}^c + \boldsymbol{\Pi}\boldsymbol{\gamma}^r)]\delta \\ \qquad + \mathbf{E}^c\boldsymbol{\gamma}^c\delta + \mathbf{E}^r\boldsymbol{\gamma}^r\delta + \boldsymbol{\epsilon}^r \end{cases} \tag{C1}$$

the following identities hold:

$$\begin{aligned} \mathrm{var}(\mathbf{y^c}) = {} & \mathrm{cov}(\mathbf{y^c}, \mathbf{x^c}\beta^c) + \\ & + \mathrm{cov}(\mathbf{y^c}, \mathbf{x^r}\beta^r) + \mathrm{cov}(\mathbf{y^c}, \mathbf{x^c}(\boldsymbol{\xi}^c \circ \boldsymbol{\gamma}^{cT})\mathbf{1}_k) + \\ & + \mathrm{cov}(\mathbf{y^c}, \mathbf{x^c}(\boldsymbol{\xi}^c \circ \boldsymbol{\gamma}^{rT} \circ \boldsymbol{\Pi}^T)\mathbf{1}_k) + \\ & + \mathrm{cov}(\mathbf{y^c}, \mathbf{x^r}(\boldsymbol{\xi}^r \circ \boldsymbol{\gamma}^{cT})\mathbf{1}_k) + \mathrm{cov}(\mathbf{y^c}, \mathbf{x^r}(\boldsymbol{\xi}^r \circ \boldsymbol{\gamma}^{rT} \circ \boldsymbol{\Pi}^T)\mathbf{1}_k) + \\ & + \mathrm{cov}(\mathbf{y^c}, \mathbf{E}^c\boldsymbol{\gamma}^c) + \mathrm{cov}(\mathbf{y^c}, \mathbf{E}^r\boldsymbol{\gamma}^r) + \mathrm{cov}(\mathbf{y^c}, \boldsymbol{\epsilon}^c) \\ \mathrm{var}(\mathbf{y^r}) = {} & \mathrm{cov}(\mathbf{y^r}, \delta\mathbf{x^c}\beta^c) + \mathrm{cov}(\mathbf{y^r}, \delta\mathbf{x^r}\beta^r) + \mathrm{cov}(\mathbf{y^r}, \delta\mathbf{x^c}(\boldsymbol{\xi}^c \circ \boldsymbol{\gamma}^{cT})\mathbf{1}_k) + \\ & + \mathrm{cov}(\mathbf{y^r}, \delta\mathbf{x^c}(\boldsymbol{\xi}^c \circ \boldsymbol{\gamma}^{rT} \circ \boldsymbol{\Pi}^T)\mathbf{1}_k) + \mathrm{cov}(\mathbf{y^r}, \delta\mathbf{x^r}(\boldsymbol{\xi}^r \circ \boldsymbol{\gamma}^{cT})\mathbf{1}_k) + \\ & + \mathrm{cov}(\mathbf{y^r}, \delta\mathbf{x^r}(\boldsymbol{\xi}^r \circ \boldsymbol{\gamma}^{rT} \circ \boldsymbol{\Pi}^T)\mathbf{1}_k) + \\ & + \mathrm{cov}(\mathbf{y^r}, \mathbf{E}^c\boldsymbol{\gamma}^c\delta) + \mathrm{cov}(\mathbf{y^r}, \mathbf{E}^r\boldsymbol{\gamma}^r\delta) + \mathrm{cov}(\mathbf{y^c}, \boldsymbol{\epsilon}^r) \end{aligned} \tag{C2}$$

after noticing that:

$$\mathrm{cov}(\mathbf{y^c}, \mathbf{1}\alpha^c + \mathbf{1A}^c\boldsymbol{\gamma}^c + \mathbf{1A}^r\boldsymbol{\gamma}^r + \mathbf{1A}^c\boldsymbol{\Pi}\boldsymbol{\gamma}^r) = 0$$
$$\mathrm{cov}(\mathbf{y^c}, \mathbf{1}\alpha^r + \mathbf{1}\alpha^c\delta + \mathbf{1A}^c\boldsymbol{\gamma}^c\delta + \mathbf{1A}^r\boldsymbol{\gamma}^r\delta + \mathbf{1A}^c\boldsymbol{\Pi}\boldsymbol{\gamma}^r\delta) = 0$$

where cov(.) and var(.) indicates the covariance and variance operators, $\circ$ denotes the Hadamard product whereas $\mathbf{1}_k$ is a $k \times 1$ vector of all ones. The following properties hold:

$$\begin{bmatrix} \text{cov}(\mathbf{y^c}, \mathbf{x^c}\beta^c) + \text{cov}(\mathbf{y^c}, \mathbf{x^r}\beta^r)+ \\ \text{cov}(\mathbf{y^c}, \mathbf{x^c}(\boldsymbol{\xi}^c \circ \boldsymbol{\gamma}^{cT})\mathbf{1}_k) + \text{cov}(\mathbf{y^c}, \mathbf{x^c}(\boldsymbol{\xi}^c \circ \boldsymbol{\gamma}^{rT} \circ \boldsymbol{\Pi}^T)\mathbf{1}_k)+ \\ \text{cov}(\mathbf{y^c}, \mathbf{x^r}(\boldsymbol{\xi}^r \circ \boldsymbol{\gamma}^{cT})\mathbf{1}_k) + \text{cov}(\mathbf{y^c}, \mathbf{x^r}(\boldsymbol{\xi}^r \circ \boldsymbol{\gamma}^{rT} \circ \boldsymbol{\Pi}^T)\mathbf{1}_k)+ \\ \text{cov}(\mathbf{y^c}, \mathbf{E}^c\boldsymbol{\gamma}^c) + \text{cov}(\mathbf{y^c}, \mathbf{E}^r\boldsymbol{\gamma}^r) + \text{cov}(\mathbf{y^c}, \boldsymbol{\epsilon}^c) \end{bmatrix} \left[\text{var}(\mathbf{y^c})\right]^{-1} = 1$$

$$\begin{bmatrix} \text{cov}(\mathbf{y^r}, \delta\mathbf{x^c}\beta^c) + \text{cov}(\mathbf{y^r}, \delta\mathbf{x^r}\beta^r)+ \\ \text{cov}(\mathbf{y^r}, \delta\mathbf{x^c}(\boldsymbol{\xi}^c \circ \boldsymbol{\gamma}^{cT})\mathbf{1}_k) + \text{cov}(\mathbf{y^r}, \delta\mathbf{x^c}(\boldsymbol{\xi}^c \circ \boldsymbol{\gamma}^{rT} \circ \boldsymbol{\Pi}^T)\mathbf{1}_k)+ \\ \text{cov}(\mathbf{y^r}, \delta\mathbf{x^r}(\boldsymbol{\xi}^r \circ \boldsymbol{\gamma}^{cT})\mathbf{1}_k) + \text{cov}(\mathbf{y^r}, \delta\mathbf{x^r}(\boldsymbol{\xi}^r \circ \boldsymbol{\gamma}^{rT} \circ \boldsymbol{\Pi}^T)\mathbf{1}_k)+ \\ \text{cov}(\mathbf{y^r}, \mathbf{E}^c\boldsymbol{\gamma}^c\delta) + \text{cov}(\mathbf{y^r}, \mathbf{E}^r\boldsymbol{\gamma}^r\delta) + \text{cov}(\mathbf{y^r}, \boldsymbol{\epsilon}^r) \end{bmatrix} \cdot \left[\text{var}(\mathbf{y^r})\right]^{-1} = 1 \tag{C3}$$

$$\begin{bmatrix} \text{cov}(\mathbf{y^c}, \mathbf{x^c}\beta^c) + \text{cov}(\mathbf{y^c}, \mathbf{x^r}\beta^r)+ \\ \text{cov}(\mathbf{y^c}, \mathbf{x^c}(\boldsymbol{\xi}^c \circ \boldsymbol{\gamma}^{cT})\mathbf{1}_k) + \text{cov}(\mathbf{y^c}, \mathbf{x^c}(\boldsymbol{\xi}^c \circ \boldsymbol{\gamma}^{rT} \circ \boldsymbol{\Pi}^T)\mathbf{1}_k)+ \\ \text{cov}(\mathbf{y^c}, \mathbf{x^r}(\boldsymbol{\xi}^r \circ \boldsymbol{\gamma}^{cT})\mathbf{1}_k) + \text{cov}(\mathbf{y^c}, \mathbf{x^r}(\boldsymbol{\xi}^r \circ \boldsymbol{\gamma}^{rT} \circ \boldsymbol{\Pi}^T)\mathbf{1}_k)+ \\ \text{cov}(\mathbf{y^c}, \mathbf{E}^c\boldsymbol{\gamma}^c) + \text{cov}(\mathbf{y^c}, \mathbf{E}^r\boldsymbol{\gamma}^r) \end{bmatrix} \cdot \left[\text{var}(\mathbf{y^c})\right]^{-1} \approx \omega_1$$

$$\begin{bmatrix} \text{cov}(\mathbf{y^r}, \delta\mathbf{x^c}\beta^c) + \text{cov}(\mathbf{y^r}, \delta\mathbf{x^r}\beta^r)+ \\ \text{cov}(\mathbf{y^r}, \delta\mathbf{x^c}(\boldsymbol{\xi}^c \circ \boldsymbol{\gamma}^{cT})\mathbf{1}_k) + \text{cov}(\mathbf{y^r}, \delta\mathbf{x^c}(\boldsymbol{\xi}^c \circ \boldsymbol{\gamma}^{rT} \circ \boldsymbol{\Pi}^T)\mathbf{1}_k)+ \\ \text{cov}(\mathbf{y^r}, \delta\mathbf{x^r}(\boldsymbol{\xi}^r \circ \boldsymbol{\gamma}^{cT})\mathbf{1}_k) + \text{cov}(\mathbf{y^r}, \delta\mathbf{x^r}(\boldsymbol{\xi}^r \circ \boldsymbol{\gamma}^{rT} \circ \boldsymbol{\Pi}^T)\mathbf{1}_k)+ \\ \text{cov}(\mathbf{y^r}, \mathbf{E}^c\boldsymbol{\gamma}^c\delta) + \text{cov}(\mathbf{y^r}, \mathbf{E}^r\boldsymbol{\gamma}^r\delta) \end{bmatrix} \cdot \left[\text{var}(\mathbf{y^r})\right]^{-1} \approx \omega_2 \tag{C4}$$

