OTTO VON GUERICKE
UNIVERSITÄT
MAGDEBURG

Rainer Schwabe

Cross-sectional Versus Longitudinal Design: Does it Really Matter?

Cross-sectional Versus Longitudinal Design: Does it Really Matter?

Rainer Schwabe, Ulrike Graßhoff

Otto von Guericke University Magdeburg

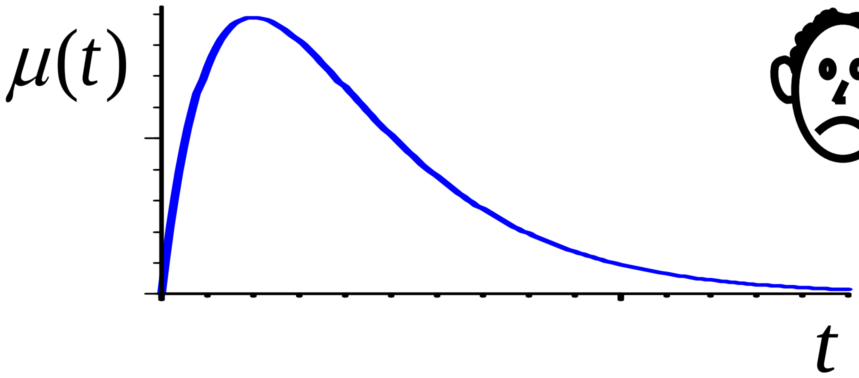
Isaac Newton Institute for Mathematical Sciences

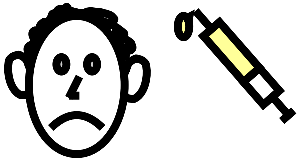
Outline

Prologue: Motivation

- <u>starting point: pharmacokinetics</u>

  measure the concentration of a drug in someone's blood over time

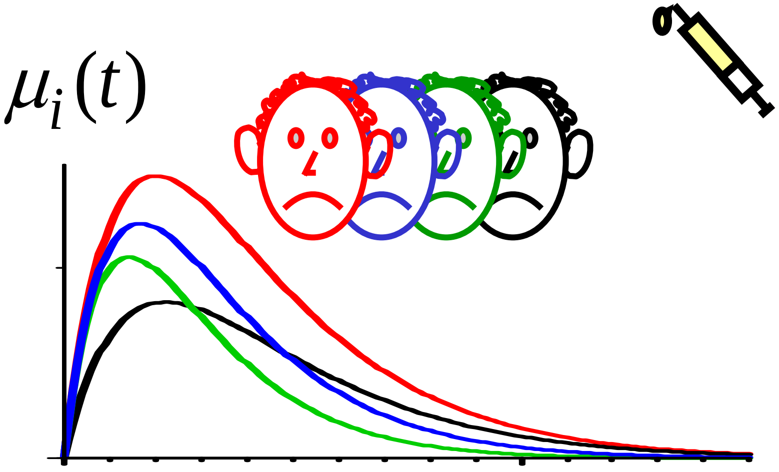




- observations

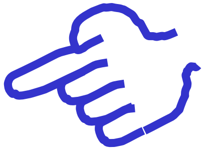
$$Y_j(t_j) = \mu(t_j) + \varepsilon_j$$

Random coefficients



- each individual
  has its **own** curve

- population parameters
  ⇒ “typical” curve



1 Standard Model

- **linear** mean response $\mu(x) = \mathbf{f}(x)^{\mathsf{T}} \boldsymbol{\beta}$

$$Y_i(x_i) = \mathbf{f}(x_i)^{\mathsf{T}} \boldsymbol{\beta} + \varepsilon_i$$

random error

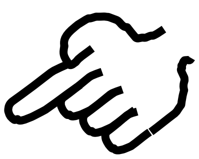
observation $i = 1,\ldots,n$

explanatory variable

ε_i i.i.d.

