

# README Simulation Study

Generalized Expectile Regression with Flexible Response Function

*Elmar Spiegel & Thomas Kneib & Petra von Gablenz & Fabian Otto-Sobotka*

*06 12 2020*

## Contents

<b>1</b>	<b>General</b>	<b>1</b>
<b>2</b>	<b>Sim_small_n</b>	<b>1</b>
<b>3</b>	<b>Sim_large_n</b>	<b>2</b>
<b>4</b>	<b>SessionInfo</b>	<b>3</b>
4.1	Sim_small_n . . . . .	3
4.2	Figure_1_2 . . . . .	4
4.3	Sim_large_n . . . . .	4
4.4	Figure_3 . . . . .	5
4.5	Plot_Sim_large_n_Covariates_more.R . . . . .	5

## 1 General

- This folder contains the code of the simulation study of the paper.
- In the paper we discuss two types of simulation studies, one for showing the improvement on the goodness of fit (small n), the other for proving identifiability (large n).
- The structure of the code mimics this, with the directories *Sim\_large\_n* and *Sim\_small\_n*.
- The intermediate results of the simulation study are too big to be shared generally. If you need them, please contact Elmar Spiegel.
- Software used: R version 3.6.3
- Main package FlexER\_0.04-06

## 2 Sim\_small\_n

- Simulation study for small sample size but 100 replications.
- Implemented to show the goodness of generalized expectile regression with flexible response function (FlexER) compared to classical Expectile Regression and generalized expectile regression with fixed response function (Fixed\_ER).
- .R files: (Only *Sim\_small\_n\_...R* and *Figure\_1\_2.R* are executed. Other files are sources called in *Sim\_small\_n\_...R*)
  - *Figure\_1\_2.R*: File to generate *Figure1.eps* and *Figure2.eps*
  - *Si\_Build\_Data\_Simulation\_5.R*: File used to build the data sets in the simulation study.

- `Si_Load_functions_small.R`: File used to load the other source files and the libraries. Defined to easier maintain of the code versions.
  - `Si_Main_simulation_function_2.R`: File which really runs the simulation, dependent on the parameters specified in each `Sim_small_n_...R` file.
  - `Si_Plot_all_results_3.R`: Aggregates the different plot functions (`Si_Plot_...R`) and executes the plotting
  - `Si_Plot_covariates_2.R`: File containing the functions to plot the estimated covariate effects.
  - `Si_Plot_link_3.R`: File containing the functions to plot the estimated response functions.
  - `Si_Plot_MSE_3.R`: File containing the functions to plot the PMWSE.
  - `Si_Set_options_small.R`: File to determine global options of the simulation study.
  - `Sim_small_n_...R`: File setting the parameters of the simulation study. Loading the libraries, and source files (see above). Starts the simulation study for one parameter setting.
  - `Sim_small_n_Fast_run.R`: File to try reproducibility (Duration approx. 45min)
- Output:
    - `Figure1.eps` & `Figure2.eps` Figures of the Paper
    - Other PDFs: Used in the Supplementary materials.
    - `.RData` files: Simulation study output
    - `.Rout` files: Console-Text written to file during the simulation study.
  - Ordering of execution:
    1. Run all `Sim_small_n_...R` (Duration up to 1 day for each parameter setting, resp. `Sim_small_n_...R` file)
    2. Run `Figure_1_2.R`
  - For checking the reproducibility try `Sim_small_n_Fast_run.R`, which is the same as `Sim_small_n_G_10_log.R`, but only with 6 repetitions. (Duration approx. 45min)

### 3 `Sim_large_n`

- Simulation Study to show the identifiability of the approach.
- `.R` files: (Only `Sim_large_n_...R` and `Figure_3.R` and `Plot_Sim_large_n_Covariates_more.R` are executed. Other files are sources called in `Sim_large_n_...R`)
  - `Figure_3.R`: Code to build `Figure3.eps`
  - `Plot_Sim_large_n_Covariates_more.R`: File to build the plots of the supplementary materials, where the estimated covariate effects and estimated response functions are shown.
  - `Si_Build_Data_Simulation_5.R`: File containing the functions, which generate the data sets. Same as for `small_n`
  - `Si_Compare_models.R`: File to define functions to build the difference between predicted expectiles based on models and marginal expectiles based on grid. Definition of plot functions.
  - `Si_Load_functions_large.R`: File to load the libraries and other source files. Defined to easier maintain the code with different versions.
  - `Si_Set_options_large.R`: File to determine global options of the simulation study.