where $\omega_1 = \|\mathbf{y^c} - \mathbf{y^{c*}}\|^2 / \|\mathbf{y^c} - \overline{\mathbf{y^c}}\|^2$ whereas $\omega_2 = \|\mathbf{y^r} - \mathbf{y^{r*}}\|^2 / \|\mathbf{y^r} - \overline{\mathbf{y^r}}\|^2$. Note that $\bar{\mathbf{y}}^{\mathbf{c}}$ and $\bar{\mathbf{y}}^{\mathbf{r}}$ are $n \times 1$ vectors containing the mean values of $\mathbf{y^c}$ and $\mathbf{y^r}$ whereas $\mathbf{y^{c*}}$ and $\mathbf{y^{r*}}$ refers to the estimated reduced equations in $\mathcal{S}'_2$ without considering the residual terms $\boldsymbol{\epsilon}^c$ and $\boldsymbol{\epsilon}^r$. Note also that the terms in the right side of Eq. C4 denotes the variance explained by the reduced system $\mathcal{S}'_2$.

Appendix D: Bias-corrected and accelerated (BCa) bootstrap procedure

Bias-corrected and accelerated (BCa) is a powerful bootstrap procedure usually adopted in mediation analysis. More precisely, Q samples (with $Q \geq 1000$) of size n are row-wise randomly drawn (with replacement) from the original matrices $\mathbf{M^c}$, $\mathbf{M^r}$ and original vectors $\mathbf{y^c}$, $\mathbf{y^r}$. For each $q-$th sample, the mediation parameters are estimated by applying the IMedA procedure on the sample matrices $\mathbf{M^c}_q$, $\mathbf{M^r}_q$ and vectors $\mathbf{y^c}_q$,$\mathbf{y^r}_q$. These steps are repeated for Q times. The ensuing sample parameter distributions are then used for computing the standard errors or BCa based confidence intervals (95% CIs) for every estimated parameter in the model. In what follows we briefly describe how the BCa based CIs can be obtained. For the sake of generality, considering the i-th parameter of $\widehat{\boldsymbol{\gamma}^c}_i$ with $\widehat{\boldsymbol{\gamma}^c}_i^* = \widehat{\boldsymbol{\gamma}^c}_{i1}, \widehat{\boldsymbol{\gamma}^c}_{i2}, \ldots, \widehat{\boldsymbol{\gamma}^c}_{iQ}$ denoting its empirical distribution. The 95% BCa based confidence interval for such parameter takes the form of $[\ \widehat{\boldsymbol{\gamma}^c}_{i_g}^*, \widehat{\boldsymbol{\gamma}^c}_{i_v}^*\ ]$, where g and v are, in turn, computed as follows:

$$g = n \cdot \Phi_{\mathcal{N}}\left(z_{\phi_2} + \frac{z_{\phi_2} + z_{\alpha/2}}{1 - \phi_1(z_{\phi_2} + z_{\alpha/2})}\right) \qquad v = n \cdot \Phi_{\mathcal{N}}\left(z_{\phi_2} + \frac{z_{\phi_2} + z_{1-\alpha/2}}{1 - \phi_1(z_{\phi_2} + z_{1-\alpha/2})}\right)$$

where $\Phi_{\mathcal{N}}$ is the cumulative normal distribution function, $z_{\alpha/2} = -1.96$ and $z_{1-\alpha/2} = 1.96$ for $\alpha = 0.05$ whereas $z_{\phi_2} = \Phi_{\mathcal{N}}^{-1}(\phi_1, 0, 1)$, with $\Phi_{\mathcal{N}}^{-1}$ being the inverse cumulative normal distribution function. Note that the term z_{ϕ_2} measures the median bias of the bootstrap distribution $\widehat{\gamma}_i^*$ where ϕ_2 is computed as follows:

$$\phi_2 = \frac{1}{Q} \sum_{\widehat{\boldsymbol{\gamma}^c}_i^* < \underline{\widehat{\boldsymbol{\gamma}^c}}} \widehat{\boldsymbol{\gamma}^c}_i^* \qquad \text{where } \underline{\widehat{\boldsymbol{\gamma}^c}} = \frac{\sum_{i=1}^{Q} \widehat{\boldsymbol{\gamma}^c}_i^*}{Q}$$

whereas, on the contrary, the acceleration term ϕ_1, which measures the rate of change of the standard deviation of $\widehat{\gamma}_i^*$, is computed as follows:

$$\phi_1 = \sum_{i=1}^{n} (\widehat{\boldsymbol{\gamma}^c}_i^* - \underline{\widehat{\boldsymbol{\gamma}^c}})^3 \cdot \left[\left(6 \sum_{i=1}^{n} (\widehat{\boldsymbol{\gamma}^c}_i^* - \underline{\widehat{\boldsymbol{\gamma}^c}})^2\right)^{\frac{3}{2}}\right]^{-1}$$

Note that, when $\phi_1 = \phi_2 = 0$ the BCa based CIs simply reduce to the standard percentile CIs.

References

Alarcon GM (2011) A meta-analysis of burnout with job demands, resources, and attitudes. J Vocat Behav 79(2):549–562

Alkhamisi M (2010) Simulation study of new estimators combining the sur ridge regression and the restricted least squares methodologies. Stat Pap 51(3):651–672

Alwin DF, Hauser RM (1975) The decomposition of effects in path analysis. Am Sociol Rev 40:37–47

Arndt S, Turvey C, Coryell WH, Dawson JD, Leon AC, Akiskal HS (2000) Charting patients' course: a comparison of statistics used to summarize patient course in longitudinal and repeated measures studies. J Psychiatr Res 34(2):105–113

Augustin T (2002) Expected utility within a generalized concept of probabilitya comprehensive framework for decision making under ambiguity. Stat Pap 43(1):5–22

Avanzi L, van Dick R, Fraccaroli F, Sarchielli G (2012) The downside of organizational identification: relations between identification, workaholism and well-being. Work Stress 26(3):289–307

Avanzi L, Balducci C, Fraccaroli F (2013) Contribution to the italian validation of the copenhagen burnout inventory (cbi). Psicol Della Salute 2:120–135

Baron RM, Kenny DA (1986) The moderator-mediator variable distinction in social psychological research: conceptual, strategic, and statistical considerations. J Pers Soc Psychol 51(6):1173–1182

Billard L, Diday E (2002) Symbolic regression analysis. In: Jajuga K, Sokołowski A, Bock HH (eds) Classification, clustering, and data analysis: recent advances and applications. Springer, Berlin, pp 281–288

Billard L, Diday E (2003) From the statistics of data to the statistics of knowledge: symbolic data analysis. J Am Stat Assoc 98(462):470–487

Blanco-Fernández A, Winker P (2016) Data generation processes and statistical management of interval data. AStA Adv Stat Anal 100(4):475–494

Blanco-Fernández A, Colubi A, García-Bárzana M (2013a) A set arithmetic-based linear regression model for modelling interval-valued responses through real-valued variables. Inf Sci 247:109–122

Blanco-Fernández A, Colubi A, González-Rodríguez G (2013b) Linear regression analysis for interval-valued data based on set arithmetic: a review. Towards advanced data analysis by combining soft computing and statistics. Springer, Berlin, pp 19–31

Bliese PD, Castro CA (2000) Role clarity, work overload and organizational support: multilevel evidence of the importance of support. Work Stress 14(1):65–73

Bollen KA, Stine R (1990) Direct and indirect effects: classical and bootstrap estimates of variability. Sociol Methodol 20(1):15–140

Caffo B, Chen S, Stewart W, Bolla K, Yousem D, Davatzikos C, Schwartz BS (2008) Are brain volumes based on magnetic resonance imaging mediators of the associations of cumulative lead dose with cognitive function? Am J Epidemiol 167(4):429–437

Calcagnì A, Lombardi L (2014) Dynamic fuzzy rating tracker (dyfrat): a novel methodology for modeling real-time dynamic cognitive processes in rating scales. Appl Soft Comput 24:948–961

Calcagnì A, Lombardi L, Sulpizio S (2017) Analyzing spatial data from mouse tracker methodology: an entropic approach. Behav Res Methods 1–19. doi:10.3758/s13428-016-0839-5

Carpita M, Ciavolino E (2017) A generalized maximum entropy estimator to simple linear measurement error model with a composite indicator. Adv Data Anal Classif 11(1):139–158

Choi JY, Lee MJ (2016) Regression discontinuity: review with extensions. Stat Pap, pp 1–30

Claessens BJ, Van Eerde W, Rutte CG, Roe RA (2004) Planning behavior and perceived control of time at work. J Organ Behav 25(8):937–950

Couso I, Dubois D (2014) Statistical reasoning with set-valued information: Ontic vs. epistemic views. Int J Approx Reason 55(7):1502–1518

Diday E, Noirhomme-Fraiture M et al (2008) Symbolic data analysis and the SODAS software. Wiley, New York

Edwards JA, Webster S, Van Laar D, Easton S (2008) Psychometric analysis of the uk health and safety executive's management standards work-related stress indicator tool. Work Stress 22(2):96–107

Edwards JR, Lambert LS (2007) Methods for integrating moderation and mediation: a general analytical framework using moderated path analysis. Psychol Methods 12(1):1–22

Everitt B (1995) The analysis of repeated measures: a practical review with examples. Stat 44(1):113–135

Fairchild AJ, MacKinnon DP, Taborga MP, Taylor AB (2009) R2 effect-size measures for mediation analysis. Behav Res Methods 41(2):486–498

