Supplementary Material for "Robust Estimation of fixed effect parameters and variances of Linear Mixed Models: The Minimum Density Power Divergence Approach"

Giovanni Saraceno[1], Abhik Ghosh[2], Ayanendranath Basu[2], and Claudio Agostinelli[1]

[1]Department of Mathematics, University of Trento, Trento, Italy
[2]Interdisciplinary Statistical Research Unit, Indian Statistical Institute, Kolkata, India

Supplementary Material

The Supplementary Material reports the assumptions needed to prove the asymptotic normality of the MDPDE estimator in Section SM–1. The simplification of the sensitivity measures in case of balanced data are shown in Section SM–2 while the equivariance properties of the estimator in Section SM–3. Section SM–4 shows an illustrative example about the robustness achieved by the predictions of random effects using the proposed estimator, while Section SM–5 reports the missing plots about theoretical quantities. Further results obtained from the Monte Carlo experiments are presented in Section SM–6, whereas Section SM–7 contains complete results from the study of the real-data example on orthodontic measures. Finally, Section SM–8 reports the analysis of the real life data on foveal and extrafoveal vision acuity (and crowding) studying their interrelationships with one's reading performances.

SM–1 Asymptotic properties

Continuing with the notation and set-up of Section 2, let $\mathcal{F}_{i,\boldsymbol{\theta}}$ be the parametric model. Assume that there exists a best fitting parameter of $\boldsymbol{\theta}$ which is independent of the index i of the different densities and let us denote it by $\boldsymbol{\theta}^g$. We also assume that all the true densities g_i belong to the model family, i.e. $g_i = f_i(\cdot;\boldsymbol{\theta})$ for some common $\boldsymbol{\theta}$, and in that case the best fitting parameter is nothing but the true parameter $\boldsymbol{\theta}$. We define, for each $i = 1, \ldots, n$, the $p \times p$ matrix $\boldsymbol{J}^{(i)}$ as

$$\boldsymbol{J}^{(i)} = (1+\alpha)\left[\int u_i(\boldsymbol{y};\boldsymbol{\theta}^g)u_i^\top(\boldsymbol{y};\boldsymbol{\theta}^g)f_i^{1+\alpha}(\boldsymbol{y};\boldsymbol{\theta}^g)dy \right.$$
$$\left. - \int\{\nabla u_i(\boldsymbol{y};\boldsymbol{\theta}^g) + \alpha u_i(\boldsymbol{y};\boldsymbol{\theta}^g)u_i^\top(\boldsymbol{y};\boldsymbol{\theta}^g)\}\{g_i(\boldsymbol{y}) - f_i(\boldsymbol{y};\boldsymbol{\theta}^g)\}f_i(\boldsymbol{y};\boldsymbol{\theta}^g)^\alpha dy\right].$$

We further define the quantities $\boldsymbol{\Psi}_n = \frac{1}{n}\sum_{i=1}^n \boldsymbol{J}^{(i)}$, and

$$\boldsymbol{\Omega}_n = (1+\alpha)^2\frac{1}{n}\sum_{i=1}^n\left[\int u_i(\boldsymbol{y};\boldsymbol{\theta}^g)u_i^\top(\boldsymbol{y};\boldsymbol{\theta}^g)f_i(\boldsymbol{y};\boldsymbol{\theta}^g)^{2\alpha}g_i(\boldsymbol{y})dy - \boldsymbol{\xi}_i\boldsymbol{\xi}_i^\top\right],$$

where $\boldsymbol{\xi}_i = \int u_i(\boldsymbol{y};\boldsymbol{\theta}^g) f_i(\boldsymbol{y};\boldsymbol{\theta}^g)^\alpha g_i(\boldsymbol{y}) dy$.

We will make the following assumptions to establish the asymptotic properties of the MDPDE under this general set-up of independent but non-homogeneous observations Ghosh and Basu [2013].

(A1) The support $\chi = \{\boldsymbol{y} | f_i(\boldsymbol{y};\boldsymbol{\theta}) > 0\}$ is independent of i and $\boldsymbol{\theta}$ for all i; the true distributions G_i are also supported for all i.

(A2) There is an open subset ω of the parameter space Θ, containing the best fitting parameter $\boldsymbol{\theta}^g$ such that for almost all $\boldsymbol{y} \in \chi$, and all $\boldsymbol{\theta} \in \Theta$, all $i = 1, \ldots, n$, the density $f_i(\boldsymbol{y};\boldsymbol{\theta})$ is thrice differentiable with respect to $\boldsymbol{\theta}$ and the third partial derivatives are continuous with respect to $\boldsymbol{\theta}$.

(A3) For $i = 1, \ldots, n$, the integrals $\int f_i(\boldsymbol{y};\boldsymbol{\theta})^{1+\alpha} dy$ and $\int f_i(\boldsymbol{y};\boldsymbol{\theta})^\alpha g_i(\boldsymbol{y}) dy$ can be differentiated thrice with respect to $\boldsymbol{\theta}$, and the derivatives can be taken under the integral sign.

(A4) For each $i = 1, \ldots, n$, the matrices $\boldsymbol{J}^{(i)}$ are positive definite and

$$\lambda_0 = \inf_n [\text{min eigenvalue of } \boldsymbol{\Psi}_n] > 0$$

(A5) There exists a function $M_{jkl}^{(i)}(\boldsymbol{y})$ such that

$$|\nabla_{jkl} H_i(\boldsymbol{y},\boldsymbol{\theta})| \leq M_{jkl}^{(i)}(\boldsymbol{y}) \qquad \forall \boldsymbol{\theta} \in \Theta, \ \forall i$$

where

$$\frac{1}{n}\sum_{i=1}^n \mathrm{E}_{g_i}[M_{jkl}^{(i)}(\boldsymbol{Y})] = \mathcal{O}(1) \qquad \forall j,k,l$$

(A6) For all j, k, we have

$$\lim_{N\to\infty} \sup_{n>1} \left\{ \frac{1}{n}\sum_{i=1}^n \mathrm{E}_{g_i}[|\nabla_j H_i(\boldsymbol{Y},\boldsymbol{\theta})| I(|\nabla_j H_i(\boldsymbol{Y},\boldsymbol{\theta})| > N)] \right\} = 0 \tag{1}$$

$$\lim_{N\to\infty} \sup_{n>1} \left\{ \frac{1}{n}\sum_{i=1}^n \mathrm{E}_{g_i}[|\nabla_{jk} H_i(\boldsymbol{Y},\boldsymbol{\theta}) - \mathrm{E}_{g_i}(\nabla_{jk} H_i(\boldsymbol{Y},\boldsymbol{\theta}))| \times I(|\nabla_{jk} H_i(\boldsymbol{Y},\boldsymbol{\theta}) - \mathrm{E}_{g_i}(\nabla_{jk} H_i(\boldsymbol{Y},\boldsymbol{\theta}))| > N)] \right\} = 0 \tag{2}$$

where $I(B)$ denotes the indicator variable of the event B.

(A7) For all $\epsilon > 0$, we have

$$\lim_{n\to\infty} \left\{ \frac{1}{n}\sum_{i=1}^n \mathrm{E}_{g_i} \left[\|\boldsymbol{\Omega}_n^{-1/2} \nabla H_i(\boldsymbol{Y},\boldsymbol{\theta})\|^2 I(\|\boldsymbol{\Omega}_n^{-1/2} \nabla H_i(\boldsymbol{Y},\boldsymbol{\theta})\| > \epsilon\sqrt{n}) \right] \right\} = 0. \tag{3}$$

Ghosh and Basu [2013] proved that, under assumptions (A1)-(A7), the MDPDE $\hat{\boldsymbol{\theta}}$ is consistent for $\boldsymbol{\theta}$ and asymptotically normal. In particular, the asymptotic distribution of $\boldsymbol{\Omega}_n^{-1/2} \boldsymbol{\Psi}_n [\sqrt{n}(\boldsymbol{\theta}_n - \boldsymbol{\theta}^g)]$ is p-dimensional normal with (vector) mean 0 and covariance matrix $\boldsymbol{I}_p$, the p-dimensional identity matrix.

SM–2 Sensitivity measures

Gross Error Sensitivity

Considering the influence function of the functional $T_\alpha^{\boldsymbol{\beta}}$, we have that

$$\begin{aligned}\gamma_{i_0}^u(T_\alpha^{\boldsymbol{\beta}}, G_1, \ldots, G_n) =& \sup_{\boldsymbol{t}} \left\{ \left\| (\boldsymbol{X'}^\top \boldsymbol{X'})^{-1} \boldsymbol{X}_{i_0}^\top \boldsymbol{V}_{i_0}^{-1} (\boldsymbol{t}_{i_0} - \boldsymbol{X}_{i_0}\boldsymbol{\beta}) f_{i_0}(\boldsymbol{t}_{i_0}; \boldsymbol{\theta})^\alpha \right\| \right\} \\ =& \frac{\sup_{\boldsymbol{Z}} \left\{ \left\| (\boldsymbol{X'}^\top \boldsymbol{X'})^{-1} \boldsymbol{X}_{i_0}^\top \boldsymbol{V}_{i_0}^{-1/2} \boldsymbol{Z} \right\| e^{-\frac{\boldsymbol{Z}^\top \boldsymbol{Z}}{2}} \right\}}{\sqrt{\alpha}(2\pi)^{\frac{n_i \alpha}{2}} |\boldsymbol{V}_{i_0}|^{\frac{\alpha}{2}}},\end{aligned}$$

where $\boldsymbol{Z} = \sqrt{\alpha}\boldsymbol{V}_{i_0}^{-1/2}(\boldsymbol{t} - \boldsymbol{X}_{i_0}\boldsymbol{\beta})$. Denoting $\boldsymbol{A} = (\boldsymbol{X'}^\top \boldsymbol{X'})^{-1} \boldsymbol{X}_{i_0}^\top \boldsymbol{V}_{i_0}^{-1/2}$, we have to find the sup of the function $\|\boldsymbol{AZ}\| \, e^{\frac{\boldsymbol{Z}^\top \boldsymbol{Z}}{2}}$ with respect to $\boldsymbol{Z}$. Then, we compute the derivative with respect to $\boldsymbol{Z}$ obtaining

$$\frac{\partial((\boldsymbol{Z}^\top \boldsymbol{A}^\top \boldsymbol{AZ})^{1/2} e^{-\frac{\boldsymbol{Z}^\top \boldsymbol{Z}}{2}})}{\partial \boldsymbol{Z}} = \frac{\boldsymbol{A}^\top \boldsymbol{AZ} e^{-\frac{\boldsymbol{Z}^\top \boldsymbol{Z}}{2}}}{(\boldsymbol{Z}^\top \boldsymbol{A}^\top \boldsymbol{AZ})^{1/2}} - (\boldsymbol{Z}^\top \boldsymbol{A}^\top \boldsymbol{AZ})^{1/2} e^{-\frac{\boldsymbol{Z}^\top \boldsymbol{Z}}{2}} \boldsymbol{Z} = 0.$$

Finally, multiplying the above equation by $\boldsymbol{Z}^\top$ we have

$$\begin{aligned}&(\boldsymbol{Z}^\top \boldsymbol{A}^\top \boldsymbol{AZ})^{1/2} e^{-\frac{\boldsymbol{Z}^\top \boldsymbol{Z}}{2}} - (\boldsymbol{Z}^\top \boldsymbol{A}^\top \boldsymbol{AZ})^{1/2} e^{-\frac{\boldsymbol{Z}^\top \boldsymbol{Z}}{2}} \boldsymbol{Z}^\top \boldsymbol{Z} = 0 \\ &(\boldsymbol{Z}^\top \boldsymbol{A}^\top \boldsymbol{AZ})^{1/2} e^{-\frac{\boldsymbol{Z}^\top \boldsymbol{Z}}{2}} \left[1 - \boldsymbol{Z}^\top \boldsymbol{Z} \right] = 0.\end{aligned}$$

A solution is given by $\boldsymbol{Z}$ such that $\boldsymbol{Z}^\top \boldsymbol{Z} = 1$, that is $\boldsymbol{Z} = \frac{\boldsymbol{k}}{\|\boldsymbol{k}\|}$ with $\boldsymbol{k} \in \mathbb{R}^{n_{i_0}}$. Hence

$$\begin{aligned}\sup_{\boldsymbol{Z}} \left\{ \|\boldsymbol{AZ}\| \, e^{-\frac{\boldsymbol{Z}^\top \boldsymbol{Z}}{2}} \right\} =& e^{-1/2} \sup_{\boldsymbol{k}} \left\{ \frac{\|\boldsymbol{Ak}\|}{\|\boldsymbol{k}\|} \right\} \\ =& e^{-1/2} \sup_{\boldsymbol{k}} \left\{ \frac{\boldsymbol{k}^\top \boldsymbol{A}^\top \boldsymbol{Ak}}{\boldsymbol{k}^\top \boldsymbol{k}} \right\}^{\frac{1}{2}}.\end{aligned}$$

We know that $\sup_{\boldsymbol{z}} \left\{ \frac{\boldsymbol{z}^\top \boldsymbol{Az}}{\boldsymbol{z}^\top \boldsymbol{z}} \right\} = \lambda_{max}(\boldsymbol{A})$, where $\lambda_{max}(\boldsymbol{A})$ is the maximum eigenvalue of the matrix $\boldsymbol{A}$. Using this general results in our case , we obtain

$$\gamma_{i_0}^u(T_\alpha^{\boldsymbol{\beta}}, G_1, \ldots, G_n) = \frac{(\lambda_{max}((\boldsymbol{X'}^\top \boldsymbol{X'})^{-2} \boldsymbol{X}_{i_0}^\top \boldsymbol{V}_{i_0}^{-1} \boldsymbol{X}_{i_0}))^{\frac{1}{2}}}{\sqrt{e}\sqrt{\alpha}(2\pi)^{\frac{n_i \alpha}{2}} |\boldsymbol{V}_{i_0}|^{\frac{\alpha}{2}}}$$

Special Case: When $n_i = p$, for all $i = 1, \ldots, n$, the matrix $\boldsymbol{X'}^\top \boldsymbol{X'}$ can be rewritten as

$$\begin{aligned}\boldsymbol{X'}^\top \boldsymbol{X'} =& \sum_{i=1}^n \frac{\boldsymbol{X}_i^\top \boldsymbol{V}^{-1} \boldsymbol{X}_i}{(1+\alpha)^{\frac{p}{2}+1} (2\pi)^{\frac{p\alpha}{2}} |\boldsymbol{V}|^{\frac{\alpha}{2}}} \\ =& \frac{1}{(1+\alpha)^{\frac{p}{2}+1} (2\pi)^{\frac{p\alpha}{2}} |\boldsymbol{V}|^{\frac{\alpha}{2}}} \sum_{i=1}^n \boldsymbol{X}_i^\top \boldsymbol{V}^{-1} \boldsymbol{X}_i\end{aligned}$$

Substituting this simpler form in the formula of the gross error sensitivity we get Equation (24).

