M-Estimation in INARCH Models with a special focus on small means

Hanan El-Saied, Roland Fried

Department of Statistics

TU Dortmund

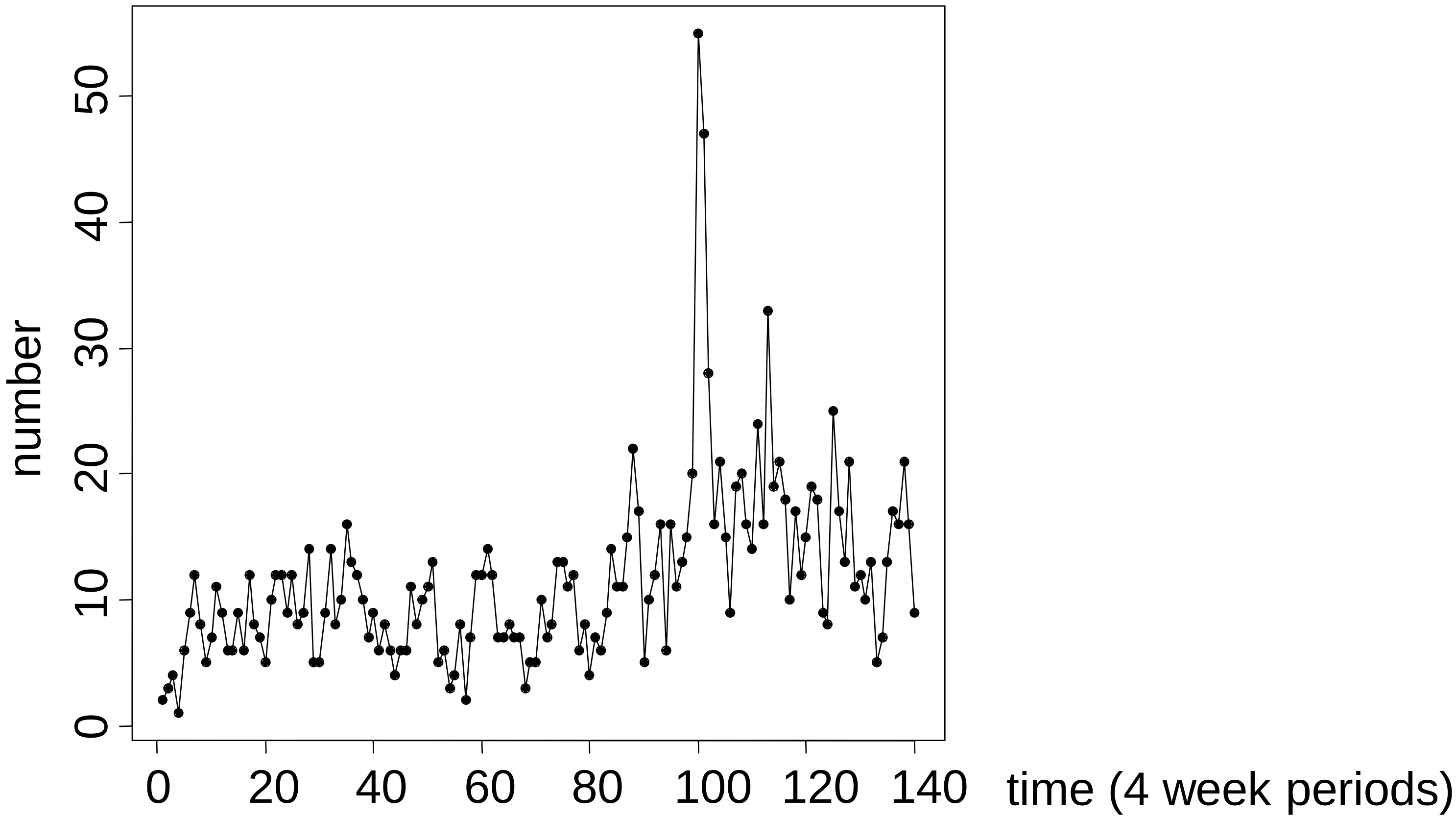
Germany

SFB 823

Contents

- Motivation: Outliers in IN(G)ARCH models
- M-estimation for i.i.d. Poisson data
- M-estimation for INARCH-model
- Bias correction
- Outlook

Motivation: Number of Campylobacterosis Infections



INGARCH-model: $Y_t \big| (Y_s, s<t) \sim \mathrm{Poi}(\mu_t)$ *Ferland, Latour, Oraichi (2006)*

$$\mu_t = \beta_0 + \alpha_1 \mu_{t-13} + \beta_1 Y_{t-1}$$

Level shift at time 84, outlier pattern at time 100 *Fokianos, F. (2010)*

• M-estimation of location μ for i.i.d. data

Minimize $$\sum_{t=1}^{n} \rho\left(\frac{y_t - \mu}{\sigma}\right) \quad \Leftrightarrow \quad \sum_{t=1}^{n} \psi\left(\frac{y_t - \mu}{\sigma}\right) = 0$$

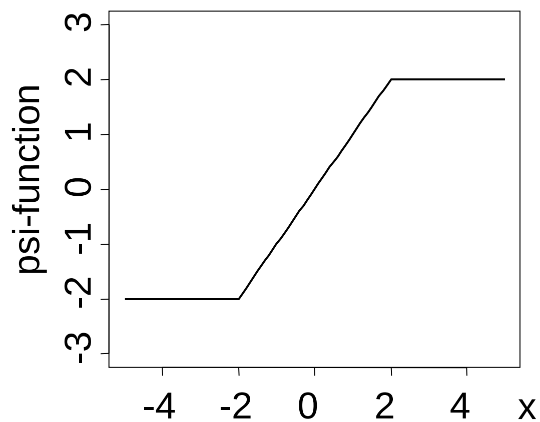
e.g. $\rho = -\log f$ gives ML-estimator