Fields GS (2003) Accounting for income inequality and its change: a new method, with application to the distribution of earnings in the united states. Res Labor Econ 22:1–38
Fishburn PC (1973) Interval representations for interval orders and semiorders. J Math Psychol 10(1):91–105
Fisher CD, To ML (2012) Using experience sampling methodology in organizational behavior. J Organ Behav 33(7):865–877
Frison L, Pocock SJ (1992) Repeated measures in clinical trials: analysis using mean summary statistics and its implications for design. Stat Med 11(13):1685–1704
Gómez G, Calle ML, Oller R (2004) Frequentist and bayesian approaches for interval-censored data. Stat Pap 45(2):139–173
Halff HM, Ortony A, Anderson RC (1976) A context-sensitive representation of word meanings. Memory Cogn 4(4):378–383
Hayes AF, Preacher KJ (2010) Quantifying and testing indirect effects in simple mediation models when the constituent paths are nonlinear. Multivar Behav Res 45(4):627–660
Imai K, Van Dyk DA (2004) Causal inference with general treatment regimes. J Am Stat Assoc 99(467):854–866
Imai K, Keele L, Yamamoto T (2010) Identification, inference and sensitivity analysis for causal mediation effects. Stat Sci 25(1):51–71
Johnson A, Mulder B, Sijbinga A, Hulsebos L (2012) Action as a window to perception: measuring attention with mouse movements. Appl Cogn Psychol 26(5):802–809
Judd CM, Kenny DA (1981) Process analysis estimating mediation in treatment evaluations. Eval Rev 5(5):602–619
Kiers HA (2002) Setting up alternating least squares and iterative majorization algorithms for solving various matrix optimization problems. Comput Stat Data Anal 41(1):157–170
Kim H, Park H (2007) Sparse non-negative matrix factorizations via alternating non-negativity-constrained least squares for microarray data analysis. Bioinformatics 23(12):1495–1502
Kristal AR, Glanz K, Tilley BC, Li S (2000) Mediating factors in dietary change: understanding the impact of a worksite nutrition intervention. Health Educ Behav 27(1):112–125
Kristensen TS, Borritz M, Villadsen E, Christensen KB (2005) The copenhagen burnout inventory: a new tool for the assessment of burnout. Work Stress 19(3):192–207
Lima Neto EdA, de Carvalho FdA (2008) Centre and range method for fitting a linear regression model to symbolic interval data. Comput Stat Data Anal 52(3):1500–1515
Lima Neto EdA, de Carvalho FdA (2010) Constrained linear regression models for symbolic interval-valued variables. Comput Stat Data Anal 54(2):333–347
Little R (1993) Statistical analysis of masked data. J Off Stat 9(2):407–426
Luce RD (1956) Semiorders and a theory of utility discrimination. Econometrica 24(2):178–191
Luo P, Geng Z (2016) Causal mediation analysis for survival outcome with unobserved mediator-outcome confounders. Comput Stat Data Anal 93:336–347
MacKinnon D (2008) Introduction to statistical mediation analysis. Routledge, New York
MacKinnon DP, Fairchild AJ (2009) Current directions in mediation analysis. Curr Dir Psychol Sci 18(1):16–20
Mood A, Graybill F (1950) Introduction to the theory of statistics. McGraw-Hill, New York
Moore RE (1966) Interval analysis. Prentice-Hall Englewood Cliffs, New York
Nkurunziza S, Ejaz Ahmed S (2011) Estimation strategies for the regression coefficient parameter matrix in multivariate multiple regression. Stat Neerl 65(4):387–406
Parchami A, Taheri SM, Mashinchi M (2012) Testing fuzzy hypotheses based on vague observations: a p-value approach. Stat Pap 53(2):469–484
Preacher KJ, Hayes AF (2008) Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. Behav Res Methods 40(3):879–891
Preacher KJ, Kelley K (2011) Effect size measures for mediation models: quantitative strategies for communicating indirect effects. Psychol Methods 16(2):93
Rosseel Y (2012) Lavaan: an R package for structural equation modeling. J Stat Softw 48(2):1–36
Salicone S (2007) Measurement uncertainty: an approach via the mathematical theory of evidence. Springer, New York
Sawyer JE (1992) Goal and process clarity: specification of multiple constructs of role ambiguity and a structural equation model of their antecedents and consequences. J Appl Psychol 77(2):130

Seibold DR, McPhee RD (1979) Commonality analysis: a method for decomposing explained variance in multiple regression analyses. Human Commun Res 5(4):355–365

Senn S, Stevens L, Chaturvedi N (2000) Repeated measures in clinical trials: simple strategies for analysis using summary measures. Stat Med 19(6):861–877

Sobel ME (1982) Asymptotic confidence intervals for indirect effects in structural equation models. Sociol Methodol 13(1982):290–312

Stolzenberg RM (1980) The measurement and decomposition of causal effects in nonlinear and nonadditive models. Sociol Methodol 11:459–488

Sutton S (1998) Predicting and explaining intentions and behavior: How well are we doing? J Appl Soc Psychol 28(15):1317–1338

Takane Y, Young FW, De Leeuw J (1977) Nonmetric individual differences multidimensional scaling: an alternating least squares method with optimal scaling features. Psychometrika 42(1):7–67

Taris TW, de Lange AH, Kompier MA (2010) Research methods in occupational health psychology. In: Houdmont J, Leka S (eds) Occupational health psychology. Wiley-Blackwell, Hoboken, pp 269–297

Taylor AB, MacKinnon D, Tein JY (2008) Tests of the three-path mediated effect. Organ Res Methods 11(2):241–269

Timmerman ME, Kiers HA (2002) Three-way component analysis with smoothness constraints. Comput Stat Data Anal 40(3):447–470

Toderi S, Balducci C, Edwards JA, Sarchielli G, Broccoli M, Mancini G (2013) Psychometric properties of the uk and italian versions of the hse stress indicator tool. Eur J Psychol Assess 29(1):72–79

Wardle J, Carnell S, Haworth CM, Farooqi IS, O'Rahilly S, Plomin R (2008) Obesity associated genetic variation in fto is associated with diminished satiety. J Clin Endocrinol Metab 93(9):3640–3643

Yahya W, Olaifa J (2014) A note on ridge regression modeling techniques. Electron J Appl Stat Anal 7(2):343–361

Yuan KH, Cheng Y, Maxwell S (2013) Moderation analysis using a two-level regression model. Psychometrika 79(4):701–732

Zhang Z, Wang L (2013) Methods for mediation analysis with missing data. Psychometrika 78(1):154–184

Supplementary material for: Multiple Mediation Analysis for Interval-Valued Data

Contents

1 Simulation Study for the $m = 2$ mediation case

This section contains the tabular results of the simulation study for the $m = 2$ case (see Section 6 of the manuscript).

Simulation design. As described in the simulation design for the $m = 1$ case.

Data Generation. As described for the previous case except as follows. The mediator variables $\mathbf{M^c}$ and $\mathbf{M^r}$ as well as the dependent variables $\mathbf{y^c}$ and $\mathbf{y^r}$ are obtained by applying the interval model depicted in Figure 2-B of the manuscript with the following parameters: $\mathbf{A^c} = \text{diag}(4.8, 2.1)$, $\mathbf{A^r} = \text{diag}(3.1, -8.6)$, $\text{vec}(\mathbf{\Xi}) = (2.7, -0.98, 4.1, 2.3)^T$, $\mathbf{\Pi} = \text{diag}(2.04, 1.1)$, $\alpha^{\mathbf{c}} = 3.0$, $\alpha^{\mathbf{r}} = -5.3$, $\beta = (2.3, 1.9)^T$, $\gamma^{\mathbf{c}} = (1.9, 0.9)$, $\gamma^{\mathbf{r}} = (2.1, -4.3)$, and $\delta = -3.25$. Note that, diag(.) is the operator that transforms a vector into a diagonal matrix whereas vec(.) transforms a $kn \times 1$ vector into a $n \times k$ matrix.

Outcome measures. As described in the simulation design for the $m = 1$ case.

Results. The results are provided in the Table 1

n, ϵ	IMedA-ALS		SEM-ML		SEM-WLS		2SMA	
	amse	PA	amse	PA	amse	PA	amse	PA
$n = 50$								
ϵ_1	0.41	96.11	0.64	80.37	37.42	58.93	0.64	80.42
ϵ_2	0.43	94.54	0.67	78.46	50.23	59.97	0.66	78.88
ϵ_3	0.46	93.00	0.71	77.18	60.93	52.87	0.70	77.68
ϵ_4	0.48	91.99	0.71	77.22	46.97	53.90	0.69	78.10
$n = 250$								
ϵ_1	0.40	96.52	0.46	93.46	37.52	67.88	0.46	93.53
ϵ_2	0.40	96.27	0.48	92.88	48.71	68.02	0.46	93.26
ϵ_3	0.41	96.05	0.49	92.03	56.76	66.46	0.47	92.68
ϵ_4	0.42	95.78	0.50	91.72	83.22	66.84	0.47	92.79
$n = 500$								
ϵ_1	0.40	96.52	0.43	94.99	41.23	72.24	0.43	95.07
ϵ_2	0.40	96.39	0.45	94.47	34.91	72.66	0.43	94.86
ϵ_3	0.40	96.37	0.46	93.98	39.25	69.86	0.44	94.64
ϵ_4	0.40	96.23	0.47	93.62	51.15	71.05	0.44	94.70
$n = 1000$								
ϵ_1	0.40	96.52	0.42	95.84	32.84	77.52	0.42	95.91
ϵ_2	0.40	96.50	0.43	95.38	8.45	77.66	0.42	95.77
ϵ_3	0.40	96.46	0.45	94.92	18.36	74.91	0.42	95.59
ϵ_4	0.40	96.40	0.46	94.60	49.70	77.24	0.42	95.68