Self-Standardized Sensitivity

Starting from the definition of the self-standardized sensitivity, we have that

$$\gamma^s_{i_0}(T^{\boldsymbol{\beta}}_\alpha, G_1, \ldots, G_n) = \frac{1}{n} \sup_{\boldsymbol{t}} \left\{ (\boldsymbol{t} - \boldsymbol{X}_{i_0}\boldsymbol{\beta})^\top \boldsymbol{V}_{i_0}^{-1} \boldsymbol{X}_{i_0} (\boldsymbol{X}^{*\top}\boldsymbol{X}^*)^{-1} \boldsymbol{X}_{i_0}^\top \boldsymbol{V}_{i_0}^{-1} (\boldsymbol{t} - \boldsymbol{X}_{i_0}\boldsymbol{\beta}) f_i^{2\alpha}(\boldsymbol{t}, \boldsymbol{\theta}) \right\}^{\frac{1}{2}}.$$

Let $\boldsymbol{Z} = \left(\frac{\boldsymbol{V}_{i_0}}{2\alpha}\right)^{-1/2} (\boldsymbol{t} - \boldsymbol{X}_{i_0}\boldsymbol{\beta})$, then the above equation has the the form

$$\begin{aligned}
\gamma^s_{i_0}(T^{\boldsymbol{\beta}}_\alpha, G_1, \ldots, G_n) =& \frac{1}{n\sqrt{2\alpha}(2\pi)^{\frac{n_i\alpha}{2}} |\boldsymbol{V}_{i_0}|^{\frac{\alpha}{2}}} \sup_{\boldsymbol{Z}} \left\{ \boldsymbol{Z}^\top \boldsymbol{V}_{i_0}^{-1/2} \boldsymbol{X}_{i_0} (\boldsymbol{X}^{*\top}\boldsymbol{X}^*)^{-1} \boldsymbol{X}_{i_0}^\top \boldsymbol{V}_{i_0}^{-1/2} \boldsymbol{Z} e^{-\frac{\boldsymbol{Z}^\top\boldsymbol{Z}}{2}} \right\}^{\frac{1}{2}} \\
=& \frac{1}{n\sqrt{2\alpha}(2\pi)^{\frac{n_i\alpha}{2}} |\boldsymbol{V}_{i_0}|^{\frac{\alpha}{2}}} \sup_{\boldsymbol{Z}} \left\{ \boldsymbol{Z}^\top \boldsymbol{A} \boldsymbol{Z} e^{-\frac{\boldsymbol{Z}^\top\boldsymbol{Z}}{2}} \right\}^{\frac{1}{2}},
\end{aligned}$$

where $\boldsymbol{A} = \boldsymbol{V}_{i_0}^{-1/2} \boldsymbol{X}_{i_0} (\boldsymbol{X}^{*\top}\boldsymbol{X}^*)^{-1} \boldsymbol{X}_{i_0}^\top \boldsymbol{V}_{i_0}^{-1/2}$. In order to find this sup, we compute the derivative with respect to $\boldsymbol{Z}$, obtaining

$$\begin{aligned}
\frac{\partial(\boldsymbol{Z}^\top \boldsymbol{A}\boldsymbol{Z} e^{-\frac{\boldsymbol{Z}^\top\boldsymbol{Z}}{2}})}{\partial \boldsymbol{Z}} =& \boldsymbol{Z}^\top \left[2\boldsymbol{A}\boldsymbol{Z} e^{-\frac{\boldsymbol{Z}^\top\boldsymbol{Z}}{2}} - \boldsymbol{Z}^\top \boldsymbol{A}\boldsymbol{Z} e^{-\frac{\boldsymbol{Z}^\top\boldsymbol{Z}}{2}} \boldsymbol{Z} \right] \\
=& \boldsymbol{Z}^\top \boldsymbol{A}\boldsymbol{Z} e^{-\frac{\boldsymbol{Z}^\top\boldsymbol{Z}}{2}} [2 - \boldsymbol{Z}^\top \boldsymbol{Z}] = 0.
\end{aligned}$$

A solution is given by $\boldsymbol{Z}$ such that $\boldsymbol{Z}^\top \boldsymbol{Z} = 2$, then $\boldsymbol{Z} = \frac{\sqrt{2}\boldsymbol{k}}{\|\boldsymbol{k}\|}$, with $\boldsymbol{k} \in \mathbb{R}^{n_{i_0}}$. Note that we multiplied by $\boldsymbol{Z}^\top$ in the same way done for the gross error sensitivity. Hence

$$\begin{aligned}
\gamma^s_{i_0}(T^{\boldsymbol{\beta}}_\alpha, G_1, \ldots, G_n) =& \frac{\sqrt{2}}{n\sqrt{2\alpha}(2\pi)^{\frac{n_i\alpha}{2}} |\boldsymbol{V}_{i_0}|^{\frac{\alpha}{2}}} \sup_{\boldsymbol{k}} \left\{ \frac{\boldsymbol{k}^\top \boldsymbol{A} \boldsymbol{k}}{\|\boldsymbol{k}\|^2} \right\}^{\frac{1}{2}} \\
=& \frac{\left[\lambda_{max} \left((\boldsymbol{X}^{*\top}\boldsymbol{X}^*)^{-1} \boldsymbol{X}_{i_0}^\top \boldsymbol{V}_{i_0}^{-1} \boldsymbol{X}_{i_0} \right) \right]^{1/2}}{n\sqrt{\alpha}(2\pi)^{\frac{n_i\alpha}{2}} |\boldsymbol{V}_i|^{\frac{\alpha}{2}} e^{1/2}},
\end{aligned}$$

which corresponds to equation (22).
Special Case: When $n_i = p$, for all $i = 1, \ldots, n$, the matrix $\boldsymbol{X}^{*\top}\boldsymbol{X}^*$ can be rewritten as

$$\begin{aligned}
\boldsymbol{X}^{*\top}\boldsymbol{X}^* =& \sum_{i=1}^n \frac{\boldsymbol{X}_i^\top \boldsymbol{V}^{-1} \boldsymbol{X}_i}{(1+\alpha)^{\frac{p}{2}+1} (2\pi)^{p\alpha} |\boldsymbol{V}|^\alpha} \\
=& \frac{1}{(1+\alpha)^{\frac{p}{2}+1} (2\pi)^{p\alpha} |\boldsymbol{V}|^\alpha} \sum_{i=1}^n \boldsymbol{X}_i^\top \boldsymbol{V}^{-1} \boldsymbol{X}_i
\end{aligned}$$

Substituting this simpler form in the formula of the self-standardized sensitivity we get equation (25).

SM–3 Equivariance

Here we show the equivariance properties of the MDPDE. Recall that, the estimator $\hat{\boldsymbol{\theta}} = (\hat{\boldsymbol{\beta}}, \sigma_j, j = 0, \ldots, r)$ is as defined in equation (7) and $\boldsymbol{y}_i$ is a n_i-variate random vector such that $\boldsymbol{y}_i \sim N_{n_i}(\boldsymbol{x}_i\boldsymbol{\beta}, \boldsymbol{V}_i)$ where $\boldsymbol{x}_i$ is a $n_i \times k$ model matrix. Note that, the case $\alpha = 0$ corresponds to the MLE which satisfies the equivariance properties. Then we show that the same hold for $\alpha > 0$.

Regression equivariance

Let $\mathring{\boldsymbol{y}}_i = \boldsymbol{y}_i + \boldsymbol{x}_i\boldsymbol{\delta}$ for some vector $\boldsymbol{\delta}$. Then $\mathring{\boldsymbol{y}}_i \sim N_{n_i}(\boldsymbol{x}_i(\boldsymbol{\beta}-\boldsymbol{\delta}), \boldsymbol{V}_i)$. An estimate $\hat{\boldsymbol{\beta}}(\boldsymbol{x}, \boldsymbol{y})$ for the fixed terms is *regression equivariant* if $\mathring{\boldsymbol{\beta}} = \hat{\boldsymbol{\beta}}(\boldsymbol{x}, \mathring{\boldsymbol{y}}) = \hat{\boldsymbol{\beta}}(\boldsymbol{x}, \boldsymbol{y}) + \boldsymbol{\delta}$.

We have that

$$\begin{aligned}
\hat{\boldsymbol{\beta}}(\boldsymbol{x}_i, \mathring{\boldsymbol{y}}_i) &= \arg\min_{\boldsymbol{\beta}} \frac{1}{n}\sum_{i=1}^{n}\left[\eta_{i\alpha}(\alpha+1)^{-\frac{n_i}{2}} - \left(1+\frac{1}{\alpha}\right)w_i((\boldsymbol{\beta}-\boldsymbol{\delta}), \boldsymbol{V}_i)\right] \\
&= \arg\min_{(\boldsymbol{\beta}-\boldsymbol{\delta})} \frac{1}{n}\sum_{i=1}^{n}\left[\eta_{i\alpha}(\alpha+1)^{-\frac{n_i}{2}} - \left(1+\frac{1}{\alpha}\right)w_i((\boldsymbol{\beta}-\boldsymbol{\delta}), \boldsymbol{V}_i)\right] + \boldsymbol{\delta} \\
&= \hat{\boldsymbol{\beta}}(\boldsymbol{x}_i, \boldsymbol{y}_i) + \boldsymbol{\delta}.
\end{aligned}$$

This shows that $\hat{\beta}$ is regression equivariant. Finally, note that the estimates of the variance components remains unchanged, hence $\hat{\sigma}_0^2$ and $\hat{\gamma}_j$ are regression invariant, i.e. $\mathring{\sigma}_0^2 = \hat{\sigma}_0^2$ and $\mathring{\gamma}_j = \hat{\gamma}_j$.

Affine equivariance

Let $\boldsymbol{A}$ be a non-singular square matrix and $\mathring{\boldsymbol{x}}_i = \boldsymbol{x}_i\boldsymbol{A}$. An estimate is *affine equivariant* if $\mathring{\boldsymbol{\beta}} = \hat{\boldsymbol{\beta}}(\mathring{\boldsymbol{x}}, \boldsymbol{y}) = \boldsymbol{A}^{-1}\hat{\boldsymbol{\beta}}(\boldsymbol{x}, \boldsymbol{y})$.

Considering the estimate for $\boldsymbol{\beta}$, we have that

$$\begin{aligned}
\hat{\boldsymbol{\beta}}(\mathring{\boldsymbol{x}}_i, \boldsymbol{y}_i) &= \arg\min_{\boldsymbol{\beta}} \frac{1}{n}\sum_{i=1}^{n}\left[\eta_{i\alpha}(\alpha+1)^{-\frac{n_i}{2}} - \left(1+\frac{1}{\alpha}\right)w_i(\boldsymbol{A}\boldsymbol{\beta}, \boldsymbol{V}_i)\right] \\
&= \boldsymbol{A}^{-1}\arg\min_{\boldsymbol{A}\boldsymbol{\beta}} \frac{1}{n}\sum_{i=1}^{n}\left[\eta_{i\alpha}(\alpha+1)^{-\frac{n_i}{2}} - \left(1+\frac{1}{\alpha}\right)w_i(\boldsymbol{A}\boldsymbol{\beta}, \boldsymbol{V}_i)\right] \\
&= \boldsymbol{A}^{-1}\hat{\boldsymbol{\beta}}(\boldsymbol{x}_i, \boldsymbol{y}_i) + \boldsymbol{\delta}.
\end{aligned}$$

This proves that the estimate $\hat{\boldsymbol{\beta}}$ is affine equivariant. As before, the estimation of the variance components parameters is not affected, hence $\hat{\sigma}_0^2$ and $\hat{\gamma}_j$ are affine invariant, i.e. $\mathring{\sigma}_0^2 = \hat{\sigma}_0^2$ and $\mathring{\gamma}_j = \hat{\gamma}_j$.

Scale equivariance

Let $\mathring{\boldsymbol{y}}_i = \delta\boldsymbol{y}_i$ for some scalar $\boldsymbol{\delta}$. Then $\mathring{\boldsymbol{y}}_i \sim N_{n_i}(\boldsymbol{x}_i(\delta\boldsymbol{\beta}), \mathring{\boldsymbol{V}}_i)$ with $\mathring{\boldsymbol{V}}_i = \mathring{\sigma}_0^2\,(\boldsymbol{I}_{n_i} + \sum_{j=1}^{r}\boldsymbol{Z}_{ij}\boldsymbol{Z}_{ij}^\top\gamma_j)$ where $\mathring{\sigma}_0^2 = \delta^2\sigma_0^2$.

