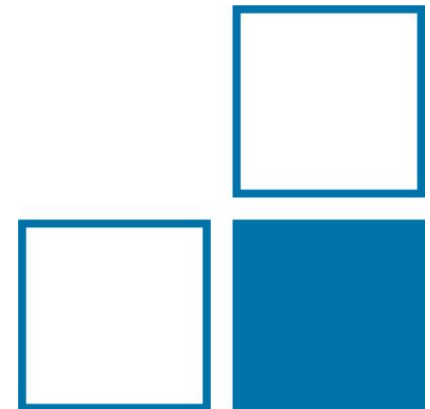


Gegenüberstellung von linearen Regressionsverfahren beim Methodenvergleich

Steffen Martens, Katy Klauenberg
und Clemens Elster

AG 8.42 Datenanalyse und Messunsicherheit

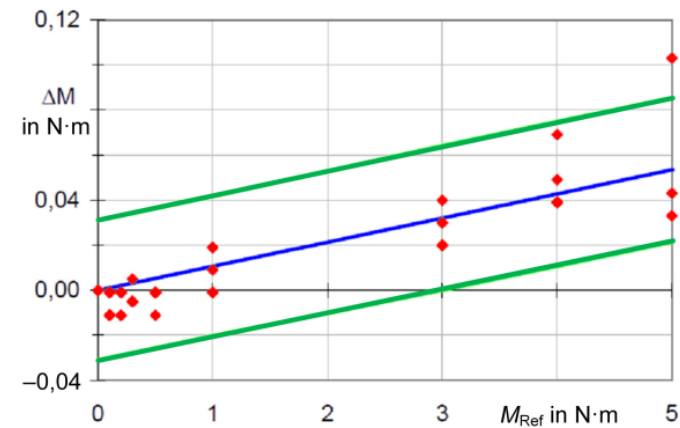
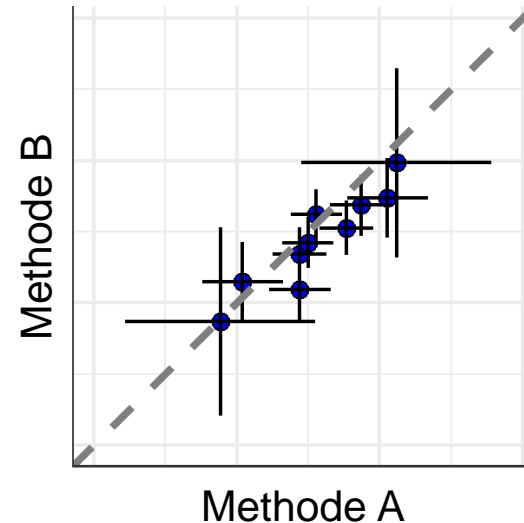
9. VDI-Fachtagung Messunsicherheit 2019
Erfurt, 13.11.19



Lineare Regressionsverfahren bei Errors-in-Variables (EIV) Modellen

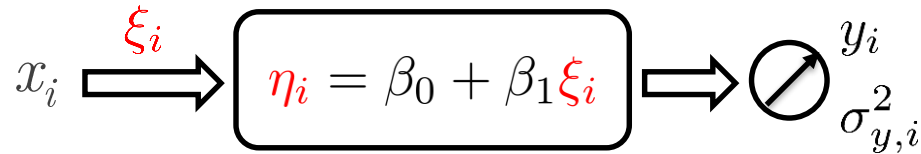
Regression von lineare Zusammenhängen ist ein häufig auftretendes Problem in der Metrologie, z.B.

- Methodenvergleich gegenüber Goldstandard
- Kalibrierung



Quelle: Bild B10 VDI/VDE 2600 Blatt 2
Kalibrierung eines Dehnungsmessstreifens

Abhängige und unabhängige Größen
sind mit Unsicherheiten behaftet



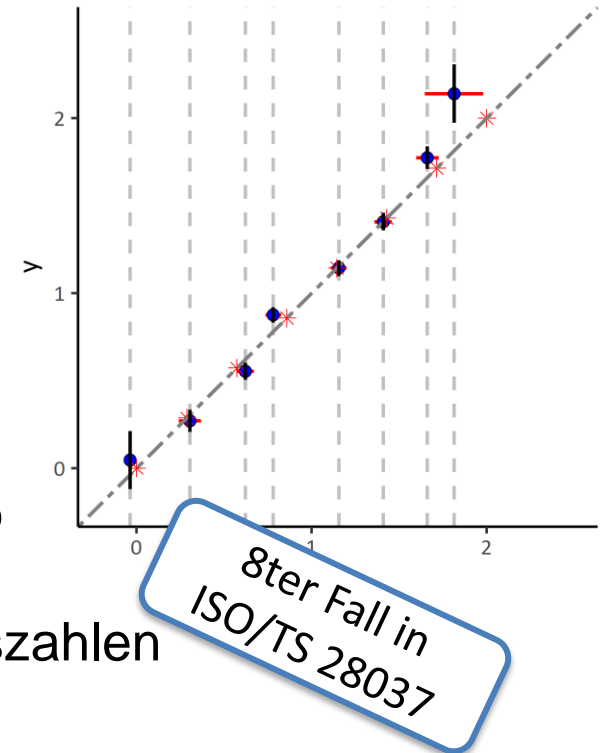
- N Paare $(x_i, y_i)^T$ mit unabhängiger x_i und abhängiger Eingangsgröße y_i

(1a) $\xi_i = x_i + \epsilon_{x,i}$

(1b) $y_i = \beta_0 + \beta_1 \xi_i + \epsilon_{y,i}$

- Annahmen:

- 1) wahre Response η_i hängt linear von ξ_i ab
- 2) Störterme der i -ten Messung beschrieben durch multivariate, normal verteilte Zufallszahlen mit Mittelwert 0 und i -ten Kovarianz Σ_i
- 3) Kovarianzen Σ_i sind bekannt



- **Ziel:** Bestimmung der Schätzwerte und deren Unsicherheiten

$$\hat{\beta}_0, \hat{\beta}_1, \hat{\xi}_i, u_{\hat{\beta}_0}, u_{\hat{\beta}_1}, u_{\hat{\xi}_i}$$

- zahlreiche Methoden existieren

- Methode der kleinsten Quadrate (LS)^[1,2]

- gewichtetes TLS (WTLS)
 - Deming Regression^[3]
 - gewöhnliches LS (OLS)

- Bayes'sche Regression^[4,5]

- Maximum likelihood Schätzer^[4]

- Instrumentale Variablen^[6]

- Momenten-Methode^[7]

und viele mehr

[1] Adcock *The Analyst* **4**, 183 (1877); **5**, 53 (1878), [2] Pearson *Philos Mag.* **2**, 559 (1901)

[3] W. E. Deming „Statistical adjustment of data“ (1943), [4] Zellner „An Introduction to Bayesian Inference Econometrics“ (1971)

[5] Carroll et al. „Measurement errors in Nonlinear models“ (2006), [6] 9 M. Y. Wong *Biometrika* **76**, 141 (1989),

[7] Pal *J. Econometrics* **14**, 349 (1980)

Zielstellung und offene Fragen

- **Ziel:** Bestimmung der Schätzwerte und deren Unsicherheiten

$$\hat{\beta}_0, \hat{\beta}_1, \hat{\xi}_i, u_{\hat{\beta}_0}, u_{\hat{\beta}_1}, \nu$$

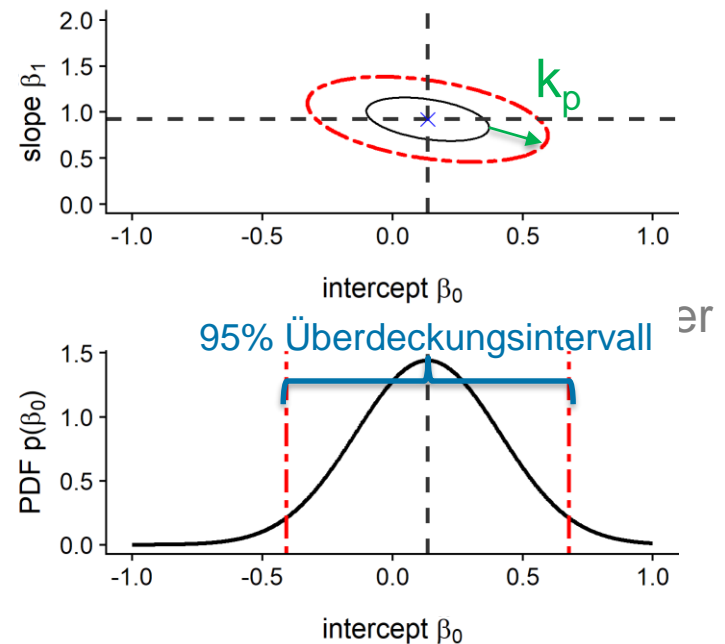
- zahlreiche Methoden existieren

- Methode der kleinsten Quadrate (LS)