Table 1. Second Monte Carlo study: Percentage of agreement (PA) index and average root mean square errors (AMSE) for the array of parameters of the multiple mediation model ($m = 2$)

2 2SMA algorithm and Lavaan R-code

The two-steps mediation analysis (2SMA) is based on a set of OLS regressions that are hierarchically estimated in order to guarantee the identification of all the IMedA's parameters. In particular, given the IMedA model:

$$\mathrm{S}_1 : \begin{cases} \mathbf{M^c} = \mathbf{1}\mathbf{A}^c + \mathbf{X}\boldsymbol{\Xi} + \mathbf{E}^c \\ \mathbf{M^r} = \mathbf{1}\mathbf{A}^r + (\mathbf{1}\mathbf{A}^c + \mathbf{X}\boldsymbol{\Xi})\boldsymbol{\Pi} + \mathbf{E^r} \end{cases} \quad \mathcal{S}_2 : \begin{cases} \mathbf{y^c} = \mathbf{1}\alpha^c + \mathbf{X}\boldsymbol{\beta} + \mathbf{M^c}\boldsymbol{\gamma^c} + \mathbf{M^r}\boldsymbol{\gamma^r} + \boldsymbol{\epsilon}^c \\ \mathbf{y^r} = \mathbf{1}\alpha^r + (\mathbf{1}\alpha^c + \mathbf{X}\boldsymbol{\beta} + \mathbf{M^c}\boldsymbol{\gamma^c} + \mathbf{M^r}\boldsymbol{\gamma^r})\delta + \boldsymbol{\epsilon}^r \end{cases} \tag{1}$$

the parameters are estimated in two main steps, one for the system $\mathcal{S}_1$ and another one for $\mathcal{S}_2$, as follows:

I Step: *estimate* $\widehat{\boldsymbol{\Xi}} = (\mathbf{X}^T\mathbf{X})^{-1}\ \mathbf{X}^T\mathbf{M^c}$

$\widehat{\mathbf{A}}^c = \overline{\mathbf{M}}^{\mathbf{c}} - \overline{\mathbf{X}}\widehat{\boldsymbol{\Xi}}$

compute $\mathbf{M}^{\mathbf{c}^*} = \mathbf{1}\widehat{\mathbf{A}}^c + \mathbf{X}\widehat{\boldsymbol{\Xi}}$

estimate $\widehat{\boldsymbol{\Pi}} = (\mathbf{M}^{\mathbf{c}^*T}\mathbf{M}^{\mathbf{c}^*})^{-1}\ \mathbf{M}^{\mathbf{c}^*T}\mathbf{M^r}$

$\widehat{\mathbf{A}}^r = \overline{\mathbf{M}}^{\mathbf{r}} - \overline{\mathbf{M}}^{\mathbf{c}^*}\widehat{\boldsymbol{\Pi}}$

compute $\mathbf{M}^{\mathbf{r}^*} = \mathbf{1}\widehat{\mathbf{A}}^r + \mathbf{M}^{\mathbf{c}^*}\widehat{\boldsymbol{\Pi}}$

II Step: *estimate* $\widehat{\boldsymbol{\beta}} = (\mathbf{X}^T\mathbf{X})^{-1}\ \mathbf{X}^T\mathbf{y^c}$

$\widehat{\gamma}^{\mathbf{c}} = (\mathbf{M^c}^T\mathbf{M^c})^{-1}\ \mathbf{M^c}^T\mathbf{y^c}$

$\widehat{\gamma}^{\mathbf{r}} = (\mathbf{M^r}^T\mathbf{M^r})^{-1}\ \mathbf{M^r}^T\mathbf{y^c}$

$\widehat{\alpha}^c = \overline{\mathbf{y}}^{\mathbf{c}} - \overline{\mathbf{X}}\widehat{\boldsymbol{\beta}} - \overline{\mathbf{M}}^{\mathbf{c}}\widehat{\gamma}^{\mathbf{c}} - \overline{\mathbf{M}}^{\mathbf{r}}\widehat{\gamma}^{\mathbf{r}}$

compute $\mathbf{y}^{\mathbf{c}^*} = \mathbf{1}\widehat{\alpha}^c + \mathbf{X}\widehat{\boldsymbol{\beta}} + \mathbf{M^c}\widehat{\gamma}^{\mathbf{c}} + \mathbf{M^r}\widehat{\gamma}^{\mathbf{r}}$

estimate $\widehat{\delta} = (\mathbf{y}^{\mathbf{c}^*T}\mathbf{y}^{\mathbf{c}^*})^{-1}\ \mathbf{y}^{\mathbf{c}^*T}\mathbf{y^r}$

$\widehat{\alpha}^r = \overline{\mathbf{y}}^{\mathbf{r}} - \overline{\mathbf{y}}^{\mathbf{c}^*}\widehat{\delta}$

compute $\mathbf{y}^{\mathbf{r}^*} = \mathbf{1}\widehat{\alpha}^r + \mathbf{y}^{\mathbf{c}^*}\widehat{\delta}$

Note that unlike IMedA-ALS, 2SMA does not involve an alternating gradient-descendent approach in minimizing the loss function associated to the regression model. On the contrary, it adopts several gradient-descent procedures that separately minimize the loss function. In this way, the estimation of a given subset of parameters does not affect the estimation of another subset. As a consequence, each subset of parameters satisfies the convergence of the algorithm toward a proper local/global stationary point whereas a global convergence is not allowed in this context. Therefore, in some circumstances, this estimation procedure may possibly yield biased results as the parameters are independently estimated.

The following R-syntax has been used to estimate mediation paths with SEM-ML and SEM-WLS approaches.

```
### How to define model in Lavaan
# model <- "mc ~ 1 + xc + xl
#           ml ~ 1 + mc
#           yc ~ 1 + xc + xl + mc + ml
#           yl ~ 1 + yc"
# I equation: MC
# II equation: ML
# III equation: yc
# IV equation: yl
###

library("lavaan")
# Note: xc, xl, yc, yl, MC1, ML1 are column-vectors of real data generated by Matlab
# and passed through RCMD BATCH connection.
Data <- data.frame(xc,xl,yc,yl,MC1,ML1)
names(Data) <- c("xc","xl","yc","yl","MC1","ML1")
model <- " MC1 ~ 1 + xc + xl
    ML1 ~ 1 + MC1
    yc ~ 1 + xc + xl + MC1 + ML1
    yl ~ 1 + yc"

data.estimator = "ML" #or "WLS"
fit <- sem(model, data = Data, estimator=data.estimator,likelihood="normal",
    std.ov=FALSE,fixed.x=TRUE,orthogonal=TRUE, control=list(iter.max=500))
summary(fit)
convergence <- lavInspect(fit, what = "converged")
print(convergence)
est.pars <- parameterEstimates(fit)
print(est.pars)
```

3 Scenario analysis for IMedA-ALS

In this section we describe the results of a short scenario analysis carried out to evaluate the flexibility of the IMedA model in capturing all the possible linear relations among the observed variables. In particular, we test the ability of the proposed model to recover the true structure of the observed data over an extended set of possible scenarios. The IMedA model is defined as:

$$\mathrm{S}_1 : \begin{cases} \mathbf{M^c} = \mathbf{1}\mathbf{A}^c + \mathbf{X}\boldsymbol{\Xi} + \mathbf{E}^c \\ \mathbf{M^r} = \mathbf{1}\mathbf{A}^r + (\mathbf{1}\mathbf{A}^c + \mathbf{X}\boldsymbol{\Xi})\boldsymbol{\Pi} + \mathbf{E^r} \end{cases} \qquad \mathcal{S}_2 : \begin{cases} \mathbf{y^c} = \mathbf{1}\alpha^c + \mathbf{X}\boldsymbol{\beta} + \mathbf{M^c}\boldsymbol{\gamma^c} + \mathbf{M^r}\boldsymbol{\gamma}^r + \boldsymbol{\epsilon}^c \\ \mathbf{y^r} = \mathbf{1}\alpha^r + (\mathbf{1}\alpha^c + \mathbf{X}\boldsymbol{\beta} + \mathbf{M^c}\boldsymbol{\gamma^c} + \mathbf{M^r}\boldsymbol{\gamma}^r)\delta + \boldsymbol{\epsilon}^r \end{cases} \tag{2}$$

that can be considered as a constrained version of the more general model:

$$\overline{\mathcal{S}}_1 : \begin{cases} \mathbf{M^c} = \mathbf{1}\mathbf{A}^c + \mathbf{X}\boldsymbol{\Xi} + \mathbf{E}^c \\ \mathbf{M^r} = \mathbf{1}\mathbf{A}^r + \mathbf{X}\boldsymbol{\Xi}_2 + \mathbf{E^r} \end{cases} \quad \overline{\mathcal{S}}_2 : \begin{cases} \mathbf{y^c} = \mathbf{1}\alpha^c + \mathbf{X}\boldsymbol{\beta} + \mathbf{M^c}\boldsymbol{\gamma^c} + \mathbf{M^r}\boldsymbol{\gamma}^r + \boldsymbol{\epsilon}^c \\ \mathbf{y^r} = \mathbf{1}\alpha^r + \mathbf{X}\boldsymbol{\beta_2} + \mathbf{M^c}\boldsymbol{\gamma_2^c} + \mathbf{M^r}\boldsymbol{\gamma_2^r} + \boldsymbol{\epsilon}^r \end{cases} \tag{3}$$