Note that $\mathring{\boldsymbol{V}}_i = \delta^2\boldsymbol{V}_i$, then $\mathring{\eta}_{i\alpha} = |\delta|^{-\alpha}\eta_{i\alpha}$ and

$$\begin{aligned}
\mathring{w}_i &= \mathring{\eta}_{i\alpha}\exp\left\{-\frac{\alpha}{2}(\delta\boldsymbol{y}_i - \boldsymbol{x}_i\boldsymbol{\beta}\delta)^\top\,\mathring{\boldsymbol{V}}_i^{-1}\,(\delta\boldsymbol{y}_i - \boldsymbol{x}_i\boldsymbol{\beta}\delta)\right\} \\
&= |\delta|^{-\alpha}\eta_{i\alpha}\exp\left\{-\frac{\alpha}{2}(\boldsymbol{y}_i - \boldsymbol{x}_i\boldsymbol{\beta})^\top\boldsymbol{V}_i^{-1}(\boldsymbol{y}_i - \boldsymbol{x}_i\boldsymbol{\beta})\right\} \\
&= |\delta|^{-\alpha}w_i.
\end{aligned}$$

Then

$$\hat{\boldsymbol{\theta}}(\boldsymbol{x}_i, \mathring{\boldsymbol{y}}_i) = \arg\min_{\boldsymbol{\theta}} \frac{1}{n} \sum_{i=1}^{n} |\delta|^{-\alpha} \left[\eta_{i\alpha}(\alpha+1)^{-\frac{n_i}{2}} - \left(1 + \frac{1}{\alpha}\right) w_i \right]$$

Hence the estimates $\hat{\beta}$ and $\hat{\gamma}_j$ are scale invariant, i.e. $\mathring{\boldsymbol{\beta}} = \hat{\boldsymbol{\beta}}$ and $\mathring{\gamma}_j = \hat{\gamma}_j$, while $\mathring{\sigma}_0^2 = \delta^2 \hat{\sigma}_0^2$.

SM–4 Example: robustness of predicted random effects

Here, we present an illustrative example to show the robustness achieved by the predicted random effects $\boldsymbol{u}_i$, $i = 1, \ldots, n$ for each observation.

Consider the model setting

$$y_{fgh} = \boldsymbol{x}_{fgh}^\top \boldsymbol{\beta}_0 + a_f + b_g + c_{fg} + e_{fgh},$$

used in the simulation study introduced in Section 5 of the manuscript. We also consider the complete contamination scenario with $\varepsilon = 15\%$, $\phi_0 = 1$ and $\omega_0 = 3$. At first, we compute the estimates of fixed effects parameters and variances through the proposed robust MDPDE with $\alpha = \frac{1}{6}$ and by the MLE. Then, the predictions of random effects are derived for the MDPDE using the formula given in equation 23, while for MLE the standard formulation is used.

Table 1: True values and estimates of fixed effects parameters and variances for the MDPDE with $\alpha = 1/6$ and MLE under 15% of outlier contamination.

Method	$\boldsymbol{\beta}$						$\sigma_{\boldsymbol{a}}$	$\sigma_{\boldsymbol{b}}$	$\sigma_{\boldsymbol{c}}$	$\sigma_{\boldsymbol{e}}$
true	0	2	2	2	2	2	0.06	0.06	0.12	0.25
MLE	0.52	2.04	2.01	2.02	2.02	1.99	0.60	0.43	1.96	0.53
MDPDE	0.10	2.01	1.98	1.97	2.00	1.97	0.07	0.06	0.08	0.28

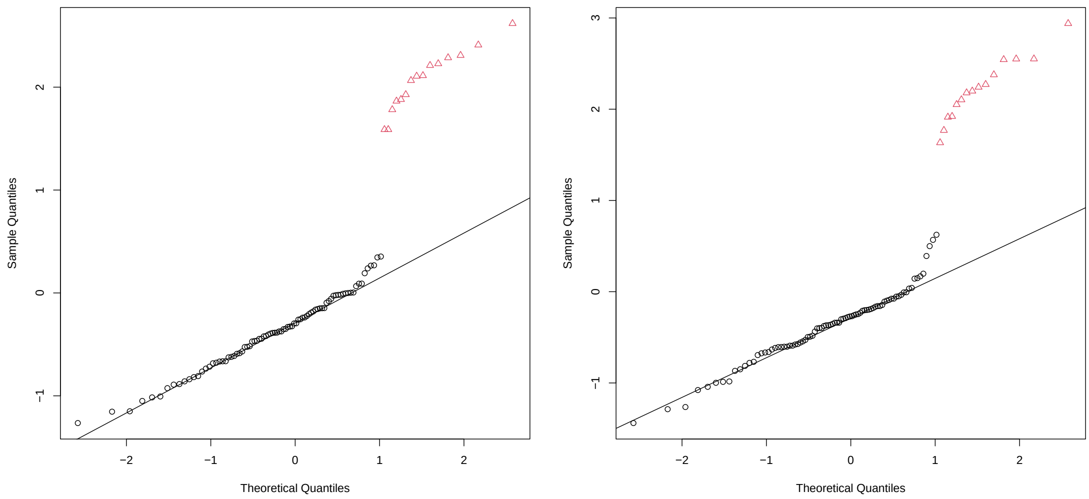
Table 1 reports the obtained estimates. The variance components estimated by the MLE are highly affected by the contamination in comparison with the MDPDE. On the other hand, the fixed effect estimates are quite similar, except for the intercept $\beta_0 = 0$. Figures 1, 2 and 3 show the QQ-plots of predictions of random terms $\boldsymbol{a}$, $\boldsymbol{b}$ and $\boldsymbol{c}$, respectively, for the MLE and the MDPDE with $\alpha = 1/6$. Values corresponding to outlying observations are displayed as red triangles while the rest of observations as black circles.

These Figures show that the predictions obtained using the robust formulation given in equation 23 clearly identify the outlying observations, due to the use of robust estimates given by the MDPDE. On the other hand, the MLE predictions tend to hide the outliers.

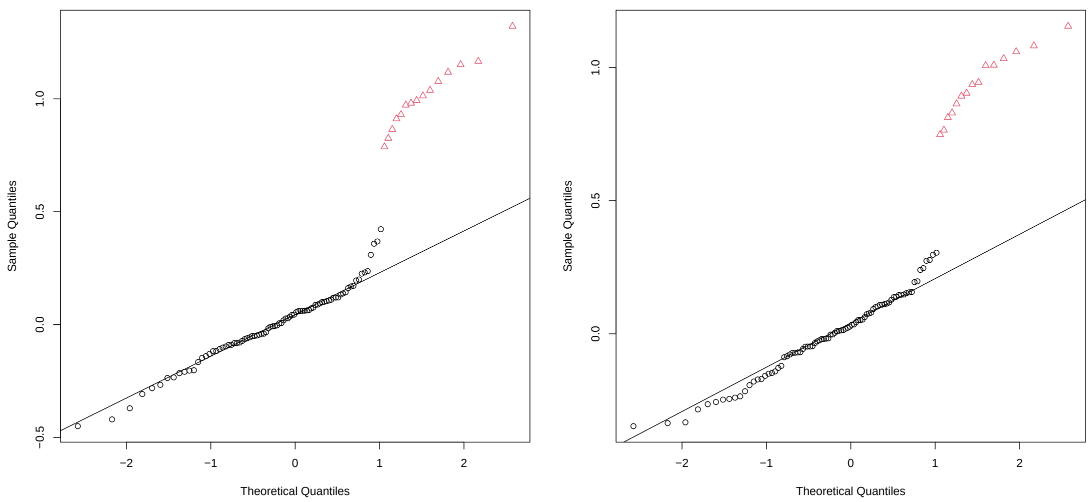
SM–5 Additional Numerical Results

In this section we provide the plots of Asymptotic Relative Efficiency and Influence functions, corresponding to the Example in Section 4.4.

Figure 4 shows the Asymptotic Relative Efficiency with respect to α for β_0, σ_0^2 and σ_1^2, respectively. In Figure 5 we can see the influence function for the estimators $T_\alpha^{\beta_1}$, $T_\alpha^{\sigma_0^2}$ and $T_\alpha^{\sigma_2^2}$ for $\alpha = 0, 0.05$. While Figure 6 presents the same quantities for $\alpha = \alpha^* = \frac{1}{11}$ and $\alpha = \bar{\alpha} = 0.2$.

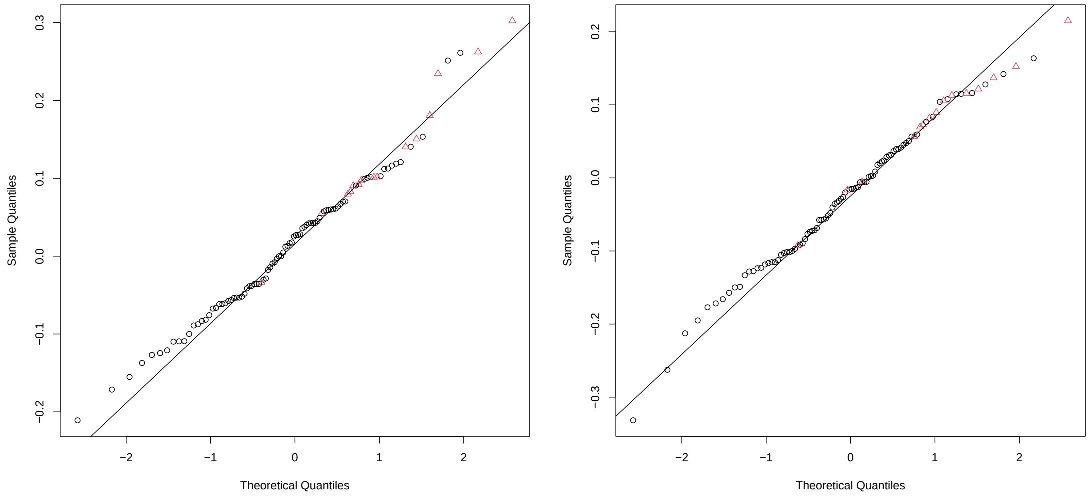


(a) MLE

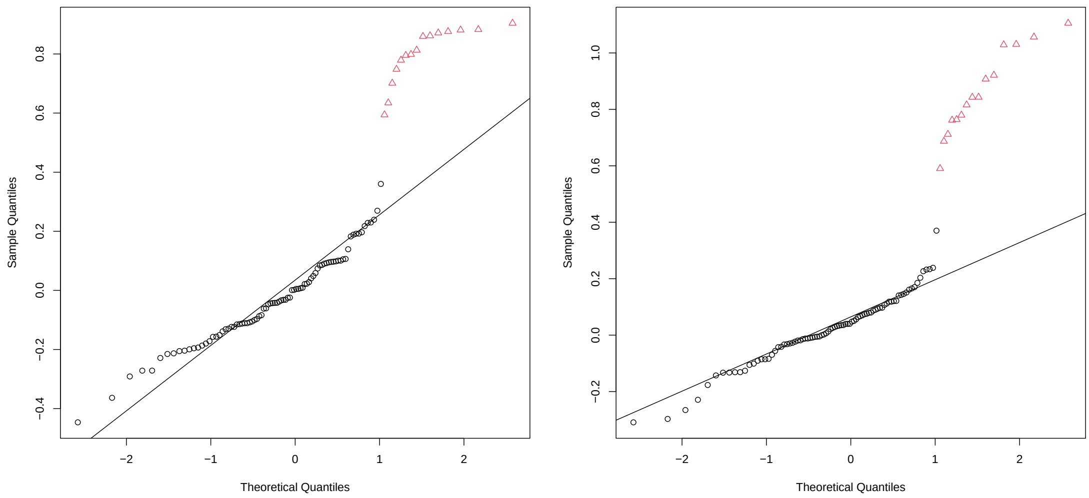


(b) MDPDE

Figure 1: QQ-plots of the random term $\boldsymbol{a}$ predicted by the MLE (top) and the MDPD-estimator (bottom) for $\alpha = \bar{\alpha} = 1/6$. Values corresponding to outlying observations are displayed as red triangles while the rest of observations as black circles.