● M-estimation of location μ for i.i.d. data

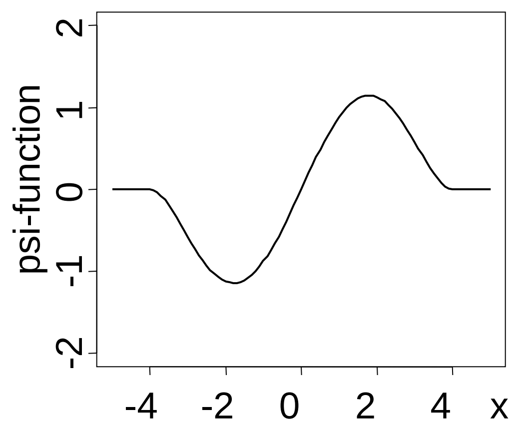
Minimize $\sum_{t=1}^{n} \rho\left(\frac{y_t - \mu}{\sigma}\right) \quad \Leftrightarrow \quad \sum_{t=1}^{n} \psi\left(\frac{y_t - \mu}{\sigma}\right) = 0$

e.g. $\rho = -\log f$ gives ML-estimator

Huber ψ-function



Tukey ψ-function



$$\psi_k(x) = \begin{cases} x, & |x| < k \\ k \cdot \mathrm{sign}(x), & |x| \geq k \end{cases}$$

$$\psi_k(x) = x\left(1 - \left(\frac{x}{k}\right)^2\right)^2 \cdot I(|x| < k)$$

M-estimation for i.i.d. Poisson data

Modified Huber ψ-function with bias correction

$$\psi_{k,a}(y,\mu) = \begin{cases} \dfrac{y-\mu-a}{\sqrt{\mu}}, & |y-\mu-a| < k \\ k \cdot \mathrm{sign}(y-\mu-a), & |y-\mu-a| \geq k \end{cases}$$

with $a=a(\mu)$ such that $E_\mu\left(\psi_{k,a}(Y_1,\mu)\right) = 0$ *Simpson et al. (1987)*

M-estimation for i.i.d. Poisson data

Modified Huber ψ-function with bias correction

$$\psi_{k,a}(y,\mu) = \begin{cases} \dfrac{y-\mu-a}{\sqrt{\mu}}, & |y-\mu-a| < k \\ k \cdot \mathrm{sign}(y-\mu-a), & |y-\mu-a| \geq k \end{cases}$$

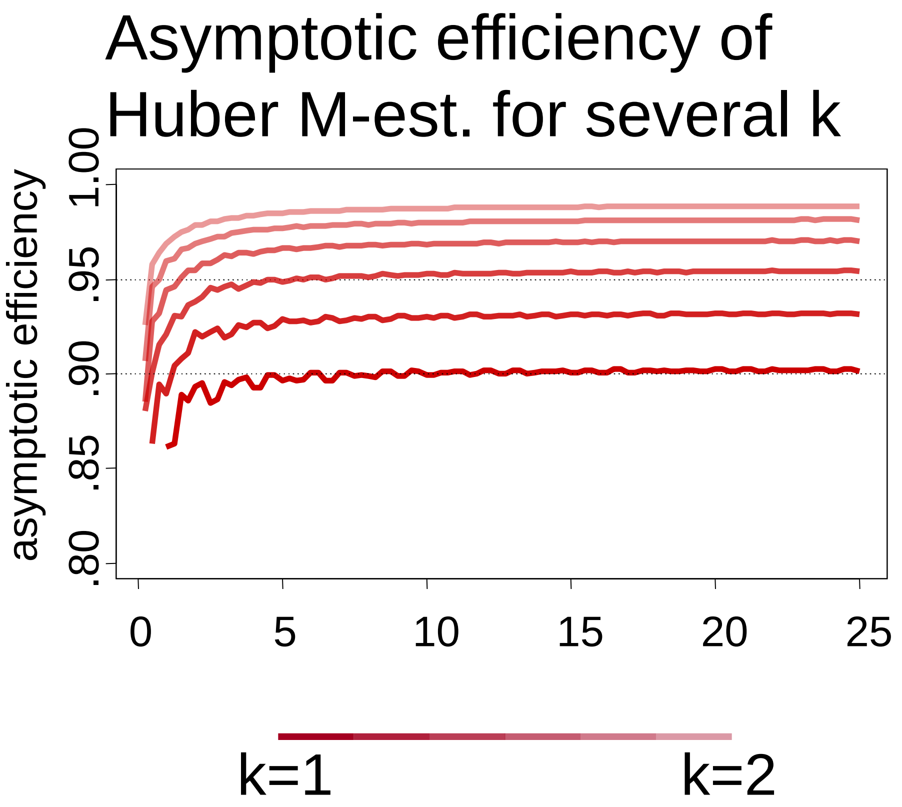
with $a=a(\mu)$ such that $E_\mu\left(\psi_{k,a}(Y_1,\mu)\right)=0$ *Simpson et al. (1987)*