- gewichtetes TLS (WTLS)
 - Deming Regression
 - gewöhnliches LS (OLS)

- Bayes'sche Regression

- GUM Dokumente erlauben Unsicherheitsbestimmung nach
 - 1) Fortpflanzung der Unsicherheiten (**GUM**^[1], GUM-S2^[2])



Zielstellung und offene Fragen

- **Ziel:** Bestimmung der Schätzwerte und deren Unsicherheiten

$$\hat{\beta}_0, \hat{\beta}_1, \hat{\xi}_i, u_{\hat{\beta}_0}, u_{\hat{\beta}_1}, \nu$$

- zahlreiche Methoden existieren

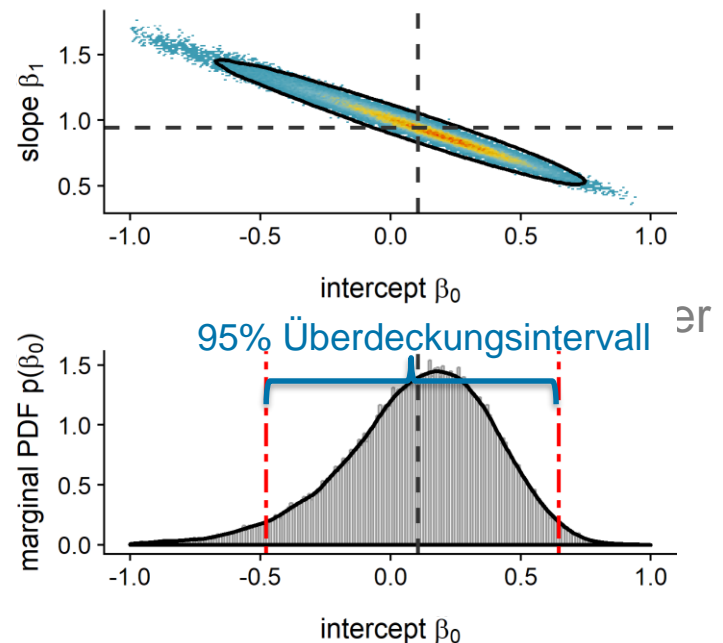
- Methode der kleinsten Quadrate (LS)

- gewichtetes TLS (WTLS)
 - Deming Regression
 - gewöhnliches LS (OLS)

- Bayes'sche Regression

- GUM Dokumente erlauben Unsicherheitsbestimmung nach
 - 1) Fortpflanzung der Unsicherheiten (**GUM**^[1], GUM-S2^[2])
 - 2) Fortpflanzung von Verteilungen (**GUM-S1**^[3], GUM-S2^[2])

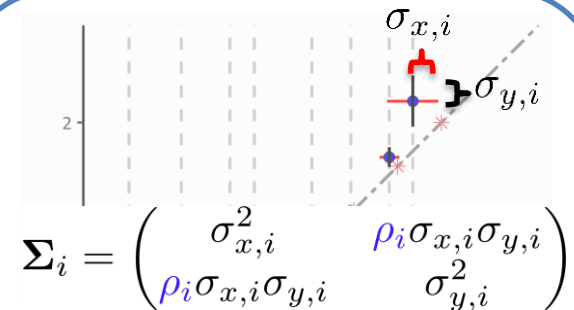
- GUM Dok. geben **keine** direkte Hilfestellung für Regressionsprobleme



- z.B. ISO/TS 28037:2010 & DIN EN ISO 6143:2004 empfehlen WTLS

$$\begin{pmatrix} \hat{\beta} \\ \hat{\xi} \end{pmatrix} = \underset{\beta, \xi}{\operatorname{argmin}} \sum_{i=1}^N \mathbf{v}_i^\top \Sigma_i^{-1} \mathbf{v}_i \quad \text{with } \mathbf{v}_i = \begin{pmatrix} x_i - \xi_i \\ y_i - \beta_0 - \beta_1 \xi_i \end{pmatrix}, \beta = \begin{pmatrix} \beta_0 \\ \beta_1 \end{pmatrix}$$

- im Allgemeinen, nichtlineares Funktional nur numerisch minimierbar^[1,2]
 - Algorithmen sind kostenlos erhältlich
 - Unsicherheiten der Schätzwerte^[2,3]
- Liefert eine Unsicherheitsbestimmung nach der die gleichen Ergebnisse für die Schätzwerte



- **oft:** bekannte Unsicherheiten der Referenzwerte werden ignoriert („ $\sigma_{x,i}$ ist klein im Verhältnis zu $\sigma_{y,i}$ “)

gewöhnliche Methode der kleinsten Quadrate (OLS)

[1] C. A. Cantrell, *Atmos. Chem. Phys.* **8**, 5477 (2008); [2] ISO/TS 28037:2010

[3] M. Krystek and M. Anton, *Meas. Sci. Technol.* **22**, 035101 (2011); [4] T. Isobe et al. *Astrophys. Journal*, **364**, 104 (1990)

[5] ISO 11095:1996

- z.B. ISO/TS 28037:2010 & DIN EN ISO 6143:2004 empfehlen WTLS

$$\begin{pmatrix} \hat{\beta} \\ \hat{\xi} \end{pmatrix} = \underset{\beta, \xi}{\operatorname{argmin}} \sum_{i=1}^N \mathbf{v}_i^\top \boldsymbol{\Sigma}_i^{-1} \mathbf{v}_i \quad \text{with } \mathbf{v}_i = \begin{pmatrix} x_i - \xi_i \\ y_i - \beta_0 - \beta_1 \xi_i \end{pmatrix}, \boldsymbol{\beta} = \begin{pmatrix} \beta_0 \\ \beta_1 \end{pmatrix}$$

- im Allgemeinen, nichtlineares Funktional nur numerisch minimierbar^[1,2]
 - Algorithmen sind kostenlos erhältlich
 - Unsicherheiten der Schätzwerte^[2,3]
- Liefert eine Unsicherheitsbestimmung nach GUM und GUM-S1
die gleichen Ergebnisse für die Schätzwerte?

- **oft:** bekannte Unsicherheiten der Referenzmethode oder des –signals werden ignoriert („ $\sigma_{x,i}$ ist klein im Verhältnis zu $\sigma_{y,i}$ “^[5])

gewöhnliche Methode der kleinsten Quadrate (OLS)

[1] C. A. Cantrell, *Atmos. Chem. Phys.* **8**, 5477 (2008); [2] ISO/TS 28037:2010

[3] M. Krystek and M. Anton, *Meas. Sci. Technol.* **22**, 035101 (2011); [4] T. Isobe et al. *Astrophys. Journal*, **364**, 104 (1990)

[5] ISO 11095:1996

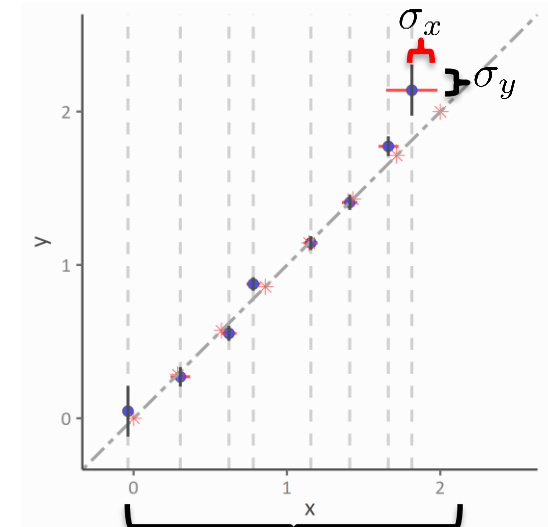
Gültigkeit von OLS in EIV Modellen

- OLS Punktschätzer sind *nicht erwartungstreu* und *inkonsistent*
- bei konstanter Messunsicherheit ($\sigma_x = \sigma_{x,i}$, $\sigma_y = \sigma_{y,i}$, $\rho = \rho_i$), Punktschätzer ist *asyp. normal verteilt*^[1]

$$E\left(\hat{\beta}_1^{\text{OLS,EIV}}\right) = \beta_1 + \frac{\sigma_x^2}{\sigma_\xi^2} \frac{\rho \sigma_y / \sigma_x - \beta_1}{1 + \sigma_x^2 / \sigma_\xi^2},$$

$$\text{Var}\left(\hat{\beta}_1^{\text{OLS,EIV}}\right) = \frac{\sigma_y^2}{N \sigma_\xi^2} \left[1 - 2\rho \frac{\sigma_x}{\sigma_y} E\left(\hat{\beta}_1^{\text{OLS,EIV}}\right) + \left(\frac{\sigma_x}{\sigma_y} E\left(\hat{\beta}_1^{\text{OLS,EIV}}\right) \right)^2 \right]$$

