Online Appendices to "Density Deconvolution with Additive Measurement Errors using Quadratic Programming"

A1 Derivation of Eq. (9)

Recall that $\widehat{\mathbf{f}}_Y$ and $\widehat{\mathbf{f}}_Y^0$ are independent random vectors with a common mean $\mathbf{C}\mathbf{f}_X$. We can view $\mathbf{C}\widehat{\mathbf{f}}_{X,\lambda}$, with $\widehat{\mathbf{f}}_{X,\lambda}$ the estimator from the QP approach, to be an estimator of this common mean. We write the SURE criterion as:

$$
\begin{aligned}
\mathrm{E}[Err] &= \mathrm{E}\left[\|\widehat{\mathbf{f}}_Y^0 - \mathbf{C}\widehat{\mathbf{f}}_{X,\lambda}\|^2\right] \\
&= \mathrm{E}\left[\|(\widehat{\mathbf{f}}_Y^0 - \mathbf{C}\mathbf{f}_X) - (\widehat{\mathbf{f}}_Y - \mathbf{C}\mathbf{f}_X) + (\widehat{\mathbf{f}}_Y - \mathbf{C}\widehat{\mathbf{f}}_{X,\lambda})\|^2\right] \\
&= \mathrm{E}\left[\|\widehat{\mathbf{f}}_Y^0 - \mathbf{C}\mathbf{f}_X\|^2\right] + \mathrm{E}\left[\|\widehat{\mathbf{f}}_Y - \mathbf{C}\mathbf{f}_X\|^2\right] + \mathrm{E}\left[\|\widehat{\mathbf{f}}_Y - \mathbf{C}\widehat{\mathbf{f}}_{X,\lambda}\|^2\right] \\
&\quad -2\mathrm{E}\left[(\widehat{\mathbf{f}}_Y^0 - \mathbf{C}\mathbf{f}_X)^T(\widehat{\mathbf{f}}_Y - \mathbf{C}\mathbf{f}_X)\right] + 2\mathrm{E}\left[(\widehat{\mathbf{f}}_Y^0 - \mathbf{C}\mathbf{f}_X)^T(\widehat{\mathbf{f}}_Y - \mathbf{C}\widehat{\mathbf{f}}_{X,\lambda})\right] \\
&\quad -2\mathrm{E}\left[(\widehat{\mathbf{f}}_Y - \mathbf{C}\mathbf{f}_X)^T(\widehat{\mathbf{f}}_Y - \mathbf{C}\widehat{\mathbf{f}}_{X,\lambda})\right]. \qquad \text{(A1.16)}
\end{aligned}
$$

Term by term in (A1.16), we have $\mathrm{E}\left[\|\widehat{\mathbf{f}}_Y^0 - \mathbf{C}\mathbf{f}_X\|^2\right] = \mathrm{E}\left[\|\widehat{\mathbf{f}}_Y - \mathbf{C}\mathbf{f}_X\|^2\right]$ (because $\widehat{\mathbf{f}}_Y^0$ is defined as a random draw from the same distribution as $\widehat{\mathbf{f}}_Y$); $\mathrm{E}\left[\|\widehat{\mathbf{f}}_Y - \mathbf{C}\widehat{\mathbf{f}}_{X,\lambda}\|^2\right] = \mathrm{E}[err]$ (by definition of err); and $\mathrm{E}\left[(\widehat{\mathbf{f}}_Y^0 - \mathbf{C}\mathbf{f}_X)^T(\widehat{\mathbf{f}}_Y - \mathbf{C}\mathbf{f}_X)\right] = \mathrm{E}\left[(\widehat{\mathbf{f}}_Y^0 - \mathbf{C}\mathbf{f}_X)^T(\widehat{\mathbf{f}}_Y - \mathbf{C}\widehat{\mathbf{f}}_{X,\lambda})\right] = 0$ (by the independence of $\widehat{\mathbf{f}}_Y^0$ and $\widehat{\mathbf{f}}_Y$). Consequently, (A1.16) reduces to:

$$
\begin{aligned}
\mathrm{E}[Err] &= \mathrm{E}[err] + 2\left\{\mathrm{E}\left[(\widehat{\mathbf{f}}_Y - \mathbf{C}\mathbf{f}_X)^T(\widehat{\mathbf{f}}_Y - \mathbf{C}\mathbf{f}_X)\right] - \mathrm{E}\left[(\widehat{\mathbf{f}}_Y - \mathbf{C}\mathbf{f}_X)^T(\widehat{\mathbf{f}}_Y - \mathbf{C}\widehat{\mathbf{f}}_{X,\lambda})\right]\right\} \\
&= \mathrm{E}[err] + 2\mathrm{E}\left[(\widehat{\mathbf{f}}_Y - \mathbf{C}\mathbf{f}_X)^T(\mathbf{C}\widehat{\mathbf{f}}_{X,\lambda} - \mathbf{C}\mathbf{f}_X)\right] \\
&= \mathrm{E}[err] + 2\mathrm{tr}\left[\mathrm{COV}(\mathbf{C}\widehat{\mathbf{f}}_{X,\lambda}, \widehat{\mathbf{f}}_Y)\right].
\end{aligned}
$$

The last equality follows even if $\mathrm{E}[\widehat{\mathbf{f}}_{X,\lambda}] \neq \mathbf{f}_X$, because $\widehat{\mathbf{f}}_Y - \mathbf{C}\mathbf{f}_X$ is zero-mean.

