Supplementary material: Quantile regression for longitudinal data via the multivariate generalized hyperbolic distribution

Alvaro J. Flórez
DSI, Universiteit Hasselt, Belgium, and
School of statistics, Universidad del Valle, Colombia

Ingrid Van Keilegom
ORSTAT, KU Leuven, Belgium

Geert Molenberghs
DSI, Universiteit Hasselt, Belgium, and
I-BioStat, KU Leuven, Belgium

Anneleen Verhasselt
DSI, Universiteit Hasselt, Belgium

March 15, 2021

A Example of a multivariate longitudinal model

To fix ideas, assume that for each patient four longitudinal profiles are collected: diastolic and systolic blood pressure, and then low density as well as high density lipoprotein, denoted by DBP, SBP, LDL, and HDL, respectively.

For patient $i = 1, \ldots, N$, the following model is postulated:

$$\begin{aligned} DBP_{ij} &= \beta_1 + \beta_3 t_{ij} + \beta_{11} g_i, \\ SBP_{ij} &= (\beta_1 + \beta_2) + (\beta_3 + \beta_4) t_{ij} + \beta_{12} g_i, \\ LDL_{ij} &= \beta_5 + \beta_7 t_{ij} + \beta_9 t_{ij}^2, \\ HDL_{ij} &= (\beta_5 + \beta_6) + (\beta_7 + \beta_8) t_{ij} + (\beta_9 + \beta_{10}) t_{ij}^2, \end{aligned}$$

where $j = 1, \ldots, T$ refers to the sequencing of the measurement occasions, T is the number of measurements planned for subject i, t_{ij} is the time at which the jth set of measurements is collected for patient i. We refer to a 'set of measurements' because at each occasion the $q = 4$ dimensional vector $(DBP_{ij}, SBP_{ij}, LDL_{ij}, HDL_{ij})'$ is collected. Finally, g_i is the sex of patient i. Clearly, the parameter vector takes the form $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_{12})'$, and is of length $p = 12$.

In the above example, a linear trend for the blood pressure variables is assumed, with a shift between sexes. For the cholesterol variables, a sex-invariant quadratic evolution is postulated. In both cases, the model is written so as to include parameters that quantify the difference between the pairs of sequences.

In spite of the multivariate longitudinal nature of data and corresponding model, we can efficiently write the model in vector and matrix form in the following way. Stack all measurements in a single column vector:

$$\begin{aligned} \mathbf{Y}_i &= (DBP_{i1}, \ldots, DBP_{iT}, SBP_{i1}, \ldots, SBP_{iT}, \\ &\qquad LDL_{i1}, \ldots, LDL_{iT}, HDL_{i1}, \ldots, HDL_{iT})'. \end{aligned}$$

The length of the vector is $n = T \cdot q$. This notation is coherent with (2), upon realizing that this multivariate longitudinal design can be written as an $n \times p = (Tq) \times p$ matrix :

$$\mathbf{X}_i = \left(\begin{array}{cccc|cccccc|cc}
1 & 0 & t_{i1} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & g_i & 0 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
1 & 0 & t_{iT} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & g_i & 0 \\
\hline
1 & 1 & t_{i1} & t_{i1} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & g_i \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
1 & 1 & t_{iT} & t_{iT} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & g_i \\
\hline
0 & 0 & 0 & 0 & 1 & 0 & t_{i1} & 0 & t_{i1}^2 & 0 & 0 & 0 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
0 & 0 & 0 & 0 & 1 & 0 & t_{iT} & 0 & t_{iT}^2 & 0 & 0 & 0 \\
\hline
0 & 0 & 0 & 0 & 1 & 1 & t_{i1} & t_{i1} & t_{i1}^2 & t_{i1}^2 & 0 & 0 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
0 & 0 & 0 & 0 & 1 & 1 & t_{iT} & t_{iT} & t_{iT}^2 & t_{iT}^2 & 0 & 0
\end{array}\right).$$

For reading comfort, the horizontal lines separate the blocks pertaining to the four longitudinal sequences. The first vertical block refers to the time curve for the blood pressure measurements, the second vertical block does the same for the lipoprotein measurements, and the third vertical block is for the sex effect on blood pressure.

Note that time is (the most trivial example of) a time-varying covariate, while sex is time-invariant. For $n = 1$, a purely multivariate outcome vector follows, and for $q = 1$ a single longitudinal outcome is investigated. Generally, a very wide class of multivariate longitudinal sequences can be written in this format.

B Simulation: MAL distribution settings

We consider a bivariate asymmetric Laplace distribution for the outcome, i.e.,

$$\mathbf{Y}_i = (Y_{i1}, Y_{i2})' \sim \mathrm{MAL}_2\left(\boldsymbol{\mu}_i, \boldsymbol{\Delta\xi}, \boldsymbol{\Delta\Sigma\Delta}\right), \text{ for } i = 1, \dots, N. \tag{B.1}$$

The entries of the location vector ($\boldsymbol{\mu}_i$) are:

$$\mu_{ij} = \beta_0 + t_j\beta_1 + t_jT_i\beta_2, \text{ for } j = 1, 2, \tag{B.2}$$

with $t_j = j - 1$ indicating the measurement time, and T_i representing a Bernoulli variable with success probability $\pi = 0.5$. The matrices $\mathbf{\Delta}$ and $\mathbf{\Psi}$ are defined as:

$$\mathbf{\Delta} = \begin{pmatrix} 0.25 & 0 \\ 0 & 0.5 \end{pmatrix} \text{ and } \mathbf{\Psi} = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix},$$

respectively. We set $\boldsymbol{\beta} = (\beta_0, \beta_1, \beta_2)' = (4, 2, 1)'$. The vector $\boldsymbol{\xi}$ and the matrix $\mathbf{\Lambda}$ are known quantities determined by τ.

Parameters of interest and estimators

We are interested in the whole vector of parameters, i.e., $\boldsymbol{\theta} = (\boldsymbol{\beta}, \mathbf{\Delta}, \mathbf{\Psi})$. These parameters are estimated using the maximum likelihood estimator (MLE) with $\epsilon = \{0, 0.01\}$. As a reference, we compare the MLE with the method-of-moments estimator (MME) defined in Section 3.

The estimators are compared using the relative bias (RB) and the relative efficiency (RE) of $\widehat{\beta}_j$, $j = 1, \ldots, p$. The latter is defined as:

$$RE_j = \frac{\text{MSE}_{\text{MME}}\big(\widehat{\beta}_j\big)}{\text{MSE}_{\text{MLE}}\big(\widehat{\beta}_j\big)},$$

If $RE_j > 1$, it indicates that the MLE is more efficient than MME.

Results

Tables B.1 and B.2 exhibit the relative bias (in percentage) and efficiency of the MLE for the bivariate asymmetric Laplace scenarios, respectively. The bias of the MLE is negligible for the location parameters ($\boldsymbol{\beta}$). Furthermore, it is not considerably affected by ϵ. However, there is an increment of the bias for the variance components (δ_1, δ_2, and ψ). On the other hand, this increment is negligible for $\boldsymbol{\beta}$. Regarding the efficiency, the MLE is considerably more efficient than the MME. The RE is larger than 1 in all cases, and increases as τ departs from 0.5 and N increases. Furthermore, the efficiency of the MLE for $\boldsymbol{\beta}$ is improved by using the estimator with $\epsilon = 0.01$.

C Simulation: Cauchy distribution

Table C.4 shows the relative bias and efficiency of the MLE for the bivariate Cauchy settings. Contrary to the previous cases, the MLE shows poor performance. It is biased and

Table B.1: Bivariate asymmetric Laplace case. Relative bias (%) of the MLE with $\epsilon = \{0, 0.01\}$ for different values of τ, ρ, and N.