$\mathrm{Var}(\varepsilon_i) = \sigma^2$

- regression functions $\mathbf{f} = (f_1,\ldots,f_p)$
- parameter $\boldsymbol{\beta} = (\beta_1,\ldots,\beta_p)$



Vector notation and estimation

$$\begin{pmatrix} Y_1 \\ \vdots \\ Y_n \end{pmatrix} \rightarrow \quad \mathbf{Y} = \mathbf{F}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \quad \leftarrow \begin{pmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_n \end{pmatrix}$$

design matrix

$$\mathbf{F} = \left(\mathbf{f}(x_1) \ \dots \ \mathbf{f}(x_n)\right)^\top$$

$$\mathrm{Cov}(\boldsymbol{\varepsilon}) = \sigma^2 \mathbf{I}_n$$

- estimator (BLUE)

$$\hat{\boldsymbol{\beta}} = (\mathbf{F}^\top \mathbf{F})^{-1} \mathbf{F}^\top \mathbf{Y}$$

- covariance

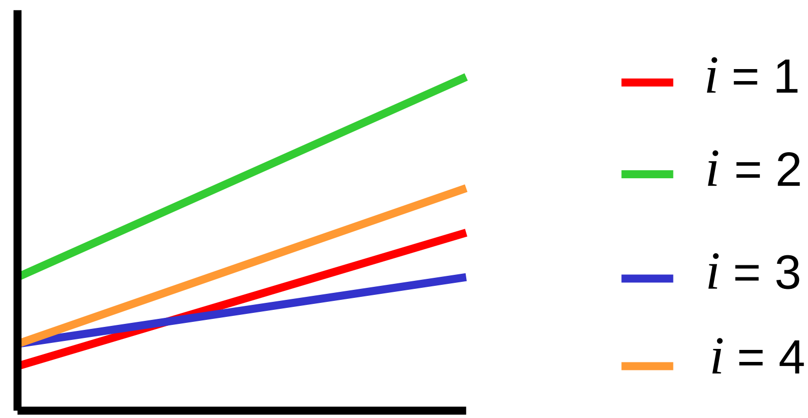
$$\mathrm{Cov}(\hat{\boldsymbol{\beta}}) = \sigma^2 \left(\mathbf{F}^\top \mathbf{F}\right)^{-1}$$

But !

- each individual has its **own** response curve

octodon degus

$$\mu_i(x) = \mathbf{f}(x)^\top \boldsymbol{\beta}_i$$



- individual responses follow a common model

2 Hierarchical Model

- individual level

individual "parameter"

$$Y_{ij} = \mathbf{f}(x_{ij})^{\top} \boldsymbol{\beta}_i + \varepsilon_{ij}$$

individual i = 1,...,n

replication j = 1,...,m_i

explanatory variable

error

$$\varepsilon_{ij} \sim \mathrm{N}(0, \sigma^2)$$

- population level

$$\boldsymbol{\beta}_i \sim \mathrm{N}_p(\boldsymbol{\beta}, \mathbf{D})$$

independent

population parameter

Individual observational vector

$$\mathbf{Y}_i = \mathbf{F}_i \boldsymbol{\beta}_i + \boldsymbol{\varepsilon}_i = \mathbf{F}_i \boldsymbol{\beta} + \mathbf{F}_i \mathbf{b}_i + \boldsymbol{\varepsilon}_i$$

$$\begin{pmatrix} Y_{i1} \\ \vdots \\ Y_{im_i} \end{pmatrix}$$

individual design matrix

$$\mathbf{F}_i = \left(\mathbf{f}(x_{i1}) \ldots \mathbf{f}(x_{im_i})\right)^{\mathsf{T}}$$

$$\begin{pmatrix} \varepsilon_{i1} \\ \vdots \\ \varepsilon_{im_i} \end{pmatrix}$$

$$\mathrm{Cov}(\boldsymbol{\varepsilon}_i) = \sigma^2 \mathbf{I}_{m_i}$$

- individual effect

$$\mathbf{b}_i = \boldsymbol{\beta}_i - \boldsymbol{\beta}, \quad \mathrm{E}(\mathbf{b}_i) = \mathbf{0}, \quad \mathrm{Cov}(\mathbf{b}_i) = \mathbf{D}$$

Individual covariance structure

$$\mathbf{Y}_i = \mathbf{F}_i \boldsymbol{\beta} + \mathbf{F}_i \mathbf{b}_i + \boldsymbol{\varepsilon}_i$$