A2 Additional Figures

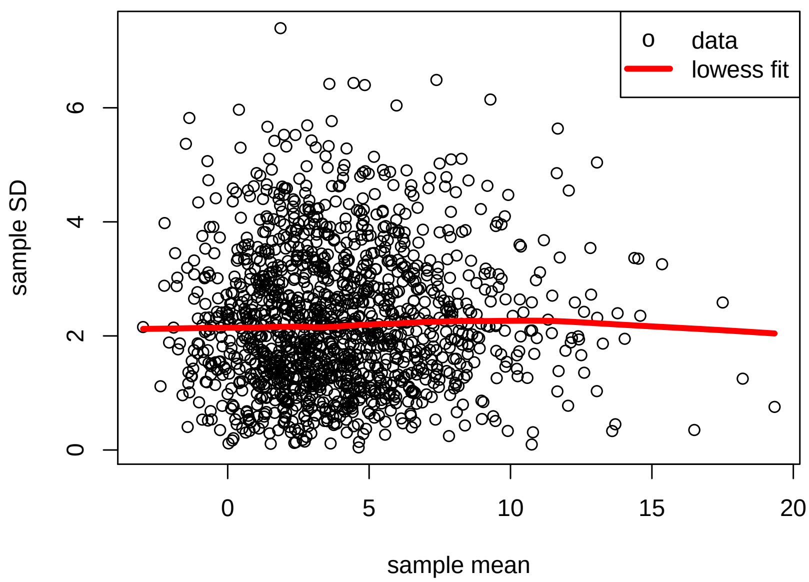

Figure A1: Plot of the sample standard deviations and the sample means in the sodium concentration example.

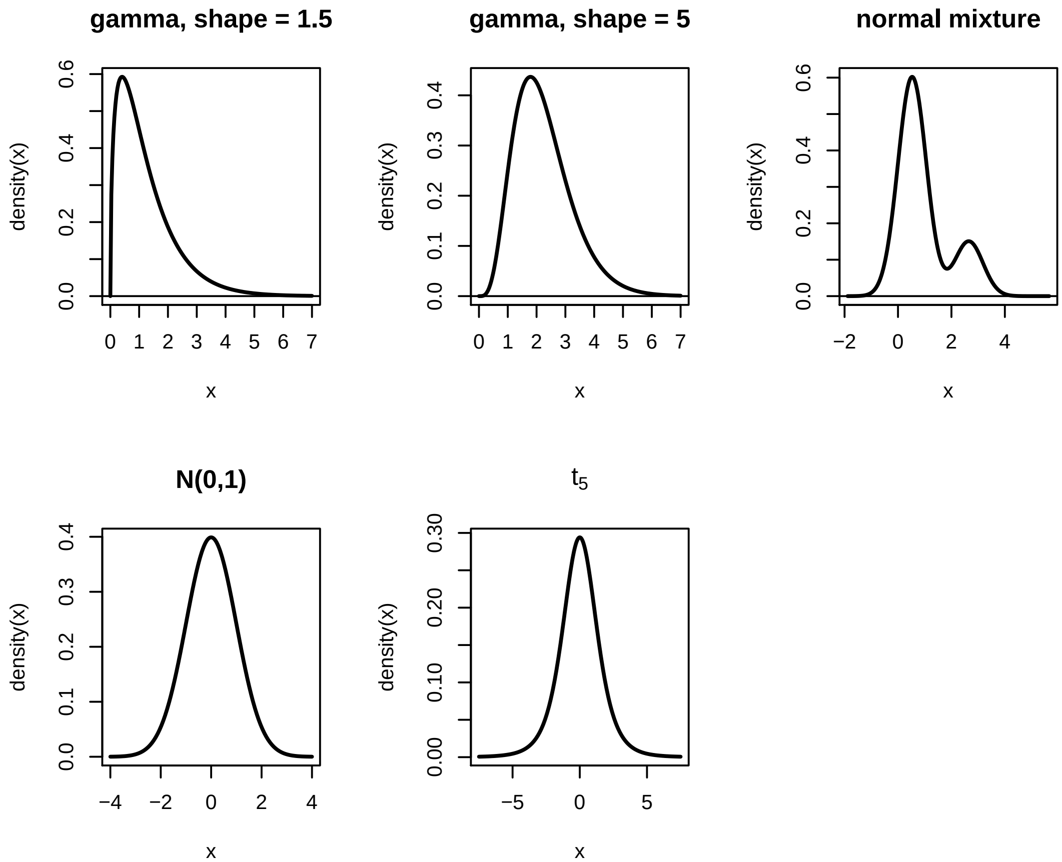

Figure A2: Densities of densities of X used in the simulation study. All of the distributions have variance equal to 1.