		MLE with $\epsilon = 0$						MLE with $\epsilon = 0.01$					
		$\rho = 0.50$			$\rho = 0.90$			$\rho = 0.50$			$\rho = 0.90$		
τ	Parm	$N = 50$	$N = 100$	$N = 200$	$N = 50$	$N = 100$	$N = 200$	$N = 50$	$N = 100$	$N = 200$	$N = 50$	$N = 100$	$N = 200$
0.25	β_0	0.08	0.10	0.06	0.29	0.35	0.27	-0.12	-0.06	-0.12	-0.14	-0.06	-0.12
	β_1	0.10	-0.56	-0.01	0.50	0.22	0.25	0.01	-0.47	-0.15	-0.14	-0.28	-0.23
	β_2	0.40	0.85	0.40	-0.02	-0.03	0.18	-0.17	0.27	0.59	-0.25	0.04	0.24
	δ_1	-0.79	-0.55	-0.41	-0.66	-0.50	-0.37	-2.63	-2.38	-2.27	-2.49	-2.30	-2.19
	δ_2	-1.43	-0.67	-0.40	-0.85	-0.36	-0.20	-3.27	-2.61	-2.37	-2.72	-2.33	-2.19
	ψ	-1.78	-0.15	-0.25	-0.18	-0.08	-0.06	-1.64	-0.26	-0.36	-0.11	-0.03	-0.02
0.50	β_0	0.03	0.01	0.00	0.09	0.02	0.01	0.00	0.02	-0.01	0.00	0.02	-0.01
	β_1	0.25	-0.26	0.09	0.31	-0.00	0.05	0.16	-0.29	0.09	0.08	-0.12	0.03
	β_2	-0.63	-0.06	0.26	-0.34	-0.28	0.17	-0.82	-0.13	0.23	-0.41	-0.07	0.12
	δ_1	-1.02	-0.61	-0.21	-1.06	-0.66	-0.23	-2.85	-2.43	-2.03	-2.85	-2.43	-2.03
	δ_2	-1.28	-0.76	-0.52	-1.27	-0.69	-0.32	-3.03	-2.57	-2.34	-2.98	-2.47	-2.12
	ψ	-1.63	-0.47	-0.06	-0.32	-0.13	-0.01	-1.54	-0.41	-0.02	-0.28	-0.10	0.00
0.90	β_0	-0.80	-0.17	-0.08	-1.08	-0.40	-0.35	-0.10	0.24	0.31	-0.06	0.24	0.33
	β_1	-0.04	0.08	-0.05	-1.63	-0.56	-0.41	0.36	0.28	0.54	0.00	0.36	0.69
	β_2	-2.65	-1.90	0.79	-0.40	-0.46	0.16	-2.50	-1.44	0.27	-0.70	-0.22	0.08
	δ_1	-1.94	-0.71	-0.30	-1.64	-0.53	-0.21	-3.83	-2.67	-2.31	-3.55	-2.49	-2.18
	δ_2	-1.66	-0.63	-0.38	-1.52	-0.41	-0.19	-3.83	-2.73	-2.47	-3.57	-2.47	-2.25
	ψ	-2.87	-0.33	-0.00	-0.38	-0.07	0.01	-2.83	-0.56	-0.32	-0.23	-0.02	0.05

Table B.2: Bivariate asymmetric Laplace case. Relative efficiency of the MLE with $\epsilon = \{0, 0.01\}$ for different values of τ, ρ, and N.

		MLE with $\epsilon = 0$						MLE with $\epsilon = 0.01$					
		$\rho = 0.50$			$\rho = 0.90$			$\rho = 0.50$			$\rho = 0.90$		
τ	Parm	$N = 50$	$N = 100$	$N = 200$	$N = 50$	$N = 100$	$N = 200$	$N = 50$	$N = 100$	$N = 200$	$N = 50$	$N = 100$	$N = 200$
0.25	β_0	1.46	1.66	2.06	1.48	1.62	1.88	1.98	2.23	2.46	1.95	2.21	2.50
	β_1	2.84	3.72	3.60	6.45	8.65	8.78	3.41	4.79	4.22	8.31	10.88	10.49
	β_2	3.67	4.23	4.18	13.61	15.94	16.74	4.21	5.37	4.77	17.16	20.17	20.14
	d_1	1.50	1.66	1.65	1.50	1.63	1.60	1.52	1.62	1.56	1.51	1.61	1.52
	d_2	1.63	1.67	1.89	1.50	1.63	1.74	1.64	1.64	1.76	1.51	1.60	1.65
	ψ	1.75	2.25	2.10	5.12	7.45	10.16	1.78	2.24	2.09	5.24	7.50	10.21
0.50	β_0	1.50	1.87	1.82	1.58	1.87	1.82	2.03	2.32	2.30	2.03	2.32	2.30
	β_1	1.74	2.32	2.25	3.44	4.42	4.40	1.99	2.73	2.57	4.30	5.47	5.50
	β_2	2.02	2.65	2.53	7.40	8.97	9.13	2.31	3.19	2.85	9.01	11.48	11.16
	d_1	1.22	1.23	1.27	1.22	1.23	1.27	1.22	1.21	1.22	1.22	1.21	1.22
	d_2	1.34	1.49	2.00	1.27	1.46	2.04	1.35	1.46	1.86	1.28	1.43	1.93
	ψ	1.24	1.52	1.56	5.18	8.38	12.60	1.26	1.52	1.57	5.26	8.44	12.69
0.9	β_0	4.45	5.12	6.15	3.89	4.67	5.06	6.04	6.75	7.24	5.76	6.58	6.81
	β_1	5.96	6.33	8.18	11.72	13.83	16.35	7.28	7.79	10.18	16.36	19.41	21.35
	β_2	10.81	13.38	14.51	41.75	49.97	56.14	13.58	16.13	17.41	53.67	66.52	65.53
	d_1	1.78	1.88	1.91	1.75	1.85	1.85	1.75	1.83	1.80	1.75	1.82	1.75
	d_2	1.78	1.90	2.04	1.76	1.84	1.91	1.75	1.85	1.87	1.75	1.81	1.79
	ψ	2.26	2.56	2.62	3.76	4.49	5.84	2.22	2.54	2.60	3.87	4.51	5.87

Table B.3: Bivariate asymmetric Laplace case. Coverage of the 95% confidence intervals based on the MLE with $\epsilon = 0.01$ for different values of τ, ρ, and N.

		$\rho = 0.50$			$\rho = 0.90$		
τ	parm	$N = 50$	$N = 100$	$N = 200$	$N = 50$	$N = 100$	$N = 200$
0.25	β_0	0.92	0.94	0.95	0.92	0.95	0.96
	β_1	0.89	0.94	0.95	0.91	0.94	0.95
	β_2	0.92	0.94	0.95	0.93	0.94	0.95
0.50	β_0	0.92	0.95	0.95	0.92	0.95	0.95
	β_1	0.88	0.94	0.95	0.91	0.94	0.95
	β_2	0.92	0.95	0.95	0.92	0.95	0.95
0.90	β_0	0.91	0.95	0.94	0.91	0.94	0.94
	β_1	0.88	0.93	0.95	0.90	0.95	0.94
	β_2	0.91	0.94	0.94	0.92	0.94	0.95

highly inefficient. This is because the Cauchy distribution does not have finite moments. Note that the MLE is based on the MGH distribution, in which the moments are well defined. On the other hand, the quantile regression provides better estimates because the quantile function of the Cauchy distribution exists.