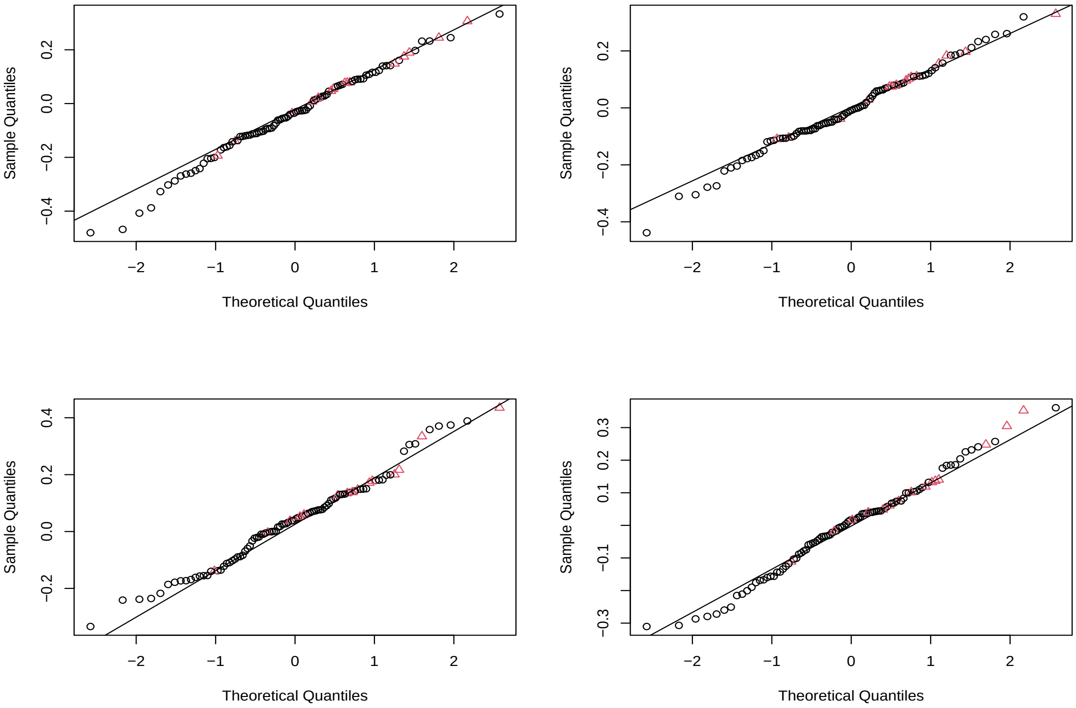


(a) MLE

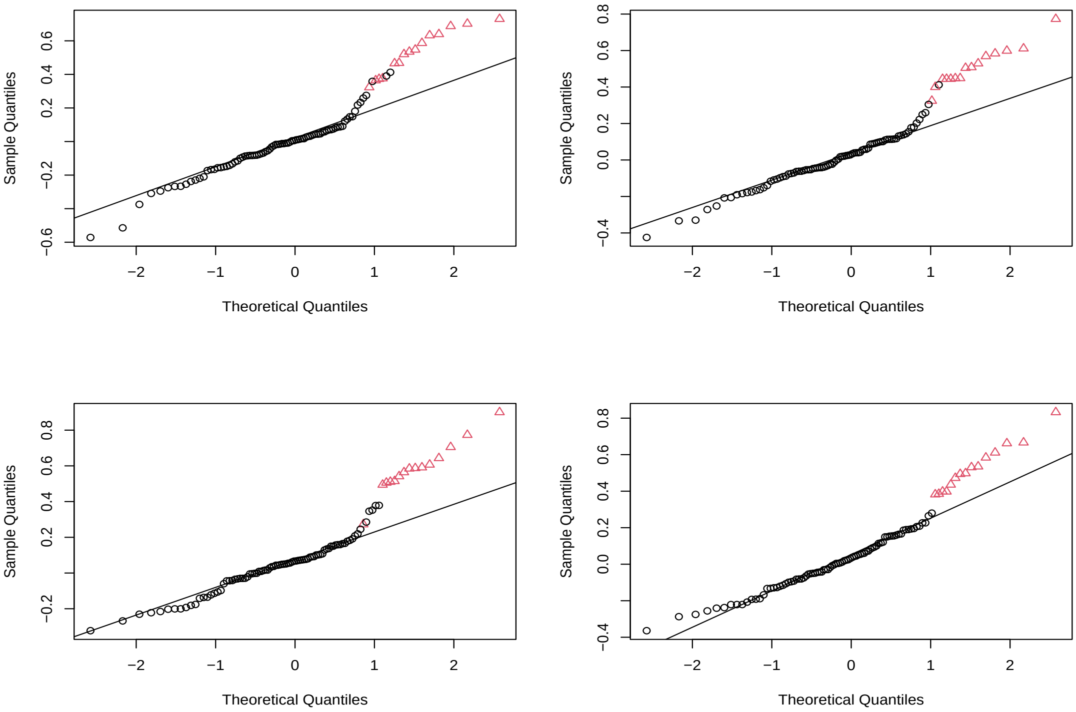


(b) MDPDE

Figure 2: QQ-plots of the random term $\boldsymbol{b}$ predicted by the MLE (top) and the MDPD-estimator (bottom) for $\alpha = \bar{\alpha} = 1/6$. Values corresponding to outlying observations are displayed as red triangles while the rest of observations as black circles.



(a) MLE



(b) MDPDE

Figure 3: QQ-plots of the random term $\boldsymbol{c}$ predicted by the MLE (top) and the MDPD-estimator (bottom) for $\alpha = \bar{\alpha} = 1/6$. Values corresponding to outlying observations are displayed as red triangles while the rest of observations as black circles.

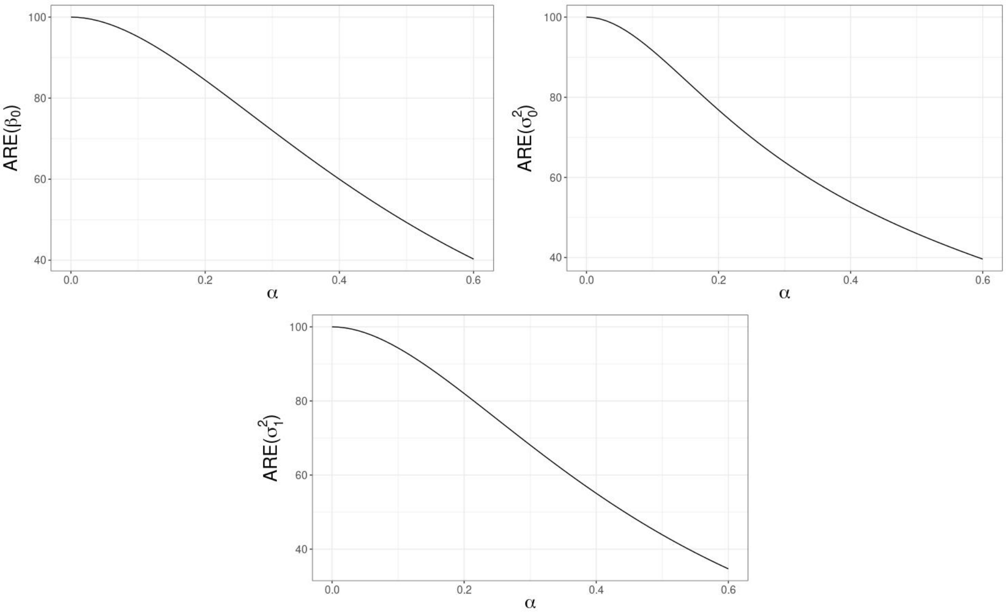


Figure 4: Asymptotic Relative Efficiency with respect to α for β_0, σ_0^2 and σ_1^2, respectively.

SM–6 Additional Results from Monte Carlo experiment

Here, we report the complete results of the Monte Carlo experiments with respect to the contamination levels considered.

The MSMD and MKLD performances of the CVFS-, SMDM-estimators, Composite τ-estimators and for MDPDE for different values of α with 5% of contamination are presented in Figure 7a and 7b, respectively, for the complete contamination. The corresponding results for the separated contamination scenarios are displayed in Figure 8a and 8b

Figure 9a and 9b display the MSMD and MKLD performances of the CVFS-, SMDM-estimators, Composite τ-estimators and for MDPDE for different values of α with 10% of contamination for the complete contamination, while the separated contamination are given in Figure 10a and 10b.

Finally, the MSMD and MKLD performances for the complete contamination of the CVFS-, SMDM-, Composite τ-estimators and for MDPDE for different values of the parameter α in case of 15% are shown in Figures 11a and 11b, whereas Figures 12a and 12b report the results in case of separated contamination.

Tables 2 and 4 report the maximum values of MSMD and MKLD of the CVFS-, SMDM-, Composite τ-estimators and for MDPDE considering different values of α in case of complete contamination and separated contamination, respectively, for 5% of contamination percentage. The results for 15% of contamination are reported in Tables 3 and 5.

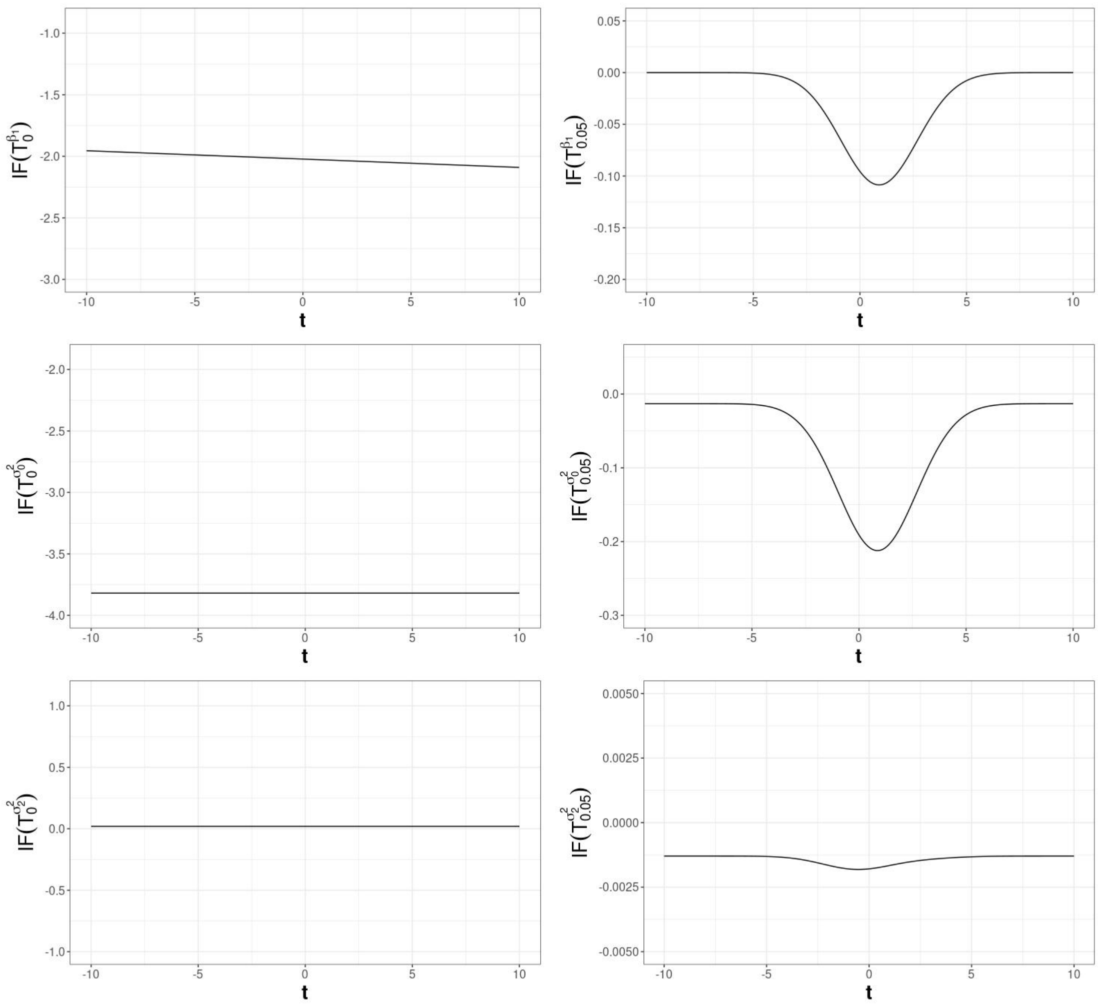


Figure 5: Influence function for the estimators $T_\alpha^{\beta_1}$, $T_\alpha^{\sigma_0^2}$ and $T_\alpha^{\sigma_2^2}$ with respect to $\alpha = 0$ (on the left) and $\alpha = 0.05$ (on the right).

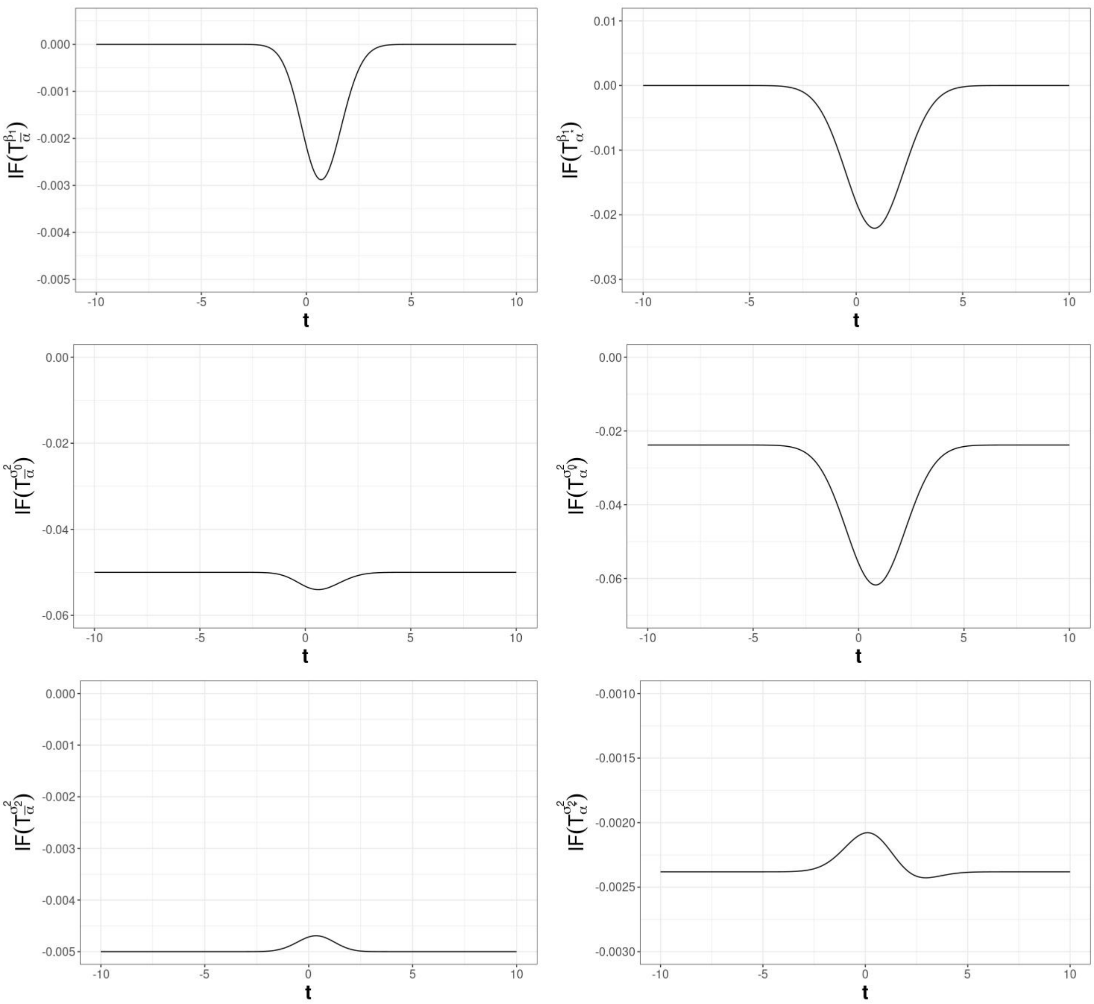
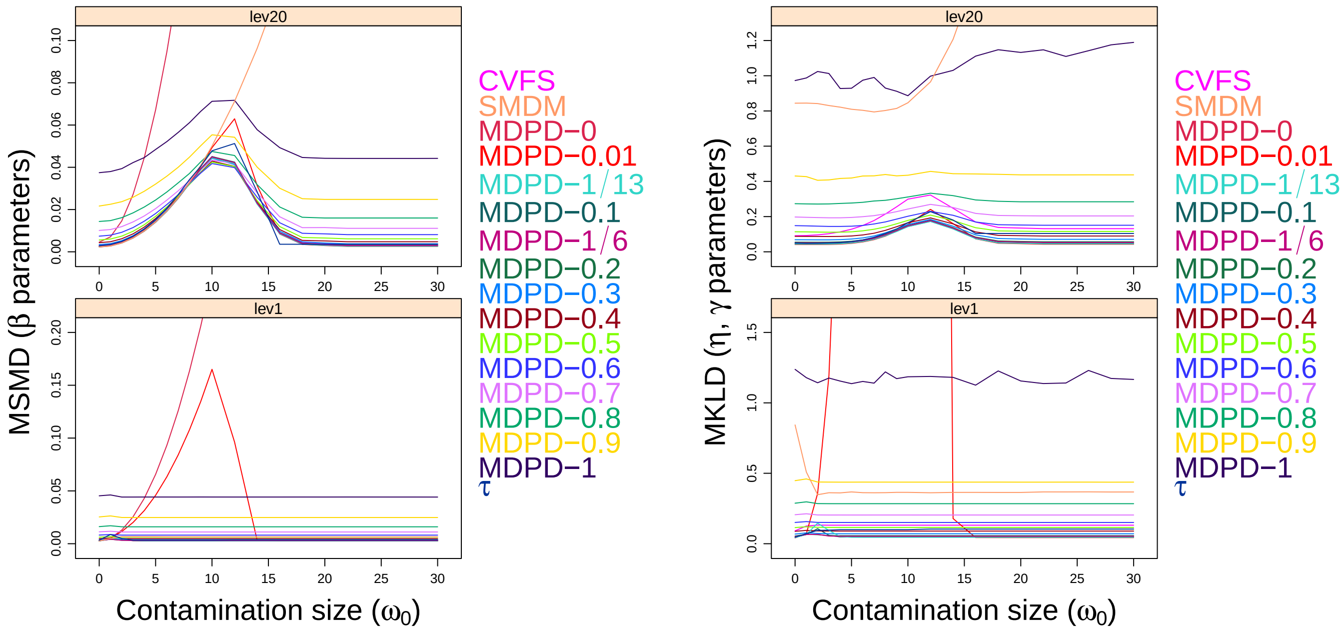