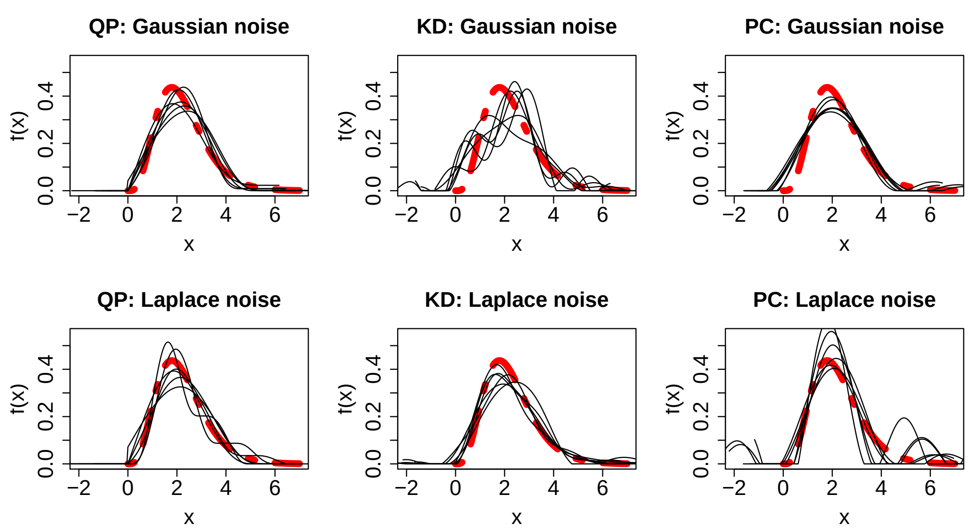

Figure A3: Simulations of Gamma(5), $\sigma_z = 1$, $n = 250$. The true density is the dashed curve.

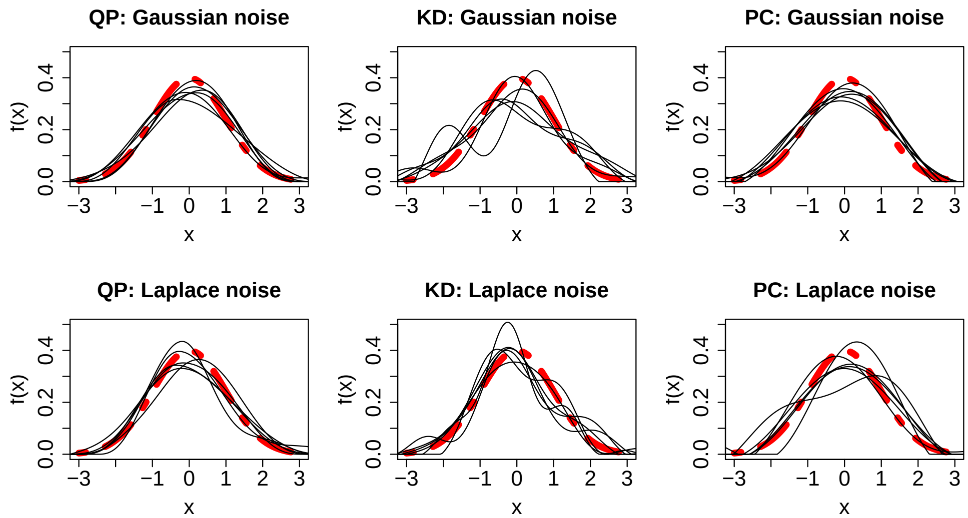

Figure A4: Simulations of $N(0,1)$, $\sigma_z = 1$, $n = 250$. The true density is the dashed curve.