D Extended settings

Settings

The bivariate normal and Student-t cases are extended to a setting with two longitudinal outcomes with two measures each. Here, the data-generating model takes the form:

$$
\begin{aligned}
Y_{1,ij} =& \alpha_{10} + t_{ij}\alpha_{11} + T_i t_{ij}\alpha_{12} + (\gamma_{10} + t_j\gamma_{11} + t_j T_i \gamma_{12})\varepsilon_{1,ij}, \\
Y_{2,ij} =& \alpha_{20} + t_{ij}\alpha_{21} + T_i t_{ij}\alpha_{22} + (\gamma_{20} + t_j\gamma_{21} + t_j T_i \gamma_{22})\varepsilon_{2,ij},
\end{aligned}
$$

for $i = 1, \ldots, N$ and $j = 1, 2$. For $\boldsymbol{\varepsilon}_i = (\varepsilon_{1,i1}, \varepsilon_{1,i2}, \varepsilon_{2,i1}, \varepsilon_{2,i2})'$, we consider two distributions:

$$
\boldsymbol{\varepsilon}_i \sim N(\mathbf{0}, \mathbf{S}), \text{ and } \boldsymbol{\varepsilon}_i \sim t_3\left(\mathbf{0}, \frac{1}{3}\mathbf{S}\right).
$$

We set $\boldsymbol{\alpha} = (\alpha_{10}, \alpha_{11}, \alpha_{12}, \alpha_{20}, \alpha_{21}, \alpha_{22})' = (4, 2, 1, 5, -3, 1)'$ and:

$$
\mathbf{S} = \begin{pmatrix} 1 & \rho & \rho & \rho^2 \\ & 1 & \rho^2 & \rho \\ & & 1 & \rho \\ & & & 1 \end{pmatrix}.
$$

Here, the location parameter $\boldsymbol{\beta}$ takes the form:

$$\boldsymbol{\beta} = \boldsymbol{\alpha} + Q_\varepsilon(\tau)\boldsymbol{\gamma},$$

where $\boldsymbol{\gamma} = (\gamma_{10}, \gamma_{11}, \gamma_{12}, \gamma_{20}, \gamma_{21}, \gamma_{22})'$ is $(1, 1, 0, 1, 1, 0)$ and $(\sqrt{1/3}, \sqrt{1/3}, 0, \sqrt{1/3}, \sqrt{1/3}, 0)$ for the four-dimensional normal and Student-t cases, respectively.

As in the previous simulations, we consider $N = \{50, 100, 200\}$, $\tau = \{0.25, 0.5, 0.9\}$, and $\rho = \{0.25, 0.5, 0.9\}$. Moreover, in each scenario, 1000 datasets are generated.

Results

The relative bias and efficiency of the MLE for each parameter for the four-dimensional normal and Student-t cases are presented in Table D.5. These results are somewhat the same as the ones observed for the bivariate settings. The MLE is more efficient for all parameters, except for the intercepts. The efficiency increases with the correlation of the outcomes, and it is larger in the four-dimensional Student-t settings.

The coverage of the 95% confidence intervals based on the MLE with $\epsilon = 0.01$ for the four-dimensional settings are exhibited in Table D.6. Generally, the coverage gets closer to 0.95 as N increases. However, it does not happen for the intercepts. Particularly for $\tau = 0.9$, the coverage for some parameters seems to decrease.

E Simulations: Comparison of the MLE with $\epsilon = 0$ and $\epsilon = 0.01$

A comparison of the MLE with $\epsilon = 0$ and $\epsilon = 0.01$ is presented. To do so, we compute the relative efficiency as follows:

$$RE_j = \frac{MSE_{\text{MLE}_{\epsilon=0}}(\beta_j)}{MSE_{\text{MLE}_{\epsilon=0.01}}(\beta_j)}.$$

If $R_j > 1$, the MLE with $\epsilon = 0.01$ is more efficient for β_j than the MLE with $\epsilon = 0$. These results are exhibited in Table E.7. In most of the cases, the inclusion of ϵ reduces the mean square error of the MLE.

F Leuven diabetes project: sensitivity analysis of the MLE

The parameter estimates and standard errors of the MLE using different values of ϵ are shown in Table F.8 and Table F.9, respectively. Furthermore, the gradient of the log-likelihood for each parameter at the ML estimate is exhibited in Table F.10. The tables show that none of these quantities change considerably as ϵ increases. These findings suggest that no computational issues are encountered in the LDP analysis.

In contrast, Table F.11 presents the same quantities for a simulated data set with the four-dimensional normal setting with $N = 200$ and $\rho = 0.5$. Here, the estimates do not change drastically with ϵ. However, the standard errors and the gradients are heavily affected. Note that the gradients are considerably large when ϵ ranges from 0 to 0.01. For $\epsilon \geq 0.1$, they are close to zero, and consequently, the standard errors are more stable.

G The multivariate asymmetric Laplace distribution

The multivariate asymmetric Laplace distribution (MAL) is defined as follows:

$$\mathbf{Y}_i \sim \mathrm{MAL}_n\big(\mathbf{X}_i\boldsymbol{\beta}, \boldsymbol{\Delta}\boldsymbol{\xi}, \boldsymbol{\Delta}\boldsymbol{\Sigma}\boldsymbol{\Delta}\big),$$

with density

$$\begin{aligned} f(\mathbf{y}|\mathbf{X}_i;\boldsymbol{\theta}) :&= f\big(\mathbf{y}|\mathbf{X}_i\boldsymbol{\beta}, \boldsymbol{\Delta}\boldsymbol{\xi}, \boldsymbol{\Delta}\boldsymbol{\Sigma}\boldsymbol{\Delta}\big) \\ &= \frac{2\exp\big[(\mathbf{y}-\mathbf{X}_i\boldsymbol{\beta})'\boldsymbol{\Delta}^{-1}\boldsymbol{\Sigma}^{-1}\boldsymbol{\xi}\big]}{(2\pi)^{n/2}|\boldsymbol{\Delta}\boldsymbol{\Sigma}\boldsymbol{\Delta}|^{1/2}}\Big(\frac{m_i}{2+d}\Big)^{\nu/2}K_\nu\Big[\sqrt{(2+d)m_i}\Big], \end{aligned} \tag{G.3}$$

where $\mathbf{y} = (y_1, \dots, y_n)'$, $(\mathbf{X}_i\boldsymbol{\beta})_j = Q_{Y_{ij}}(\tau|\mathbf{X}_i)$ is the τth conditional quantile of Y_{ij} (the location parameter vector, for $j = 1, \dots, n$), $\boldsymbol{\Delta}\boldsymbol{\xi}$ is the scale (or skewness) parameter vector and $\boldsymbol{\Sigma} = \boldsymbol{\Lambda}\boldsymbol{\Psi}\boldsymbol{\Lambda}$ a positive definite matrix. Here, we use the following notation: $\boldsymbol{\Delta} = \mathrm{diag}(\delta_1, \dots, \delta_n)$, $\delta_j > 0$ (for $j = 1, \dots, n$), $\boldsymbol{\xi} = (\xi_1, \dots, \xi_n)'$, $\xi_j = \frac{1-2\tau}{\tau(1-\tau)}$ for $j = 1, \dots, n$, $\boldsymbol{\Lambda} = \mathrm{diag}(\lambda_1, \dots, \lambda_n)$, $\lambda_j^2 = \frac{2}{\tau(1-\tau)}$ for $j = 1, \dots, n$, and $\boldsymbol{\Psi}$ is a correlation matrix. Further $m_i = (\mathbf{y}-\mathbf{X}_i\boldsymbol{\beta})'(\boldsymbol{\Delta}\boldsymbol{\Sigma}\boldsymbol{\Delta})^{-1}(\mathbf{y}-\mathbf{X}_i\boldsymbol{\beta})$, $d = \boldsymbol{\xi}'\boldsymbol{\Sigma}\boldsymbol{\xi}$, K_ν is the modified Bessel function of the third kind with index parameter $\nu = (2-n)/2$ and $\boldsymbol{\theta} = (\boldsymbol{\beta}, \boldsymbol{\Delta}, \boldsymbol{\Psi})$ is the vector of unknown parameters. Note that if $\mathbf{Y}_i \sim \mathrm{MAL}_n\big(\mathbf{X}_i\boldsymbol{\beta}, \boldsymbol{\Delta}\boldsymbol{\xi}, \boldsymbol{\Delta}\boldsymbol{\Sigma}\boldsymbol{\Delta}\big)$, then $\boldsymbol{\beta}$ captures the location of $\mathbf{Y}_i$, $\boldsymbol{\Psi}$ characterizes the dependence between the components of $\mathbf{Y}_i$, whereas the variability of $Y_{i1}, \dots, Y_{in}$ is captured by the variances $\delta_1, \dots, \delta_n$ in the matrix $\boldsymbol{\Delta}$. Also note that for reasons of notational simplicity, we do not add an index τ to the parameters of the model, although it is clear that the parameters will change with τ.