In particular, the systems $\mathcal{S}_1$ and $\mathcal{S}_2$ are more parsimonious than $\overline{\mathcal{S}}_1$ and $\overline{\mathcal{S}}_2$ as they require 12m parameters against 16m, respectively. In order to proceed with the scenario analysis, we decide to analyse the characteristics of the systems $\mathcal{S}_2$ and $\overline{\mathcal{S}}_2$, as those ones related to the other two systems can be easily obtained just by generalizing the ensuing results (note that $\mathcal{S}_1$ and $\overline{\mathcal{S}}_1$ are formally the same of $\mathcal{S}_2$ and $\overline{\mathcal{S}}_2$). For the sake of simplicity, $\mathcal{S}_2$ and $\overline{\mathcal{S}}_2$ can be re-written in terms of their *regression cores*, as follows:

$$\mathrm{S}_2^* : \begin{cases} \mathbf{y^c} = \mathbf{X}\boldsymbol{\beta} \\ \mathbf{y^r} = \mathbf{X}\boldsymbol{\beta}\delta \end{cases} \qquad \overline{\mathcal{S}}_2^* : \begin{cases} \mathbf{y^c} = \mathbf{X}\boldsymbol{\beta} \\ \mathbf{y^r} = \mathbf{X}\boldsymbol{\beta_2} \end{cases} \tag{4}$$

that consist of a set of two equations modeling centers and ranges by the matrix of interval-valued independent variables. Note that, because $\mathcal{S}_2^* \subset \mathcal{S}_2$ and $\overline{\mathcal{S}}_2^* \subset \overline{\mathcal{S}}_2$, we can study the more simple regression systems as proxies for the more complex ones. By looking at their structures we can immediately notice how $\overline{\mathcal{S}}_2^*$ is able to model all the relations among the observed variables $\mathbf{y^c}$, $\mathbf{y^r}$, $\mathbf{x^c}$, and $\mathbf{x^r}$ (remember that $\mathbf{X} = (\mathbf{x^c}, \mathbf{x^r})$) and, therefore, can be considered the *golden rule* in terms of flexibility: its four parameters are capable to capture all the possible linear relationships among the observed variables. As a consequence, we can assume $\overline{\mathcal{S}}_2^*$ as the "reference point" for any comparison of $\overline{\mathcal{S}}_1^*$.

Simulation design. With regards to the interval-valued variables involved in the models' representations, we can define the following covariance matrix:

$$\begin{array}{c} \\ \mathbf{x^c} \\ \mathbf{x^r} \\ \mathbf{y^c} \\ \mathbf{y^r} \end{array} \begin{array}{c} \begin{array}{cccc} \mathbf{x^c} & \mathbf{x^r} & \mathbf{y^c} & \mathbf{y^r} \end{array} \\ \begin{pmatrix} 1 & & & \\ 0 & 1 & & \\ a & b & 1 & \\ c & d & e & 1 \end{pmatrix} \end{array}$$

that represent all the possible linear relationships among them. Because the covariance is a linear operator, the five parameters a, b, c, d, e can take just three subsets of real values (i.e., $\mathbb{R}_0^+$, $\mathbb{R}_0^-$, $\{0\}$). As a consequence, all the possible linear relations among the variables are $3^5 = 243$. However, not all scenarios can be tested in our context: indeed, because of the linear constraints realized by the regression systems, the admissible scenarios simply reduce to 81 (e.g., constant models are here meaningless). Finally, we selected a set of meaningful scenarios through which we test $\overline{\mathcal{S}}_1^*$ and $\overline{\mathcal{S}}_2^*$ (see Tables 2-3).

	$\{a,b\} > 0$		$\{a,b\} < 0$		$\{a > 0, b < 0\}$		$\{a < 0, b > 0\}$
1	$\{c,d\} > 0$	5	$\{c,d\} > 0$	9	$\{c,d\} > 0$	13	$\{c,d\} > 0$
2	$\{c,d\} < 0$	6	$\{c,d\} < 0$	10	$\{c,d\} < 0$	14	$\{c,d\} < 0$
3	$\{c > 0, d < 0\}$	7	$\{c > 0, d < 0\}$	11	$\{c > 0, d < 0\}$	15	$\{c > 0, d < 0\}$
4	$\{c < 0, d > 0\}$	8	$\{c < 0, d > 0\}$	12	$\{c < 0, d > 0\}$	16	$\{c < 0, d > 0\}$

Table 2. Scenario analysis: admissible scenarios (panel I)

	$\{a = 0, b > 0\}$		$\{a = 0, b < 0\}$		$\{a > 0, b = 0\}$		$\{a < 0, b = 0\}$
17	$\{c = 0, d > 0\}$	19	$\{c = 0, d > 0\}$	21	$\{c > 0, d = 0\}$	23	$\{c > 0, d = 0\}$
18	$\{c = 0, d < 0\}$	20	$\{c = 0, d < 0\}$	22	$\{c < 0, d = 0\}$	24	$\{c < 0, d = 0\}$

Table 3. Scenario analysis: admissible scenarios (panel II)

Note that Tables 2-3 reports the covariance parameters that refer to the twenty-four relationships between exogenous ($\mathbf{x^c}$, $\mathbf{x^r}$) and endogenous ($\mathbf{y^c}$, $\mathbf{y^r}$) variables. The covariance parameter e of the relationship between $\mathbf{y^c}$ and $\mathbf{y^r}$ can be considered by simply replicating panels I and II for three times, obtaining so eighty-one scenarios. At this point, the scenario analysis is carried out by generating new datasets for each scenario according to a pre-fixed covariance matrix, by applying the regression systems $\overline{\mathcal{S}}_1^*$ and $\overline{\mathcal{S}}_2^*$, and by evaluating the final results in terms of reconstruction of the original data structures.

Data generation. We created 81 semi-positive defined covariance matrices by guaranteeing that the covariance parameters are in the desiderated natural ranges: for values in $\mathbb{R}_0$ the parameters a, b, c, d, e take as large values as possible whereas for values in $\{0\}$ the parameters take values in the interval $[0-\epsilon, 0+\epsilon]$ with ϵ being a small positive quantity closed to zero. Each covariance matrix defines a specific scenario. Next, for each covariance matrix, 1000 $n \times 4$ datasets are drawn from a normal multivariate distribution by constraining the sample covariance matrix to be as close as possible to the fixed covariance matrix.[1] For the sake of simplicity, we used $n = 50$. In order to guarantee the positiveness of ranges, the second and fourth columns of the generated data matrix are linearly rescaled to lie in $\mathbb{R}^+$. Finally, on each new dataset we ran the two regression systems $\overline{\mathcal{S}}_1^*$ and $\overline{\mathcal{S}}_2^*$ and saved the obtained results (i.e., model's parameters and predicted values). Overall, the simulation procedure generated 1000 x 81 = 81000 new datasets as well as an equivalent number of parameters and reconstructed datasets.

Outcome measures. Each simulation was evaluated considering the ability of the models $\overline{\mathcal{S}}_1^*$ and $\overline{\mathcal{S}}_2^*$ in recovering the drawn data sample. Particularly, we computed *means*, *variances*, *RMSE*, and *AoR* on the sample data $\mathbf{y^c_q}$, $\mathbf{y^r_q}$, $\widetilde{\mathbf{Y}}_\mathbf{q} = [\mathbf{y^c_q} - \mathbf{y^r_q}, \mathbf{y^c_q} + \mathbf{y^r_q}]$ and the corresponding reconstructed data $\mathbf{y^{c^*}_q}$, $\mathbf{y^{r^*}_q}$, $\widetilde{\mathbf{Y}}_\mathbf{q} = [\mathbf{y^{c^*}_q} - \mathbf{y^{r^*}_q}, \mathbf{y^{c^*}_q} + \mathbf{y^{r^*}_q}]$ (with $q = 1...1000$). Formally, the root mean squares error (RMSE) is computed as:

$$\text{RMSE} = \sqrt{n^{-1} \cdot \|\mathbf{y} - \mathbf{y}^*\|^2}$$

where $\mathbf{y} \in \{\mathbf{y^c}, \mathbf{y^r}, \widetilde{\mathbf{Y}}\}$ whereas $\mathbf{y}^* \in \{\mathbf{y^{c^*}}, \mathbf{y^{r^*}}, \widetilde{\mathbf{Y}^*}\}$. Note that RMSE gives information about the amount of error of reconstructed data in resembling the sample data. By contrast, the amount of reconstruction (AoR) index and gives information about the *amount of reconstruction* performed by the two regressions systems. The index is defined as follows:

$$\text{AoR} = \frac{\min\{\|\mathbf{y}^+\|, \|\mathbf{y}\ \|\}}{\max\{\|\mathbf{y}^+\|, \|\mathbf{y}\ \|\}} \qquad \text{with} \quad \mathbf{y}^+ = \frac{\mathbf{y}^{*T}\mathbf{y}}{\|\mathbf{y}\|} \cdot \frac{\mathbf{y}}{\|\mathbf{y}\|}$$

where $\mathbf{y}$ and $\mathbf{y}^*$ are defined as above. Note that AoR takes values in $[0, 1]$ with 0 indicating that no reconstruction occurred whereas 1 means that complete reconstruction occurred. In this context, variances gives information about the rigidity/flexibility of the $\overline{\mathcal{S}}_1^*$ and $\overline{\mathcal{S}}_2^*$ systems in order to reconstruct the sample data. In particular, when $\text{var}(\mathbf{y}^*{}_{\overline{\mathcal{S}}_1^*}) < \text{var}(\mathbf{y}^*{}_{\overline{\mathcal{S}}_2^*})$ we state the system $\overline{\mathcal{S}}_1^*$ less flexible than $\overline{\mathcal{S}}_2^*$ in recovering $\mathbf{y}^*$.

