

# SIMEX

## die Simulation-Extrapolation Methode zur Messfehlerkorrektur

Andreas Hölzl

Seminarvortrag zum Thema Statistische  
Herausforderungen im Umgang mit fehlenden bzw.  
fehlerbehafteten Daten

Erklärung zu  
SIMEX

Intuitive Einführung  
zu SIMEX  
der  
SIMEX-Algorithmus  
Theorie zu SIMEX  
Simulationsstudien zu  
SIMEX

Messfehler

Kalibrierung  
Messwiederholung  
Modellbasiert

SIMEX bei semi-  
parametrischer  
Regression

Erklärung SIMEX bei  
semiparametrischer  
Regression  
Probleme bei SIMEX  
für semiparametrische  
Regression  
Anwendung auf echte  
Daten

Ausblick

## Erklärung zu SIMEX

- Intuitive Einführung zu SIMEX
- der SIMEX-Algorithmus
- Theorie zu SIMEX
- Simulationsstudien zu SIMEX

## Messfehler

- Kalibrierung
- Messwiederholung
- Modellbasiert

## SIMEX bei semi-parametrischer Regression

- Erklärung SIMEX bei semiparametrischer Regression
- Probleme bei SIMEX für semiparametrische Regression
- Anwendung auf echte Daten

## Ausblick

### Erklärung zu SIMEX

- Intuitive Einführung zu SIMEX
- der SIMEX-Algorithmus
- Theorie zu SIMEX
- Simulationsstudien zu SIMEX

### Messfehler

- Kalibrierung
- Messwiederholung
- Modellbasiert

### SIMEX bei semi-parametrischer Regression

- Erklärung SIMEX bei semiparametrischer Regression
- Probleme bei SIMEX für semiparametrische Regression
- Anwendung auf echte Daten

### Ausblick

## Erklärung zu SIMEX

Intuitive Einführung zu SIMEX

der SIMEX-Algorithmus

Theorie zu SIMEX

Simulationsstudien zu SIMEX

## Messfehler

Kalibrierung

Messwiederholung

Modellbasiert

## SIMEX bei semi-parametrischer Regression

Erklärung SIMEX bei semiparametrischer Regression

Probleme bei SIMEX für semiparametrische Regression

Anwendung auf echte Daten

## Ausblick

Erklärung zu  
SIMEX

**Intuitive Einführung  
zu SIMEX**

der  
SIMEX-Algorithmus

Theorie zu SIMEX

Simulationsstudien zu  
SIMEX

Messfehler

Kalibrierung

Messwiederholung

Modellbasiert

SIMEX bei semi-  
parametrischer  
Regression

Erklärung SIMEX bei  
semiparametrischer  
Regression

Probleme bei SIMEX  
für semiparametrische  
Regression

Anwendung auf echte  
Daten

Ausblick

# Einfluss eines Messfehlers in der unabhängigen Variable

- ▶ Bias in Parameterschätzung bei Messfehler in unabhängiger Variable
- ▶ beim linearen Regressionsmodell ergibt sich

$$\beta_{est} = \frac{\beta_{true}}{1 + \frac{\sigma}{\sigma_x}}$$

- ▶ lineares Modell als Beispiel aber allgemeiner anwendbar

Erklärung zu  
SIMEX

**Intuitive Einführung  
zu SIMEX**

der  
SIMEX-Algorithmus  
Theorie zu SIMEX  
Simulationsstudien zu  
SIMEX

Messfehler

Kalibrierung  
Messwiederholung  
Modellbasiert

SIMEX bei semi-  
parametrischer  
Regression

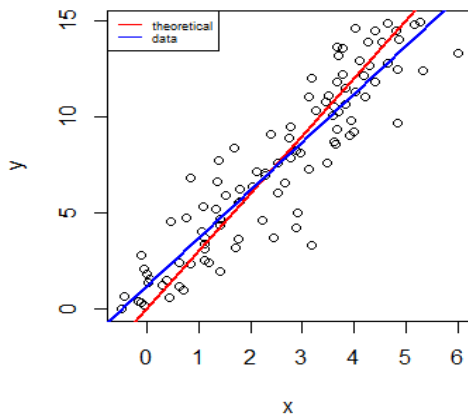
Erklärung SIMEX bei  
semiparametrischer  
Regression