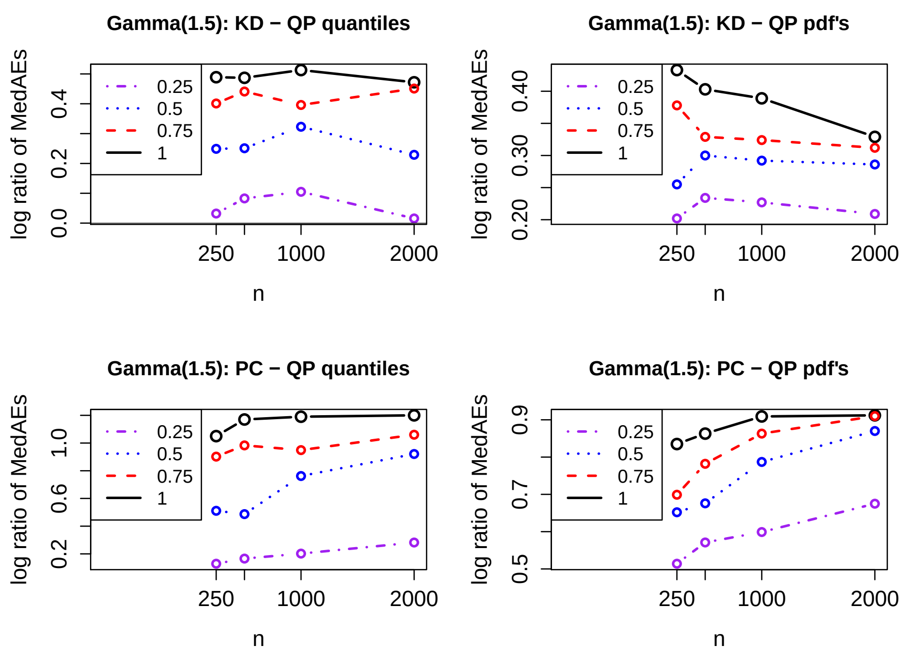

Figure A5: Log ratios for MedAEquant and MedAEpdf with f_x the Gamma(1.5) distribution and Laplace noise. σ_z is given in the legend.

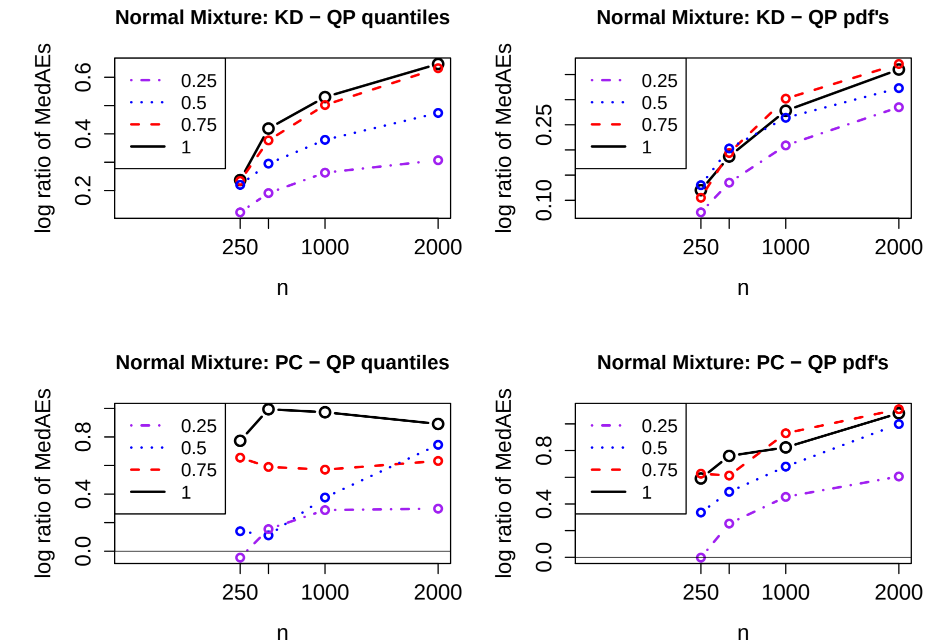

Figure A6: Log ratios for MedAEquant and MedAEpdf with a normal mixture distribution and f_Z Laplace. σ_z is given in the legend.

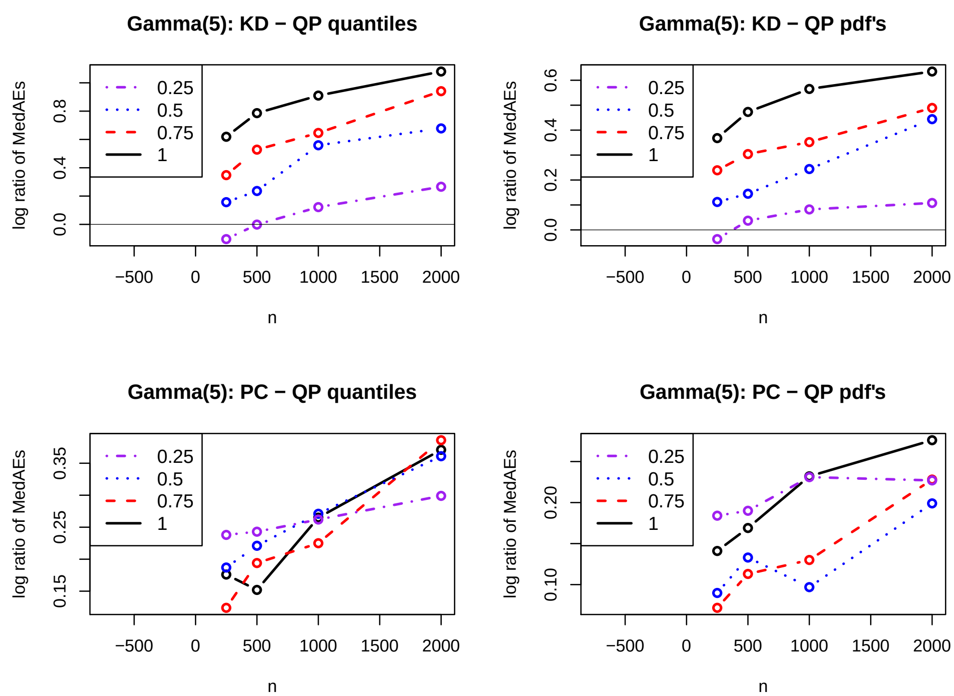

Figure A7: Log ratios for MedAEquant and MedAEpdf with f_X Gamma(5) and f_Z Gaussian. σ_z is given in the legend.