Modified Tukey ψ-function with bias correction

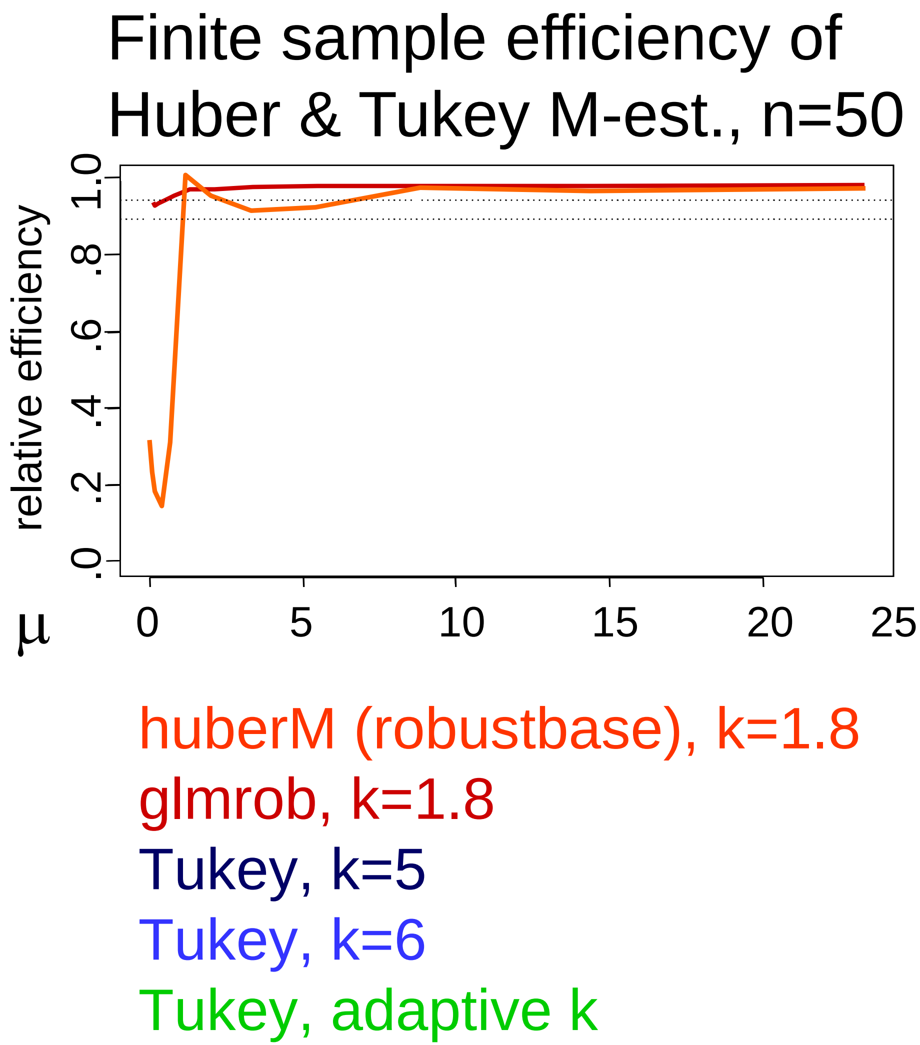
$$\psi_{k,a}(y,\mu) = \left(\frac{y-\mu}{\sqrt{\mu}} - a\right) \cdot \left(k^2 - \left(\frac{y-\mu}{\sqrt{\mu}} - a\right)^2\right)^2 \cdot I\left(\left|\frac{y-\mu}{\sqrt{\mu}} - a\right| < k\right)$$

Initialization by sample median or by estimating $P_\mu(Y=0)$

Efficiencies: asymptotic and sample size n=50

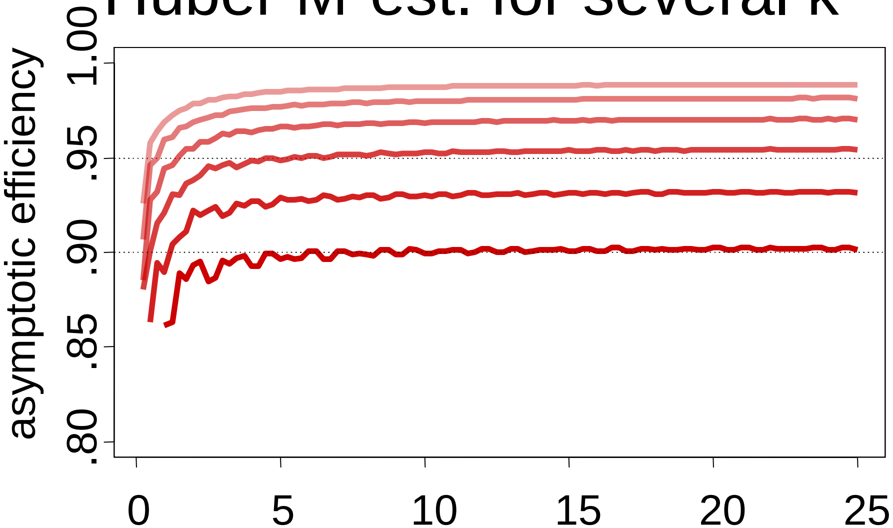


Cadigan & Chen (2001)



Efficiencies: asymptotic and sample size n=50

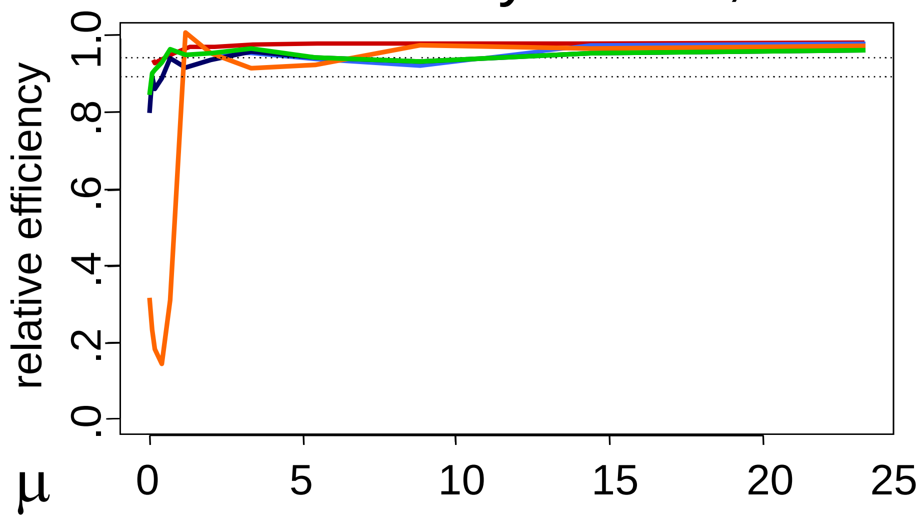
Asymptotic efficiency of
Huber M-est. for several k



k=1 k=2

Cadigan & Chen (2001)