- individual covariance matrix

$$\mathrm{Cov}(\mathbf{Y}_i) = \mathbf{F}_i \mathbf{D} \mathbf{F}_i^{\mathsf{T}} + \sigma^2 \mathbf{I}_{m_i}$$

observations are correlated

Example: linear regression

$$Y_{ij} = \beta_{i0} + \beta_{i1} x_{ij} + \varepsilon_{ij}$$

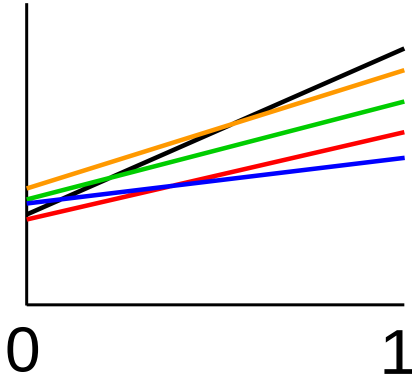
$$\mathbf{f}(x) = (1, x)^{\mathsf{T}}$$

- uncorrelated coefficients $\mathbf{D} = \begin{pmatrix} d_0 & 0 \\ 0 & d_1 \end{pmatrix}$

baseline regression

$0 \le x \le 1$

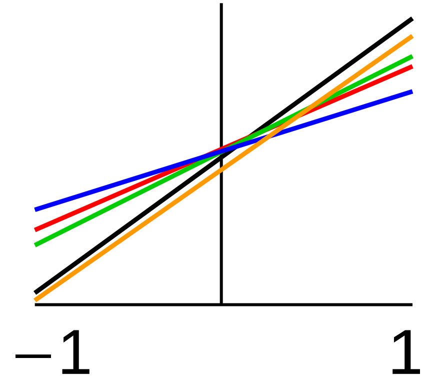
BL



$d_0 << d_1$

regression to the mean

$-1 \le x \le 1$



R2M

3 Longitudinal Design

- all individuals
  at the same experimental settings

$$Y_{ij} = \mathbf{f}(x_j)^{\mathsf{T}} \boldsymbol{\beta}_i + \varepsilon_{ij}$$

$$m_i \equiv m$$
$$x_{ij} \equiv x_j$$

$$\mathbf{F}_i \equiv \mathbf{F}$$
$$\mathrm{Cov}(\mathbf{Y}_i) \equiv \sigma^2 \mathbf{I}_m + \mathbf{F}\,\mathbf{D}\,\mathbf{F}^{\mathsf{T}}$$

Estimation

- BLUE

$$\hat{\boldsymbol{\beta}} = (\mathbf{F}^\top \mathbf{F})^{-1} \mathbf{F}^\top \overline{\mathbf{Y}}$$

$$\overline{\mathbf{Y}} = \frac{1}{n} \sum_{i=1}^{n} \mathbf{Y}_i$$

does not require the knowledge of $\mathbf{D}$ (WLS=OLS)

- covariance

$$\mathrm{Cov}(\hat{\boldsymbol{\beta}}) = \frac{1}{n}\left(\sigma^2 \left(\mathbf{F}^\top \mathbf{F}\right)^{-1} + \mathbf{D}\right)$$

Design

- aim: choose $x_1,\ldots,x_m$ to minimize $\mathrm{Cov}(\hat{\boldsymbol{\beta}})$

- approximate design

within individuals

$$\xi = \begin{pmatrix} x_1 & \cdots & x_k \\ w_1 & \cdots & w_k \end{pmatrix}$$

settings

proportions

- standardized covariance per observation

$$\mathbf{C}_{\mathrm{LG}}(\xi) = \left(\sum_{j=1}^{k} w_j \, \mathbf{f}(x_j)\mathbf{f}(x_j)^{\mathsf{T}}\right)^{-1} + m\mathbf{D}/\sigma^2$$

D-optimality

- ξ^* D-optimal, if ξ^* minimizes $\det \mathbf{C}_{\mathrm{LG}}(\xi)$

- equivalence theorem:

ξ^* D-optimal $\Leftrightarrow$

$$\left(\mathbf{f}(x) - \mathbf{a}(\xi^*)\right)^\top \mathbf{A}(\xi^*)\left(\mathbf{f}(x) - \mathbf{a}(\xi^*)\right) \le c(\xi^*)$$