The MAL density in (G.3) for $n \geq 2$ is unbounded, and in particular it diverges to infinity when $\mathbf{y}$ tends to $\mathbf{X}_i\boldsymbol{\beta}$. This is because the Bessel function $K_\nu(u)$ is proportional to $u^{-\nu}$ for u close to zero (see the Appendix of Kotz et al. (2001)), and hence $m_i^{\nu/2}K_\nu\big(\sqrt{(2+d)m_i}\big)$ grows unboundedly for $\mathbf{y}$ approaching $\mathbf{X}_i\boldsymbol{\beta}$, except when $\nu = 1/2$, i.e. when $n = 1$. This is a crucial observation, which has some profound consequences. In fact, it shows that the log-likelihood corresponding to the MAL density will be equal to infinity whenever one of the data points $\mathbf{Y}_i$ equals exactly $\mathbf{X}_i\boldsymbol{\beta}$. Since the system of equations $\mathbf{Y}_i = \mathbf{X}_i\boldsymbol{\beta}$ is solvable for all $i = 1, \ldots, N$ whenever the rank of the matrix $\mathbf{X}_i$ is at least equal to n (which is usually the case in practice), there will be an abundance of local maxima in that case for which the log-likelihood equals infinity. On the other hand, in the univariate case it is well known (see Koenker (2005)) that the MAL density reduces to

$$f(y|\mathbf{X}_i;\boldsymbol{\theta}) = \gamma^{-1}\exp[-|y-\mathbf{X}_i\boldsymbol{\beta}|\delta^{-2}\lambda^{-2}\{\gamma - \xi\,\mathrm{sgn}(y-\mathbf{X}_i\boldsymbol{\beta})\}], \qquad \text{(G.4)}$$

with $\gamma = \sqrt{2\delta^2\lambda^2 + \xi^2}$. Note that this expression reduces to (1.1) when $\delta = 1$. It is easily seen that (G.4) is bounded (albeit non-smooth) in y and it reaches its maximal value γ^{-1} when $y = \mathbf{X}_i\boldsymbol{\beta}$.

H Asymptotic properties of the MLE

Suppose that the true density of $\mathbf{Y}_i$ given $\mathbf{X}_i$ is $g(\cdot|\mathbf{X}_i)$. If $f_\epsilon(\cdot|\mathbf{X}_i;\boldsymbol{\theta})$ defined in (G.3) contains the true structure $g(\cdot|\mathbf{X}_i)$, then $g(\cdot|\mathbf{X}_i) = f_\epsilon(\cdot|\mathbf{X}_i;\boldsymbol{\theta}_0)$ for some $\boldsymbol{\theta}_0$. The following regularity conditions are required for the asymptotic results. Here and in what follows, expectations are taken with respect to the true distribution.

(C1) The data $(\mathbf{Y}_i, \mathbf{X}_i)(i = 1, \ldots, N)$ have common joint distribution function G on a measurable Euclidean space, with Radon-Nikodým density g given by $g(\mathbf{Y}_i, \mathbf{X}_i) = g(\mathbf{Y}_i|\mathbf{X}_i)f_X(\mathbf{X}_i)$.

(C2) $E(\log g(\mathbf{Y}_i|\mathbf{X}_i)) = \int \log g(\mathbf{y}|\mathbf{x})\,dG(\mathbf{y},\mathbf{x})$ exists and $|\log f_\epsilon(\mathbf{y}|\mathbf{x};\boldsymbol{\theta})| \leq M(\mathbf{y}|\mathbf{x})$ for all $\boldsymbol{\theta} \in \Theta$, where M is integrable with respect to G and $\mathbf{x} = (\mathbf{x}_1', \ldots, \mathbf{x}_n')$ with $\mathbf{x}_j = (x_{j1}, \ldots, x_{jp})'$, for $j = 1, \ldots, n$. Moreover,

$$E(\log[g(\mathbf{Y}_i|\mathbf{X}_i)/f_\epsilon(\mathbf{Y}_i|\mathbf{X}_i;\boldsymbol{\theta})]) = \int \log g(\mathbf{y}|\mathbf{x})\,dG(\mathbf{y},\mathbf{x}) - \int \log f_\epsilon(\mathbf{y}|\mathbf{x};\boldsymbol{\theta})\,dG(\mathbf{y},\mathbf{x})$$

has a unique minimum at $\boldsymbol{\theta}_* = (\boldsymbol{\beta}_*, \boldsymbol{\Delta}_*, \boldsymbol{\Psi}_*) \in \Theta$.

(C3) The functions $|\partial^2 \log f_\epsilon(\mathbf{y}|\mathbf{x};\boldsymbol{\theta})/\partial\boldsymbol{\theta}_j\partial\boldsymbol{\theta}_l|$ and $|\partial \log f_\epsilon(\mathbf{y}|\mathbf{x};\boldsymbol{\theta})/\partial\boldsymbol{\theta}_j \cdot \partial \log f_\epsilon(\mathbf{y}|\mathbf{x};\boldsymbol{\theta})/\partial\boldsymbol{\theta}_l|$ $(j, l = 1, \ldots, k)$ are dominated by functions that are integrable with respect to G for all $\mathbf{y} \in I\!R^n$, $\mathbf{x} \in I\!R^{n\times p}$ and $\boldsymbol{\theta} \in \Theta$.

(C4) (a) $\boldsymbol{\theta}_*$ lies in the interior of Θ.
(b) The matrix

$$\mathbf{B}(\boldsymbol{\theta}) = \Big(E\Big[\frac{\partial}{\partial\boldsymbol{\theta}_j}\log f_\epsilon(\mathbf{Y}_i|\mathbf{X}_i;\boldsymbol{\theta})\frac{\partial}{\partial\boldsymbol{\theta}_l}\log f_\epsilon(\mathbf{Y}_i|\mathbf{X}_i;\boldsymbol{\theta})\Big]\Big)_{j,l=1}^k$$

is non-singular.
(c) The matrix

$$\mathbf{A}(\boldsymbol{\theta}) = \Big(E\Big[\frac{\partial^2}{\partial\boldsymbol{\theta}_j\partial\boldsymbol{\theta}_l}\log f_\epsilon(\mathbf{Y}_i|\mathbf{X}_i;\boldsymbol{\theta})\Big]\Big)_{j,l=1}^k$$

has constant rank in some open neighborhood of $\boldsymbol{\theta}_*$.

Proof of Theorem 3.1.

The results in White (1982) are valid under a set of 'primitive' or 'high-level' assumptions, that need to be verified when these results are applied to specific models. They are denoted by A1–A6 in White (1982)'s paper. Note that assumptions A1, A3, A5 and A6 correspond respectively to our conditions (C1)–(C4). For verifying assumptions A2 and A4, we need to show that the partial derivatives of $\log f_\epsilon(\mathbf{y}|\mathbf{x};\boldsymbol{\theta})$ with respect to the components of $\boldsymbol{\theta}$ exist and are continuously differentiable functions of $\boldsymbol{\theta}$ for all $(\mathbf{y}, \mathbf{x})$.