- bester linearer erwartungstreuer Schätzer (BLUE)
aber nur, wenn $\sigma_x / \sigma_y \rightarrow 0$, $\sigma_x / \sigma_\xi \rightarrow 0$
und $\varepsilon_{y,i}$ unkorreliert sind



^[1] Gleser et al., *Ann. Stat.* **15**, 220-233 (1987)

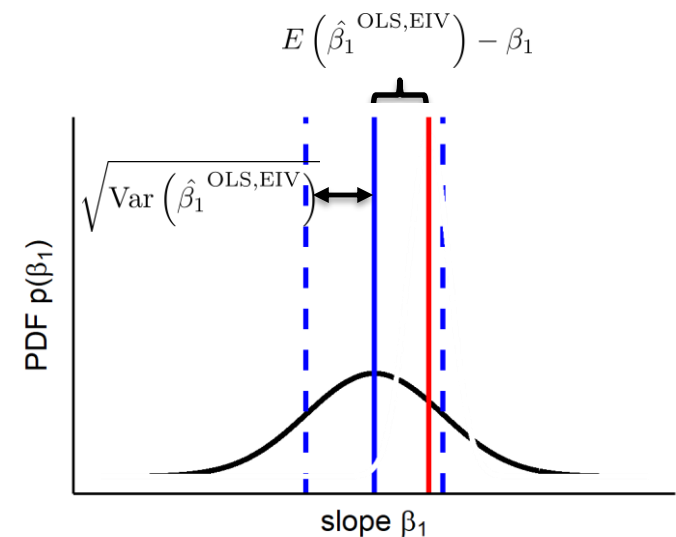
$$\sigma_\xi^2 = N^{-1} \sum_{i=1}^N (\xi_i - \bar{\xi})^2$$

Kann man OLS trotzdem anwenden?

Bedingungen für die Nutzung von OLS:

1. Abweichung vom wahren Wert muss mit der Unsicherheit des Schätzers verträglich sein.

$$(2a) \quad \text{Var} \left(\hat{\beta}_1^{\text{OLS,EIV}} \right) \gg \left(E \left(\hat{\beta}_1^{\text{OLS,EIV}} \right) - \beta_1 \right)^2$$



Kann man OLS trotzdem anwenden?

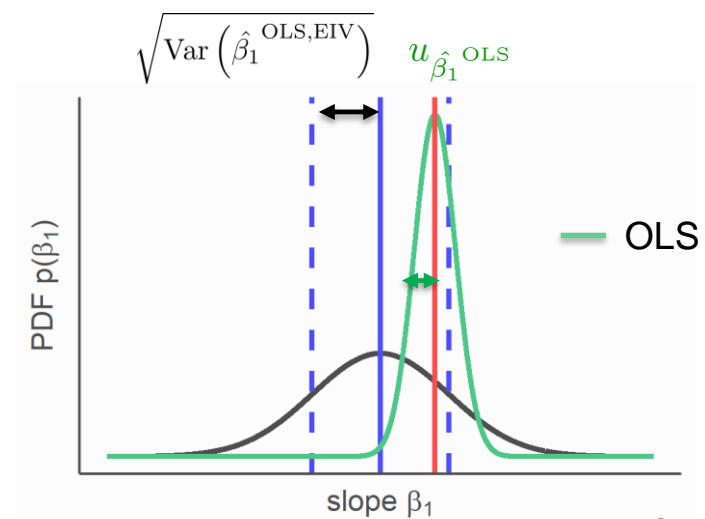
Bedingungen für die Nutzung von OLS:

1. Abweichung vom wahren Wert muss mit der Unsicherheit des Schätzers verträglich sein.

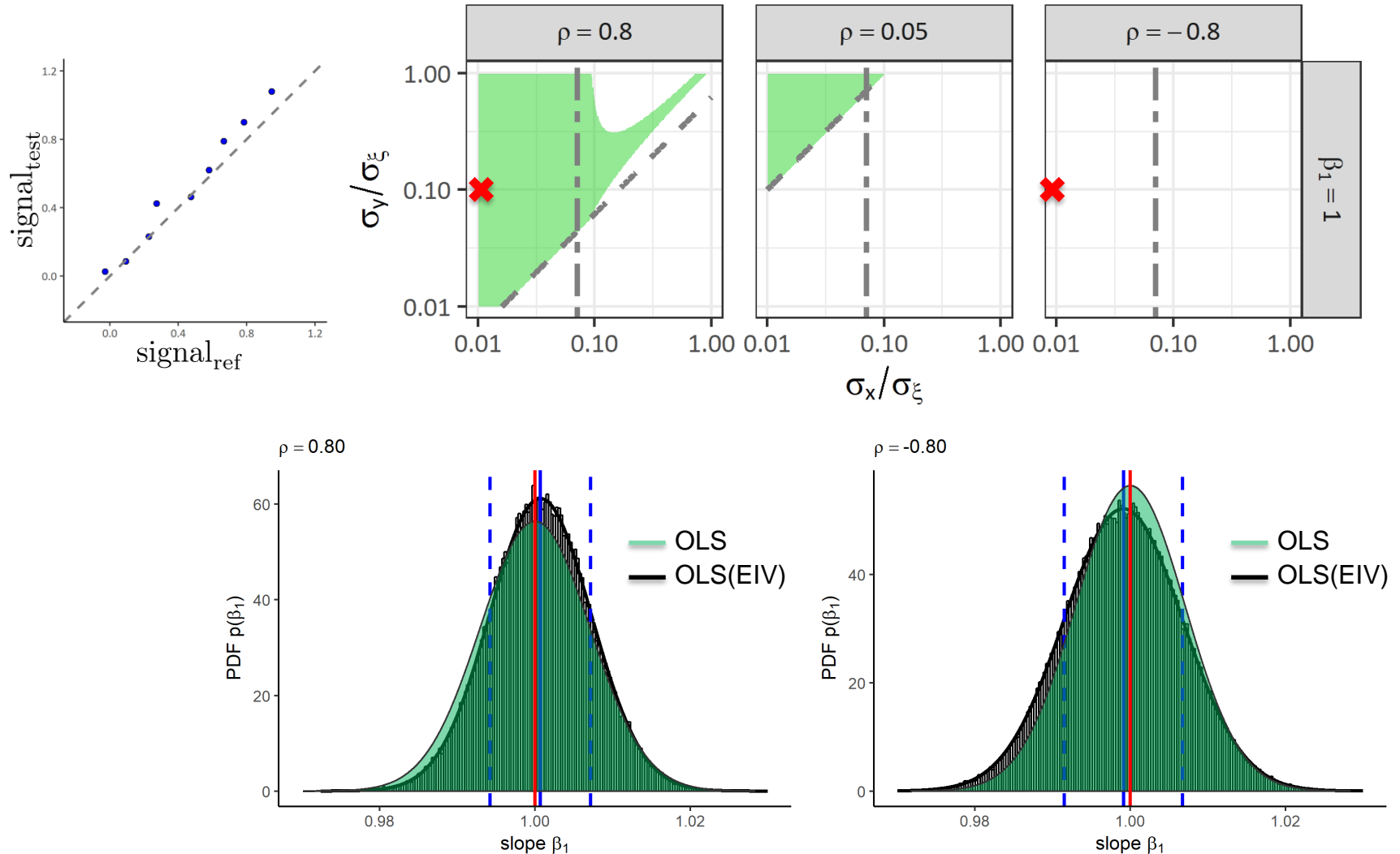
$$(2a) \quad \text{Var} \left(\hat{\beta}_1^{\text{OLS,EIV}} \right) \gg \left(E \left(\hat{\beta}_1^{\text{OLS,EIV}} \right) - \beta_1 \right)^2$$