- with equality at the settings of ξ^*

Example: linear regression

$$Y_{ij} = \beta_{i0} + \beta_{i1} x_j + \varepsilon_{ij}$$

- regression to the mean $\quad -1 \le x_j \le 1$

$$\xi^* = \begin{pmatrix} -1 & 1 \\ 1/2 & 1/2 \end{pmatrix}$$

$$\mathbf{C}_{\mathrm{LG}}(\xi^*) = \begin{pmatrix} 1+\delta_0 & 0 \\ 0 & 1+\delta_1 \end{pmatrix}$$

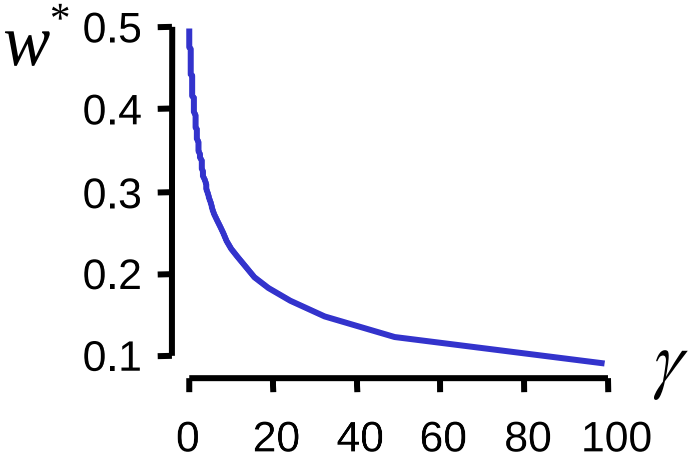
$$\delta_i = md_i / \sigma^2$$

Example: linear regression

- baseline regression $0 \le x_j \le 1$

$$\xi^* = \xi_{w^*} = \begin{pmatrix} 0 & 1 \\ 1 - w^* & w^* \end{pmatrix}$$

$$w^* = \left(\sqrt{1+\gamma} - 1\right)/\gamma \quad \text{where} \quad \gamma = d_1/(d_0 + \sigma^2/m)$$



$w^* \to 0$

for $\gamma \to \infty$

4 Cross-sectional Design

- each individual

  only one experimental setting

$$Y_{ij} = \mathbf{f}(x_i)^\top \boldsymbol{\beta}_i + \varepsilon_{ij}$$

- model with constant term: $\mathrm{f}_1(x) \equiv 1$

$$m_i \equiv m$$

$$x_{ij} \equiv x_i$$

$$\mathbf{F}_i \equiv \mathbf{1}_m \mathbf{f}(x_i)^\top$$

$$\mathrm{Cov}(\mathbf{Y}_i) \equiv \sigma^2 \mathbf{I}_m + v(x_i) \mathbf{1}_m \mathbf{1}_m^\top$$

- “variance function” $\mathrm{v}(x) = \mathbf{f}(x)^\top \mathbf{D} \mathbf{f}(x)$

Estimation

- BLUE

$$\hat{\boldsymbol{\beta}} = \left(\sum_{i=1}^{n} \mathbf{F}_i^{\top} \mathbf{V}_i^{-1} \mathbf{F}_i\right)^{-1} \sum_{i=1}^{n} \mathbf{F}_i^{\top} \mathbf{V}_i^{-1} \mathbf{Y}_i$$

requires the knowledge of $\mathbf{D}$ through v

- covariance

$$\mathrm{Cov}(\hat{\boldsymbol{\beta}}) = \left(\sum_{i=1}^{n} \frac{m}{\sigma^2 + m v(x_i)} \mathbf{f}(x_i) \mathbf{f}(x_i)^{\top}\right)^{-1}$$

Design

- aim: choose $x_1,\ldots,x_n$ to minimize $\mathrm{Cov}(\hat{\boldsymbol{\beta}})$