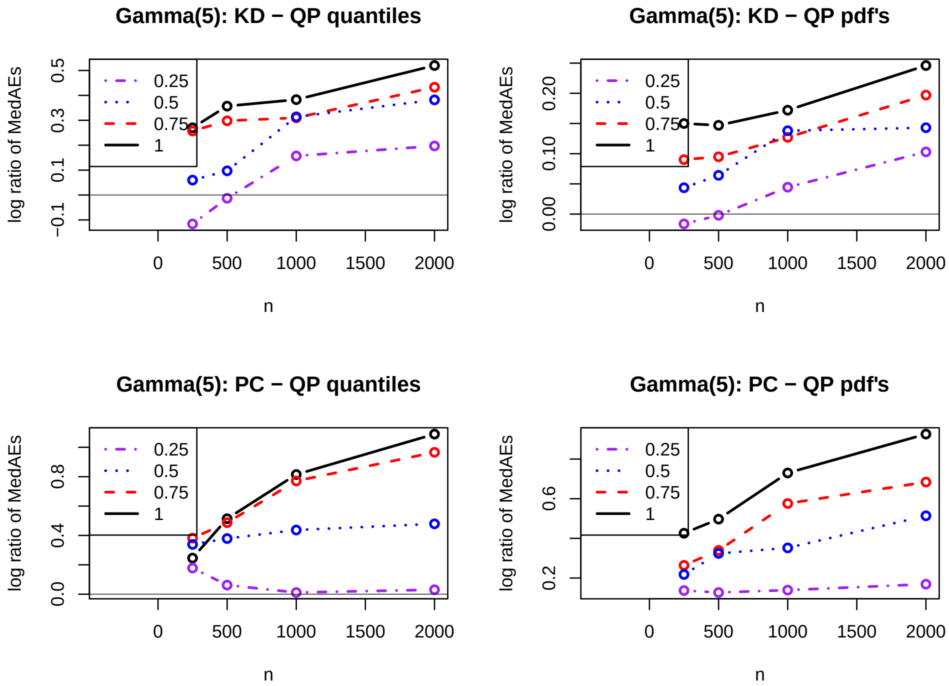

Figure A8: Log ratios for MAEquant and MAEpdf with f_X Gamma(5) and f_Z Laplace. σ_z is given in the legend.

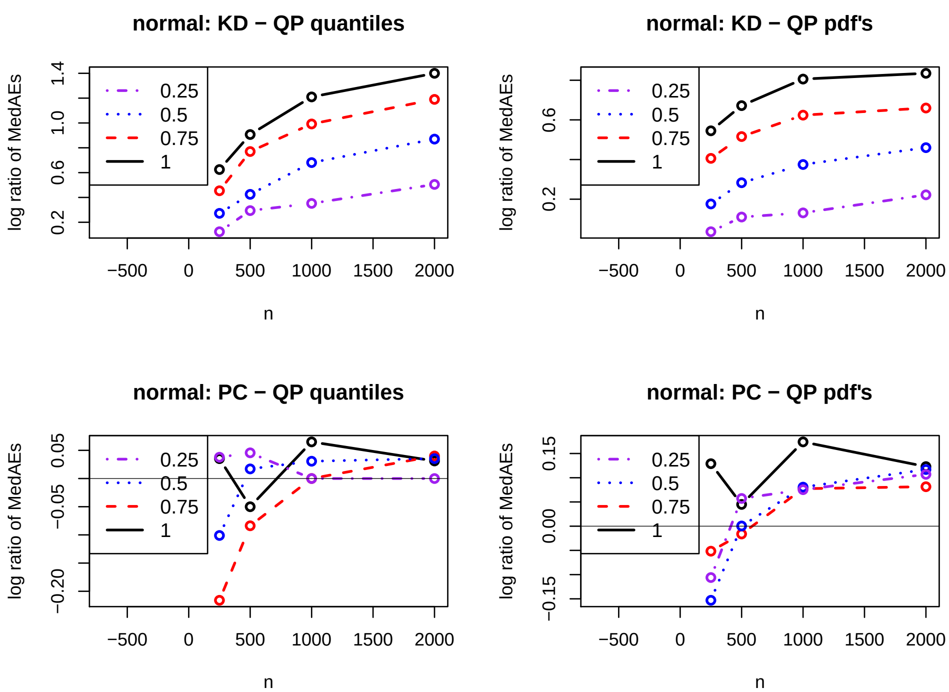

Figure A9: Log ratios for MedAEquant and MedAEpdf with f_X and f_Z both Gaussian. σ_z is given in the legend.