Results. Tables 4-10 show the obtained results. In particular, Tables 4, 6, 8 report the AoR and RMSE measures whereas Tables 5, 7, 9 show the mean values and variances. Table 10 shows the regression parameters for the equation $\mathbf{y^r}$ computed for both the $\overline{\mathcal{S}}_1^*$ and $\overline{\mathcal{S}}_2^*$ systems. Overall, the results show how $\overline{\mathcal{S}}_1^*$ shows the same performances of $\overline{\mathcal{S}}_2^*$ in almost all scenarios. In particular, $\overline{\mathcal{S}}_1^*$ reconstructs well the data interval-valued structures $\widetilde{\mathbf{Y}}$ like $\overline{\mathcal{S}}_1^*$ does. Moreover, the variances computed on the reconstructed data nicely shows how $\overline{\mathcal{S}}_1^*$ is flexible enough to show equivalent

[1] Several dissimilarities measures can be used in comparing two covariance matrices $\mathbf{A}$ and $\mathbf{B}$. In our context, we resort to use the following measure: $\mathcal{D}(\mathbf{A}, \mathbf{B}) = 1 - [\text{Tr}(\mathbf{A} \cdot \mathbf{B}) \cdot (\|\mathbf{A}\|_f \cdot \|\mathbf{B}\|_f)^{-1}$ with f being the Frobenius norm. Note that $\mathcal{D}(\mathbf{A}, \mathbf{B}) \in [0, 1]$. Among others, such matrix distance yields more strong results in terms of stability and robustness. For further details, see: Herdin, M., Czink, N., zcelik, H., Bonek, E. (2005, June). Correlation matrix distance, a meaningful measure for evaluation of non-stationary MIMO channels. In *Vehicular Technology Conference, 2005. VTC 2005-Spring. 2005 IEEE 61st* (Vol. 1, pp. 136-140). IEEE.

performances of $\overline{\mathcal{S}}_2^*$. This is partially guaranteed by the fact that δ tends to be closed - in absolute value - to the mean value between β_2^c and β_r^2. By contrast, considering the interval-valued components $\mathbf{y^c}$ and $\mathbf{y^r}$ the system $\overline{\mathcal{S}}_2^*$ outperforms $\overline{\mathcal{S}}_1^*$ in terms of AoR, RMSE, and variance just in 24 cases on 81 (note that in Tables such cases are in gray). In particular, in these cases the mean amount of error of recovering (i.e., $1 - AoR$) is almost equal to 5% whereas the error of estimation (RMSE) increases considerably. As a consequence, $\overline{\mathcal{S}}_1^*$ becomes progressively less flexible as also stated by the rapid decrease of the variances. Formally, this is due to the behavior of δ that tends to approximate the mean value between β_2^c and β_r^2. Indeed, in the aforementioned cases δ tends to be closed to zero and consequently the component $\mathbf{y^r}$ proportionally tends to be under-recovered. Nevertheless, as a consequence of the general least squares properties, in such a case $\mathbf{y^r}^*$ tends to assume the mean values of $\mathbf{y^r}$. This is the reason why $\overline{\mathcal{S}}_1^*$ is still comparable with $\overline{\mathcal{S}}_2^*$ in recovering $\widetilde{\mathbf{Y}}$.

scenario	AoR$_{\overline{\mathcal{S}}_1^*}$			AoR$_{\overline{\mathcal{S}}_2^*}$			RMSE$_{\overline{\mathcal{S}}_1^*}$			RMSE$_{\overline{\mathcal{S}}_2^*}$		
	$\mathbf{y^c}$	$\mathbf{y^r}$	$\widetilde{\mathbf{Y}}$	$\mathbf{y^c}$	$\mathbf{y^r}$	$\widetilde{\mathbf{Y}}$	$\mathbf{y^c}$	$\mathbf{y^r}$	$\widetilde{\mathbf{Y}}$	$\mathbf{y^c}$	$\mathbf{y^r}$	$\widetilde{\mathbf{Y}}$
1	1.000	0.996	1.000	1.000	1.000	1.000	0.060	0.276	0.282	0.003	0.000	0.003
2	0.815	0.929	0.825	0.815	0.931	0.825	6.015	1.104	6.117	6.015	1.085	6.113
3	0.943	0.916	0.941	0.943	0.985	0.947	3.234	1.201	3.453	3.232	0.515	3.273
4	0.932	0.913	0.930	0.932	0.983	0.936	3.545	1.234	3.756	3.544	0.542	3.586
5	0.816	0.927	0.825	0.816	0.928	0.826	6.015	1.123	6.121	6.015	1.119	6.120
6	1.000	0.993	0.999	1.000	0.993	0.999	0.135	0.344	0.370	0.135	0.341	0.367
7	0.936	0.911	0.934	0.936	0.983	0.940	3.460	1.242	3.679	3.458	0.538	3.501
8	0.940	0.917	0.939	0.941	0.984	0.944	3.321	1.201	3.534	3.319	0.518	3.360
9	0.957	0.911	0.953	0.957	0.970	0.958	2.851	1.245	3.114	2.849	0.729	2.942
10	0.975	0.915	0.970	0.975	0.978	0.975	2.177	1.214	2.496	2.175	0.623	2.264
11	0.985	0.992	0.985	0.985	0.994	0.986	1.662	0.363	1.702	1.661	0.312	1.691
12	0.795	0.930	0.807	0.795	0.931	0.807	6.427	1.094	6.520	6.426	1.088	6.519
13	0.968	0.914	0.964	0.968	0.975	0.969	2.442	1.216	2.732	2.441	0.654	2.528
14	0.964	0.910	0.960	0.964	0.973	0.965	2.588	1.247	2.876	2.587	0.681	2.676
15	0.802	0.931	0.813	0.802	0.931	0.813	6.329	1.096	6.424	6.329	1.095	6.424
16	1.000	1.000	1.000	1.000	1.000	1.000	0.092	0.003	0.092	0.092	0.000	0.092
17	0.944	0.961	0.945	0.944	0.961	0.945	3.185	0.823	3.291	3.185	0.823	3.291
18	0.831	0.923	0.840	0.832	0.924	0.840	5.525	1.150	5.646	5.525	1.149	5.645
19	0.838	0.921	0.846	0.838	0.922	0.846	5.407	1.167	5.533	5.407	1.166	5.533
20	0.971	0.992	0.973	0.971	0.992	0.973	2.302	0.375	2.333	2.299	0.373	2.329
21	0.998	0.992	0.998	0.998	0.992	0.998	0.542	0.369	0.657	0.541	0.362	0.653
22	0.833	0.921	0.841	0.833	0.921	0.841	5.514	1.164	5.637	5.514	1.163	5.637
23	0.835	0.921	0.843	0.835	0.922	0.843	5.489	1.162	5.612	5.489	1.162	5.612
24	0.966	0.999	0.969	0.966	0.999	0.969	2.472	0.126	2.476	2.472	0.124	2.475

Table 4. Scenario analysis ($d > 0$): AoR and RMSE for both the $\overline{\mathcal{S}}_1^*$ and $\overline{\mathcal{S}}_2^*$ systems. Note that in gray are represented the scenarios where $\overline{\mathcal{S}}_1^*$ largely differs from $\overline{\mathcal{S}}_2^*$.

scenario	$\text{MVs}_{\overline{\mathcal{S}}_1^*}$			$\text{MVs}_{\overline{\mathcal{S}}_2^*}$			$\text{Vs}_{\overline{\mathcal{S}}_1^*}$			$\text{Vs}_{\overline{\mathcal{S}}_2^*}$		
	$\mathbf{y^c}$	$\mathbf{y^r}$	$\widetilde{\mathbf{Y}}$	$\mathbf{y^c}$	$\mathbf{y^r}$	$\widetilde{\mathbf{Y}}$	$\mathbf{y^c}$	$\mathbf{y^r}$	$\widetilde{\mathbf{Y}}$	$\mathbf{y^c}$	$\mathbf{y^r}$	$\widetilde{\mathbf{Y}}$
1	12.012	3.995	12.012	12.011	3.996	12.005	40.131	1.724	553.594	40.153	1.802	551.338
2	12.021	3.991	12.021	12.020	3.991	12.049	17.528	0.522	73.050	17.541	0.563	72.654
3	11.988	3.984	11.988	11.988	3.984	11.983	33.273	0.314	83.630	33.292	1.528	77.257
4	11.928	4.023	11.928	11.927	4.023	11.877	31.936	0.253	64.825	31.963	1.518	59.740
5	12.033	4.001	12.033	12.033	4.000	12.002	18.207	0.515	74.786	18.216	0.522	74.729
6	11.950	3.994	11.950	11.951	3.994	11.986	39.913	1.678	536.203	39.931	1.679	536.221
7	12.053	4.008	12.053	12.054	4.008	12.033	32.220	0.221	56.994	32.246	1.514	52.446
8	12.010	4.007	12.010	12.010	4.007	12.061	33.040	0.333	88.181	33.058	1.542	81.475
9	12.007	4.020	12.007	12.007	4.020	11.949	38.911	0.236	73.496	38.933	1.284	69.359
10	11.987	4.004	11.987	11.988	4.004	12.011	41.789	0.300	100.673	41.819	1.418	95.278
11	11.948	3.995	11.948	11.948	3.995	11.940	38.308	1.671	512.425	38.328	1.706	511.459
12	12.022	3.980	12.022	12.023	3.980	12.082	18.474	0.585	86.175	18.481	0.599	86.079
13	12.009	3.999	12.009	12.008	3.999	11.973	40.912	0.294	96.581	40.943	1.377	91.442
14	11.964	3.995	11.964	11.964	3.995	11.912	40.426	0.216	70.080	40.449	1.340	66.001
15	12.026	4.026	12.026	12.026	4.026	12.036	20.607	0.560	92.062	20.615	0.562	92.082
16	11.986	4.005	11.986	11.986	4.006	12.007	39.910	1.812	578.486	39.926	1.811	578.543
17	11.951	3.986	11.951	11.950	3.987	11.913	29.935	1.115	267.672	29.963	1.114	267.706
18	12.000	4.002	12.000	12.001	4.003	12.014	8.825	0.460	32.404	8.829	0.465	32.639
19	11.960	4.010	11.960	11.960	4.010	12.003	10.114	0.411	33.128	10.115	0.414	33.274
20	12.017	4.004	12.017	12.018	4.005	11.994	35.302	1.661	470.070	35.361	1.661	470.556
21	11.927	4.015	11.927	11.927	4.015	11.940	40.100	1.677	537.489	40.105	1.682	537.405
22	12.002	3.980	12.002	12.002	3.980	11.983	9.095	0.418	30.173	9.096	0.420	30.315
23	12.048	3.995	12.048	12.049	3.995	12.139	9.323	0.430	31.956	9.320	0.435	32.189
24	12.014	3.995	12.014	12.013	3.995	12.009	33.593	1.810	486.395	33.598	1.811	486.423