- `Sim_large_n_...R`: Files setting the parameters of the simulation study. Loading the libraries, and source files (see above). Starts the simulation study for parameter setting. Also calculating the marginal expectiles and comparing them with the estimated expectiles.
- `Sim_large_n_Fast_run.R`: File to try reproducibility.
- Output:
  - `Figure3.eps`: Figure of the Paper
  - Other PDFs: used in the supplementary materials.
  - `.RData` files: Simulation study output
  - `.Rout` files: Console-Text written to file during the simulation study.
- Ordering of execution:
  1. Run all `Sim_large_n_...R` (Duration up to 12h for each parameter setting, resp. `Sim_large_n_...R` file)
  2. Run `Figure_3.R`
  3. Run `Plot_Sim_large_n_Covariates_more.R`
- For checking the reproducibility try `Sim_large_n_Fast_run.R`, which is the same as `Sim_large_n_G_10_log.R`, but with lower sample size. (Duration approx. 20min)

## 4 SessionInfo

### 4.1 Sim\_small\_n

```
> sessionInfo()
R version 3.6.3 (2020-02-29)
Platform: x86_64-pc-linux-gnu (64-bit)
Running under: Ubuntu 20.04.1 LTS

Matrix products: default
BLAS: /usr/lib/x86_64-linux-gnu/blas/libblas.so.3.9.0
LAPACK: /usr/lib/x86_64-linux-gnu/lapack/liblapack.so.3.9.0

locale:
 [1] LC_CTYPE=en_US.UTF-8      LC_NUMERIC=C
 [3] LC_TIME=en_US.UTF-8      LC_COLLATE=en_US.UTF-8
 [5] LC_MONETARY=en_US.UTF-8  LC_MESSAGES=en_US.UTF-8
 [7] LC_PAPER=en_US.UTF-8     LC_NAME=C
 [9] LC_ADDRESS=C             LC_TELEPHONE=C
[11] LC_MEASUREMENT=en_US.UTF-8 LC_IDENTIFICATION=C

attached base packages:
[1] splines      stats      graphics  grDevices  utils      datasets  methods
[8] base

other attached packages:
[1] FlexER_0.04-06 foreign_0.8-75 vioplot_0.3.5 zoo_1.8-8      sm_2.2-5.6
[6] mvtnorm_1.0-10 mgcv_1.8-31    nlme_3.1-148 Matrix_1.2-18 MASS_7.3-51.6

loaded via a namespace (and not attached):
```

```
[1] Rcpp_1.0.3          lattice_0.20-41    grid_3.6.3
[4] RcppEigen_0.3.3.5.0 compiler_3.6.3     colorspace_1.4-1
[7] tcltk_3.6.3
```

## 4.2 Figure\_1\_2

```
> sessionInfo()
R version 3.6.3 (2020-02-29)
Platform: x86_64-pc-linux-gnu (64-bit)
Running under: Ubuntu 20.04.1 LTS

Matrix products: default
BLAS: /usr/lib/x86_64-linux-gnu/blas/libblas.so.3.9.0
LAPACK: /usr/lib/x86_64-linux-gnu/lapack/liblapack.so.3.9.0

locale:
 [1] LC_CTYPE=en_US.UTF-8      LC_NUMERIC=C
 [3] LC_TIME=en_US.UTF-8      LC_COLLATE=en_US.UTF-8
 [5] LC_MONETARY=en_US.UTF-8  LC_MESSAGES=en_US.UTF-8
 [7] LC_PAPER=en_US.UTF-8     LC_NAME=C
 [9] LC_ADDRESS=C             LC_TELEPHONE=C
[11] LC_MEASUREMENT=en_US.UTF-8 LC_IDENTIFICATION=C

attached base packages:
[1] splines  stats    graphics  grDevices  utils      datasets  methods
[8] base

other attached packages:
[1] FlexER_0.04-06 foreign_0.8-75 mvtnorm_1.0-10 mgcv_1.8-31  nlme_3.1-148
[6] Matrix_1.2-18 MASS_7.3-51.6

loaded via a namespace (and not attached):
[1] colorspace_1.4-1  compiler_3.6.3    RcppEigen_0.3.3.5.0
[4] Rcpp_1.0.3        grid_3.6.3        lattice_0.20-41
```