Figure 6: Influence function for the estimators $T_{\alpha}^{\beta_1}$, $T_{\alpha}^{\sigma_0^2}$ and $T_{\alpha}^{\sigma_2^2}$ with respect to $\alpha = \bar{\alpha}(0.2)$ (on the left) and $\alpha = \alpha^*(1/11)$ (on the right).



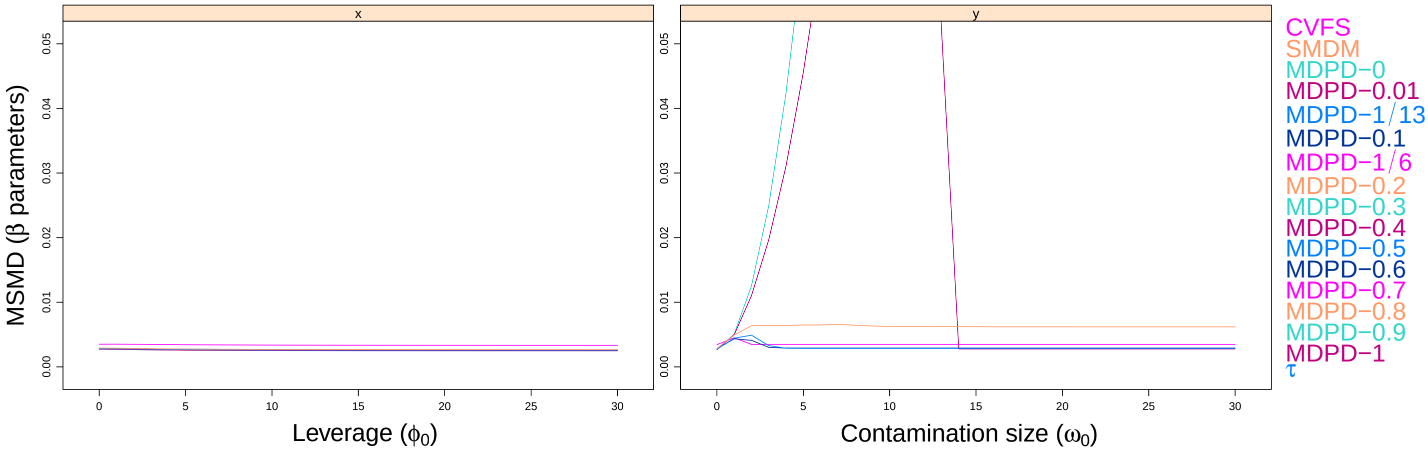
(a) MSMD performance of the MDPD-estimators of $\boldsymbol{\beta}$

(b) MKLD performance of the MDPD-estimators of (η, γ)

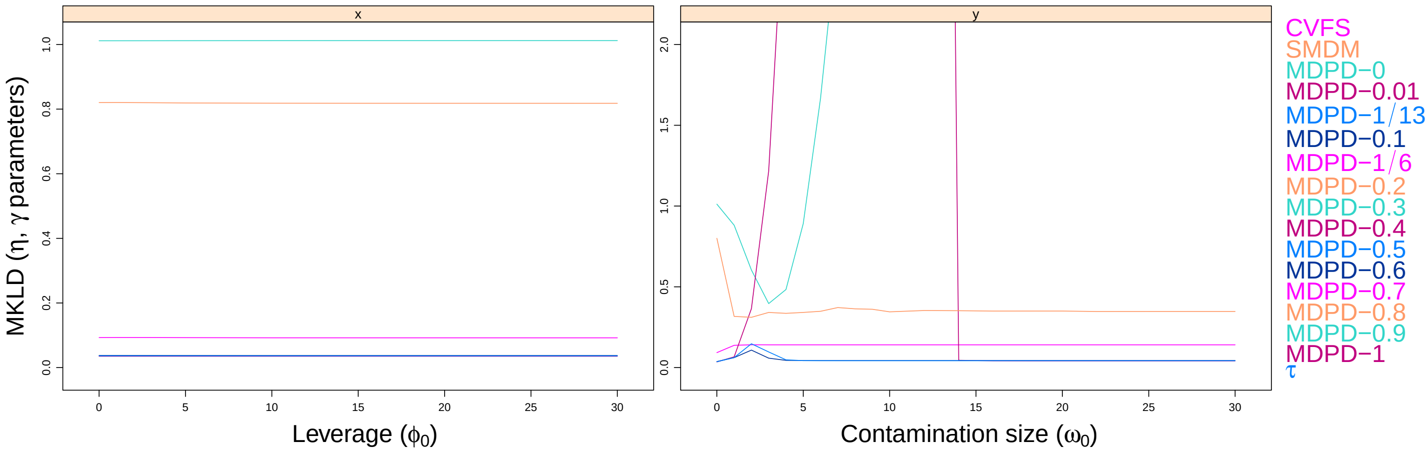
Figure 7: Complete contamination. Performance of the MDPD-estimators of $\boldsymbol{\beta}$ and (η, γ) considering different values of α (including $\alpha^* = 1/13$ and $\bar{\alpha} = 1/6$) compared to the CVFS-, SMDM- and Composite τ-estimators, under 5% outlier contamination.

Table 2: Complete contamination. Maximum values of MSMD and MKLD for the CVFS-, SMDM-, Composite τ-estimators and for MDPDE considering different values of α under 5% of outlier contamination.

		MSMD		MKLD	
Method	(α)	lev1	lev20	lev1	lev20
CVFS	-	0.005	0.045	0.132	0.322
SMDM	-	0.007	0.444	0.843	9.347
Composite τ	-	0.009	0.051	0.104	0.228
MDPDE	0	2.255	2.255	2.276e22	8.198e25
	0.01	0.165	0.063	33.537	0.241
	$\alpha^*(\frac{1}{13})$	0.005	0.045	0.146	0.172
	0.1	0.004	0.044	0.107	0.173
	$\bar{\alpha}(\frac{1}{6})$	0.004	0.045	0.066	0.177
	0.2	0.004	0.045	0.067	0.179
	0.3	0.005	0.044	0.077	0.184
	0.4	0.005	0.043	0.095	0.192
	0.5	0.007	0.042	0.121	0.209
	0.6	0.009	0.042	0.157	0.230

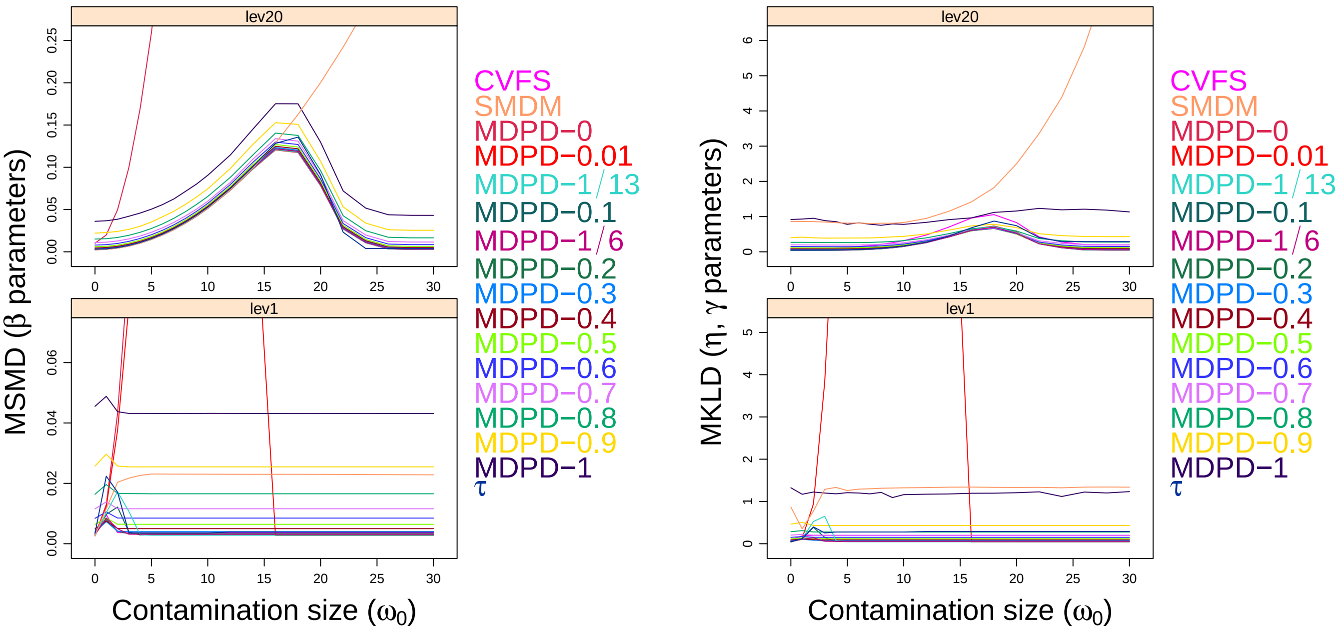


(a) MSMD performance of the MDPD-estimators of $\boldsymbol{\beta}$



(b) MKLD performance of the MDPD-estimators of $(\eta, \boldsymbol{\gamma})$

Figure 8: Contamination on $\boldsymbol{x}$ and $\boldsymbol{y}$. Performance of the MDPD-estimators of $\boldsymbol{\beta}$ and $(\eta, \boldsymbol{\gamma})$ considering different values of α (including $\alpha^* = 1/13$ and $\bar{\alpha} = 1/6$) compared to the CVFS-, SMDM- and Composite τ-estimators, under 5% outlier contamination.