- approximate design

across individuals

$$\xi = \begin{pmatrix} x_1 & \cdots & x_k \\ w_1 & \cdots & w_k \end{pmatrix}$$

← settings

← proportions

- standardized covariance per observation

$$\mathbf{C}_{\mathrm{CS}}(\xi) = \left(\sum_{j=1}^{k} w_j \frac{1}{1+mv(x_j)/\sigma^2} \mathbf{f}(x_j)\mathbf{f}(x_j)^{\mathsf{T}}\right)^{-1}$$

D-optimality

- ξ^* D-optimal, if ξ^* minimizes $\det \mathbf{C}_{\mathrm{CS}}(\xi)$

- equivalence theorem:

ξ^* D-optimal $\Leftrightarrow$

$$\mathbf{f}(x)^\mathsf{T}\mathbf{C}_{\mathrm{CS}}(\xi^*)\mathbf{f}(x) \le p\left(1 + m\,\mathbf{f}(x)^\mathsf{T}\mathbf{D}\mathbf{f}(x)/\sigma^2\right)$$

- with equality at the settings of ξ^*

Example: linear regression

$$Y_{ij} = \beta_{i0} + \beta_{i1} x_j + \varepsilon_{ij}$$

- baseline regression

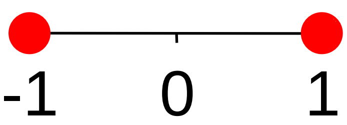
$$0 \le x_j \le 1$$

$$\xi^* = \begin{pmatrix} 0 & 1 \\ 1/2 & 1/2 \end{pmatrix}$$

- regression to the mean

$$-1 \le x_j \le 1$$

$$\xi_1 = \begin{pmatrix} -1 & 1 \\ 1/2 & 1/2 \end{pmatrix}$$



D-optimal $\Leftrightarrow$ $\gamma \le 1$

$$\gamma = d_1 / (d_0 + \sigma^2 / m)$$

R2M: $\gamma = d_1 / (d_0 + \sigma^2 / m) > 1$

- optimal two-point design

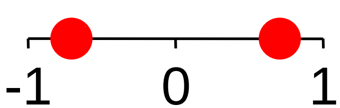
$$\xi^* = \begin{pmatrix} x_1^* & x_2^* \\ 1/2 & 1/2 \end{pmatrix} \quad \text{with} \quad x_1^* x_2^* = -\gamma$$

solution not unique

$$\mathbf{C}_{\mathrm{CS}}(\xi^*) = 2\begin{pmatrix} \sigma^2 + m d_0 & 0 \\ 0 & m d_1 \end{pmatrix} = p\left(\sigma^2 \mathbf{e}_1 \mathbf{e}_1^{\mathsf{T}} + m\,\mathbf{D}\right)$$