Table 5. Scenario analysis ($d > 0$): means (MVs) and variances (Vs) for both the $\overline{\mathcal{S}}_1^*$ and $\overline{\mathcal{S}}_2^*$ systems. Note that in gray are represented the scenarios where $\overline{\mathcal{S}}_1^*$ largely differs from $\overline{\mathcal{S}}_2^*$.

scenario	$\text{AoR}_{\overline{\mathcal{S}}_1^*}$			$\text{AoR}_{\overline{\mathcal{S}}_2^*}$			$\text{RMSE}_{\overline{\mathcal{S}}_1^*}$			$\text{RMSE}_{\overline{\mathcal{S}}_2^*}$		
	$\mathbf{y^c}$	$\mathbf{y^r}$	$\widetilde{\mathbf{Y}}$	$\mathbf{y^c}$	$\mathbf{y^r}$	$\widetilde{\mathbf{Y}}$	$\mathbf{y^c}$	$\mathbf{y^r}$	$\widetilde{\mathbf{Y}}$	$\mathbf{y^c}$	$\mathbf{y^r}$	$\widetilde{\mathbf{Y}}$
1	0.811	0.930	0.821	0.811	0.931	0.821	6.091	1.102	6.191	6.091	1.098	6.190
2	0.985	0.982	0.984	0.984	0.982	0.984	1.677	0.554	1.768	1.677	0.551	1.767
3	0.950	0.924	0.948	0.950	0.985	0.953	3.031	1.147	3.243	3.028	0.507	3.071
4	0.939	0.913	0.937	0.939	0.983	0.943	3.369	1.224	3.587	3.367	0.534	3.410
5	0.993	0.999	0.994	0.993	1.000	0.994	1.099	0.130	1.107	1.099	0.061	1.100
6	0.805	0.930	0.816	0.805	0.930	0.816	6.158	1.100	6.256	6.158	1.098	6.256
7	0.936	0.923	0.935	0.936	0.986	0.940	3.439	1.162	3.632	3.436	0.498	3.473
8	0.942	0.910	0.940	0.942	0.982	0.946	3.284	1.243	3.514	3.282	0.553	3.329
9	0.970	0.911	0.965	0.970	0.978	0.971	2.374	1.235	2.679	2.372	0.620	2.453
10	0.964	0.914	0.960	0.964	0.974	0.965	2.592	1.223	2.869	2.590	0.669	2.677
11	0.799	0.932	0.811	0.799	0.932	0.811	6.343	1.082	6.436	6.343	1.081	6.436
12	0.999	0.991	0.998	0.999	0.991	0.998	0.491	0.403	0.637	0.491	0.402	0.636
13	0.954	0.910	0.950	0.954	0.972	0.956	2.932	1.247	3.189	2.931	0.690	3.012
14	0.967	0.914	0.963	0.967	0.975	0.968	2.478	1.222	2.766	2.477	0.659	2.564
15	0.998	0.992	0.998	0.998	0.992	0.998	0.524	0.376	0.647	0.524	0.376	0.647
16	0.790	0.933	0.802	0.790	0.933	0.802	6.481	1.072	6.570	6.481	1.071	6.570
17	0.834	0.921	0.842	0.834	0.921	0.842	5.477	1.169	5.602	5.477	1.168	5.602
18	0.977	0.997	0.979	0.977	0.997	0.979	2.034	0.231	2.047	2.034	0.231	2.047
19	0.994	0.998	0.994	0.994	0.998	0.994	1.044	0.208	1.065	1.044	0.207	1.064
20	0.834	0.922	0.842	0.834	0.922	0.842	5.507	1.165	5.631	5.507	1.164	5.630
21	0.833	0.923	0.842	0.833	0.923	0.842	5.486	1.162	5.609	5.486	1.161	5.609
22	0.949	0.975	0.952	0.949	0.975	0.952	3.024	0.661	3.096	3.024	0.661	3.096
23	0.997	0.988	0.996	0.997	0.988	0.996	0.742	0.454	0.871	0.742	0.454	0.871
24	0.833	0.922	0.841	0.833	0.922	0.841	5.510	1.165	5.634	5.510	1.164	5.633

Table 6. Scenario analysis ($d < 0$): AoR and RMSE for both the $\overline{\mathcal{S}}_1^*$ and $\overline{\mathcal{S}}_2^*$ systems. Note that in gray are represented the scenarios where $\overline{\mathcal{S}}_1^*$ largely differs from $\overline{\mathcal{S}}_2^*$.

scenario	$MVs_{\overline{\mathcal{S}}_1^*}$			$MVs_{\overline{\mathcal{S}}_2^*}$			$Vs_{\overline{\mathcal{S}}_1^*}$			$Vs_{\overline{\mathcal{S}}_2^*}$		
	$\mathbf{y^c}$	$\mathbf{y^r}$	$\widetilde{\mathbf{Y}}$	$\mathbf{y^c}$	$\mathbf{y^r}$	$\widetilde{\mathbf{Y}}$	$\mathbf{y^c}$	$\mathbf{y^r}$	$\widetilde{\mathbf{Y}}$	$\mathbf{y^c}$	$\mathbf{y^r}$	$\widetilde{\mathbf{Y}}$
1	12.007	4.005	12.007	12.006	4.005	11.989	17.045	0.574	78.191	17.051	0.584	78.170
2	12.023	3.997	12.023	12.023	3.997	12.066	36.981	1.492	441.587	36.993	1.496	441.522
3	11.997	4.013	11.997	11.996	4.013	11.966	32.627	0.474	124.030	32.664	1.564	115.746
4	11.988	3.994	11.988	11.987	3.994	11.968	33.297	0.251	66.870	33.313	1.500	61.610
5	11.974	4.012	11.974	11.975	4.012	11.960	37.524	1.818	546.971	37.536	1.832	546.664
6	11.967	3.992	11.967	11.967	3.992	11.955	15.608	0.584	72.830	15.613	0.588	72.869
7	11.978	4.017	11.978	11.979	4.016	11.975	31.621	0.415	104.981	31.642	1.550	97.162
8	12.024	3.977	12.024	12.025	3.977	12.058	34.053	0.241	65.595	34.080	1.519	60.679
9	11.992	3.979	11.992	11.992	3.979	11.974	41.289	0.257	85.000	41.318	1.431	80.146
10	11.977	3.998	11.977	11.977	3.998	12.001	40.186	0.284	91.685	40.208	1.363	86.545
11	11.924	3.994	11.924	11.924	3.994	11.967	20.806	0.606	100.474	20.814	0.608	100.477
12	11.958	3.994	11.958	11.958	3.994	11.982	39.582	1.638	519.188	39.595	1.639	519.196
13	11.925	3.995	11.925	11.924	3.995	11.920	38.523	0.205	63.665	38.545	1.318	59.821
14	11.965	4.000	11.965	11.964	4.000	11.926	41.267	0.293	97.056	41.298	1.384	91.887
15	12.049	4.005	12.049	12.049	4.005	12.080	39.962	1.660	530.546	39.976	1.659	530.595
16	11.970	3.982	11.970	11.969	3.982	11.956	17.158	0.641	87.645	17.164	0.644	87.720
17	11.954	3.988	11.954	11.954	3.987	11.936	9.523	0.394	29.796	9.522	0.397	29.956
18	12.069	4.015	12.069	12.068	4.015	12.064	35.691	1.784	509.196	35.718	1.783	509.272
19	12.014	4.001	12.014	12.014	4.001	12.022	38.571	1.773	547.686	38.603	1.772	547.818
20	12.049	4.012	12.049	12.049	4.012	12.016	8.954	0.418	29.709	8.959	0.419	29.798
21	11.982	4.015	11.982	11.982	4.015	12.016	9.038	0.401	28.924	9.035	0.406	29.136
22	11.954	4.012	11.954	11.954	4.012	11.933	31.126	1.367	341.207	31.131	1.367	341.232
23	12.035	4.000	12.035	12.034	4.000	12.039	39.758	1.594	507.060	39.764	1.594	507.082
24	12.033	4.006	12.033	12.033	4.005	12.037	8.943	0.425	30.195	8.942	0.428	30.392

Table 7. Scenario analysis ($d < 0$): means (MVs) and variances (Vs) for both the $\overline{\mathcal{S}}_1^*$ and $\overline{\mathcal{S}}_2^*$ systems. Note that in gray are represented the scenarios where $\overline{\mathcal{S}}_1^*$ largely differs from $\overline{\mathcal{S}}_2^*$.