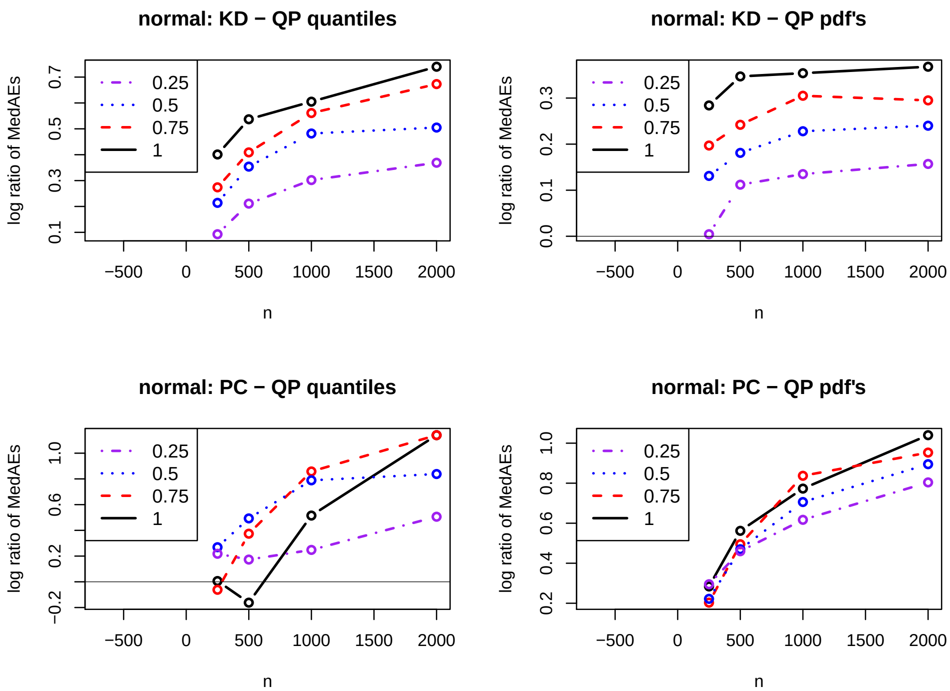

Figure A10: Log ratios for MedAEquant and MedAEpdf with f_X = Gaussian and f_Z = Laplace. σ_z is given in the legend.

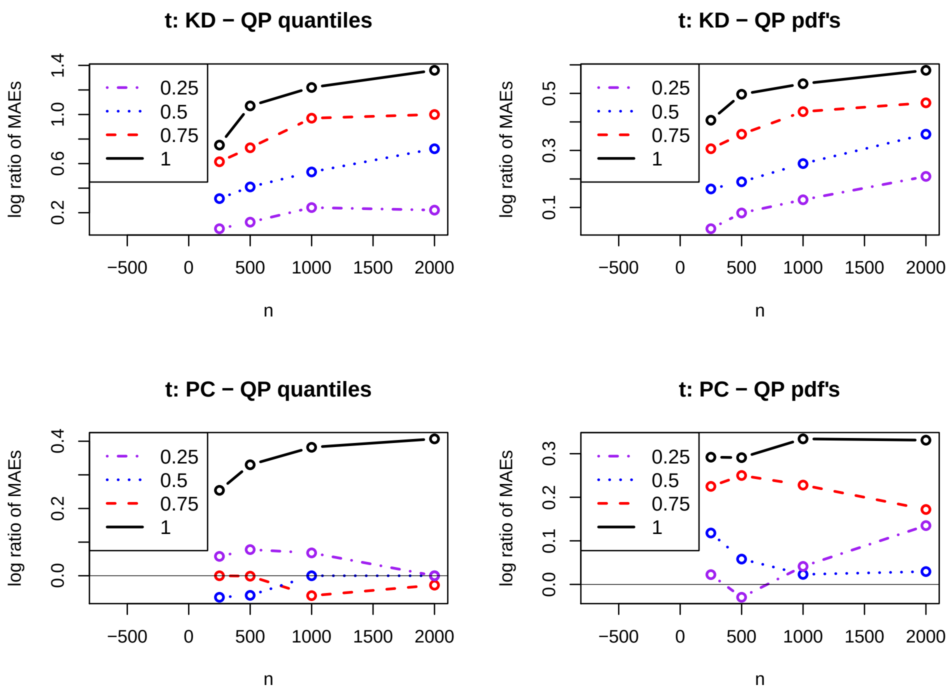

Figure A11: Log ratios for MedAEquant and MedAEpdf with $f_X = t_5$ and $f_Z =$ Gaussian. σ_z is given in the legend.