2. Die Unsicherheit des Schätzers darf nicht unterschätzt werden.

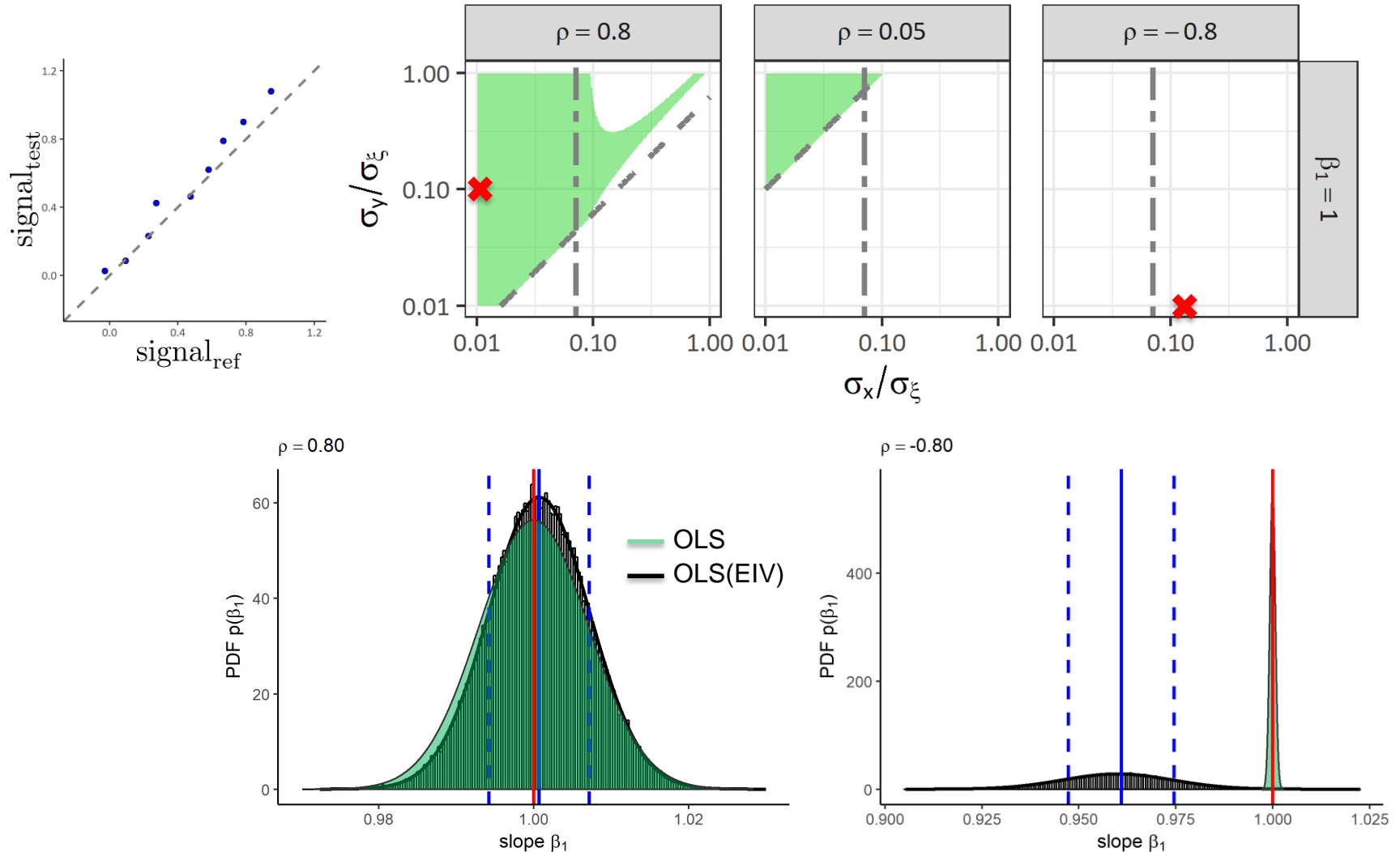
$$(2b) \quad \sqrt{\text{Var} \left(\hat{\beta}_1^{\text{OLS,EIV}} \right)} \leq u_{\hat{\beta}_1^{\text{OLS}}}$$



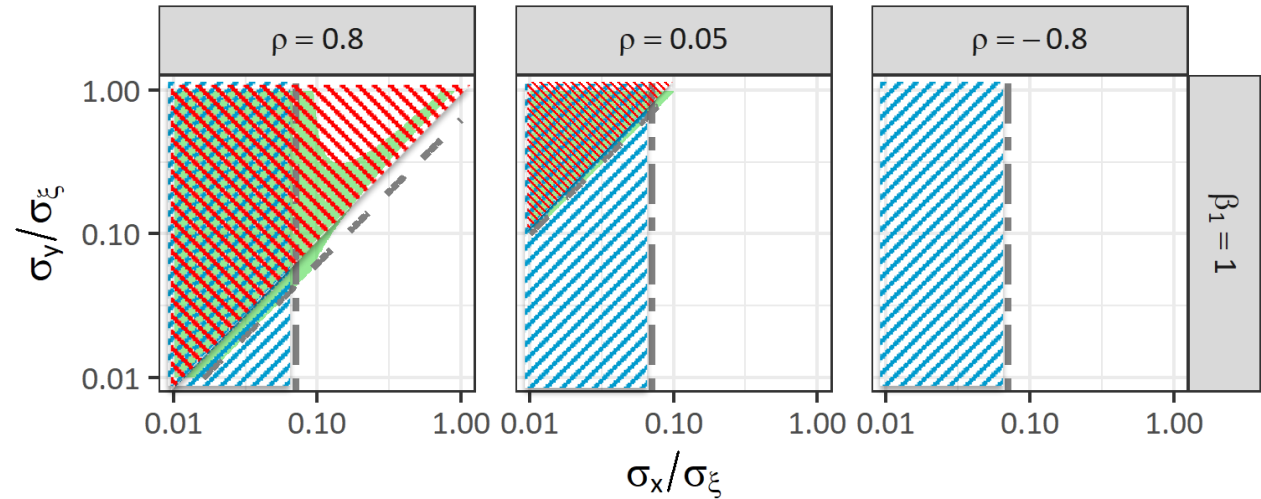
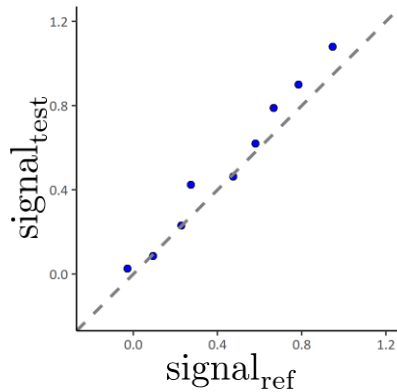
Gültigkeit von OLS in EIV Modellen