Examining the formula of the log-likelihood $\ell(\boldsymbol{\theta})$ given in (3.4), it is clear that all terms are continuously differentiable with respect to all components of $\boldsymbol{\theta}$, except possibly the last term, of which the derivative equals

$$\frac{\partial}{\partial\boldsymbol{\theta}_l}\log K_\nu\Big[\sqrt{(2+d)(\epsilon+m_i)}\Big] = \frac{K'_\nu\Big[\sqrt{(2+d)(\epsilon+m_i)}\Big]\frac{\partial}{\partial\boldsymbol{\theta}_l}\sqrt{(2+d)(\epsilon+m_i)}}{K_\nu\Big[\sqrt{(2+d)(\epsilon+m_i)}\Big]}.$$

Note that the Bessel function $K_\nu(u)$ is positive, continuous and decreasing in u for $u > 0$ and $K_\nu(-u) = K_\nu(u)$, and its derivative satisfies $K'_\nu(u) = -\frac{1}{2}[K_{\nu-1}(u) + K_{\nu+1}(u)]$, showing that $K'_\nu(u)$ exists and is continuous for $u \neq 0$ (see the Appendix of Kotz et al. (2001)). As $\sqrt{(2+d)(\epsilon+m_i)} > 0$, $K'_\nu\Big[\sqrt{(2+d)(\epsilon+m_i)}\Big]$ as well as $\frac{\partial}{\partial\boldsymbol{\theta}_l}\sqrt{(2+d)(\epsilon+m_i)}$ exist and are continous in $\boldsymbol{\theta}_l$. The proof for the differentiability with respect to the components of $\boldsymbol{\Delta}$ and $\boldsymbol{\Psi}$ is similar, and therefore omitted. □

References

Koenker, R. (2005). *Quantile regression*. Cambridge University Press.

Kotz, S., Kozubowski, T. J., and Podgorski, K. (2001). *The Laplace Distribution and Generalizations*. Springer Science/Business Media, New Jersey.

White, H. (1982). Maximum likelihood estimation of misspecified models. *Econometrica*, 50:1–25.

Table C.4: Bivariate Cauchy case. Relative bias (%) and efficiency of the MLE with $\epsilon = 0.01$ for different values of τ, ρ, and N

		Relative bias (%)					
		$\rho = 0.50$			$\rho = 0.90$		
τ	parm	$N = 50$	$N = 100$	$N = 200$	$N = 50$	$N = 100$	$N = 200$
0.25	β_0	-236.2870	-124.9192	-123.5677	-266.1739	-230.9356	-232.6405
	β_1	-281.3268	-292.3642	-918.0867	-700.5748	-548.4709	-1009.2382
	β_2	-143.8569	-23.8664	115.0147	-214.5173	-17.0145	155.9189
0.50	β_0	-26.1117	2.0663	3.6602	-25.6943	-13.2308	5.3613
	β_1	-3.2208	18.0443	-34.9988	-10.0857	-10.9045	-20.5207
	β_2	-87.1362	-88.5733	207.8706	-210.8982	-16.4909	151.6264
0.90	β_0	91.7900	37.1932	69.0483	98.1470	96.9779	107.6107
	β_1	119.6601	129.1206	224.8149	157.2951	130.5659	224.5341
	β_2	-193.8205	-59.3270	71.9711	-263.8333	-50.8552	141.7179
		Relative efficiency					
		$\rho = 0.50$			$\rho = 0.90$		
0.25	β_0	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	β_1	0.0005	0.0001	0.0000	0.0002	0.0001	0.0000
	β_2	0.0007	0.0012	0.0005	0.0004	0.0016	0.0004
0.50	β_0	0.0001	0.0020	0.0002	0.0001	0.0002	0.0002
	β_1	0.0004	0.0005	0.0002	0.0003	0.0004	0.0002
	β_2	0.0002	0.0010	0.0001	0.0001	0.0004	0.0001
0.90	β_0	0.0003	0.0005	0.0001	0.0003	0.0003	0.0001
	β_1	0.0140	0.0005	0.0001	0.0049	0.0010	0.0002
	β_2	0.0478	0.1584	0.0080	0.0186	0.1399	0.0106

Table D.5: four-dimensional normal and Student-t cases. Relative bias (%) and efficiency of the MLE with $\epsilon = 0.01$ for different values of τ, ρ, and N.

			Relative bias						Relative efficiency					
			$\rho = 0.50$			$\rho = 0.90$			$\rho = 0.50$			$\rho = 0.90$		
distr.	τ	parm	$N = 50$	$N = 100$	$N = 200$	$N = 50$	$N = 100$	$N = 200$	$N = 50$	$N = 100$	$N = 200$	$N = 50$	$N = 100$	$N = 200$
normal	0.25	β_{10}	-0.86	-0.72	-0.73	-0.78	-0.79	-0.65	1.03	1.05	0.97	1.00	1.08	1.09
		β_{11}	-0.88	-2.35	-1.87	-1.62	-2.17	-1.68	1.51	1.44	1.54	3.20	3.07	3.21
		β_{12}	0.40	0.31	-0.57	-0.34	0.66	-0.06	1.45	1.63	1.84	6.18	6.21	6.63
		β_{20}	-0.58	-0.70	-0.58	-0.58	-0.63	-0.48	0.97	1.02	0.99	0.95	1.03	1.05
		β_{21}	0.29	0.55	0.53	0.52	0.70	0.55	1.40	1.54	1.50	2.86	3.24	2.89
		β_{22}	-1.34	1.51	-0.18	-1.14	0.94	-0.07	1.49	1.70	1.75	6.17	6.65	6.26
	0.50	β_{10}	0.09	0.04	0.05	0.06	0.08	0.03	1.06	1.02	1.15	1.05	0.98	1.14
		β_{11}	0.73	-0.17	0.23	0.33	0.03	0.07	1.25	1.43	1.43	2.68	3.31	3.17
		β_{12}	-0.97	1.72	-1.67	-0.43	0.93	-0.51	1.22	1.33	1.50	4.68	5.49	5.93
		β_{20}	-0.04	-0.04	0.06	0.00	0.05	0.04	1.04	1.04	1.10	1.02	1.11	1.11
		β_{21}.1	-0.86	-0.21	-0.17	-0.35	-0.07	-0.09	1.23	1.41	1.44	3.22	3.31	3.11
		β_{22}	-4.20	1.01	-0.77	-1.71	0.93	-0.49	1.24	1.41	1.45	5.30	5.59	5.37
	0.90	β_{10}	2.97	2.53	2.70	2.10	1.89	1.96	0.74	0.64	0.46	0.89	0.84	0.65
		β_{11}	3.49	4.54	4.37	3.24	3.22	3.05	1.44	1.45	1.31	3.12	2.81	2.55
		β_{12}	0.18	-0.62	-0.24	-0.57	0.20	-0.16	2.04	2.50	2.50	9.92	11.10	10.96
		β_{20}	2.33	2.31	2.32	1.74	1.63	1.65	0.72	0.60	0.41	0.77	0.76	0.64
		β_{21}	-7.50	-8.35	-7.95	-6.68	-6.09	-5.73	1.74	1.44	1.40	3.24	2.75	2.90
		β_{22}	-0.33	-0.06	-1.23	-1.15	0.20	-0.47	2.37	2.44	2.64	10.12	11.17	11.25
Student-t	0.25	β_{10}	1.69	1.78	1.95	1.68	1.76	1.97	1.05	0.84	0.57	1.05	0.83	0.57
		β_{11}	4.77	4.48	4.37	4.22	4.28	4.50	1.71	1.72	1.55	3.92	3.41	2.61
		β_{12}	-2.38	0.15	0.14	-0.83	0.10	-0.03	2.09	2.22	2.24	9.01	9.37	9.11
		β_{20}	1.35	1.43	1.55	1.34	1.39	1.55	0.98	0.92	0.59	0.99	0.89	0.56
		β_{21}	-1.74	-2.12	-1.97	-1.83	-1.99	-2.07	1.89	1.64	1.58	3.41	3.19	2.48
		β_{22}	-1.50	0.42	0.31	-0.77	0.21	-0.00	2.09	2.28	2.16	8.08	9.26	9.58
	0.50	β_{10}	0.02	-0.07	-0.01	0.02	-0.07	-0.01	1.10	1.22	1.24	1.10	1.22	1.24
		β_{11}	0.50	0.13	-0.03	0.28	-0.05	-0.03	1.43	1.61	1.53	3.35	3.78	3.22
		β_{12}	-1.38	0.53	-0.21	-0.69	0.26	-0.11	1.48	1.63	1.65	5.46	6.37	6.05
		β_{20}	-0.03	-0.02	0.02	-0.00	-0.05	0.01	1.14	1.18	1.23	1.12	1.25	1.20
		β_{21}	-0.10	-0.19	-0.07	-0.13	-0.02	-0.02	1.50	1.49	1.50	3.21	3.58	3.22
		β_{22}	-1.50	0.54	0.01	-0.83	0.31	-0.07	1.45	1.54	1.46	5.32	6.31	5.74
	0.90	β_{10}	-1.05	-1.80	-1.95	-1.96	-2.46	-2.55	1.49	1.22	0.82	1.44	1.06	0.65
		β_{11}	-1.40	-2.79	-3.29	-2.93	-4.01	-4.22	3.69	2.48	2.72	6.19	3.57	3.21
		β_{12}	-1.45	-0.74	-0.66	-1.07	-0.12	-0.39	7.17	6.01	6.49	29.04	27.84	27.77
		β_{20}	-0.86	-1.49	-1.65	-1.61	-2.09	-2.12	1.09	1.01	0.72	1.07	0.90	0.58
		β_{21}	2.56	3.65	4.30	4.13	5.77	5.95	3.05	2.49	2.71	6.91	3.78	3.25
		β_{22}	-0.58	-0.42	-0.51	-1.06	-0.04	-0.42	5.92	6.33	6.10	27.11	28.55	29.07