## 4.3 Sim\_large\_n

```
> sessionInfo()
R version 3.6.3 (2020-02-29)
Platform: x86_64-pc-linux-gnu (64-bit)
Running under: Ubuntu 20.04.1 LTS

Matrix products: default
BLAS: /usr/lib/x86_64-linux-gnu/blas/libblas.so.3.9.0
LAPACK: /usr/lib/x86_64-linux-gnu/lapack/liblapack.so.3.9.0

locale:
 [1] LC_CTYPE=en_US.UTF-8      LC_NUMERIC=C
 [3] LC_TIME=en_US.UTF-8      LC_COLLATE=en_US.UTF-8
 [5] LC_MONETARY=en_US.UTF-8  LC_MESSAGES=en_US.UTF-8
```

```

[7] LC_PAPER=en_US.UTF-8      LC_NAME=C
[9] LC_ADDRESS=C              LC_TELEPHONE=C
[11] LC_MEASUREMENT=en_US.UTF-8 LC_IDENTIFICATION=C

attached base packages:
[1] stats      graphics  grDevices  utils      datasets  methods    base

other attached packages:
[1] FlexER_0.04-06 mgcv_1.8-31    nlme_3.1-148  Matrix_1.2-18 MASS_7.3-51.6
[6] foreign_0.8-75

loaded via a namespace (and not attached):
[1] colorspace_1.4-1  compiler_3.6.3    RcppEigen_0.3.3.5.0
[4] Rcpp_1.0.3        splines_3.6.3     grid_3.6.3
[7] lattice_0.20-41

```

#### 4.4 Figure\_3

```

> sessionInfo()
R version 3.6.3 (2020-02-29)
Platform: x86_64-pc-linux-gnu (64-bit)
Running under: Ubuntu 20.04.1 LTS

Matrix products: default
BLAS: /usr/lib/x86_64-linux-gnu/blas/libblas.so.3.9.0
LAPACK: /usr/lib/x86_64-linux-gnu/lapack/liblapack.so.3.9.0

locale:
 [1] LC_CTYPE=en_US.UTF-8      LC_NUMERIC=C
 [3] LC_TIME=en_US.UTF-8      LC_COLLATE=en_US.UTF-8
 [5] LC_MONETARY=en_US.UTF-8  LC_MESSAGES=en_US.UTF-8
 [7] LC_PAPER=en_US.UTF-8     LC_NAME=C
 [9] LC_ADDRESS=C             LC_TELEPHONE=C
[11] LC_MEASUREMENT=en_US.UTF-8 LC_IDENTIFICATION=C

attached base packages:
[1] stats      graphics  grDevices  utils      datasets  methods    base

other attached packages:
[1] FlexER_0.04-06 mgcv_1.8-31    nlme_3.1-148  Matrix_1.2-18 MASS_7.3-51.6
[6] foreign_0.8-75

loaded via a namespace (and not attached):
[1] colorspace_1.4-1  compiler_3.6.3    RcppEigen_0.3.3.5.0
[4] Rcpp_1.0.3        splines_3.6.3     grid_3.6.3
[7] lattice_0.20-41

```

#### 4.5 Plot\_Sim\_large\_n\_Covariates\_more.R

```
> sessionInfo()
R version 3.6.3 (2020-02-29)
Platform: x86_64-pc-linux-gnu (64-bit)
Running under: Ubuntu 20.04.1 LTS

Matrix products: default
BLAS: /usr/lib/x86_64-linux-gnu/blas/libblas.so.3.9.0
LAPACK: /usr/lib/x86_64-linux-gnu/lapack/liblapack.so.3.9.0

locale:
 [1] LC_CTYPE=en_US.UTF-8      LC_NUMERIC=C
 [3] LC_TIME=en_US.UTF-8      LC_COLLATE=en_US.UTF-8
 [5] LC_MONETARY=en_US.UTF-8  LC_MESSAGES=en_US.UTF-8
 [7] LC_PAPER=en_US.UTF-8     LC_NAME=C
 [9] LC_ADDRESS=C             LC_TELEPHONE=C
[11] LC_MEASUREMENT=en_US.UTF-8 LC_IDENTIFICATION=C

attached base packages:
[1] stats      graphics  grDevices  utils      datasets  methods   base

other attached packages:
[1] colorspace_1.4-1 FlexER_0.04-06 mgcv_1.8-31      nlme_3.1-148
[5] Matrix_1.2-18   MASS_7.3-51.6   foreign_0.8-75

loaded via a namespace (and not attached):
[1] compiler_3.6.3      RcppEigen_0.3.3.5.0 Rcpp_1.0.3
[4] splines_3.6.3       grid_3.6.3         lattice_0.20-41
```