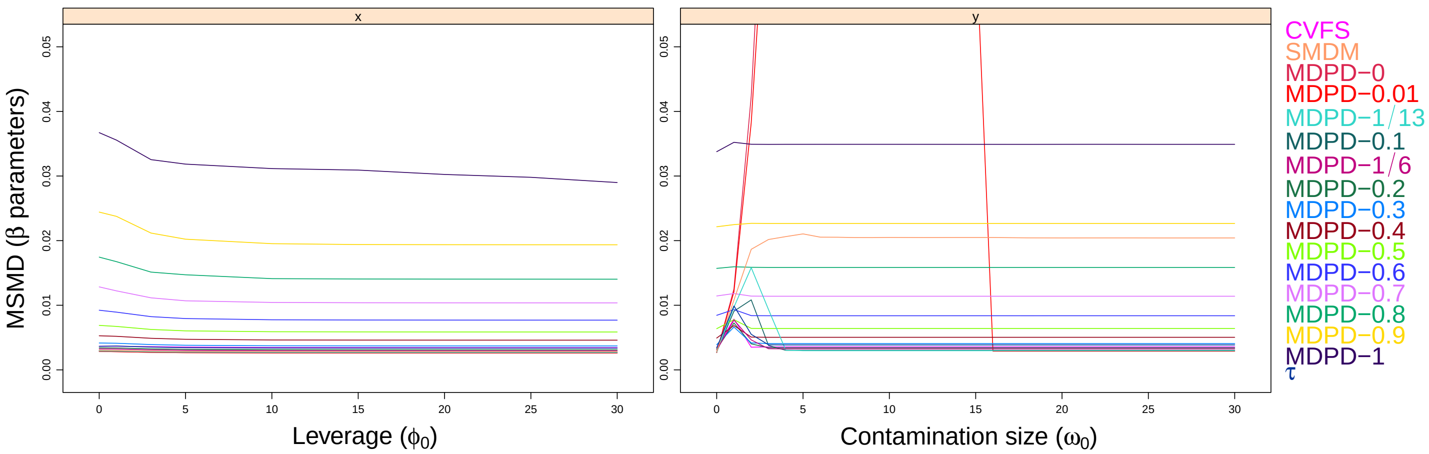
(a) MSMD performance of the MDPD-estimators of $\boldsymbol{\beta}$

(b) MKLD performance of the MDPD-estimators of $(\eta, \boldsymbol{\gamma})$

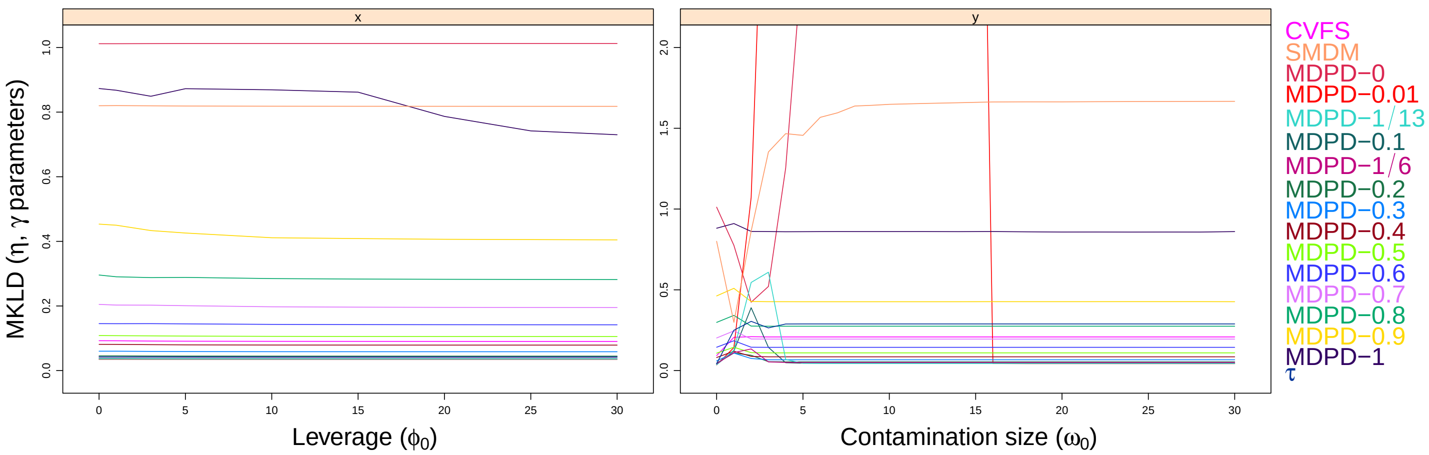
Figure 9: Complete contamination. Performance of the MDPD-estimators of $\boldsymbol{\beta}$ and $(\eta, \boldsymbol{\gamma})$ considering different values of α (including $\alpha^* = 1/13$ and $\bar{\alpha} = 1/6$) compared to the CVFS-, SMDM- and Composite τ-estimators, under 10% outlier contamination.

Table 3: Complete contamination. Maximum values of MSMD and MKLD for the CVFS-, SMDM-, Composite τ-estimators and for MDPDE considering different values of α under 15% of outlier contamination.

		MSMD		MKLD	
Method	(α)	lev1	lev20	lev1	lev20
CVFS	-	0.019	0.240	0.306	2.549
SMDM	-	1.174	0.452	225.429	8.162
Composite τ	-	0.100	0.268	1.182	2.288
MDPDE	0	20.260	20.275	1.439e23	1.606e22
	0.01	2.892	0.239	164.868	2.002
	$\alpha^*(\frac{1}{13})$	0.240	0.240	3.878	1.891
	0.1	0.036	0.240	1.516	1.852
	$\bar{\alpha}(\frac{1}{6})$	0.241	0.241	0.405	1.741
	0.2	0.016	0.242	0.232	1.689
	0.3	0.013	0.244	0.155	1.556
	0.4	0.012	0.246	0.159	1.452
	0.5	0.013	0.249	0.174	1.372
	0.6	0.014	0.255	0.203	1.308

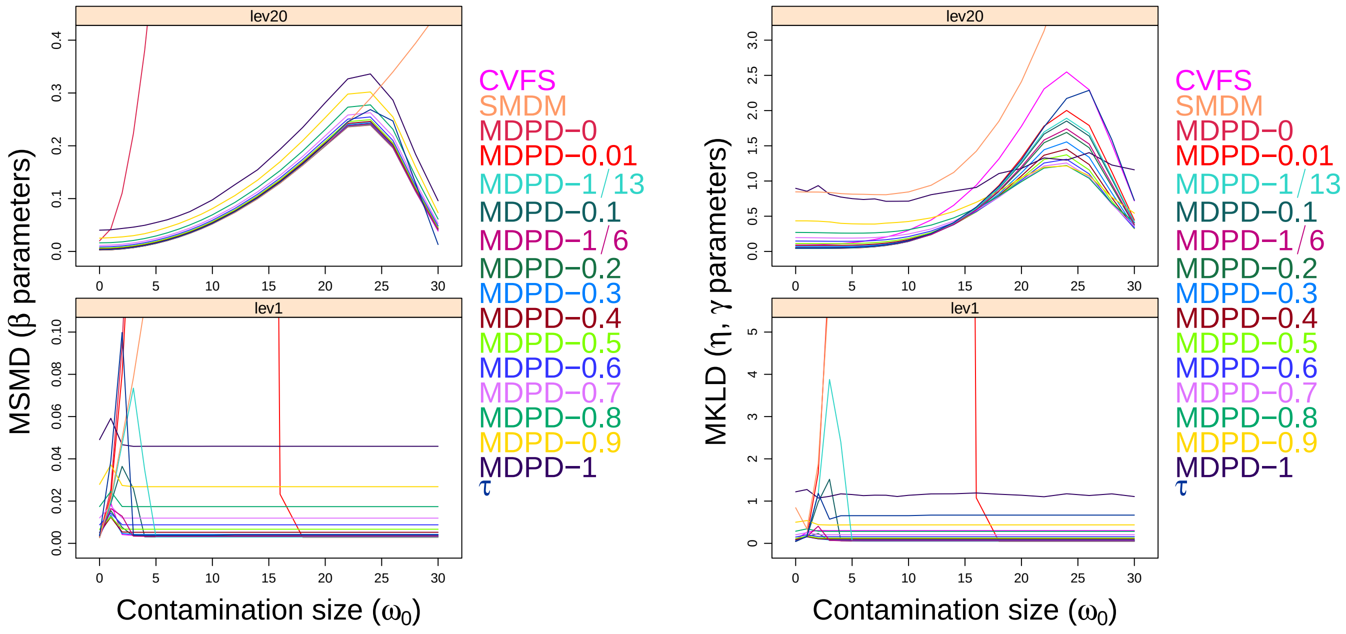


(a) MSMD performance of the MDPD-estimators of $\boldsymbol{\beta}$



(b) MKLD performance of the MDPD-estimators of $(\eta, \boldsymbol{\gamma})$

Figure 10: Contamination on $\boldsymbol{x}$ and $\boldsymbol{y}$. Performance of the MDPD-estimators of $\boldsymbol{\beta}$ and $(\eta, \boldsymbol{\gamma})$ considering different values of α (including $\alpha^* = 1/13$ and $\bar{\alpha} = 1/6$) compared to the CVFS-, SMDM- and Composite τ-estimators, under 10% outlier contamination.



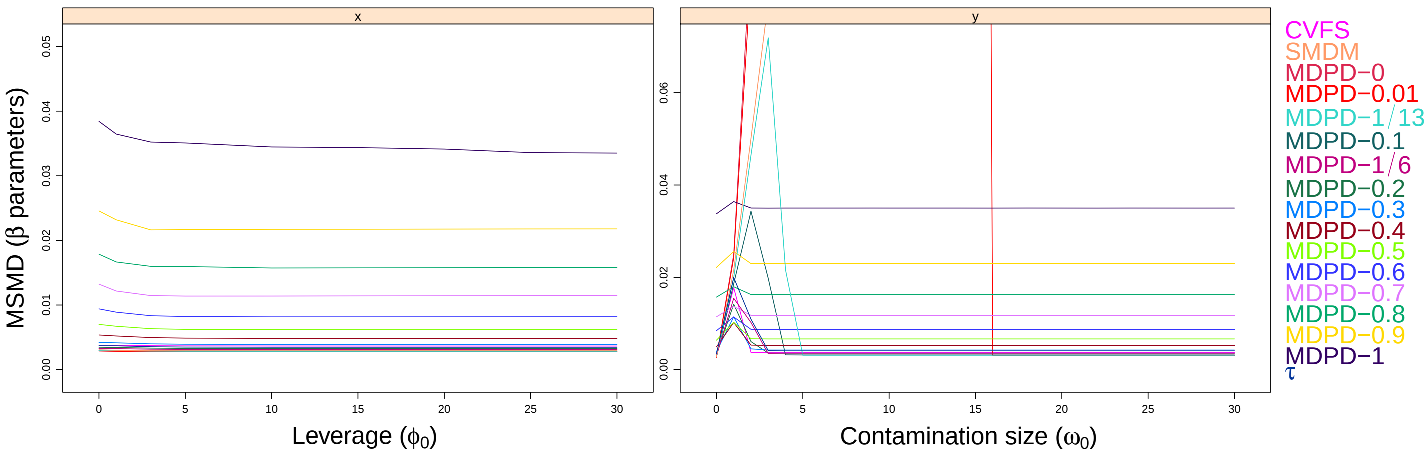
(a) MSMD performance of the MDPD-estimators of $\boldsymbol{\beta}$

(b) MKLD performance of the MDPD-estimators of $(\eta, \boldsymbol{\gamma})$

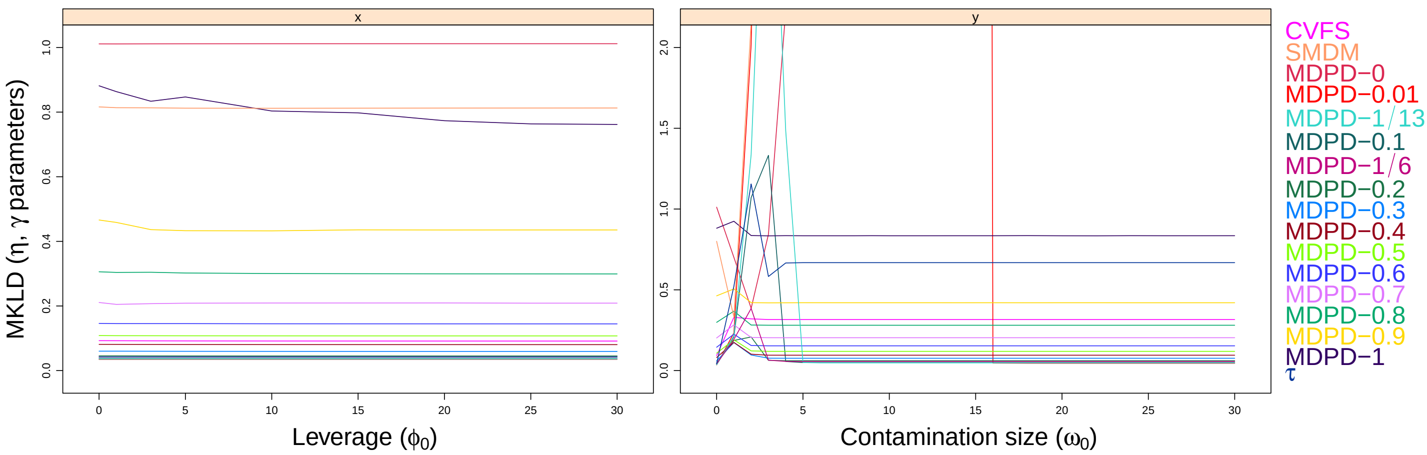
Figure 11: Complete contamination. Performance of the MDPD-estimators of $\boldsymbol{\beta}$ and $(\eta, \boldsymbol{\gamma})$ considering different values of α (including $\alpha^* = 1/13$ and $\bar{\alpha} = 1/6$) compared to the CVFS-, SMDM- and Composite τ-estimators, under 15% outlier contamination.

Table 4: Contamination on $\boldsymbol{x}$ and $\boldsymbol{y}$ separately. Maximum values of MSMD and MKLD for the CVFS-, SMDM-, Composite τ-estimators and for MDPDE considering different values of α under 5% of outlier contamination.

		MSMD		MKLD	
Method	(α)	$\boldsymbol{x}$	$\boldsymbol{y}$ $\boldsymbol{x}$	$\boldsymbol{y}$	
CVFS	-	0.003	0.004	0.093	0.141
SMDM	-	0.003	0.007	0.820	0.800
Composite τ	-	0.003	0.005	0.042	0.104
MDPDE	0	0.003	2.250	1.012	103.539
	0.01	0.003	0.165	0.035	33.547
	$\alpha^*(\frac{1}{13})$	0.003	0.005	0.036	0.147
	0.1	0.003	0.004	0.037	0.108
	$\bar{\alpha}(\frac{1}{6})$	0.003	0.004	0.042	0.064
	0.2	0.003	0.004	0.046	0.064
	0.3	0.004	0.005	0.060	0.074
	0.4	0.005	0.005	0.081	0.092
	0.5	0.007	0.007	0.108	0.119
	0.6	0.009	0.008	0.146	0.158



(a) MSMD performance of the MDPD-estimators of $\boldsymbol{\beta}$



(b) MKLD performance of the MDPD-estimators of $(\eta, \boldsymbol{\gamma})$

Figure 12: Contamination on $\boldsymbol{x}$ and $\boldsymbol{y}$. Performance of the MDPD-estimators of $\boldsymbol{\beta}$ and $(\eta, \boldsymbol{\gamma})$ considering different values of α (including $\alpha^* = 1/13$ and $\bar{\alpha} = 1/6$) compared to the CVFS-, SMDM- and Composite τ-estimators, under 15% outlier contamination.

Table 5: Contamination on $\boldsymbol{x}$ and $\boldsymbol{y}$ separately. Maximum values of MSMD and MKLD for the CVFS-, SMDM-, Composite τ-estimators and for MDPDE considering different values of α under 15% of outlier contamination.