Probleme bei SIMEX  
für semiparametrische  
Regression

Anwendung auf echte  
Daten

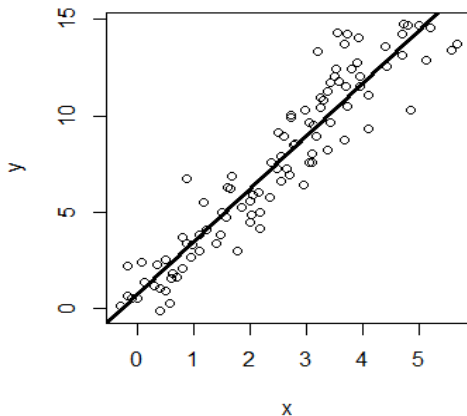
Ausblick

# Veranschaulichung des Einfluss des Messfehlers



Bias in der Parameterschätzung durch Messfehler in der unabhängigen Variable

# Schätzung des Modells aus Daten



im echten Modell ist nur der Schätzer aus den messfehlerbehafteten Daten bekannt



Erklärung zu  
SIMEX

**Intuitive Einführung  
zu SIMEX**

der  
SIMEX-Algorithmus  
Theorie zu SIMEX  
Simulationsstudien zu  
SIMEX

Messfehler

Kalibrierung  
Messwiederholung  
Modellbasiert

SIMEX bei semi-  
parametrischer  
Regression

Erklärung SIMEX bei  
semiparametrischer  
Regression  
Probleme bei SIMEX  
für semiparametrische  
Regression  
Anwendung auf echte  
Daten

Ausblick

# Bemerkung über Messfehler

- ▶ die Messfehlerverteilung als bekannt angenommen
- ▶ Später im Vortrag geht es um deren Bestimmung
- ▶ Hinzufügen von Messfehlern: aus dieser Messfehlerverteilung sampeln und auf die messfehlerbelasteten Variable addieren

Erklärung zu  
SIMEX

**Intuitive Einführung  
zu SIMEX**

der  
SIMEX-Algorithmus  
Theorie zu SIMEX  
Simulationsstudien zu  
SIMEX

Messfehler

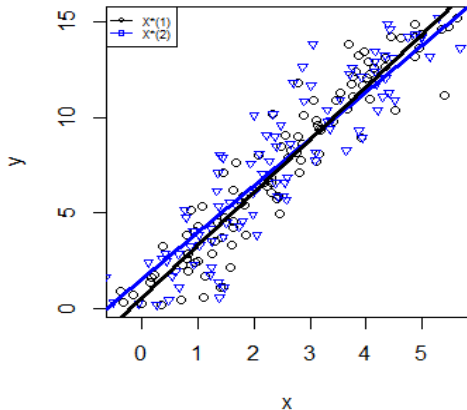
Kalibrierung  
Messwiederholung  
Modellbasiert

SIMEX bei semi-  
parametrischer  
Regression

Erklärung SIMEX bei  
semiparametrischer  
Regression  
Probleme bei SIMEX  
für semiparametrische  
Regression  
Anwendung auf echte  
Daten

Ausblick

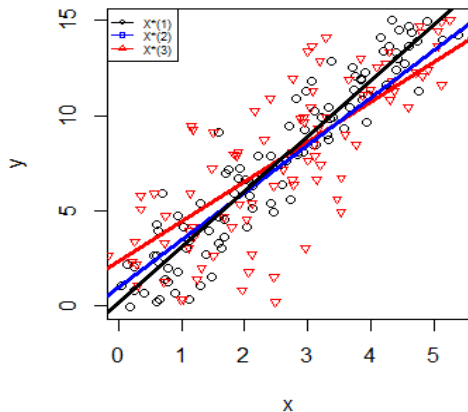
# Hinzufügen von Messfehlern - 1



Schätzung eines weiteren Modells mit weiterem hinzugefügtem Messfehler



# Veränderung der Parameter



Veränderung der Parameter bei weiteren hinzugefügten Messfehlern

# Veränderung der Parameterschätzer

Erklärung zu  
SIMEX

**Intuitive Einführung  
zu SIMEX**

der  
SIMEX-Algorithmus  
Theorie zu SIMEX  
Simulationsstudien zu  
SIMEX

Messfehler

Kalibrierung  
Messwiederholung  
Modellbasiert

SIMEX bei semi-  
parametrischer  
Regression

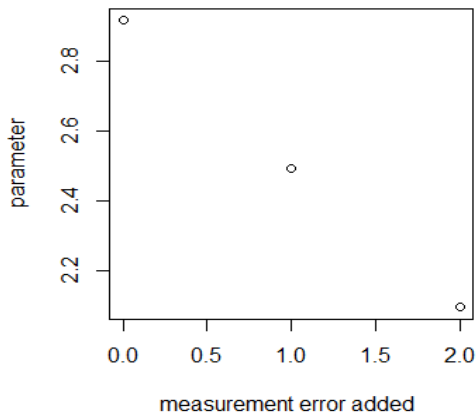
Erklärung SIMEX bei  
semiparametrischer  
Regression  
Probleme bei SIMEX  
für semiparametrische  
Regression  
Anwendung auf echte  
Daten