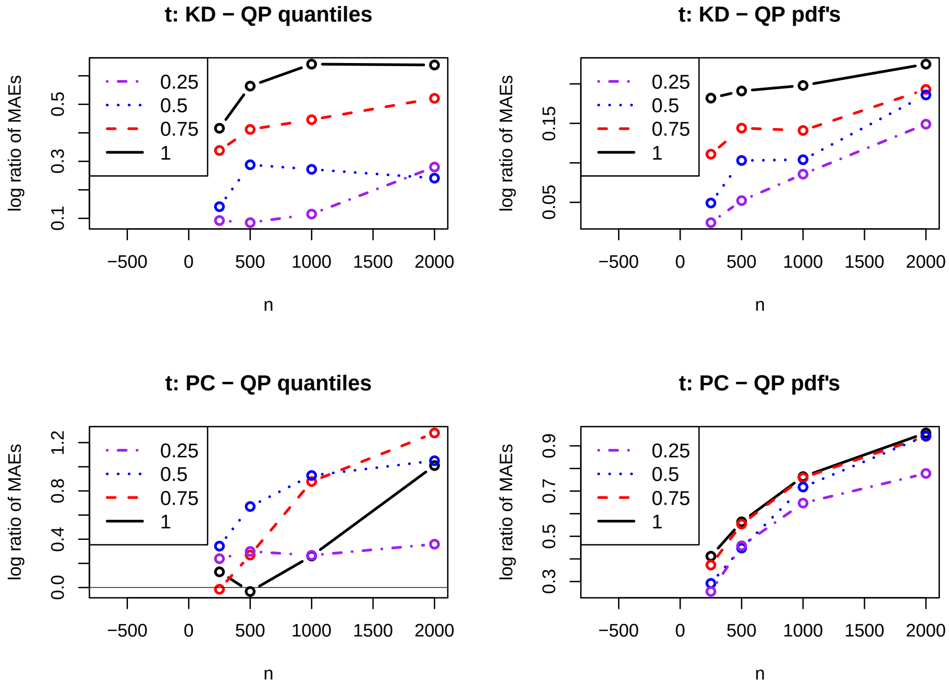

Figure A12: Log ratios for MedAEquant and MedAEpdf with $f_X = t_5$ and $f_Z =$ Laplace. σ_z is given in the legend.

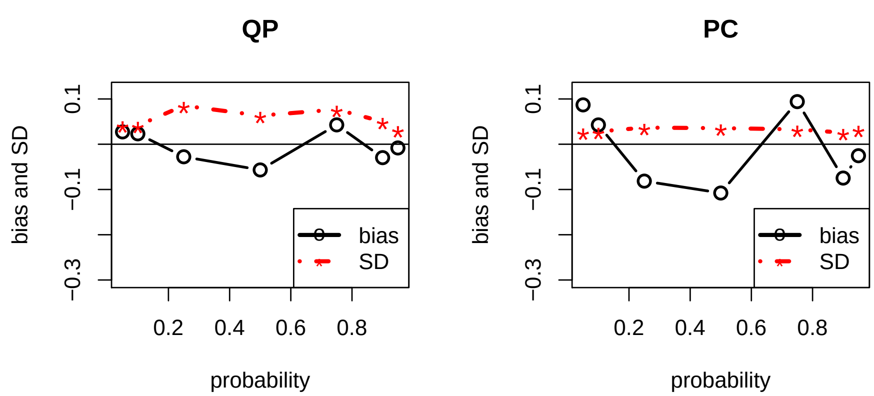

Figure A13: Bias and standard deviation of the QP and PC estimates of f_X at quantiles corresponding to the probability on the x-axis. f_X is a normal mixture. The sample size is 1000, $\sigma_z = 0.75$, and f_Z is Gaussian.

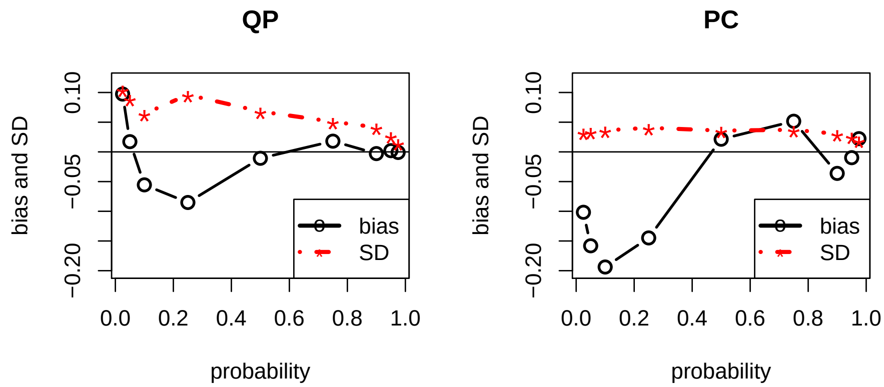

Figure A14: Bias and standard deviation of the QP and PC estimates of f_X at quantiles corresponding to the probability on the x-axis. f_X is Gamma(1.5). The sample size is 1000, $\sigma_z = 0.75$, and f_Z is Gaussian.