		MSMD		MKLD	
Method	(α)	x	y	x	y
CVFS	-	0.004	0.018	0.093	0.330
SMDM	-	0.003	1.176	0.816	224.819
Composite τ	-	0.004	0.020	0.042	1.155
MDPDE	0	0.003	20.246	1.012	291.512
	0.01	0.003	2.923	0.035	169.689
	$\alpha^*(\frac{1}{13})$	0.003	0.072	0.036	4.036
	0.1	0.003	0.034	0.037	1.332
	$\bar{\alpha}(\frac{1}{6})$	0.003	0.015	0.042	0.385
	0.2	0.004	0.014	0.046	0.208
	0.3	0.004	0.011	0.060	0.174
	0.4	0.005	0.010	0.081	0.174
	0.5	0.007	0.010	0.108	0.190
	0.6	0.009	0.011	0.146	0.225

SM–7 Results from real-data example: Orthodontic distance growth

In this Section, we report some additional results from the study of the real-data example about orthodontic measures presented in Section 6.

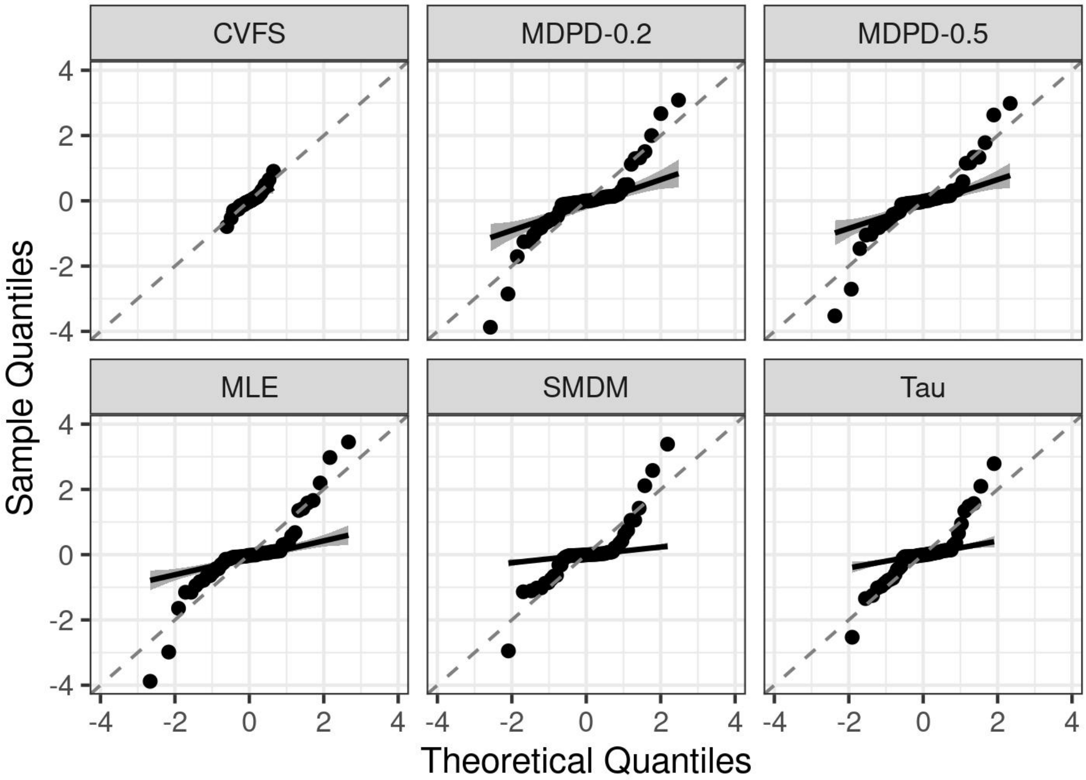
Figure 13a and 13b show the QQ-plot for the predictions of the random intercept and random slope (age), respectively, computed by the MLE, the SMDM-estimator, the CVFS-estimator, the Composite τ-estimator and the MDPDE for $\alpha = \alpha^* = 0.2$ and $\alpha = \bar{\alpha} = 0.5$.

SM–8 Real-data example: Extrafoveal Vision Acuity

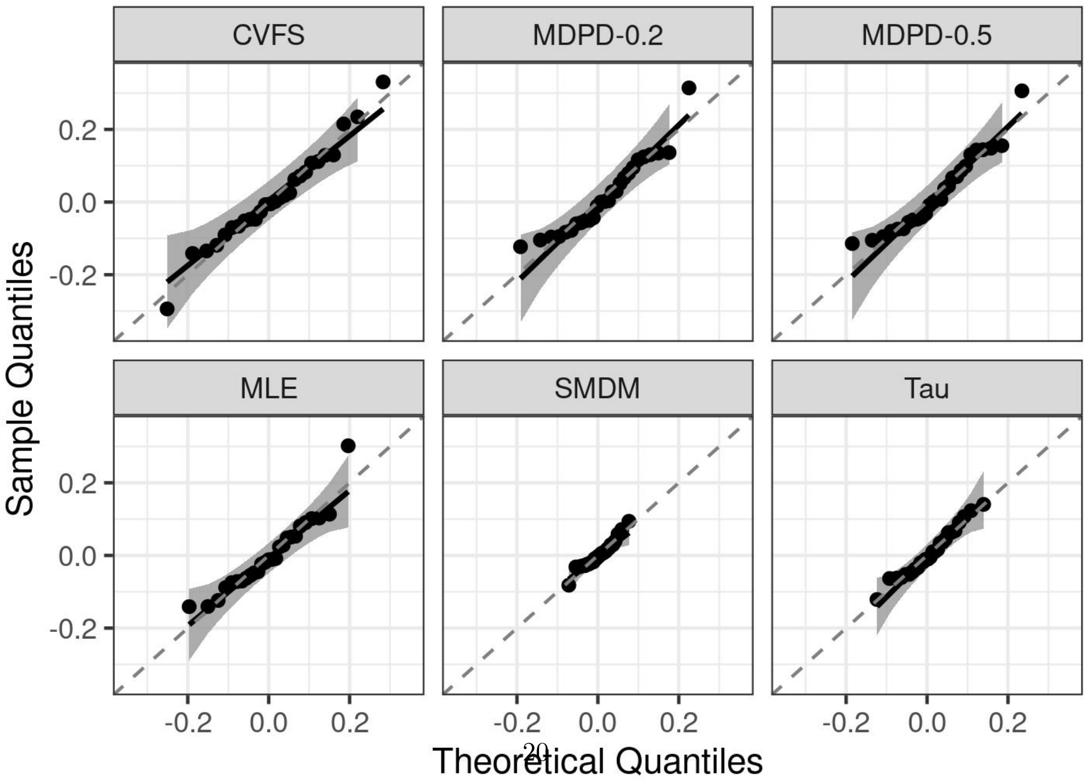
We compare the estimates obtained by the minimum DPD method with those obtained using the classical (non-robust) restricted MLE, computed using the `lmer` function in `R`, as well as the robust competitors, the SMDM-estimator and the CVFS-estimator. A very important consideration in real situations is the selection of an "optimum" value of α that applies to the given data set. We will use different values of α to highlight the behavior of the estimator seen in the simulations. In general, we are going to consider the values α^* and $\bar{\alpha}$, derived from theoretical computations, as suggested optimal values.

We consider the study conducted by Frömer et al. [2015] about the relationship between individual differences in foveal visual acuity and extrafoveal vision (acuity and crowding) and reading time measures, such as reading rate and preview benefit.

There were 40 participants in the study, with normal visual acuity measured with the adaptive computerized Freiburg Acuity Test (FrACT) Bach [1996]. The study was organized in two test sessions. During the first session, the extrafoveal vision assessment (EVA) was provided, involving a test of crowded and uncrowded extrafoveal vision. In addition, visual acuity of fovea was measured using the FrACT. The second session was taken after a week, consisting of an eye-tracking experiment with list reading followed by the EVA procedure. The EVA was performed considering four test conditions: identification of single letters and flanked letters in the left



(a) Intercept



(b) Age

Figure 13: QQ-plots of the random terms estimated by the MLE, CVFS-, SMDM-, Composite τ-estimators and the MDPD-estimators for$\alpha = \alpha^* = 0.2$ and $\alpha = \bar{\alpha} = 0.5$.

and right visual field. Here, we consider only data coming from measurements with the EVA procedure and do not deal with data related to the reading task.

This kind of data can be modeled using a Linear Mixed Model. In particular, we studied a repeated measures *Analysis of Variance* (ANOVA) of the threshold eccentricities (TE) with random effects given by extrafoveal vision (EV)(single versus crowded letter), hemifield (H)(left, right), and test repetition (T_1, T_2, T_3). Thus, combining the factors given above, we have $p = 12$ measurements for each subject (participant). The model for each subject (i-th) has the form

$$\begin{aligned} TE_i =& \beta_0 + \beta_1 EV_i + \beta_2 H_i + \beta_3 T2\_1_i + \beta_4 T3\_2_i + \beta_5 (EV * H)_i + \beta_6 (EV * T2\_1)_i + \\ & \beta_7 (EV * T3\_3)_i + \beta_8 (H * T2\_1)_i + \beta_9 (H * T3\_2)_i + \beta_{10} (EV * H * T2\_1)_i + \\ & \beta_{11} (EV * H * T3\_2)_i + u_1 + EV u_2 + H u_3 + T2\_1 u_4 + T3\_2 u_5 + (EV * H) u_6 + \\ & (EV * T2\_1) u_7 + (EV * T3\_3) u_8 + (H * T2\_1) u_9 + (H * T3\_2) u_{10} + \epsilon_i \end{aligned}$$

where $i \in \{1, \ldots, n\}$, $n = 40$, while $T2\_1$ and $T3\_2$ substitute the factor time (T_1, T_2, T_3) indicating the transitions between the first and second sessions, and between third and second sessions, respectively. Hence, we have 12 fixed effect parameters $(\beta_0, \ldots, \beta_{11})$, and 10 random effects of which we will estimate the variance components σ_j^2, $j \in \{1, \ldots, 10\}$.

Table 6: Estimates of model parameters obtained using different estimators.

Method	lmer	SMDM	CVFS	MDPDE					
(α)	-	-	-	0.05	1/13	1/6	0.2	0.4	0.6
β									
Intercept	7.007	7.034	7.041	7.040	7.060	7.123	7.151	7.196	7.197
EV	7.791	7.826	7.884	7.828	7.834	7.798	7.766	7.643	7.722
H	-0.021	-0.009	-0.001	0.000	0.002	-0.007	-0.010	-0.048	-0.030
T2_1	0.062	0.024	-0.016	0.002	-0.019	-0.068	-0.090	-0.165	-0.018
T3_2	-0.081	-0.066	-0.033	-0.040	-0.030	-0.015	-0.009	0.043	0.189
EV*H	0.439	0.454	0.495	0.459	0.449	0.372	0.333	0.205	-0.012
EV*T2_1	0.180	0.079	0.046	0.084	0.054	0.003	-0.015	-0.070	0.036
EV*T3_2	-0.204	-0.139	-0.114	-0.122	-0.109	-0.104	-0.101	-0.036	0.090
H*T2_1	0.144	0.142	0.099	0.122	0.118	0.124	0.123	0.106	0.468
H*T3_2	-0.162	-0.122	-0.092	-0.121	-0.111	-0.109	-0.112	-0.134	-0.270
EV*H*T2_1	0.261	0.243	0.153	0.215	0.195	0.148	0.121	-0.007	0.504
EV*H*T3_2	-0.476	-0.370	-0.319	-0.365	-0.335	-0.276	-0.243	-0.152	-0.396
σ^2									
Intercept	0.987	1.078	0.130	0.120	0.109	0.074	0.055	0.010	0.000
EV	1.128	1.010	0.129	0.154	0.140	0.125	0.118	0.064	0.000
H	0.613	0.618	0.071	0.061	0.063	0.071	0.075	0.088	0.000
T2_1	0.296	0.000	0.010	0.012	0.012	0.012	0.012	0.011	0.000
T3_2	0.452	0.000	0.005	0.019	0.016	0.011	0.011	0.013	0.000
EV*H	0.656	0.000	0.063	0.050	0.048	0.033	0.021	0.024	0.000
EV*T2_1	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
EV*T3_2	1.032	0.000	0.101	0.103	0.091	0.061	0.046	0.000	0.000
H*T2_1	0.000	0.000	0.000	0.000	0.000	0.000	0.006	0.058	0.000
H*T3_2	0.439	0.000	0.000	0.008	0.003	0.000	0.000	0.000	0.000
η_0	0.154	0.118	0.121	0.120	0.111	0.099	0.095	0.082	0.000

Table 6 shows the estimates of model parameters obtained using the `lmer` estimator, the

SMDM- and CVFS-estimators, the MDPDE for different values of α. For increasing α, the MDPDE's capacity to accurately estimate the variance components drop, especially for $\alpha \geq 0.4$, while the estimates of the fixed terms do not significantly change. It may be seen that the SMDS-estimator has a poor performance. On the other hand, the MDPDEs for $\alpha = 0.05$ and $\alpha = 1/13$ show similar estimates to those obtained using `lmer` and the CVFS-estimator.

Finally, we tested the `lmer` estimator, the CVFS-estimator and the MDPDE with $\alpha = 1/13$ in the case where some TE values are substituted by outlying values. In particular, we implemented an iterative procedure where, in each step, an outlying observation is added. Let $\boldsymbol{X}$ be the (40×12) matrix of the TE values. After selecting a random cell (i, j), with $i \in \{1, \ldots, 40\}$ and $j \in \{1, \ldots, 12\}$, X_{ij} is replaced by a value sampled from $N(kv_j, 0.1^2)$, where $k = 10$, $\boldsymbol{v}$ is the eigenvector corresponding to the smallest eigenvalue of the maximum likelihood estimate of the covariance matrix and v_j indicates the j-th component of the vector $\boldsymbol{v}$. Before adding the next outlying value, the estimates of `lmer`, the CVFS-estimator and MDPDE with $\alpha = 1/13$ are computed. We repeated the procedure until 9 values had been substituted. Tables 8, 9 and 10 show the estimates and the corresponding p-values obtained using the usual `lmer`, the CVFS-estimator and the proposed MDPDE, respectively, as the number of substituted cells (m) increases.