Gültigkeit von OLS in EIV Modellen



Gültigkeit von OLS in EIV Modellen

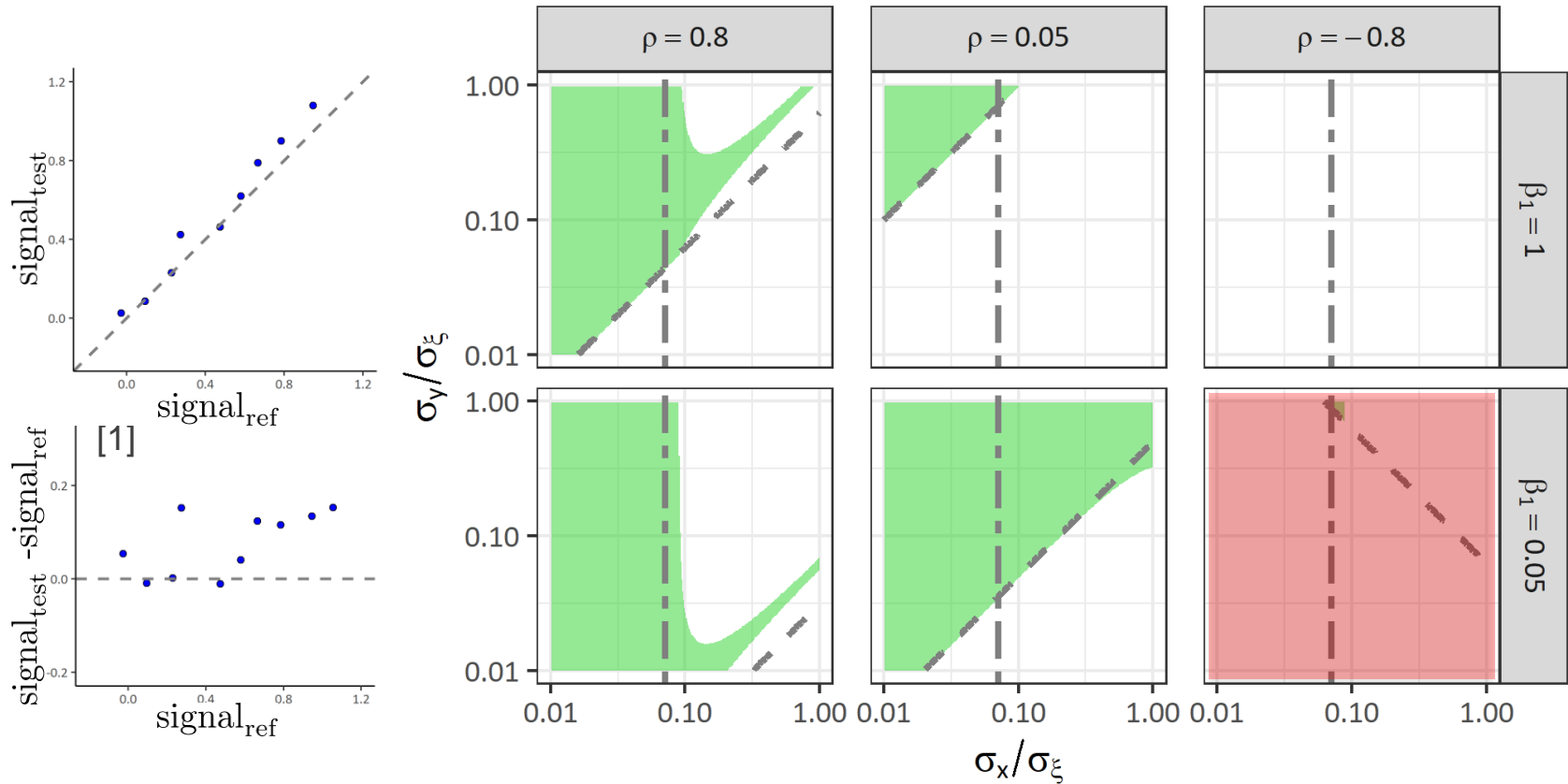


Abweichungsbedingung (2a) erfüllt



Varianzbedingung (2b) erfüllt

Gültigkeit von OLS in EIV Modellen



- „ $\sigma_{x,i}$ ist klein im Vergleich zu $\sigma_{y,i}$ “ ist nicht ausreichend
- OLS kann für EIV Modelle **nicht** empfohlen werden im Besonderen, wenn $\text{sgn}(\rho\beta_1) = -1$

[1] Bsp. H.3 in JCGM 100:2008; OIML R 111-1 e04; Bsp. 2 in VDI 2600 Blatt 2 (2019)

- z.B. ISO/TS 28037:2010 & DIN EN ISO 6143:2004 empfehlen WTLS

$$\begin{pmatrix} \hat{\beta} \\ \hat{\xi} \end{pmatrix} = \underset{\beta, \xi}{\operatorname{argmin}} \sum_{i=1}^N \mathbf{v}_i^\top \Sigma_i^{-1} \mathbf{v}_i \quad \text{with } \mathbf{v}_i = \begin{pmatrix} x_i - \xi_i \\ y_i - \beta_0 - \beta_1 \xi_i \end{pmatrix}, \beta = \begin{pmatrix} \beta_0 \\ \beta_1 \end{pmatrix}$$

- im Allgemeinen, nichtlineares Funktional nur numerisch minimierbar^[1,2]
 - Algorithmen sind kostenlos erhältlich
 - Unsicherheiten der Schätzwerte^[2,3]
- Liefert eine Unsicherheitsbestimmung nach GUM und GUM-S1 die gleichen Ergebnisse für die Schätzwerte?

- **oft:** bekannte Unsicherheiten der Referenzmethode oder des –signals werden ignoriert („ $\sigma_{x,i}$ ist klein im Verhältnis zu $\sigma_{y,i}$ “^[5])

gewöhnliche Methode der kleinsten Quadrate (OLS)

[1] C. A. Cantrell, *Atmos. Chem. Phys.* **8**, 5477 (2008); [2] ISO/TS 28037:2010

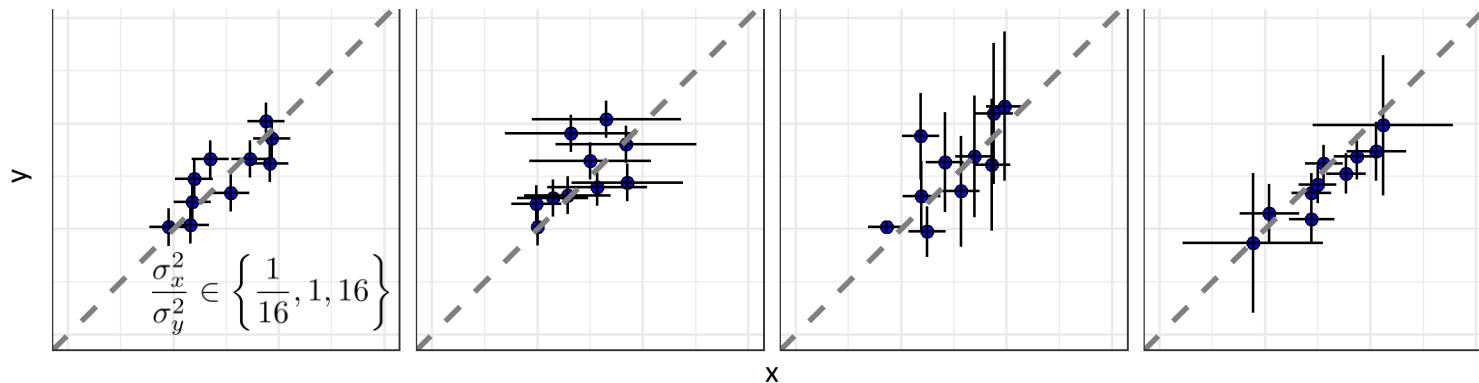
[3] M. Krystek and M. Anton, *Meas. Sci. Technol.* **22**, 035101 (2011); [4] T. Isobe et al. *Astrophys. Journal*, **364**, 104 (1990)

[5] ISO 11095:1996

Vergleich GUM von GUM-S1 für WTLS

- umfangreiche numerische Untersuchungen durchgeführt
- „synthetische Daten“ gemäß statist. Model (1a) - (1b) erzeugt

| N | ρ_i | $\sigma_0^2 / \sigma_\xi^2$ | MU Designs |
|-----------|----------------|-----------------------------|------------|
| {10, 100} | {-0.8, 0, 0.8} | {1.2%, 4.6%, 9.4%} | 6 |



- für jede Parameterkombination:
 $N_{\text{rep}} = 1000$ Datensätze + $N_{S1} = 5 \cdot 10^4$ S1 sub-samples (Monte-Carlo)

- ISO^[1] Implementierung wendet LPU auf linearisiertes Problem an

1) Überdeckungsintervalle und –wahrscheinlichkeit

- 95% Überdeckungsintervall nach GUM liefert 95% frequentistische Überdeckungswk.
- GUM-S1 führt zu etwas längeren Intervallen
 - Effekt wird für große Messunsicherheiten ($\sigma_{x,i}$, $\sigma_{y,i}$) verstärkt

2) Schätzwerte:

- GUM: Punktschätzer ist erwartungstreu
- GUM-S1 schätzt Anstieg β_1 etwas größer und β_0 etwas kleiner
 - Unterschied zwischen GUM und GUM-S1 nimmt mit N ab

^[1] ISO/TS 28037:2010 (E), *Determination and use of straight-line calibration functions*.

Vergleich GUM von GUM-S1 für WTLS

- ISO^[1] Implementierung wendet LPU auf linearisiertes Problem an

1) Überdeckungsintervalle und –wahrscheinlichkeit

- 95% Überdeckungsintervall nach GUM liefert 95% frequentistische Überdeckungswk.

2)

- befürworten die Nutzung der ISO 28037:2010 WTLS-Implementierung
- empfehlen die einfachere Unsicherheitsbestimmung nach Fortpflanzung der Unsicherheiten (**GUM**)

- GUM-S1 schätzt Anstieg β_1 etwas größer und β_0 etwas kleiner
 - Unterschied zwischen GUM und GUM-S1 nimmt mit N ab

^[1] ISO/TS 28037:2010 (E), *Determination and use of straight-line calibration functions*.