Finite sample efficiency of
Huber & Tukey M-est., n=50



huberM (robustbase), k=1.8

glmrob, k=1.8

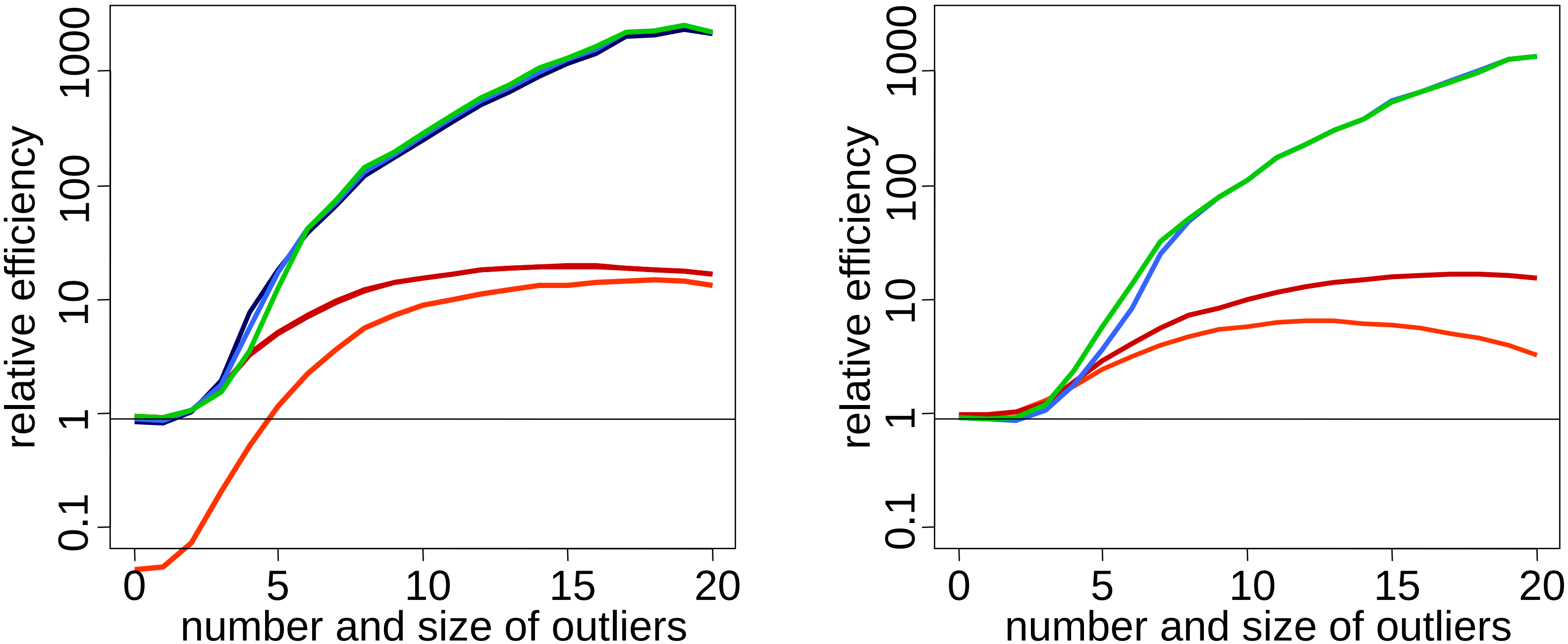
Tukey, k=5

Tukey, k=6

Tukey, adaptive k

Robustness for μ=0.5 and μ=5

Efficiency relatively to sample mean in case of increasing number of outliers of increasing size, n=50, log-scale



huberM (robustbase), k=1.8 glmrob, k=1.8

Tukey, k=5 Tukey, k=6 Tukey, adaptive k

●Conditional likelihod estimation for INARCH

INARCH-model: $Y_t \big| (Y_s, s<t) \sim \mathrm{Poi}(\mu_t), \quad \mu_t = \beta_0 + \beta_1 Y_{t-1} + \cdots + \beta_p Y_{t-p}$

Conditioning on first p observations $y_1, \ldots, y_p$:

$$\sum_{t=p+1}^{n} \frac{y_t - \mu_t}{\sqrt{\mu_t}} \frac{1}{\sqrt{\mu_t}} \begin{pmatrix} 1 \\ y_{t-1} \\ \vdots \\ y_{t-p} \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ \vdots \\ 0 \end{pmatrix}$$

•Conditional likelihod estimation for INARCH

INARCH-model: $Y_t \big| (Y_s, s<t) \sim \mathrm{Poi}(\mu_t), \quad \mu_t = \beta_0 + \beta_1 Y_{t-1} + \cdots + \beta_p Y_{t-p}$

Conditioning on first p observations $y_1, \ldots, y_p$:

$$\sum_{t=p+1}^{n} \frac{y_t - \mu_t}{\sqrt{\mu_t}} \frac{1}{\sqrt{\mu_t}} \begin{pmatrix} 1 \\ y_{t-1} \\ \vdots \\ y_{t-p} \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ \vdots \\ 0 \end{pmatrix}$$

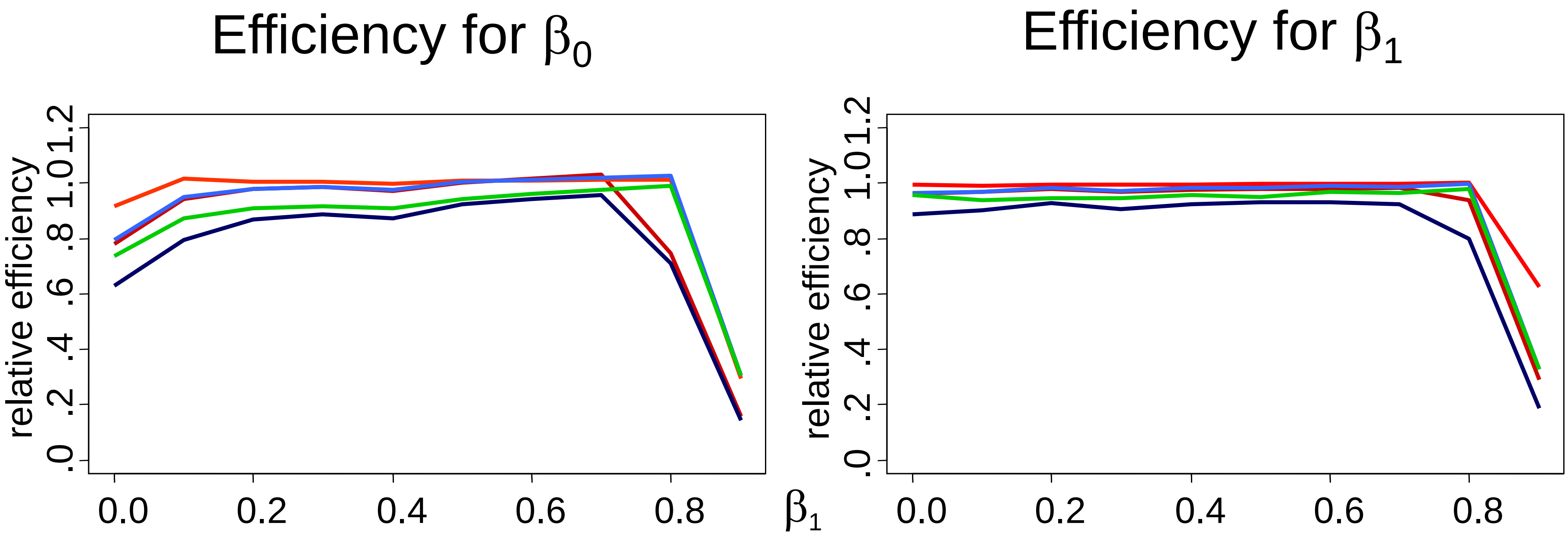
M-estimation:

$$\psi\left(\frac{y_t - \mu_t}{\sqrt{\mu_t}}\right)$$

μ, σ^2 marginal mean & variance

$$\begin{pmatrix} 1 \\ \mu + \sigma\psi\left(\frac{y_{t-1} - \mu}{\sqrt{\mu}}\right) \\ \vdots \\ \mu + \sigma\psi\left(\frac{y_{t-p} - \mu}{\sqrt{\mu}}\right) \end{pmatrix}$$

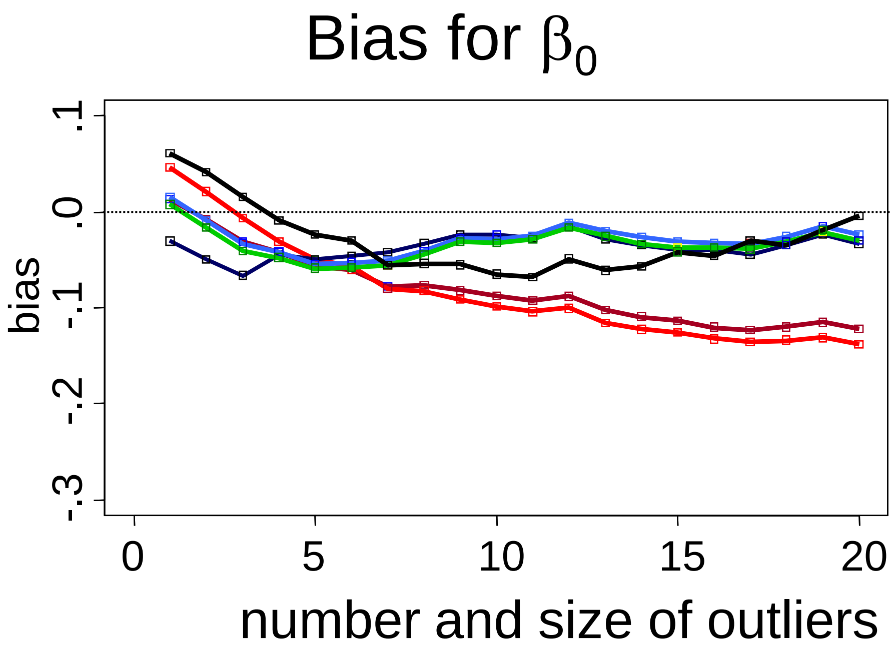
Efficiencies: INARCH(1), β_0=1,several β_1, n=100



Huber, k=1.8, Huber, k=2.5
Tukey, k=5, Tukey, k=7
Tukey, adaptive k

Robustness: INARCH(1) with β_0=1, β_1=.4

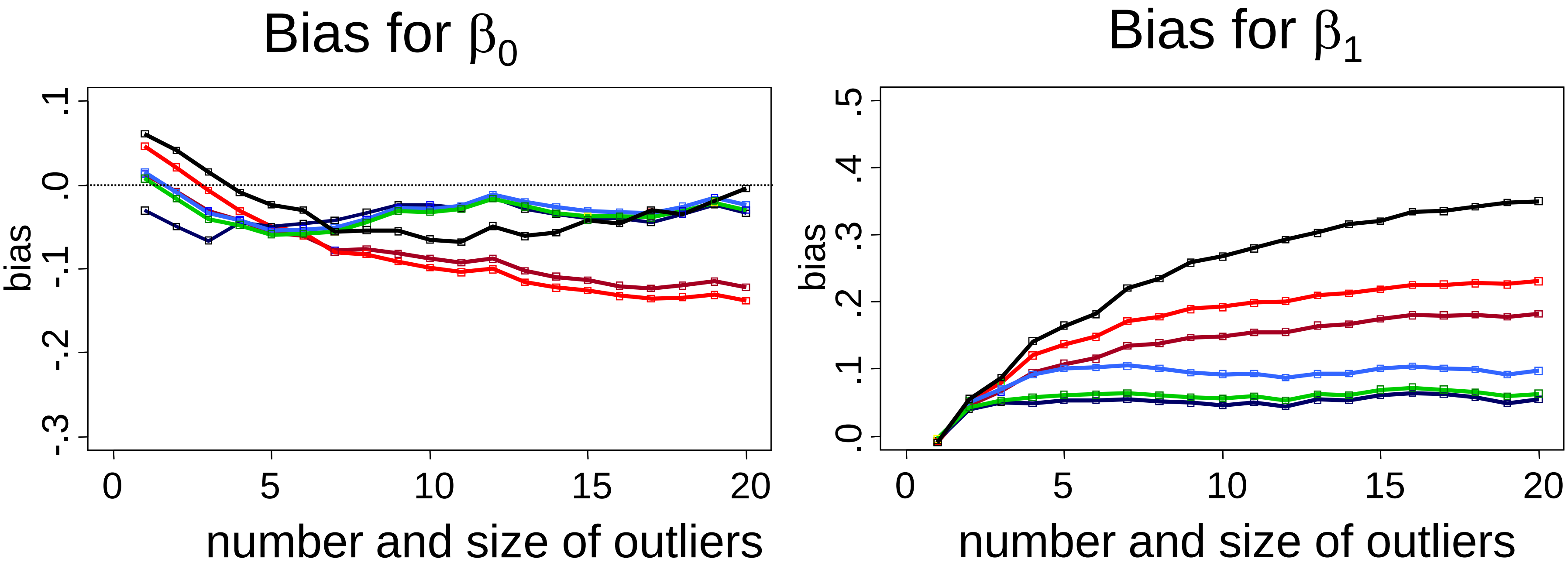
Increasing number k of outliers of size k at end of time series