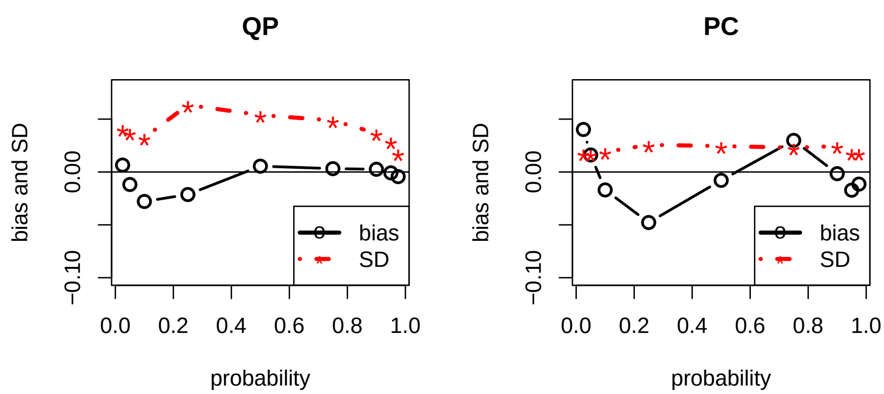

Figure A15: Bias and standard deviation of the QP and PC estimates of f_X at quantiles corresponding to the probability on the x-axis. f_X is Gamma(5). The sample size is 1000, $\sigma_z = 0.75$, and f_Z is Gaussian.

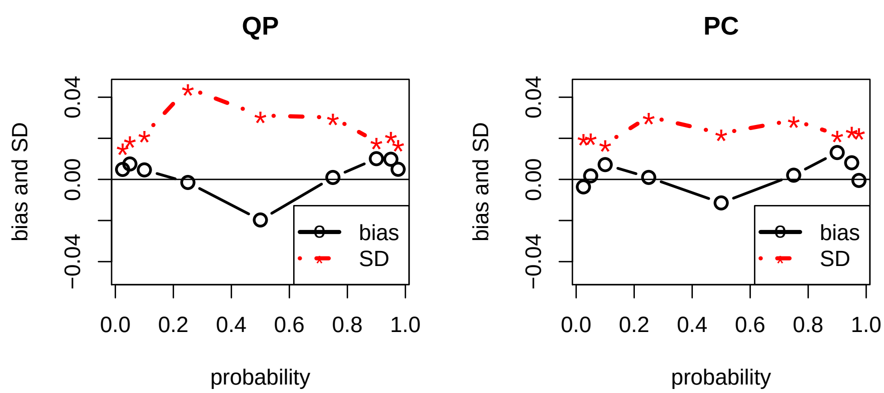

Figure A16: Bias and standard deviation of the QP and PC estimates of f_X at quantiles corresponding to the probability on the x-axis. f_X is $N(0,1)$. The sample size is 1000, $\sigma_z = 0.75$, and f_Z is Gaussian.

A3 Tables

n	σ_z	lower limit	upper limit	length
250.00	0.25	0.08	0.19	0.11
250.00	0.50	0.39	0.51	0.12
250.00	0.75	0.68	0.81	0.13
250.00	1.00	0.95	1.09	0.14
500.00	0.25	0.07	0.19	0.12
500.00	0.50	0.43	0.55	0.12
500.00	0.75	0.69	0.82	0.13
500.00	1.00	1.03	1.16	0.14
1000.00	0.25	0.09	0.20	0.11
1000.00	0.50	0.48	0.59	0.11
1000.00	0.75	0.95	1.06	0.11
1000.00	1.00	1.14	1.27	0.13
2000.00	0.25	0.16	0.28	0.12
2000.00	0.50	0.61	0.72	0.10
2000.00	0.75	0.89	1.00	0.11
2000.00	1.00	1.19	1.30	0.11

Table A1: Lower and upper limits of 95% BC$_a$ confidence intervals for the KD/PC log-ratios for quantile-MedAEs in Figure 9, upper-left panel. The intervals were created by the function **bcanon()** in the **bootstrap** package and used 5000 resamples. The final column contains the lengths of the intervals.

n	σ_z	lower limit	upper limit	length
250.00	0.25	0.21	0.27	0.06
250.00	0.50	0.45	0.53	0.08
250.00	0.75	0.65	0.74	0.09
250.00	1.00	0.80	0.89	0.10
500.00	0.25	0.22	0.27	0.05
500.00	0.50	0.47	0.53	0.06
500.00	0.75	0.64	0.73	0.09
500.00	1.00	0.81	0.91	0.10
1000.00	0.25	0.21	0.26	0.05
1000.00	0.50	0.41	0.48	0.06
1000.00	0.75	0.67	0.75	0.08
1000.00	1.00	0.79	0.88	0.09
2000.00	0.25	0.22	0.28	0.05
2000.00	0.50	0.39	0.45	0.06
2000.00	0.75	0.58	0.67	0.08
2000.00	1.00	0.74	0.83	0.09

Table A2: Lower and upper limits of 95% BC_a confidence intervals for the KD/PC log-ratios of pdf-MedAEs in Figure 9, upper-right panel. The intervals were created by the function **bcanon()** in the **bootstrap** package and used 5000 resamples. The final column contains the lengths of the intervals.