Table 7: p-values obtained from the estimates of the `lmer` estimator, the CVFS-estimator and the MDPDE with $\alpha = 1/13$, for uncontaminated data ($m = 0$) and for the data set with $m = 9$ substituted cells.

		Intercept	EV	H	T2_1	T3_2	EV*H
`lmer`	$m = 0$	0.000	0.000	0.685	0.200	0.126	0.000
	$m = 9$	0.000	0.000	0.383	0.685	0.844	0.011
CVFS	$m = 0$	0.000	0.000	0.987	0.777	0.565	0.000
	$m = 9$	0.000	0.000	0.978	0.932	0.763	0.040
MDPDE 1/13	$m = 0$	0.000	0.000	0.977	0.716	0.534	0.000
	$m = 9$	0.000	0.000	0.674	0.903	0.438	0.000
		EV*T2_1	EV*T3_2	H*T2_1	H*T3_2	EV*H*T2_1	EV*H*T3_2
`lmer`	$m = 0$	0.042	0.066	0.103	0.083	0.139	0.007
	$m = 9$	0.881	0.816	0.734	0.789	0.858	0.486
CVFS	$m = 0$	0.639	0.312	0.188	0.318	0.336	0.038
	$m = 9$	0.827	0.831	0.686	0.694	0.724	0.669
MDPDE 1/13	$m = 0$	0.493	0.263	0.139	0.189	0.252	0.027
	$m = 9$	0.724	0.347	0.075	0.217	0.193	0.009

The obtained results are summarized in Table 7, which reports the p-values of the tests, checking whether the parameters are significantly different from zero giving an idea of the importance of the corresponding variables, for uncontaminated data and when $m = 9$ cells are substituted. The estimates obtained using `lmer` are more affected than those given by the CVFS-estimator and the MDPDE. On the other hand, the MDPDE seems quite stable with respect to the corresponding p-values, while those computed using `lmer` and the CVFS-estimator show large variations.

Table 8: Estimates and the corresponding p-values obtained using `lmer` for increasing m.

		$m=1$	$m=2$	$m=3$	$m=4$	$m=5$	$m=6$	$m=7$	$m=8$	$m=9$
β	Intercept	7.001	6.994	6.987	6.963	6.940	6.917	6.896	6.890	6.882
	EV	7.801	7.815	7.830	7.783	7.736	7.691	7.649	7.662	7.676
	H	-0.032	-0.017	-0.002	0.046	-0.001	0.045	0.087	0.099	0.085
	T2_1	0.046	0.068	0.068	0.139	0.070	0.070	0.007	0.026	0.048
	T3_2	-0.065	-0.065	-0.088	-0.088	-0.018	-0.086	-0.023	-0.023	-0.023
	EV*H	0.460	0.431	0.401	0.497	0.404	0.495	0.579	0.553	0.583
	EV*T2_1	0.211	0.168	0.168	0.311	0.172	0.172	0.046	0.008	-0.036
	EV*T3_2	-0.235	-0.235	-0.190	-0.190	-0.051	-0.187	-0.061	-0.061	-0.061
	H*T2_1	0.113	0.070	0.070	-0.074	-0.212	-0.212	-0.087	-0.125	-0.081
	H*T3_2	-0.130	-0.130	-0.086	-0.086	0.053	0.189	0.063	0.063	0.063
	EV*H*T2_1	0.323	0.410	0.410	0.123	-0.155	-0.155	0.097	0.173	0.085
	EV*H*T3_2	-0.538	-0.538	-0.628	-0.628	-0.351	-0.079	-0.330	-0.330	-0.330
σ^2	Intercept	0.952	0.909	0.911	0.590	0.450	0.495	0.473	0.466	0.456
	EV	0.974	0.985	1.049	0.814	0.520	0.609	0.481	0.472	0.464
	H	0.608	0.544	0.673	0.412	0.176	0.000	0.000	0.000	0.001
	T2_1	0.009	0.407	0.518	0.557	0.216	0.136	0.003	0.001	0.000
	T3_2	0.412	0.307	0.633	0.176	0.000	0.176	0.004	0.000	0.000
	EV*H	0.603	0.990	1.153	0.961	0.225	0.690	0.617	0.644	0.618
	EV*T2_1	0.037	0.136	0.476	1.272	0.625	0.132	0.000	0.020	0.013
	EV*T3_2	0.893	1.152	1.354	0.386	0.044	0.911	0.707	0.698	0.674
	H*T2_1	0.004	0.001	0.000	1.133	0.444	0.257	0.000	0.003	0.000
	H*T3_2	0.000	0.000	0.000	0.000	0.000	0.000	0.002	0.001	0.001
	η_0	0.174	0.182	0.175	0.369	0.715	0.836	1.078	1.095	1.122
p-value	Intercept	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	EV	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	H	0.571	0.749	0.970	0.508	0.993	0.594	0.362	0.300	0.383
	T2_1	0.324	0.224	0.249	0.113	0.481	0.504	0.952	0.824	0.685
	T3_2	0.232	0.215	0.169	0.216	0.846	0.416	0.841	0.842	0.844
	EV*H	0.000	0.000	0.001	0.001	0.014	0.015	0.010	0.016	0.011
	EV*T2_1	0.025	0.086	0.096	0.095	0.408	0.405	0.842	0.971	0.881
	EV*T3_2	0.038	0.062	0.149	0.184	0.787	0.445	0.814	0.815	0.816
	H*T2_1	0.228	0.467	0.458	0.674	0.288	0.310	0.709	0.595	0.734
	H*T3_2	0.164	0.174	0.362	0.530	0.778	0.356	0.785	0.787	0.789
	EV*H*T2_1	0.085	0.034	0.030	0.652	0.683	0.706	0.834	0.712	0.858
	EV*H*T3_2	0.004	0.006	0.001	0.022	0.355	0.848	0.477	0.481	0.486

Table 9: Estimates and the corresponding p-values obtained using the CVFS-estimator for increasing m.

		$m=1$	$m=2$	$m=3$	$m=4$	$m=5$	$m=6$	$m=7$	$m=8$	$m=9$
β	Intercept	7.041	7.048	7.048	7.048	7.078	7.078	7.073	7.073	7.064
	EV	7.884	7.900	7.900	7.900	7.890	7.894	7.900	7.900	7.888
	H	-0.001	0.000	0.000	0.000	-0.005	-0.003	0.008	0.008	-0.006
	T2_1	-0.016	-0.017	-0.017	-0.017	-0.014	-0.006	0.005	0.005	0.013
	T3_2	-0.033	-0.042	-0.042	-0.042	-0.047	-0.056	-0.064	-0.064	-0.055
	EV*H	0.495	0.504	0.504	0.504	0.495	0.512	0.527	0.527	0.534
	EV*T2_1	0.046	0.033	0.033	0.033	0.051	0.070	0.066	0.066	0.068
	EV*T3_2	-0.114	-0.139	-0.139	-0.139	-0.158	-0.176	-0.171	-0.171	-0.159
	H*T2_1	0.099	0.078	0.078	0.078	0.104	0.128	0.129	0.129	0.132
	H*T3_2	-0.091	-0.077	-0.077	-0.077	-0.108	-0.126	-0.117	-0.117	-0.104
	EV*H*T2_1	0.153	0.096	0.096	0.096	0.139	0.170	0.174	0.174	0.223
	EV*H*T3_2	-0.319	-0.277	-0.277	-0.277	-0.312	-0.352	-0.338	-0.338	-0.360
σ^2	Intercept	0.130	0.138	0.138	0.138	0.111	0.120	0.129	0.129	0.138
	EV	0.129	0.128	0.128	0.128	0.133	0.142	0.151	0.151	0.178
	H	0.071	0.078	0.078	0.078	0.083	0.092	0.096	0.096	0.088
	T2_1	0.010	0.010	0.010	0.010	0.014	0.016	0.016	0.016	0.016
	T3_2	0.005	0.005	0.005	0.005	0.009	0.012	0.017	0.017	0.028
	EV*H	0.063	0.071	0.071	0.071	0.082	0.085	0.087	0.087	0.090
	EV*T2_1	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	EV*T3_2	0.101	0.070	0.070	0.070	0.080	0.089	0.104	0.104	0.134
	H*T2_1	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	H*T3_2	0.000	0.000	0.000	0.000	0.006	0.021	0.022	0.022	0.018
	η_0	0.121	0.125	0.125	0.125	0.126	0.129	0.138	0.138	0.146
p-value	Intercept.1	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	EV	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.622	0.000
	H	0.989	0.998	1.000	0.999	0.971	0.976	0.948	0.984	0.978
	T2_1	0.784	0.842	0.849	0.835	0.920	0.971	0.975	0.973	0.932
	T3_2	0.600	0.771	0.681	0.745	0.752	0.764	0.726	0.841	0.763
	EV*H	0.000	0.000	0.001	0.001	0.010	0.018	0.018	0.128	0.040
	EV*T2_1	0.672	0.802	0.847	0.878	0.903	0.820	0.825	0.931	0.827
	EV*T3_2	0.348	0.496	0.363	0.435	0.758	0.708	0.612	0.907	0.831
	H*T2_1	0.178	0.798	0.669	0.692	0.694	0.646	0.679	0.909	0.686
	H*T3_2	0.263	0.791	0.898	0.725	0.639	0.581	0.653	0.906	0.694
	EV*H*T2_1	0.359	0.894	0.821	0.820	0.782	0.766	0.775	0.812	0.724
	EV*H*T3_2	0.069	0.712	0.472	0.527	0.492	0.497	0.677	0.929	0.669

Table 10: Estimates and the corresponding p-values obtained using the proposed MDPDE with $\alpha = 1/13$ for increasing m.

		$m=1$	$m=2$	$m=3$	$m=4$	$m=5$	$m=6$	$m=7$	$m=8$	$m=9$
β	Intercept	7.062	7.062	7.055	7.040	7.059	7.060	7.082	7.078	7.070
	EV	7.836	7.836	7.835	7.843	7.837	7.838	7.828	7.841	7.856
	H	0.002	0.003	-0.003	0.012	0.006	0.007	0.004	0.018	0.026
	T2_1	-0.021	-0.022	-0.026	-0.013	-0.017	-0.017	-0.018	-0.010	0.007
	T3_2	-0.030	-0.028	-0.019	-0.023	-0.027	-0.028	-0.029	-0.033	-0.042
	EV*H	0.448	0.450	0.450	0.454	0.421	0.420	0.410	0.430	0.447
	EV*T2_1	0.051	0.048	0.033	0.032	0.026	0.025	0.034	0.048	0.030
	EV*T3_2	-0.108	-0.104	-0.091	-0.088	-0.096	-0.095	-0.107	-0.109	-0.106
	H*T2_1	0.116	0.115	0.113	0.122	0.120	0.120	0.140	0.148	0.158
	H*T3_2	-0.105	-0.105	-0.085	-0.097	-0.081	-0.081	-0.101	-0.117	-0.103
	EV*H*T2_1	0.194	0.193	0.167	0.210	0.210	0.208	0.238	0.260	0.245
	EV*H*T3_2	-0.329	-0.326	-0.326	-0.367	-0.347	-0.345	-0.371	-0.418	-0.401
σ^2	Intercept	0.107	0.107	0.106	0.104	0.095	0.095	0.074	0.077	0.081
	EV	0.137	0.137	0.142	0.146	0.151	0.150	0.150	0.151	0.154
	H	0.064	0.064	0.065	0.061	0.063	0.062	0.064	0.061	0.063
	T2_1	0.011	0.011	0.014	0.012	0.012	0.013	0.015	0.016	0.017
	T3_2	0.016	0.015	0.016	0.018	0.018	0.018	0.021	0.022	0.023
	EV*H	0.051	0.049	0.060	0.066	0.034	0.035	0.038	0.034	0.036
	EV*T2_1	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	EV*T3_2	0.094	0.089	0.102	0.112	0.113	0.114	0.121	0.131	0.138
	H*T2_1	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.001
	H*T3_2	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	η_0	0.109	0.109	0.104	0.105	0.107	0.106	0.103	0.106	0.113
p-value	Intercept	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	EV	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	H	0.968	0.961	0.952	0.840	0.912	0.907	0.951	0.756	0.674
	T2_1	0.685	0.662	0.610	0.799	0.748	0.745	0.746	0.862	0.903
	T3_2	0.543	0.566	0.689	0.635	0.590	0.587	0.578	0.534	0.438
	EV*H	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	EV*T2_1	0.514	0.536	0.670	0.686	0.746	0.758	0.674	0.563	0.724
	EV*T3_2	0.268	0.284	0.353	0.384	0.354	0.358	0.306	0.313	0.347
	H*T2_1	0.148	0.151	0.166	0.142	0.163	0.163	0.095	0.083	0.075
	H*T3_2	0.210	0.211	0.299	0.243	0.334	0.337	0.213	0.153	0.217
	EV*H*T2_1	0.256	0.259	0.328	0.218	0.232	0.236	0.172	0.150	0.193
	EV*H*T3_2	0.029	0.030	0.034	0.016	0.025	0.025	0.017	0.005	0.009

References

M. Bach. The freiburg visual acuity test—automatic measurement of visual acuity. *Optometry and Vision Science*, 73:49–53, 1996.

R. Frömer, O. Dimigen, F. Niefind, N. Krause, R. Kliegl, and W. Sommer. Are individual differences in reading speed related to extrafoveal visual acuity and crowding? *Plos One*, 10: 1–18, 03 2015.

A. Ghosh and A. Basu. Robust estimation for independent non-homogeneous observations using density power divergence with applications to linear regression. *Electronic Journal of Statistics*, 7:2420–2456, 2013.