Conditional ML
Huber, k=1.8, Huber, k=2.5
Tukey, k=5, Tukey, k=7
Tukey, adaptive k

Robustness: INARCH(1) with β_0=1, β_1=.4

Increasing number k of outliers of size k at end of time series



Conditional ML

Huber, k=1.8, Huber, k=2.5

Tukey, k=5, Tukey, k=7

Tukey, adaptive k

• Bias correction for INARCH(p) model

M-estimator with bias correction:

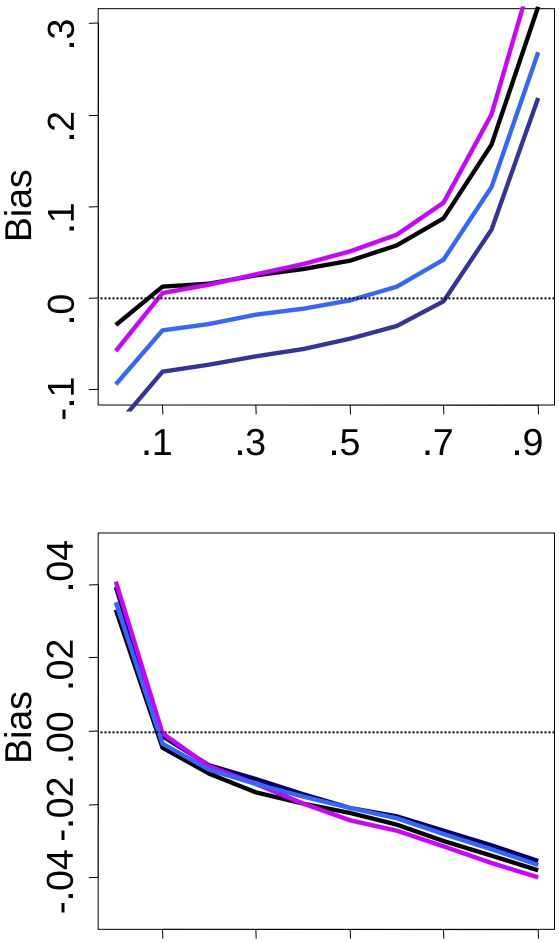
$$\sum_{t=p+1}^{n} \left(\psi\left(\frac{y_t - \mu_t}{\sqrt{\mu_t}}\right) \frac{1}{\sqrt{\mu_t}} \begin{pmatrix} 1 \\ \mu + \sigma\psi\left(\frac{y_{t-1} - \mu}{\sqrt{\mu}}\right) \\ \vdots \\ \mu + \sigma\psi\left(\frac{y_{t-p} - \mu}{\sqrt{\mu}}\right) \end{pmatrix} - \begin{pmatrix} a_0 \\ \vdots \\ \vdots \\ a_p \end{pmatrix} \right) = \begin{pmatrix} 0 \\ \vdots \\ \vdots \\ 0 \end{pmatrix}$$

with $a_o, \ldots, a_p$ depending on $\beta_0, \ldots, \beta_p$ such that expectation of left hand side equals 0.