Ausblick

Parameter/ $\lambda$	0	1	2
Intercept	0.9976851	1.969627	3.077275
Steigung	2.6683550	2.394560	1.859049

Schätzung eines weiteren Modells mit weiterem hinzugefügtem  
Messfehler

# Visualisierung der Veränderung der Parameter



Veränderung der Parameter bei weiteren hinzugefügten  
Messfehlern

Erklärung zu  
SIMEX

**Intuitive Einführung  
zu SIMEX**

der  
SIMEX-Algorithmus  
Theorie zu SIMEX  
Simulationsstudien zu  
SIMEX

Messfehler

Kalibrierung  
Messwiederholung  
Modellbasiert

SIMEX bei semi-  
parametrischer  
Regression

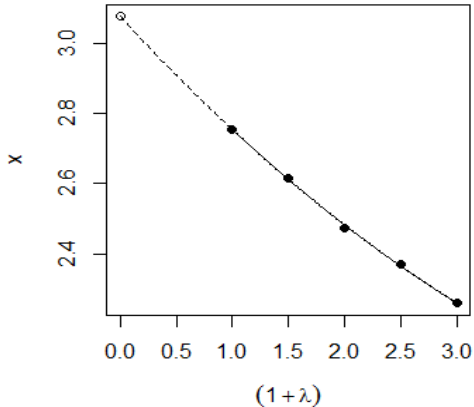
Erklärung SIMEX bei  
semiparametrischer  
Regression

Probleme bei SIMEX  
für semiparametrische  
Regression

Anwendung auf echte  
Daten

Ausblick

# Interpolation der Veränderung der Parameter



Interpolation der Veränderung der Parameter bei weiteren  
hinzugefügten Messfehlern

- ▶ Parameterschätzung von Ziehung aus Messfehlerverteilung abhängig
- ▶ Verbesserung durch mehrfache Ziehung und Mitteln des Schätzers
- ▶ hinzugefügte Messfehler nicht zwingend ganzzahlige Vielfache des geschätzten Messfehlers

## Erklärung zu SIMEX

Intuitive Einführung zu SIMEX

der SIMEX-Algorithmus

Theorie zu SIMEX

Simulationsstudien zu SIMEX

## Messfehler

Kalibrierung

Messwiederholung

Modellbasiert

## SIMEX bei semi-parametrischer Regression

Erklärung SIMEX bei semiparametrischer Regression

Probleme bei SIMEX für semiparametrische Regression

Anwendung auf echte Daten

## Ausblick

### Erklärung zu SIMEX

Intuitive Einführung zu SIMEX

der SIMEX-Algorithmus

Theorie zu SIMEX

Simulationsstudien zu SIMEX

### Messfehler

Kalibrierung

Messwiederholung

Modellbasiert

### SIMEX bei semi-parametrischer Regression

Erklärung SIMEX bei semiparametrischer Regression

Probleme bei SIMEX für semiparametrische Regression

Anwendung auf echte Daten

### Ausblick

- ▶ Simuliere Pseudo-Daten  $X_{b,i}(\lambda) = X_i^* + \sqrt{\lambda b} U_{b,i}$  mit  $U_{b,i}$  aus angenommener Messfehlerverteilung gezogen. Dann gilt  $\text{Var}(X_{b,i}) = \sigma_x^2 * (1 + \lambda)$
- ▶ Mache dies B mal
- ▶ Berechne Mittelwertschätzer für alle  $\lambda$  :  
$$\beta(\lambda) = \frac{1}{B} \sum_0^B \beta_{naive}(Y_i, X_{b,i}(\lambda))$$