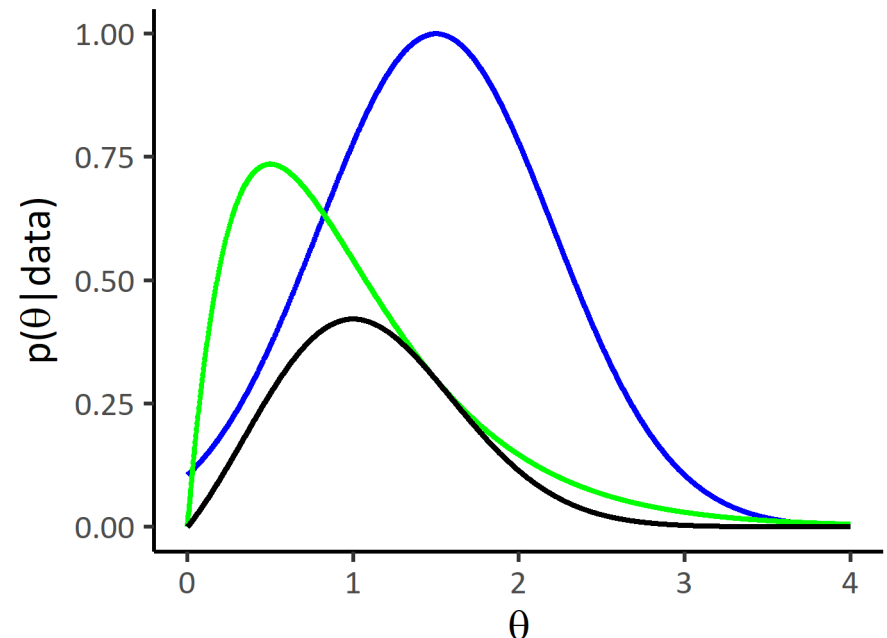
- GUM und GUM-S1 ermöglichen **nicht** die Berücksichtigung von Vorwissen
- Bayes'sche Regression ist allgemein anwendbar und flexibler
- nach Bayes Theorem, Posterior für Messgrößen $\theta = (\beta_0, \beta_1, \xi^T)$

$$p(\theta|\text{data}) \propto \pi_0(\theta) \mathcal{L}(\theta; \text{data})$$

mit Prior $\pi_0(\theta)$ und

likelihood $\mathcal{L}(\theta; \text{data})$

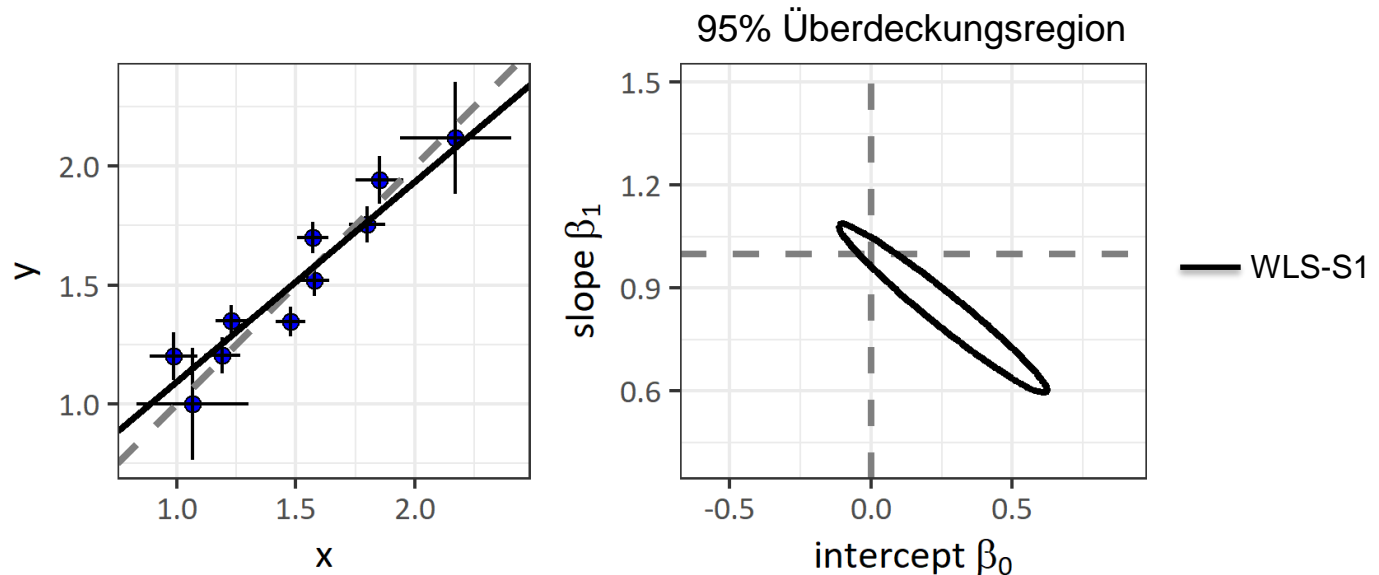
- Wann und ob überhaupt hat Bayes'sche Regression mit Prior-Wissen Vorteile im Vergleich mit GUM-S1?



Bayes'sche Regression

- wählen nicht informativen Prior für ξ und bivariat normal verteilten

Prior für β mit $(\mu_{\beta_0}, \mu_{\beta_1})^\top = (0, 1)^\top$ und $(\sigma_{\beta_0}^2, \sigma_{\beta_1}^2)^\top = q(u_{\hat{\beta}_0}^2 \text{WLS}, u_{\hat{\beta}_1}^2 \text{WLS})^\top$

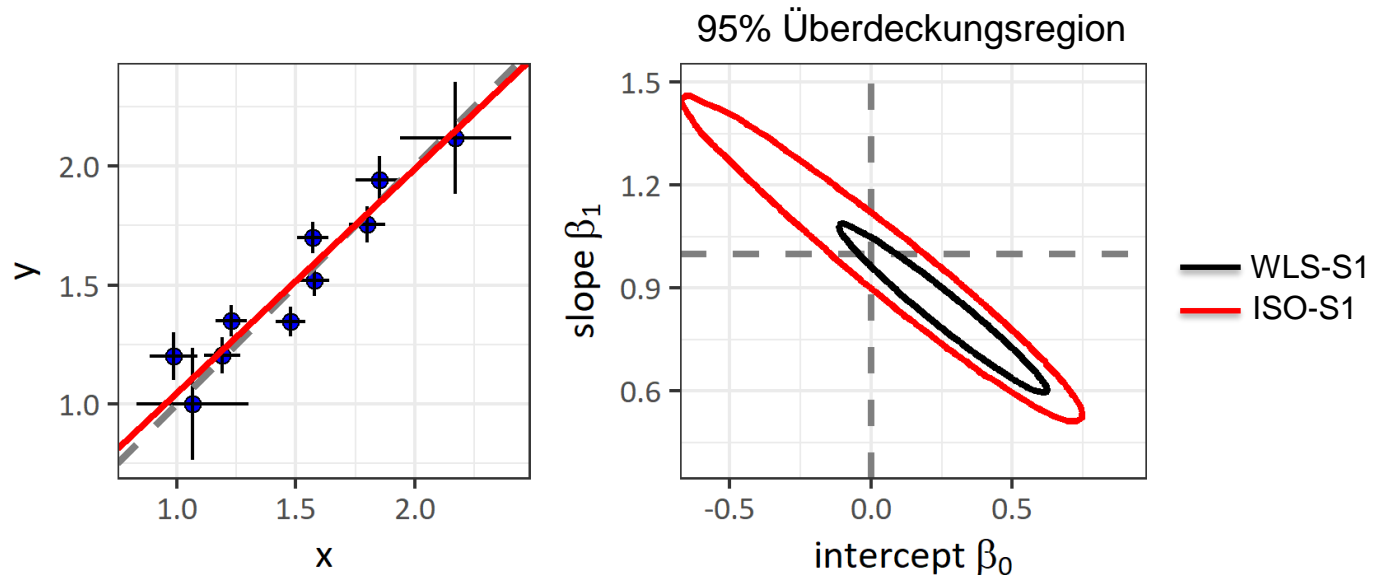


| 95% CI | least squares Schätzer | Bayes'sche Regression mit Normal Prior | | | |
|-----------|------------------------|--|--|--|--|
| | WLS – S1 | | | | |
| β_0 | (-0.03, 0.55) | | | | |
| β_1 | (0.64, 1.03) | | | | |