Bias for INARCH(1) in dependence on β_1

n=100

β_0



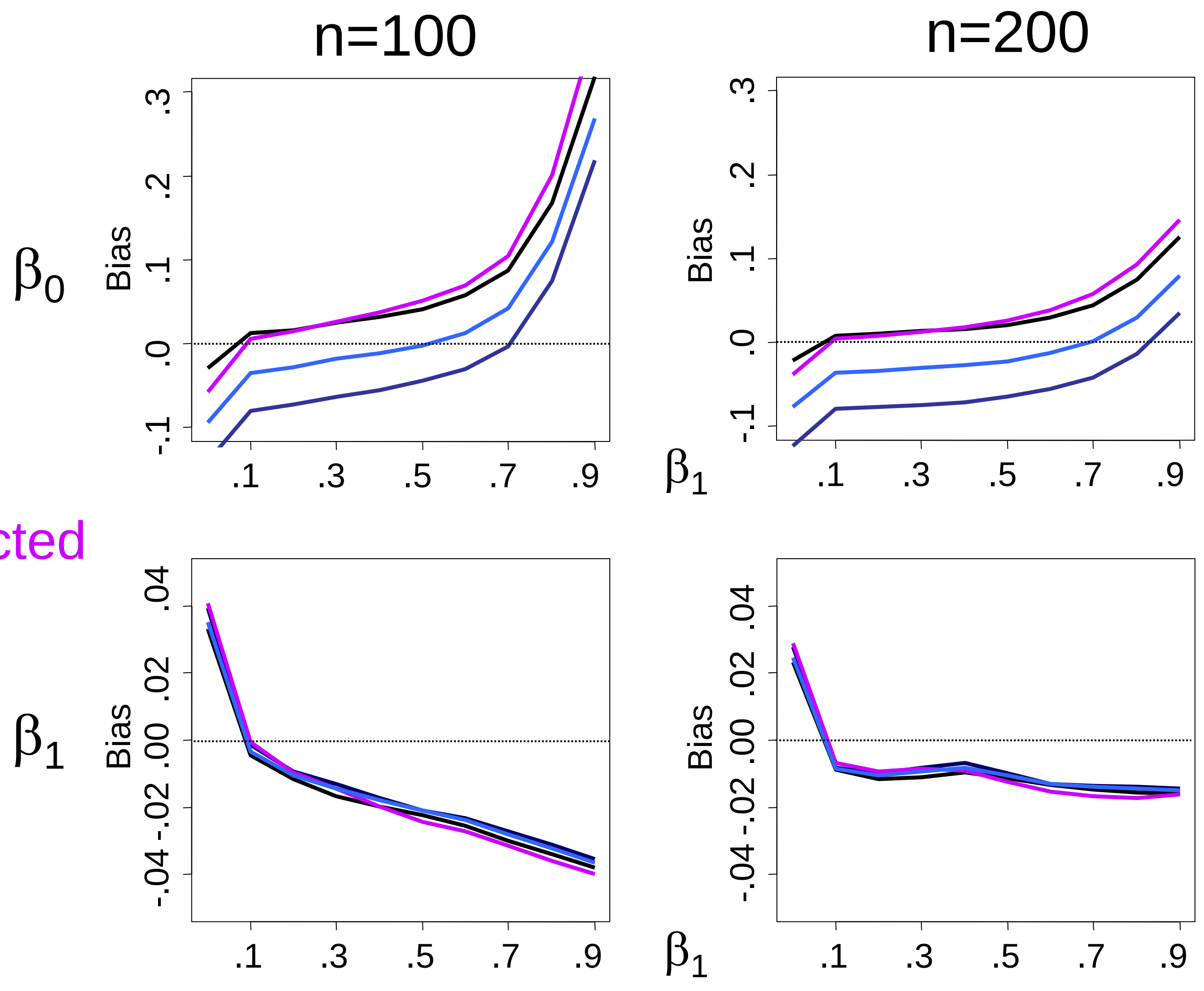
β_1

β_1

β_1

Conditional ML
Tukey, k=7
Tukey, k=5
Tukey, k=5, corrected

Bias for INARCH(1) in dependence on β_1



Conditional ML
Tukey, k=7
Tukey, k=5
Tukey, k=5, corrected

Bias correction effective only in large samples

Conclusions

Tukey M-estimators more robust against many large outliers

Needs good robust initialization - from median or $P_\mu(Y=0)$

Adaptive choice of the tuning constant k gives M-estimators
with good efficiencies irrespective of the true Poisson parameter

M-estimators provide robustness also in INARCH case

Bias correction works for long time series

Ongoing work: extend to INGARCH, prove asymptotic normality

References

Cadigan, N.G., Chen, J. (2001).
Properties of Robust M-estimators for Poisson and Negative Binomial Data.
J. Statist. Comput. Simul. 70, 273-288.

Davis, R.A., Dunsmuir, W.T.M., Street, S.B. (2003).
Observation driven models for Poisson counts.
Biometrika 90, 777-790.

Ferland, R.A., Latour, A., Oraichi, D. (2006).
Integer-valued GARCH processes.
Journal of Time Series Analysis 27, 923-942.

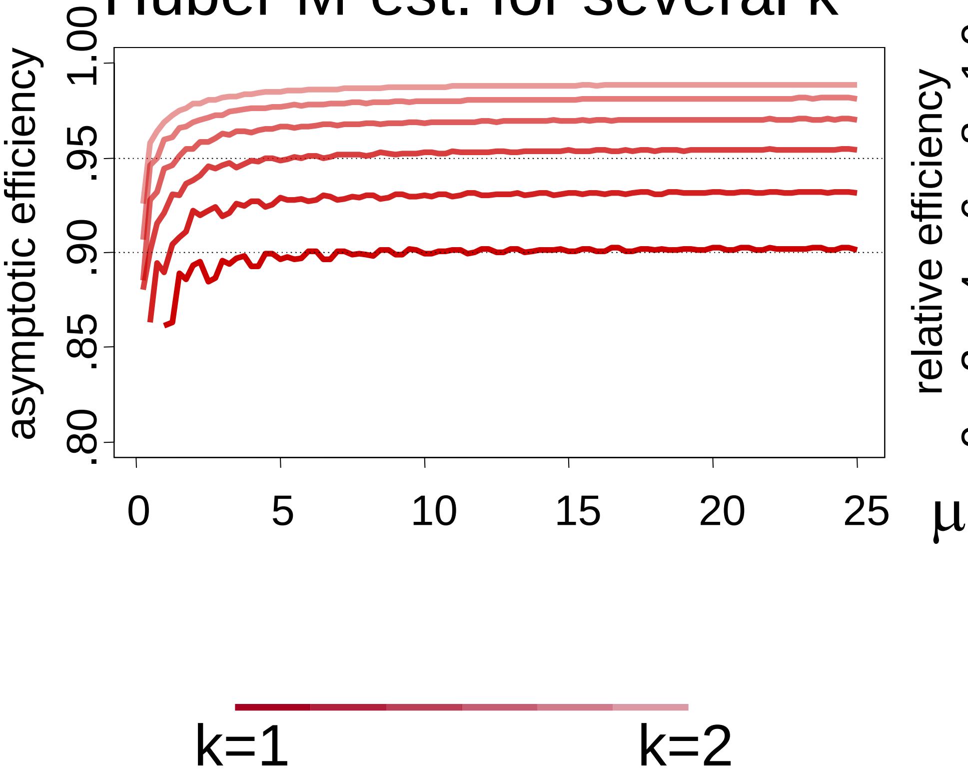
Fokianos, K., Fried, R. (2010).
Outliers in INGARCH Processes.
Journal of Time Series Analysis 31, 210-225.

Fokianos, K., Rahbek, A., Tjøstheim, D. (2009).
Poisson Autoregression.
Journal of the American Statistical Association 104, 1430-1439.