Table D.6: four-dimensional normal and Student-t cases. Coverage of the 95% confidence intervals based on the MLE using $\epsilon = 0.01$ for different values of τ, ρ, and N.

		Normal						Student-t					
		$\rho = 0.50$			$\rho = 0.90$			$\rho = 0.50$			$\rho = 0.90$		
τ	parm	$N = 50$	$N = 100$	$N = 200$	$N = 50$	$N = 100$	$N = 200$	$N = 50$	$N = 100$	$N = 200$	$N = 50$	$N = 100$	$N = 200$
0.25	β_{10}	91.60	94.80	94.50	90.10	94.00	95.70	87.50	87.30	78.80	87.90	88.20	77.30
	β_{11}	91.10	94.40	94.90	89.70	93.70	94.60	88.50	90.30	90.40	88.20	90.00	84.80
	β_{12}	94.80	93.90	94.90	93.70	94.90	95.10	93.50	93.40	94.30	93.50	95.00	94.00
	β_{20}	89.90	95.00	95.20	90.20	93.50	95.80	87.40	85.30	76.70	86.70	87.10	78.00
	β_{21}	90.20	94.30	95.10	90.50	94.00	94.70	89.20	92.00	89.10	88.50	89.40	84.50
	β_{22}	94.20	95.50	95.50	93.80	95.50	95.20	93.00	94.80	94.10	93.70	94.70	94.80
0.50	β_{10}	92.90	93.60	94.60	91.99	95.20	95.60	92.00	95.60	94.20	92.00	95.60	94.20
	β_{11}	90.50	93.70	94.00	92.69	94.40	94.70	88.90	93.00	93.20	91.90	95.20	94.10
	β_{12}	93.10	93.70	94.00	93.39	94.90	94.50	92.40	94.50	95.00	92.40	94.50	95.00
	β_{20}	91.00	95.70	94.60	92.09	94.90	94.70	91.40	95.40	94.00	91.70	95.60	94.10
	β_{21}	90.10	93.90	94.10	92.09	94.70	94.90	90.80	92.70	94.60	92.80	95.60	93.80
	β_{22}	94.00	93.90	95.70	93.69	94.50	94.30	93.70	95.00	94.60	92.80	94.90	94.80
0.90	β_{10}	87.66	88.06	77.70	78.55	90.80	85.70	83.82	84.40	79.60	77.56	77.68	69.50
	β_{11}	88.63	93.28	90.50	83.89	92.20	91.20	90.55	90.20	87.80	84.37	84.38	77.40
	β_{12}	92.27	93.98	93.50	92.45	95.20	94.90	94.27	95.10	95.40	94.69	94.99	95.30
	β_{20}	87.55	88.77	77.00	77.84	91.30	86.80	85.03	85.70	78.30	76.75	77.58	69.50
	β_{21}	89.59	92.98	92.80	83.69	91.30	91.80	90.05	91.30	88.90	84.27	84.18	77.40
	β_{22}	93.78	94.68	95.50	92.55	96.00	95.50	93.97	94.90	94.90	94.39	96.30	95.00

Table E.7: Normal and Student-t cases. Relative efficiency of the MLE with $\epsilon = 0.01$ compared to the MLE with $\epsilon = 0$ for different values of τ, ρ, and N.

		bivariate normal						bivariate Student-t					
		$\rho = 0.50$			$\rho = 0.90$			$\rho = 0.50$			$\rho = 0.90$		
τ	parm	$N = 50$	$N = 100$	$N = 200$	$N = 50$	$N = 100$	$N = 200$	$N = 50$	$N = 100$	$N = 200$	$N = 50$	$N = 100$	$N = 200$
0.25	β_0	1.22	1.34	1.34	1.22	1.21	1.22	1.09	1.12	1.03	1.09	1.14	1.00
	β_1	1.08	1.28	1.27	1.11	1.14	1.16	1.50	1.70	1.66	3.58	3.83	3.54
	β_2	1.10	1.23	1.22	1.16	1.15	1.18	1.89	2.05	1.97	7.78	8.79	8.23
0.50	β_0	1.27	1.28	1.27	1.26	1.26	1.23	1.11	1.12	1.15	1.11	1.12	1.15
	β_1	1.16	1.12	1.16	1.19	1.24	1.16	1.39	1.51	1.44	3.00	3.50	2.93
	β_2	1.12	1.13	1.14	1.19	1.21	1.18	1.42	1.53	1.48	5.04	6.05	5.51
0.90	β_0	1.13	1.24	1.18	1.13	1.15	1.15	1.45	1.38	1.34	1.60	1.38	1.34
	β_1	1.02	1.14	1.11	1.08	1.20	1.24	2.85	2.22	2.52	5.89	4.14	4.93
	β_2	1.06	1.12	1.10	1.13	1.43	1.31	5.00	4.43	4.78	24.20	22.23	23.37