- symmetric solution

$$\xi_{x^*} = \begin{pmatrix} -x^* & x^* \\ 1/2 & 1/2 \end{pmatrix}$$

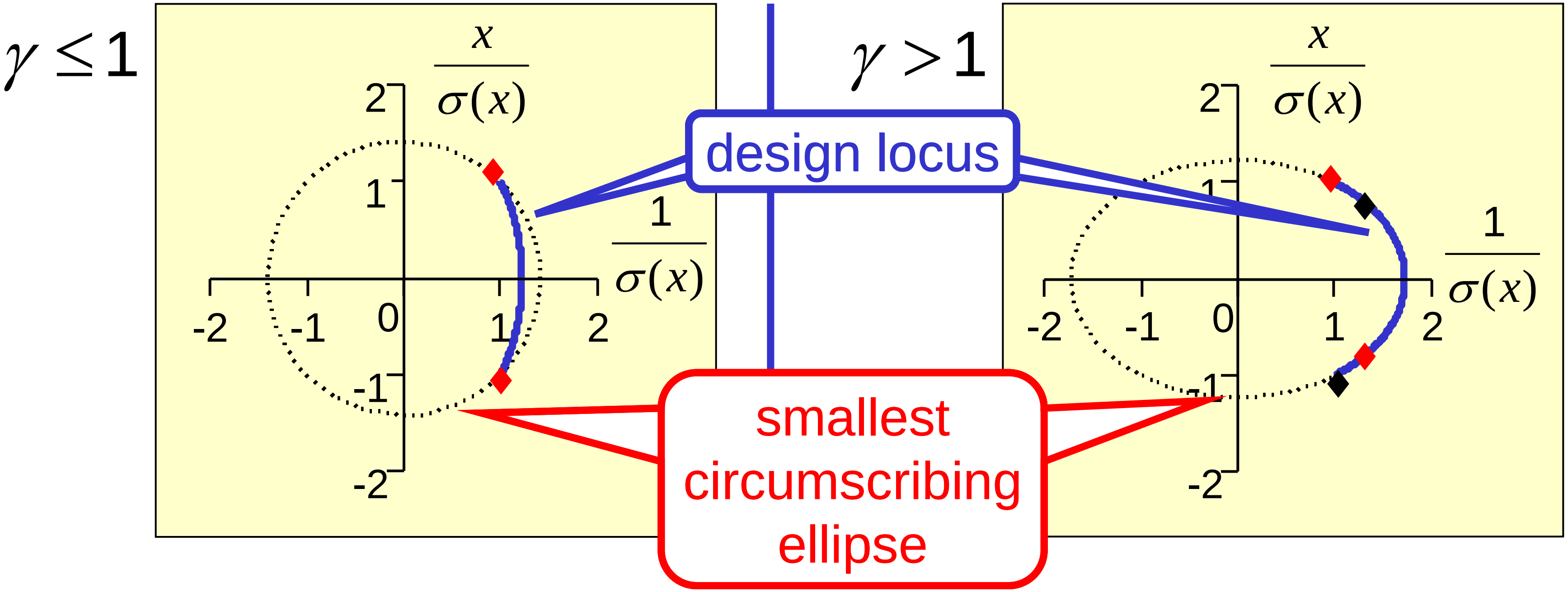


$$x^* = 1/\sqrt{\gamma} \to 0 \quad \text{for} \quad \gamma \to \infty$$

Geometric Interpretation

- standardized design locus

$$\left\{ \tilde{\mathbf{f}}(x) = \sqrt{\frac{m}{\sigma^2 + m v(x)}}\,\mathbf{f}(x); -1 \le x \le 1 \right\}$$