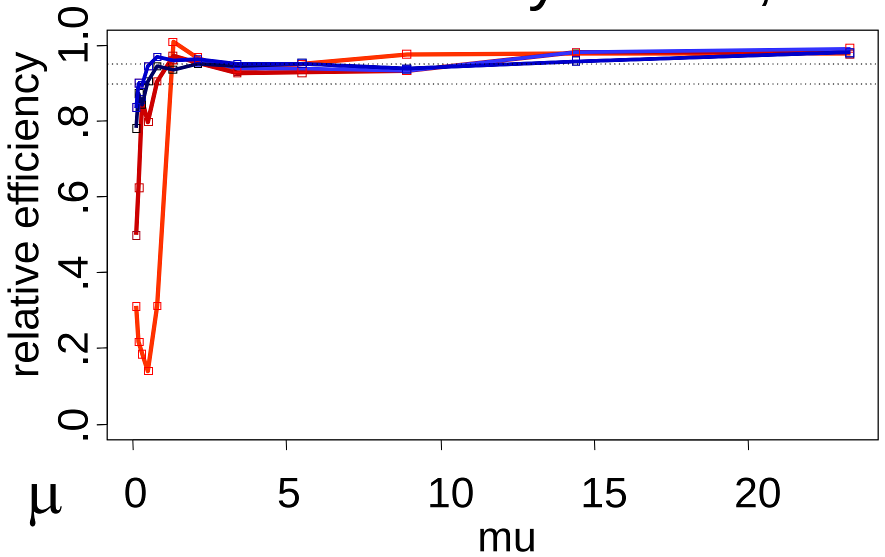
Simpson, D.G., Carroll, R.J., Ruppert, D. (1987).
M-Estimation for Discrete Data: Asymptotic Distribution Theory & Implications.
Annals of Statistics 15, 657-669.

Efficiencies: asymptotic and n=50

Asymptotic efficiency of
Huber M-est. for several k



Finite sample efficiency of
Huber & Tukey M-est., n=50



huberM (robustbase), k=1.8

glmrob, k=1.8

Tukey, k=5

Tukey, k=6

Tukey, adaptive k

Outliers in the INGARCH(p,q)-Model

Clean process: $$Y_t \big|_{(Y_s, s<t)} \sim \mathrm{Poi}(\lambda_t)$$

$$\lambda_t = \beta_0 + \sum_{i=1}^{p} \alpha_i \lambda_{t-i} + \sum_{j=1}^{q} \beta_j Y_{t-j}$$

Ergodicity for INGARCH(1,1): *Fokianos, Rahbek & Tjøstheim (2009)*

Process contaminated by outlier of size ν at time τ:

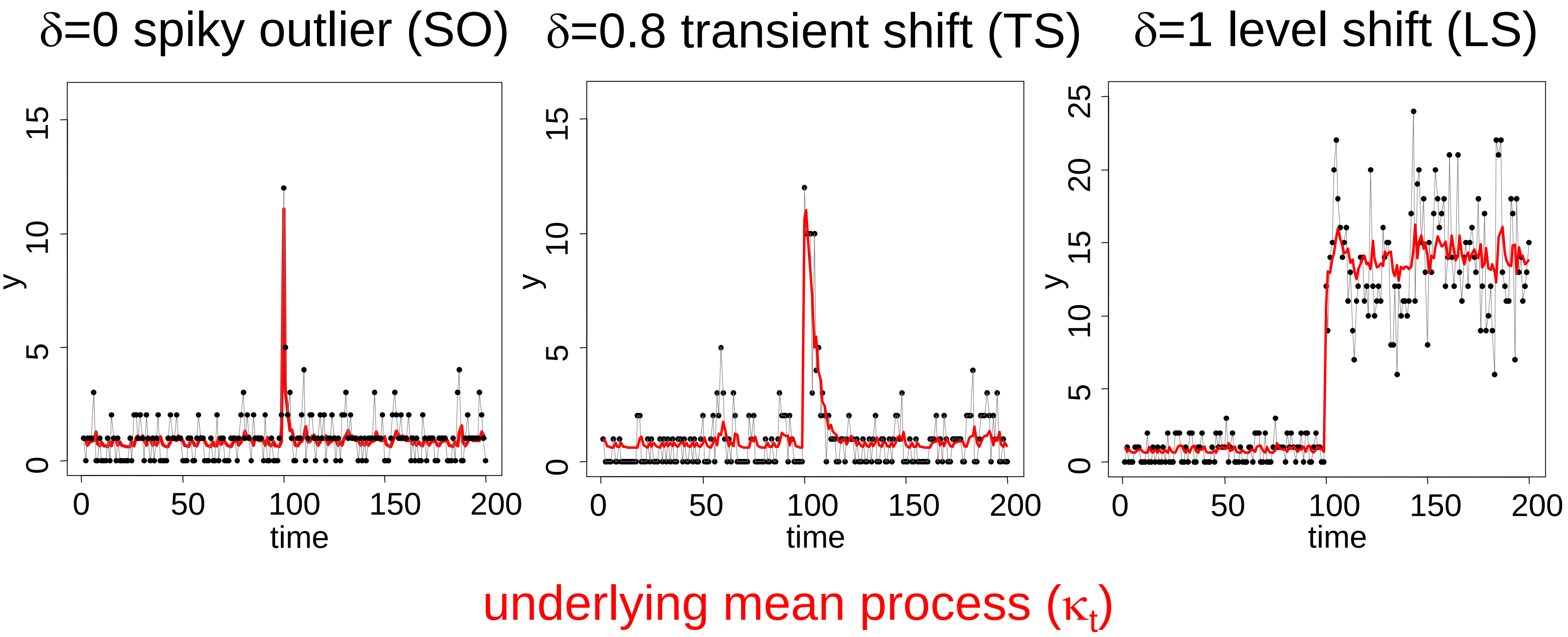
$$Z_t \big|_{(Z_s, s<t)} \sim \mathrm{Poi}(\kappa_t)$$

$$\kappa_t = \beta_0 + \sum_{i=1}^{p} \alpha_i \kappa_{t-i} + \sum_{j=1}^{q} \beta_j Z_{t-j} + \nu \delta^{t-\tau} I(t \geq \tau)$$

Equivalently $$Z_t = Y_t + C_t, \quad C_t \sim \mathrm{Poi}(\kappa_t - \lambda_t)$$

for $\nu>0$ and $t \geq \tau$: $$\kappa_t - \lambda_t = \sum_{i=1}^{p} \alpha_i (\kappa_{t-i} - \lambda_{t-i}) + \sum_{j=1}^{q} \beta_j C_{t-j} + \nu \delta^{t-\tau}$$

Different Types of Outliers



Goal: Detect and classify different outliers