		four-dimensional normal						four-dimensional Student-t					
		$\rho = 0.50$			$\rho = 0.90$			$\rho = 0.50$			$\rho = 0.90$		
τ	parm	$N = 50$	$N = 100$	$N = 200$	$N = 50$	$N = 100$	$N = 200$	$N = 50$	$N = 100$	$N = 200$	$N = 50$	$N = 100$	$N = 200$
0.25	β_{10}	1.18	1.15	1.06	1.29	1.18	1.08	1.05	0.84	0.57	1.05	0.83	0.57
	β_{11}	1.08	1.08	1.03	1.25	1.16	1.09	1.71	1.72	1.55	3.92	3.41	2.61
	β_{12}	1.12	1.08	1.05	1.25	1.18	1.09	2.09	2.22	2.24	9.01	9.37	9.11
	β_{20}	1.16	1.15	1.06	1.26	1.19	1.07	0.98	0.92	0.59	0.99	0.89	0.56
	β_{21}	1.10	1.07	1.03	1.22	1.14	1.07	1.89	1.64	1.58	3.41	3.19	2.48
	β_{22}	1.09	1.09	1.05	1.23	1.18	1.09	2.09	2.28	2.16	8.08	9.26	9.58
0.50	β_{10}	1.26	1.21	1.08	1.31	1.25	1.12	1.10	1.22	1.24	1.10	1.22	1.24
	β_{11}	1.10	1.10	1.07	1.24	1.21	1.13	1.43	1.61	1.53	3.35	3.78	3.22
	β_{12}	1.08	1.11	1.07	1.25	1.19	1.10	1.48	1.63	1.65	5.46	6.37	6.05
	β_{20}	1.24	1.22	1.09	1.29	1.23	1.12	1.14	1.18	1.23	1.12	1.25	1.20
	β_{21}	1.13	1.08	1.09	1.25	1.18	1.14	1.50	1.49	1.50	3.21	3.58	3.22
	β_{22}	1.08	1.10	1.07	1.24	1.21	1.10	1.45	1.54	1.46	5.32	6.31	5.74
0.90	β_{10}	1.04	0.98	0.95	1.06	1.06	0.98	1.49	1.22	0.82	1.44	1.06	0.65
	β_{11}	1.03	1.01	0.99	1.05	1.05	0.99	3.69	2.48	2.72	6.19	3.57	3.21
	β_{12}	1.05	1.05	1.03	1.18	1.17	1.11	7.17	6.01	6.49	29.04	27.84	27.77
	β_{20}	1.04	0.99	0.95	1.10	1.05	0.97	1.09	1.01	0.72	1.07	0.90	0.58
	β_{21}	1.01	1.02	0.99	1.07	1.05	1.00	3.05	2.49	2.71	6.91	3.78	3.25
	β_{22}	1.04	1.05	1.03	1.19	1.20	1.12	5.92	6.33	6.10	27.11	28.55	29.07

Table F.8: LDP data. ML estimates using different values of ϵ.

		LDL						HbA1c					
τ	effect	$\epsilon=0$	$\epsilon=0.001$	$\epsilon=0.01$	$\epsilon=0.1$	$\epsilon=0.2$	$\epsilon=0.5$	$\epsilon=0$	$\epsilon=0.001$	$\epsilon=0.01$	$\epsilon=0.1$	$\epsilon=0.2$	$\epsilon=0.5$
0.25	intercept	84.200	84.168	84.052	83.065	82.230	80.453	6.274	6.273	6.272	6.252	6.233	6.190
	gender	4.885	4.885	4.923	4.814	4.724	4.606	0.069	0.069	0.066	0.054	0.049	0.043
	age	0.155	0.155	0.157	0.158	0.156	0.154	-0.006	-0.006	-0.006	-0.006	-0.006	-0.006
	insulin	-13.011	-13.017	-13.058	-12.910	-12.706	-12.334	0.196	0.197	0.200	0.213	0.220	0.229
	diab.dur.	-1.237	-1.285	-1.396	-1.662	-1.765	-1.915	0.228	0.228	0.228	0.223	0.219	0.215
	time	-9.439	-9.461	-9.535	-9.762	-9.887	-10.111	-0.142	-0.142	-0.143	-0.146	-0.146	-0.142
	BMI	-0.013	-0.013	-0.008	-0.004	-0.011	-0.027	0.014	0.014	0.014	0.014	0.014	0.014
0.50	intercept	103.018	102.808	102.805	102.900	103.031	103.328	6.734	6.735	6.738	6.756	6.768	6.789
	gender	5.163	5.221	5.208	5.206	5.201	5.171	0.028	0.031	0.032	0.032	0.031	0.028
	age	0.107	0.110	0.112	0.120	0.121	0.120	-0.009	-0.009	-0.009	-0.009	-0.008	-0.008
	insulin	-10.240	-10.089	-10.001	-9.710	-9.611	-9.499	0.309	0.302	0.299	0.293	0.293	0.294
	diab.dur.	-3.162	-3.422	-3.460	-3.465	-3.402	-3.280	0.250	0.252	0.254	0.258	0.258	0.256
	time	-9.186	-9.244	-9.289	-9.588	-9.775	-10.102	-0.216	-0.216	-0.217	-0.225	-0.231	-0.241
	BMI	0.207	0.203	0.202	0.186	0.172	0.148	0.016	0.016	0.016	0.016	0.016	0.016
0.90	intercept	146.400	146.454	147.049	151.395	154.616	160.777	8.213	8.214	8.237	8.404	8.527	8.766
	gender	3.931	3.930	3.905	3.755	3.681	3.562	-0.040	-0.040	-0.042	-0.053	-0.060	-0.070
	age	0.050	0.050	0.049	0.048	0.046	0.040	-0.011	-0.011	-0.011	-0.012	-0.012	-0.012
	insulin	-7.394	-7.407	-7.446	-7.636	-7.740	-7.859	0.327	0.327	0.326	0.324	0.323	0.322
	diab.dur.	-3.041	-3.036	-3.019	-2.953	-2.912	-2.828	0.301	0.301	0.301	0.302	0.303	0.304
	time	-12.253	-12.209	-12.280	-12.767	-13.101	-13.692	-0.549	-0.549	-0.554	-0.597	-0.627	-0.685
	BMI	0.633	0.636	0.633	0.616	0.605	0.587	0.034	0.034	0.034	0.034	0.035	0.035

Table F.9: LDP data. ML standard error using different values of ϵ.

		LDL					HbA1c				
τ	effect	$\epsilon=0.001$	$\epsilon=0.01$	$\epsilon=0.1$	$\epsilon=0.2$	$\epsilon=0.5$	$\epsilon=0.001$	$\epsilon=0.01$	$\epsilon=0.1$	$\epsilon=0.2$	$\epsilon=0.5$
0.25	intercept	1.436	1.380	1.274	1.251	1.220	0.028	0.028	0.028	0.028	0.028
	gender	1.757	1.717	1.631	1.602	1.554	0.036	0.036	0.034	0.034	0.033
	age	0.090	0.089	0.086	0.084	0.080	0.002	0.002	0.002	0.002	0.002
	insulin	2.188	2.113	2.031	2.025	2.004	0.058	0.056	0.053	0.052	0.051
	diab.dur.	1.479	1.435	1.380	1.369	1.345	0.033	0.033	0.033	0.032	0.031
	time	0.819	0.772	0.749	0.752	0.765	0.018	0.017	0.017	0.017	0.017
	BMI	0.186	0.177	0.167	0.163	0.157	0.003	0.003	0.003	0.003	0.003
0.50	intercept	1.295	1.284	1.241	1.218	1.187	0.036	0.035	0.033	0.033	0.032
	gender	1.728	1.702	1.608	1.568	1.519	0.040	0.039	0.037	0.036	0.036
	age	0.082	0.080	0.074	0.072	0.070	0.002	0.002	0.002	0.002	0.002
	insulin	2.608	2.571	2.415	2.324	2.207	0.066	0.065	0.061	0.059	0.058
	diab.dur.	1.469	1.440	1.342	1.305	1.265	0.035	0.033	0.031	0.031	0.031
	time	0.839	0.822	0.769	0.755	0.746	0.018	0.018	0.018	0.018	0.018
	BMI	0.163	0.159	0.147	0.143	0.138	0.004	0.003	0.003	0.003	0.003
0.90	intercept	2.104	2.092	2.094	2.125	2.223	0.077	0.077	0.080	0.083	0.090
	gender	2.575	2.576	2.623	2.639	2.668	0.077	0.077	0.079	0.081	0.083
	age	0.118	0.119	0.122	0.123	0.124	0.003	0.003	0.003	0.004	0.004
	insulin	2.797	2.801	2.815	2.808	2.842	0.083	0.083	0.085	0.087	0.090
	diab.dur.	2.160	2.161	2.178	2.179	2.193	0.066	0.066	0.066	0.066	0.067
	time	1.413	1.421	1.497	1.571	1.734	0.056	0.057	0.061	0.065	0.074
	BMI	0.243	0.243	0.244	0.244	0.246	0.008	0.008	0.008	0.008	0.008

Table F.10: LDP data. Gradient at the ML estimates using different values of ϵ.