scenario	$\text{AoR}_{\overline{\mathcal{S}}_1^*}$			$\text{AoR}_{\overline{\mathcal{S}}_2^*}$			$\text{RMSE}_{\overline{\mathcal{S}}_1^*}$			$\text{RMSE}_{\overline{\mathcal{S}}_2^*}$		
	$\mathbf{y^c}$	$\mathbf{y^r}$	$\widetilde{\mathbf{Y}}$	$\mathbf{y^c}$	$\mathbf{y^r}$	$\widetilde{\mathbf{Y}}$	$\mathbf{y^c}$	$\mathbf{y^r}$	$\widetilde{\mathbf{Y}}$	$\mathbf{y^c}$	$\mathbf{y^r}$	$\widetilde{\mathbf{Y}}$
1	0.892	0.950	0.898	0.892	0.951	0.898	4.420	0.923	4.517	4.420	0.922	4.517
2	0.904	0.950	0.909	0.904	0.950	0.909	4.160	0.929	4.264	4.160	0.928	4.264
3	1.000	0.896	0.991	1.000	1.000	1.000	0.015	1.335	1.335	0.005	0.001	0.005
4	1.000	0.898	0.991	1.000	1.000	1.000	0.025	1.324	1.324	0.024	0.002	0.024
5	0.909	0.945	0.912	0.909	0.946	0.912	4.111	0.977	4.227	4.111	0.961	4.223
6	0.905	0.954	0.909	0.905	0.954	0.910	4.148	0.893	4.244	4.148	0.890	4.243
7	1.000	0.899	0.991	1.000	1.000	1.000	0.020	1.318	1.318	0.019	0.000	0.019
8	0.999	0.897	0.991	1.000	1.000	1.000	0.270	1.333	1.363	0.270	0.038	0.274
9	1.000	0.898	0.991	1.000	1.000	1.000	0.245	1.327	1.353	0.245	0.032	0.248
10	1.000	0.898	0.991	1.000	1.000	1.000	0.249	1.331	1.357	0.248	0.031	0.251
11	0.901	0.943	0.905	0.901	0.943	0.905	4.237	0.994	4.354	4.237	0.993	4.353
12	0.894	0.947	0.899	0.894	0.947	0.899	4.381	0.957	4.485	4.381	0.956	4.485
13	1.000	0.899	0.991	1.000	1.000	1.000	0.240	1.327	1.352	0.240	0.029	0.243
14	1.000	0.898	0.991	1.000	1.000	1.000	0.259	1.325	1.353	0.258	0.033	0.261
15	0.893	0.949	0.899	0.893	0.952	0.899	4.384	0.935	4.483	4.383	0.907	4.477
16	0.906	0.946	0.910	0.906	0.947	0.910	4.125	0.962	4.237	4.125	0.954	4.235
17	0.902	0.952	0.907	0.902	0.952	0.907	4.217	0.906	4.315	4.217	0.905	4.314
18	0.904	0.955	0.908	0.904	0.955	0.908	4.183	0.879	4.276	4.183	0.878	4.275
19	0.900	0.955	0.905	0.900	0.955	0.905	4.246	0.882	4.338	4.246	0.882	4.338
20	0.901	0.953	0.906	0.901	0.953	0.906	4.220	0.904	4.318	4.220	0.903	4.317
21	0.902	0.955	0.907	0.902	0.955	0.907	4.212	0.886	4.306	4.212	0.886	4.306
22	0.907	0.952	0.911	0.907	0.952	0.911	4.113	0.907	4.213	4.113	0.906	4.213
23	0.900	0.950	0.905	0.900	0.950	0.905	4.238	0.927	4.339	4.238	0.927	4.339
24	0.906	0.953	0.910	0.906	0.953	0.910	4.134	0.908	4.234	4.134	0.907	4.234

Table 8. Scenario analysis ($d = 0$): AoR and RMSE for both the $\overline{\mathcal{S}}_1^*$ and $\overline{\mathcal{S}}_2^*$ systems. Note that in gray are represented the scenarios where $\overline{\mathcal{S}}_1^*$ largely differs from $\overline{\mathcal{S}}_2^*$.

scenario	$MVs_{\overline{\mathcal{S}}_1^*}$			$MVs_{\overline{\mathcal{S}}_2^*}$			$Vs_{\overline{\mathcal{S}}_1^*}$			$Vs_{\overline{\mathcal{S}}_2^*}$		
	$\mathbf{y^c}$	$\mathbf{y^r}$	$\widetilde{\mathbf{Y}}$	$\mathbf{y^c}$	$\mathbf{y^r}$	$\widetilde{\mathbf{Y}}$	$\mathbf{y^c}$	$\mathbf{y^r}$	$\widetilde{\mathbf{Y}}$	$\mathbf{y^c}$	$\mathbf{y^r}$	$\widetilde{\mathbf{Y}}$
1	11.997	3.994	11.997	11.996	3.994	12.029	20.278	0.954	154.814	20.287	0.956	154.824
2	11.992	3.999	11.992	11.991	3.999	11.987	22.020	0.936	164.815	22.030	0.937	164.827
3	12.039	3.986	12.039	12.038	3.986	11.987	43.652	0.004	1.293	43.667	1.842	1.159
4	12.026	3.993	12.026	12.025	3.993	12.013	43.128	0.002	0.608	43.143	1.808	0.571
5	11.954	3.988	11.954	11.955	3.988	11.872	27.971	0.827	184.925	27.988	0.860	184.437
6	11.984	4.020	11.984	11.984	4.020	11.963	22.248	0.977	173.645	22.256	0.982	173.567
7	11.990	3.984	11.990	11.991	3.983	11.959	43.534	0.002	0.544	43.549	1.794	0.512
8	11.989	3.994	11.989	11.990	3.994	11.947	42.710	0.000	0.019	42.725	1.829	0.015
9	11.988	3.994	11.988	11.988	3.994	11.972	48.403	0.004	1.600	48.423	1.820	1.461
10	11.987	4.000	11.987	11.988	4.000	11.985	48.175	0.001	0.504	48.195	1.825	0.456
11	11.996	3.992	11.996	11.996	3.992	12.035	21.495	0.774	132.756	21.504	0.776	132.743
12	11.993	4.011	11.993	11.993	4.011	11.983	20.559	0.881	144.396	20.570	0.883	144.384
13	12.017	4.010	12.017	12.016	4.010	11.997	46.908	0.002	0.766	46.929	1.816	0.718
14	12.021	3.998	12.021	12.021	3.998	12.006	48.028	0.002	0.939	48.049	1.810	0.854
15	11.976	3.991	11.976	11.975	3.991	11.969	20.078	0.918	146.880	20.090	0.971	146.082
16	11.990	3.988	11.990	11.989	3.988	11.995	22.493	0.851	152.539	22.502	0.867	152.374
17	12.002	3.980	12.002	12.002	3.980	12.004	21.776	0.965	167.816	21.795	0.965	167.798
18	12.002	4.003	12.002	12.001	4.003	12.015	22.079	1.007	177.314	22.110	1.007	177.374
19	11.959	3.978	11.959	11.960	3.978	11.958	21.623	1.003	173.139	21.640	1.003	173.130
20	11.954	3.995	11.954	11.954	3.995	11.953	21.797	0.972	169.081	21.816	0.972	169.072
21	12.003	4.004	12.003	12.003	4.004	11.943	21.802	0.994	173.157	21.803	0.995	173.242
22	11.989	3.992	11.989	11.989	3.992	11.996	22.801	0.970	176.394	22.801	0.971	176.456
23	11.961	3.982	11.961	11.961	3.982	11.938	21.583	0.934	161.068	21.584	0.935	161.156
24	12.008	4.011	12.008	12.008	4.011	12.026	22.439	0.977	175.099	22.439	0.978	175.186

Table 9. Scenario analysis ($d = 0$): means (MVs) and variances (Vs) for both the $\overline{\mathcal{S}}_1^*$ and $\overline{\mathcal{S}}_2^*$ systems. Note that in gray are represented the scenarios where $\overline{\mathcal{S}}_1^*$ largely differs from $\overline{\mathcal{S}}_2^*$.

scenario	$d > 0$			$d < 0$			$d = 0$		
	δ	β_2^c	β_2^r	δ	β_2^c	β_2^r	δ	β_2^c	β_2^r
1	0.208	0.151	0.451	0.184	0.053	0.373	0.217	0.092	0.408
2	-0.173	-0.076	-0.240	-0.201	-0.132	-0.464	-0.206	-0.088	-0.417
3	0.097	0.183	-0.639	-0.121	0.143	-0.884	0.009	0.162	-0.803
4	0.089	-0.133	0.811	-0.087	-0.176	0.673	0.006	-0.154	0.823
5	-0.168	0.064	0.281	-0.220	0.121	0.626	-0.172	0.086	0.391
6	0.205	-0.105	-0.651	0.194	-0.061	-0.336	0.210	-0.106	-0.342
7	0.083	0.135	-0.812	-0.115	0.184	-0.614	0.006	0.154	-0.812
8	0.101	-0.181	0.662	-0.084	-0.130	0.845	0.001	-0.163	0.825
9	0.078	0.156	0.290	-0.079	0.075	0.725	0.009	0.132	0.620
10	0.085	-0.066	-0.754	-0.084	-0.155	-0.341	-0.005	-0.128	-0.631
11	0.209	0.157	-0.694	0.171	0.073	-0.413	0.190	0.115	-0.513
12	-0.179	-0.079	0.385	-0.204	-0.148	0.640	-0.208	-0.107	0.592
13	0.085	0.064	0.743	-0.073	0.151	0.338	0.007	0.121	0.662
14	0.073	-0.154	-0.351	-0.084	-0.069	-0.732	0.007	-0.135	-0.594
15	-0.165	0.068	-0.416	-0.204	0.141	-0.670	-0.215	0.132	-0.520
16	0.214	-0.146	0.721	0.194	-0.078	0.421	0.195	-0.105	0.564
17	0.193	-0.076	0.865	0.204	-0.019	0.495	0.211	-0.041	0.778
18	-0.227	0.020	-0.533	-0.224	0.088	-1.043	-0.214	0.040	-0.790
19	-0.202	-0.018	0.501	-0.215	-0.082	1.069	-0.216	-0.042	0.793
20	0.213	0.092	-1.055	0.217	0.019	-0.508	0.212	0.041	-0.783
21	0.205	0.226	-0.421	0.211	0.106	-0.098	0.214	0.166	-0.198
22	-0.215	-0.107	0.094	-0.210	-0.203	0.395	-0.207	-0.164	0.197
23	-0.216	0.109	-0.101	-0.201	0.216	-0.391	-0.209	0.161	-0.195
24	0.233	-0.231	0.411	0.219	-0.109	0.095	0.209	-0.164	0.184

Table 10. Scenario analysis: regression parameters for the equation $\mathbf{y^r}$ for both the $\overline{\mathcal{S}}_1^*$ and $\overline{\mathcal{S}}_2^*$ systems. Note that in gray are represented the scenarios where $\overline{\mathcal{S}}_1^*$ largely differs from $\overline{\mathcal{S}}_2^*$.