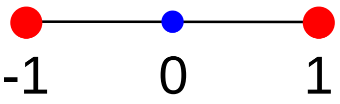


= design locus for trigonometric regression without intercept on an arc

R2M: $\gamma = d_1 / (d_0 + \sigma^2 / m) > 1$

- optimal midpoint design

$$\xi_{w^*} = \begin{pmatrix} -1 & 0 & 1 \\ w^* & 1-2w^* & w^* \end{pmatrix}$$



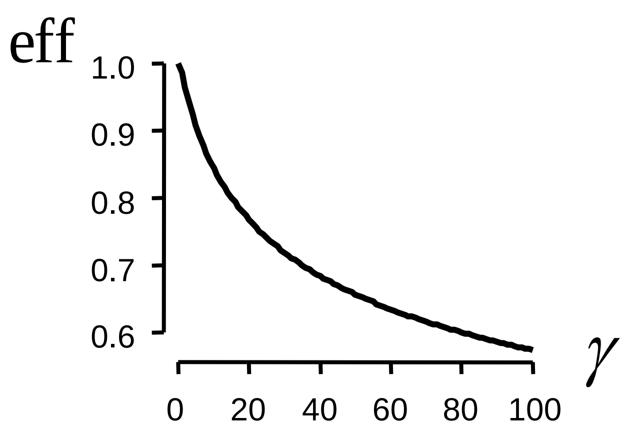
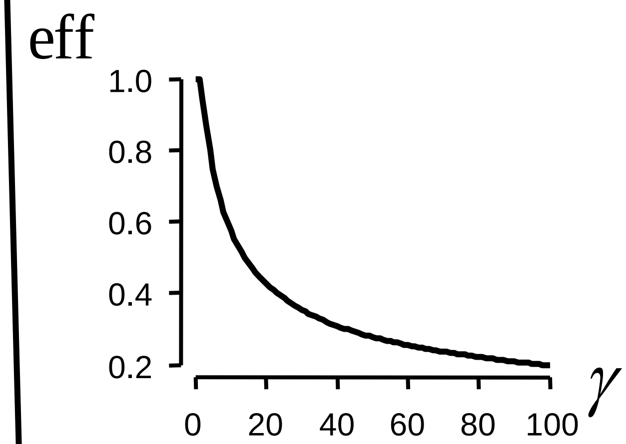
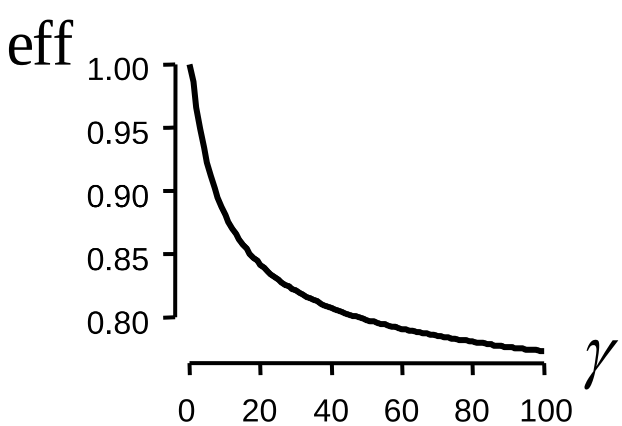
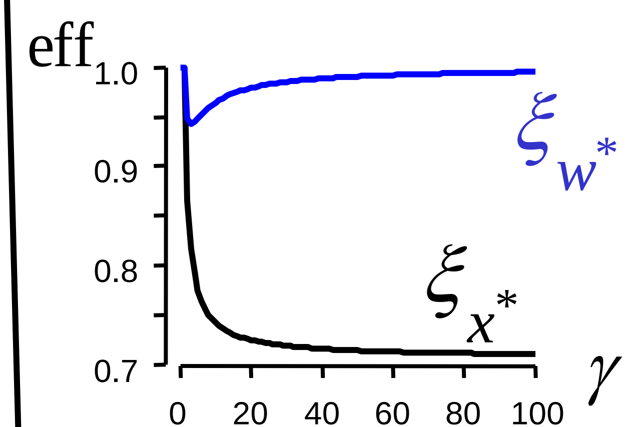
$$\mathbf{C}_{\mathrm{CS}}(\xi_{w^*}) = \mathbf{C}_{\mathrm{CS}}(\xi_{x^*})$$

$$w^* = \left(1 + \frac{1}{\gamma}\right) / 4 \to 1/4 \quad \text{for } \gamma \to \infty$$

5 Comparison: An Example

- linear regression
  - baseline regression
  - regression to the mean
- uncorrelated coefficients
- longitudinal versus cross-sectional design
- variance ratio γ large

D-efficiencies

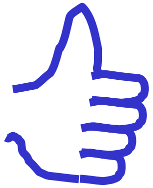
	baseline model	R2M model
longitudinal design in cross-sectional model		
cross-sectional design in longitudinal model		

D-efficiencies

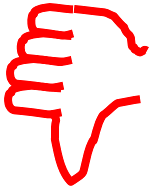
	baseline model	R2M model
longitudinal design in cross-sectional model	$2\sqrt{w^*(1-w^*)} = O(\gamma^{-1/4}) \to \underline{\underline{0}}$	$\dfrac{2\sqrt{\gamma}}{1+\gamma} \to \underline{\underline{0}}$
cross-sectional design in longitudinal model	$\mathrm{eff}_{\mathrm{LG}}(\xi^*) \geq \underline{\underline{0.70}}$	$\mathrm{eff}_{\mathrm{LG}}(\xi_{x^*}) \geq \underline{\underline{0.70}}$ $\mathrm{eff}_{\mathrm{LG}}(\xi_{w^*}) \geq \underline{\underline{0.94}}$

Conclusion

- cross-sectional designs
  may work quite well
  in a longitudinal setting
  even for large γ



- longitudinal designs
  may become very inefficient
  in a cross-sectional setting
  for large γ



Epilogue: Outlook

- cross-sectional: more predictors, general **D** ✓ algorithm
- combination of cross-sectional and longitudinal (✓) additive
- prediction ✓ Maryna Prus
- non-linear models ? ? ?

Cross-sectional versus longitudinal design:

Does it really matter?

Message
To Go:

"Yes, it does."