		LDL					HbA1c				
τ	effect	$\epsilon = 0.001$	$\epsilon = 0.01$	$\epsilon = 0.1$	$\epsilon = 0.2$	$\epsilon = 0.5$	$\epsilon = 0.001$	$\epsilon = 0.01$	$\epsilon = 0.1$	$\epsilon = 0.2$	$\epsilon = 0.5$
0.25	intercept	-0.0013	-0.0003	0.0007	0.0007	0.0018	0.0305	0.0304	-0.0096	0.0141	-0.0885
	gender	0.0003	0.0001	-0.0005	-0.0007	-0.0012	0.0319	0.0194	-0.0034	0.0178	-0.0468
	age	-0.0168	0.0010	-0.0038	0.0001	-0.0021	-0.1097	0.1642	0.0525	-0.2149	-0.0105
	insulin	-0.0005	-0.0001	0.0004	0.0001	0.0002	0.0224	0.0026	-0.0127	-0.0008	-0.0105
	diab.dur.	-0.0006	-0.0002	-0.0001	-0.0001	0.0000	-0.0143	0.0047	0.0090	-0.0021	0.0259
	time	-0.0001	-0.0005	-0.0002	0.0006	-0.0004	0.0743	0.0249	-0.0048	0.0068	-0.0751
	BMI	0.0024	-0.0009	0.0057	-0.0026	-0.0120	-0.2483	-0.1016	-0.0800	-0.1267	-0.1411
0.50	intercept	0.0023	0.0026	0.0093	0.0043	0.0029	0.0584	0.0390	0.0449	0.0809	0.0301
	gender	-0.0020	-0.0012	-0.0011	-0.0050	-0.0067	-0.0086	-0.0132	0.0272	0.0448	0.0194
	age	-0.0000	0.0045	-0.0062	0.0086	0.0018	-1.5255	-1.7590	0.4453	-0.1168	0.0500
	insulin	-0.0002	-0.0002	-0.0055	-0.0095	-0.0079	0.0241	0.0075	0.0210	0.0237	0.0128
	diab.dur.	0.0002	-0.0001	0.0021	0.0034	0.0026	0.0220	-0.0283	0.0095	0.0064	0.0061
	time	0.0007	0.0000	0.0027	-0.0018	-0.0004	-0.0316	-0.0315	0.0385	0.0714	-0.0177
	BMI	0.0166	0.0159	0.0013	-0.0045	-0.0060	0.0199	0.0794	-0.1002	0.0414	-0.0170
0.90	intercept	0.0015	0.0015	0.0023	0.0052	-0.0019	0.0236	0.0552	-0.0212	-0.1627	-0.1234
	gender	0.0007	-0.0002	-0.0007	0.0006	-0.0041	-0.0075	0.0005	-0.0284	-0.1229	-0.0867
	age	-0.0258	0.0045	-0.0357	-0.0354	0.0100	0.6859	-0.9292	0.3974	0.3108	-0.8346
	insulin	0.0010	0.0009	-0.0026	-0.0021	0.0014	0.0207	0.0235	-0.0196	-0.0192	-0.0330
	diab.dur.	-0.0006	-0.0000	0.0018	0.0004	-0.0005	-0.0073	-0.0235	0.0463	0.0295	-0.0228
	time	0.0010	0.0017	0.0005	0.0008	-0.0013	0.2048	0.2911	0.0220	-0.2205	-0.2578
	BMI	0.0004	0.0094	0.0053	0.0140	0.0031	-0.0957	-0.0091	0.0690	0.1256	0.0451

Table F.11: Simulated data. MLE using different values of ϵ.

		Estimates						Standard error					Gradient				
τ	effect	$\epsilon = 0$	$\epsilon = 0.001$	$\epsilon = 0.01$	$\epsilon = 0.1$	$\epsilon = 0.2$	$\epsilon = 0.5$	$\epsilon = 0.001$	$\epsilon = 0.01$	$\epsilon = 0.1$	$\epsilon = 0.2$	$\epsilon = 0.5$	$\epsilon = 0.001$	$\epsilon = 0.01$	$\epsilon = 0.1$	$\epsilon = 0.2$	$\epsilon = 0.5$
0.25	β_{10}	3.45	3.45	3.44	3.40	3.37	3.29	0.009	0.009	0.008	0.008	0.008	-0.0028	-0.0036	-0.0043	0.0007	-0.0128
	β_{11}	1.27	1.27	1.27	1.24	1.21	1.16	0.036	0.036	0.031	0.030	0.029	-0.0005	-0.0024	-0.0020	-0.0017	0.0009
	β_{12}	1.05	1.05	1.05	1.06	1.07	1.08	0.078	0.077	0.071	0.069	0.066	0.0000	0.0002	-0.0004	0.0007	0.0000
	β_{20}	4.43	4.43	4.42	4.39	4.36	4.29	0.010	0.010	0.009	0.009	0.009	0.0009	0.0017	0.0027	-0.0069	-0.0015
	β_{21}	-3.96	-3.96	-3.96	-3.98	-4.00	-4.05	0.026	0.026	0.026	0.027	0.028	0.0003	-0.0011	0.0006	0.0002	-0.0020
	β_{22}	1.67	1.66	1.66	1.65	1.63	1.61	0.067	0.067	0.066	0.066	0.066	-0.0001	0.0000	0.0011	-0.0015	0.0005
0.50	β_{10}	4.08	4.08	4.08	4.08	4.08	4.07	0.007	0.006	0.006	0.005	0.005	-0.0017	-0.0005	0.0009	0.0001	0.0001
	β_{11}	1.82	1.82	1.82	1.82	1.82	1.83	0.032	0.032	0.029	0.027	0.026	0.0009	-0.0004	-0.0005	-0.0001	0.0009
	β_{12}	1.20	1.20	1.20	1.21	1.21	1.21	0.097	0.094	0.081	0.074	0.067	0.0002	-0.0002	-0.0006	-0.0000	0.0001
	β_{20}	5.05	5.05	5.05	5.05	5.05	5.06	0.005	0.005	0.005	0.005	0.005	0.0036	0.0002	-0.0000	0.0000	0.0004
	β_{21}	-3.37	-3.37	-3.37	-3.36	-3.35	-3.34	0.031	0.031	0.028	0.027	0.026	-0.0001	0.0001	-0.0000	-0.0002	-0.0008
	β_{22}	1.74	1.74	1.74	1.73	1.72	1.70	0.096	0.094	0.085	0.080	0.072	-0.0002	0.0002	0.0002	-0.0000	0.0001
0.9	β_{10}	5.42	5.42	5.41	5.45	5.49	5.59	0.013	0.010	0.010	0.011	0.012	0.0002	-0.0000	0.0118	0.0051	0.0380
	β_{11}	2.99	2.99	3.00	3.06	3.10	3.18	0.070	0.070	0.071	0.071	0.070	0.0001	-0.0000	-0.0054	-0.0015	0.0162
	β_{12}	1.25	1.25	1.26	1.27	1.28	1.28	0.077	0.072	0.068	0.067	0.064	0.0000	-0.0000	-0.0038	-0.0010	0.0151
	β_{20}	6.39	6.39	6.40	6.46	6.51	6.61	0.009	0.009	0.012	0.013	0.015	-0.0001	0.0000	-0.0050	-0.0075	-0.0241
	β_{21}	-1.88	-1.88	-1.87	-1.81	-1.76	-1.66	0.081	0.081	0.081	0.081	0.080	0.0000	-0.0000	-0.0005	-0.0024	-0.0006
	β_{22}	1.39	1.39	1.41	1.45	1.46	1.47	0.135	0.108	0.086	0.083	0.078	0.0000	0.0000	-0.0006	-0.0013	-0.0013