Bayes'sche Regression

- wählen nicht informativen Prior für ξ und bivariat normal verteilten

Prior für β mit $(\mu_{\beta_0}, \mu_{\beta_1})^\top = (0, 1)^\top$ und $(\sigma_{\beta_0}^2, \sigma_{\beta_1}^2)^\top = q(u_{\hat{\beta}_0}^2 \text{WLS}, u_{\hat{\beta}_1}^2 \text{WLS})^\top$

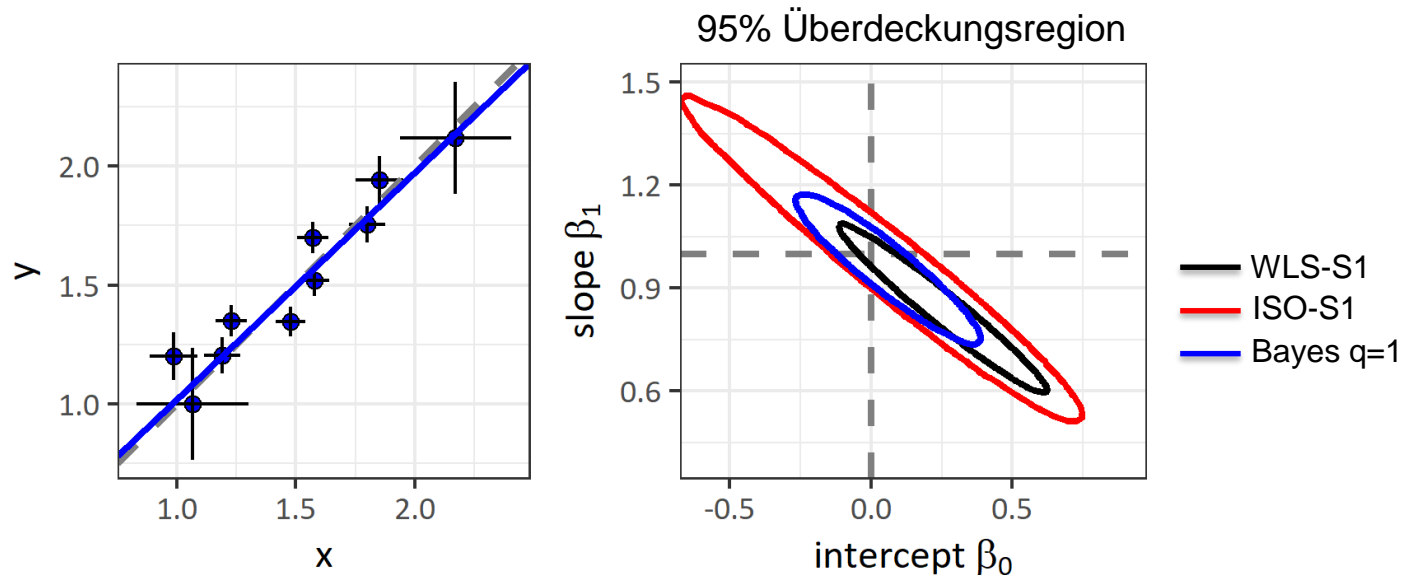


| 95% CI | least squares Schätzer | | Bayes'sche Regression mit Normal Prior | | |
|-----------|------------------------|---------------|--|--|--|
| | WLS – S1 | ISO – S1 | | | |
| β_0 | (-0.03, 0.55) | (-0.48, 0.64) | | | |
| β_1 | (0.64, 1.03) | (0.58, 1.34) | | | |

Bayes'sche Regression

- wählen nicht informativen Prior für ξ und bivariat normal verteilten

Prior für β mit $(\mu_{\beta_0}, \mu_{\beta_1})^\top = (0, 1)^\top$ und $(\sigma_{\beta_0}^2, \sigma_{\beta_1}^2)^\top = q(u_{\hat{\beta}_0}^2 \text{WLS}, u_{\hat{\beta}_1}^2 \text{WLS})^\top$

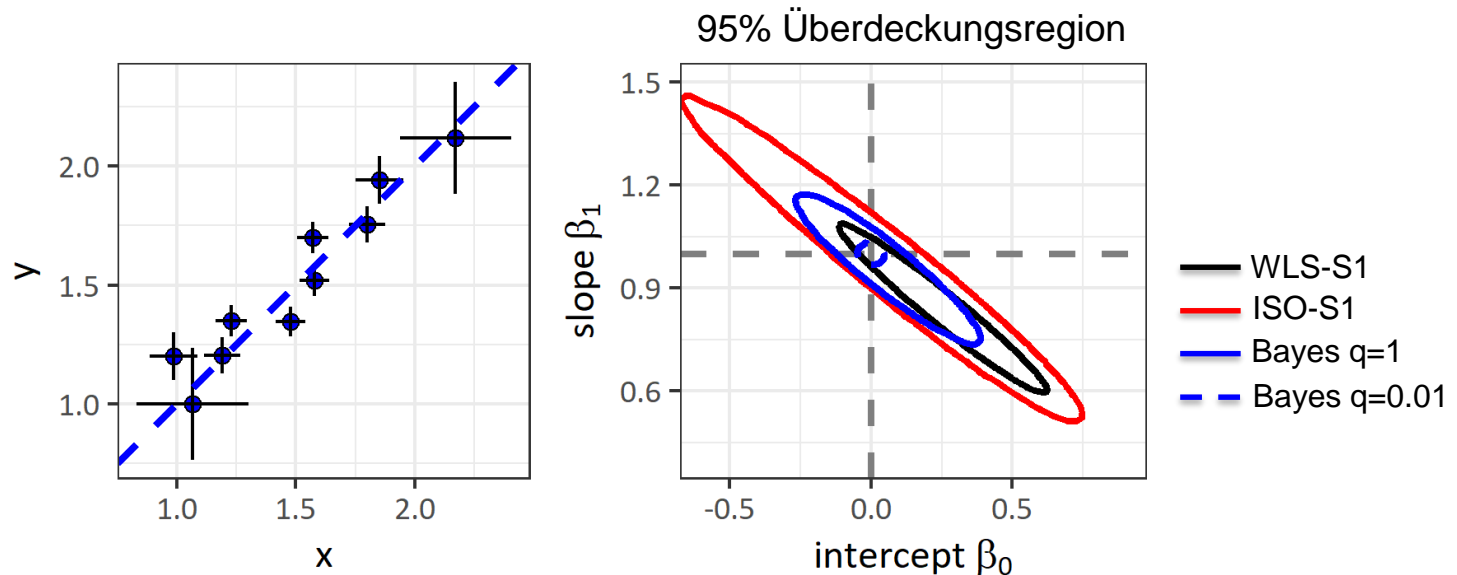


| 95% CI | least squares Schätzer | | Bayes'sche Regression mit Normal Prior | | |
|-----------|------------------------|--------------|--|--------------|--|
| | WLS – S1 | ISO – S1 | | q=1 | |
| β_0 | (-0.03,0.55) | (-0.48,0.64) | | (-0.19,0.33) | |
| β_1 | (0.64,1.03) | (0.58,1.34) | | (0.85,1.13) | |

Bayes'sche Regression

- wählen nicht informativen Prior für ξ und bivariat normal verteilten

Prior für β mit $(\mu_{\beta_0}, \mu_{\beta_1})^\top = (0, 1)^\top$ und $(\sigma_{\beta_0}^2, \sigma_{\beta_1}^2)^\top = q (u_{\hat{\beta}_0}^2 \text{WLS}, u_{\hat{\beta}_1}^2 \text{WLS})^\top$

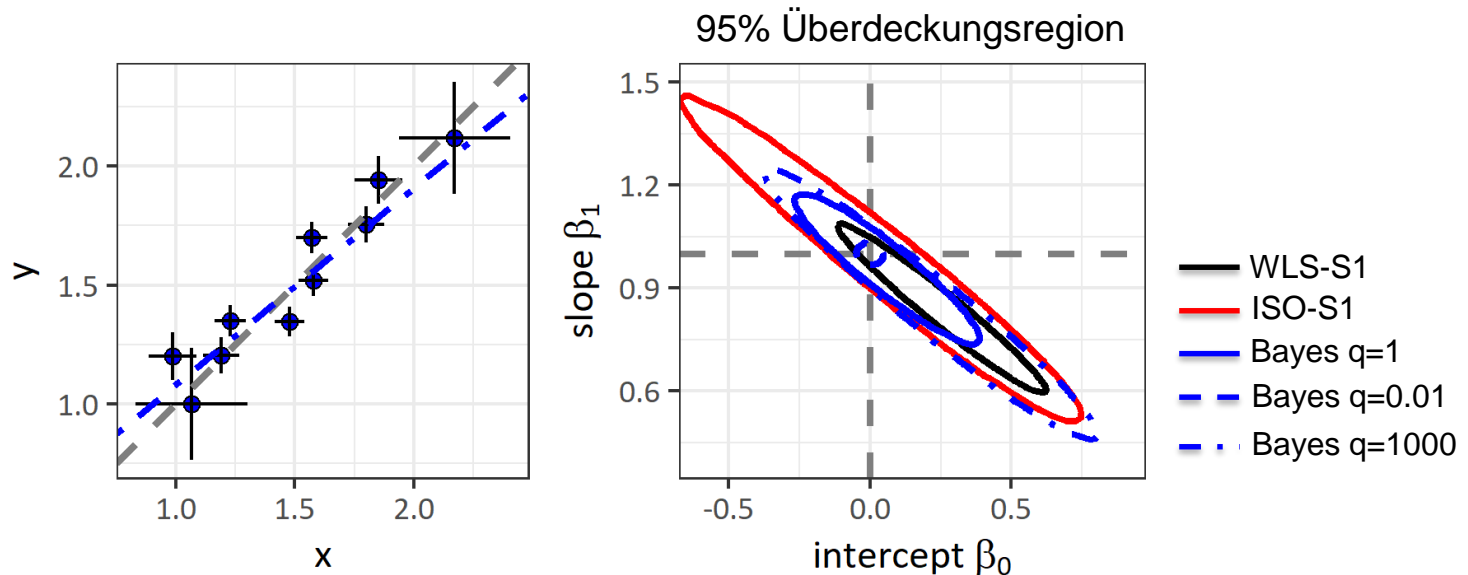


| 95% CI | least squares Schätzer | | Bayes'sche Regression mit Normal Prior | | |
|-----------|------------------------|--------------|--|--------------|--|
| | WLS – S1 | ISO – S1 | q=0.01 | q=1 | |
| β_0 | (-0.03,0.55) | (-0.48,0.64) | (-0.03,0.04) | (-0.19,0.33) | |
| β_1 | (0.64,1.03) | (0.58,1.34) | (0.98,1.03) | (0.85,1.13) | |

Bayes'sche Regression

- wählen nicht informativen Prior für ξ und bivariat normal verteilten

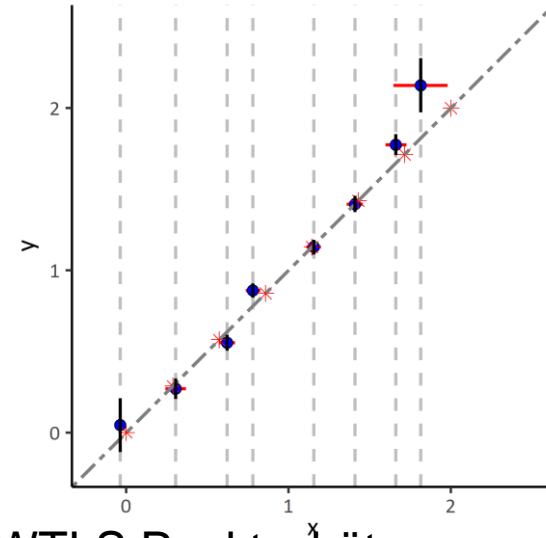
Prior für β mit $(\mu_{\beta_0}, \mu_{\beta_1})^\top = (0, 1)^\top$ und $(\sigma_{\beta_0}^2, \sigma_{\beta_1}^2)^\top = q (u_{\hat{\beta}_0}^2 \text{WLS}, u_{\hat{\beta}_1}^2 \text{WLS})^\top$



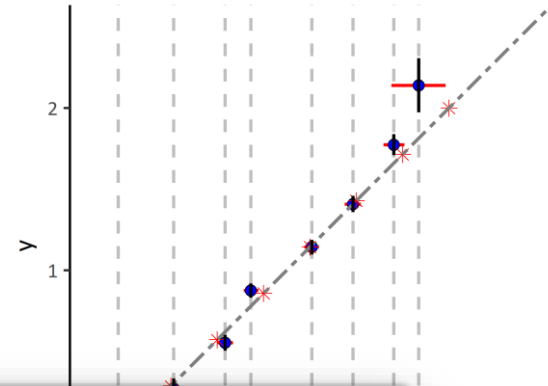
| 95% CI | least squares Schätzer | | Bayes'sche Regression mit Normal Prior | | |
|-----------|------------------------|--------------|--|--------------|--------------|
| | WLS – S1 | ISO – S1 | q=0.01 | q=1 | q=1000 |
| β_0 | (-0.03,0.55) | (-0.48,0.64) | (-0.03,0.04) | (-0.19,0.33) | (-0.24,0.71) |
| β_1 | (0.64,1.03) | (0.58,1.34) | (0.98,1.03) | (0.85,1.13) | (0.51,1.14) |

Zusammenfassung

- ☑ untersuchten lineare Regressionsverfahren bei Errors-in-Variables (EIV) Modellen
- ☑ Gültigkeit des OLS Punktschätzers
 - „ $\sigma_{x,i}$ ist klein im Vergleich zu $\sigma_{y,i}$ “ ist nicht ausreichend
 - OLS kann für EIV Modelle **nicht** empfohlen werden im Besonderen $\text{sgn}(\rho\beta_1) = -1$
- ☑ Unsicherheitsbestimmung nach GUM oder GUM-S1 für WTLS Punktschätzer
 - empfehlen die einfachere Unsicherheitsbestimmung nach GUM
 - Nutzung der ISO 28037:2010 WTLS-Implementierung
- ☑ Bayes'sche Regression mit informativen Prioren
 - bevorzugen, wenn ausreichende Vorkenntnisse vorhanden sind



- ☑ untersuchten lineare Regressionsverfahren bei Errors-in-Variables (EIV) Modellen
- ☑ Gültigkeit des OLS Punktschätzers
 - „ $\sigma_{x,i}$ ist klein im Vergleich zu $\sigma_{y,i}$ “ ist nicht ausreichend
 - OLS kann für EIV Modelle nicht empfohlen werden



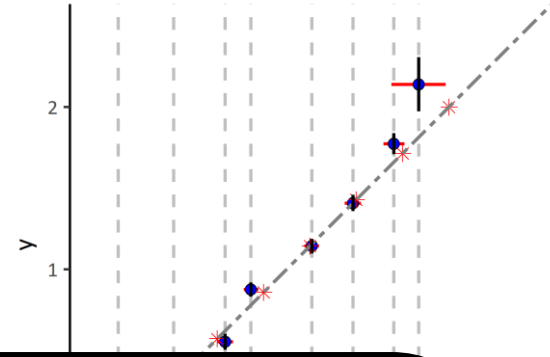
- ☑ Uns – Straight line regression in errors-in-variables models
– Comparison between the application of the GUM with its supplements and Bayesian analyses

Steffen Martens¹, Katy Klauenberg¹, Maurice G. Cox²,
Alen Bošnjaković³, John Greenwood⁴, Adriaan M. H. van
der Veen⁵, and Clemens Elster¹

- ☑ Bay submitted to Metrologia
 - bevorzugen, wenn ausreichende Vorkenntnisse vorhanden sind

Zusammenfassung

- ☑ untersuchten lineare Regressionsverfahren bei Errors-in-Variables (EIV) Modellen
- ☑ Gültigkeit des OLS Punktschätzers
 - „ $\sigma_{x,i}$ ist klein im Vergleich zu $\sigma_{y,i}$ “ ist nicht ausreichend



EMPIR



EURAMET

The EMPIR initiative is co-funded by the European Union's Horizon 2020 research and innovation programme and the EMPIR Participating States

Arbeit ist Teil des **E**xamples of **M**easurement **U**ncertainty **E**valuation Projekts und hat Mittel aus dem von den Teilnehmerstaaten kofinanzierten EMPIR-Programm und aus dem Forschungs- und Innovationsprogramm der Europäischen Union Horizon 2020 erhalten

Vielen Dank für Ihre Aufmerksamkeit!



Dr. Steffen Martens

Physikalisch-Technische Bundesanstalt
Braunschweig und Berlin
AG 8.42 Datenanalyse und Messunsicherheit

Abbestraße 2-12
10587 Berlin



steffen.martens@ptb.de