Erklärung zu  
SIMEX

Intuitive Einführung  
zu SIMEX

der  
SIMEX-Algorithmus

Theorie zu SIMEX

Simulationsstudien zu  
SIMEX

Messfehler

Kalibrierung

Messwiederholung

Modellbasiert

SIMEX bei semi-  
parametrischer  
Regression

Erklärung SIMEX bei  
semiparametrischer  
Regression

Probleme bei SIMEX  
für semiparametrische  
Regression

Anwendung auf echte  
Daten

Ausblick

- ▶ Finde ein Modell, dass den Zusammenhang zwischen  $(\lambda_k, \beta_k(\lambda_k))$  modelliert
- ▶ Mache die Vorhersage dieses Modells für  $\beta(-1)$
- ▶ lineares Modell und quadratisches Modell werden oft verwendet, aber auch andere denkbar

## Erklärung zu SIMEX

Intuitive Einführung zu SIMEX

der SIMEX-Algorithmus

Theorie zu SIMEX

Simulationsstudien zu SIMEX

## Messfehler

Kalibrierung

Messwiederholung

Modellbasiert

## SIMEX bei semi-parametrischer Regression

Erklärung SIMEX bei semiparametrischer Regression

Probleme bei SIMEX für semiparametrische Regression

Anwendung auf echte Daten

## Ausblick



## Erklärung zu SIMEX

Intuitive Einführung zu SIMEX  
der SIMEX-Algorithmus

**Theorie zu SIMEX**

Simulationsstudien zu SIMEX

## Messfehler

Kalibrierung

Messwiederholung

Modellbasiert

## SIMEX bei semi-parametrischer Regression

Erklärung SIMEX bei semiparametrischer Regression

Probleme bei SIMEX für semiparametrische Regression

Anwendung auf echte Daten

## Ausblick

### Erklärung zu SIMEX

Intuitive Einführung  
zu SIMEX

der  
SIMEX-Algorithmus

**Theorie zu SIMEX**

Simulationsstudien zu  
SIMEX

### Messfehler

Kalibrierung

Messwiederholung

Modellbasiert

### SIMEX bei semi- parametrischer Regression

Erklärung SIMEX bei  
semiparametrischer  
Regression

Probleme bei SIMEX  
für semiparametrische  
Regression

Anwendung auf echte  
Daten

### Ausblick

- ▶ von Cook und Stefanski in (Cook&Stefanski, 1994) vorgestellt, zunächst ohne formalen Beweis
- ▶ In (Carroll&Küchenhoff et al. 1996) wurde die asymptotische Verteilung der SIMEX-Schätzung bei Schätzgleichungen theoretisch hergeleitet

## Erklärung zu SIMEX

Intuitive Einführung zu SIMEX

der SIMEX-Algorithmus

**Theorie zu SIMEX**

Simulationsstudien zu SIMEX

## Messfehler

Kalibrierung

Messwiederholung

Modellbasiert

## SIMEX bei semi-parametrischer Regression

Erklärung SIMEX bei semiparametrischer Regression

Probleme bei SIMEX für semiparametrische Regression

Anwendung auf echte Daten

## Ausblick

## Erklärung zu SIMEX

Intuitive Einführung zu SIMEX  
der SIMEX-Algorithmus

Theorie zu SIMEX

Simulationsstudien zu SIMEX

## Messfehler

Kalibrierung

Messwiederholung

Modellbasiert

## SIMEX bei semi-parametrischer Regression

Erklärung SIMEX bei semiparametrischer Regression

Probleme bei SIMEX für semiparametrische Regression

Anwendung auf echte Daten

## Ausblick

### Erklärung zu SIMEX

Intuitive Einführung zu SIMEX  
der SIMEX-Algorithmus  
Theorie zu SIMEX

### Simulationsstudien zu SIMEX

### Messfehler

Kalibrierung  
Messwiederholung  
Modellbasiert

### SIMEX bei semi-parametrischer Regression

Erklärung SIMEX bei semiparametrischer Regression  
Probleme bei SIMEX für semiparametrische Regression  
Anwendung auf echte Daten

### Ausblick

- ▶ (Shang, 2012) Simex zur Messfehlerbehebung in Quantilsregression
- ▶ (Hu&Liang, 2012) Simex bei longitudinalem multi-state Modell
- ▶ (Mallick&Fung, 2002) Simex funktioniert nicht beim Cox-Modell

## Erklärung zu SIMEX

Intuitive Einführung zu SIMEX  
der SIMEX-Algorithmus  
Theorie zu SIMEX  
**Simulationsstudien zu SIMEX**

## Messfehler

Kalibrierung  
Messwiederholung  
Modellbasiert

## SIMEX bei semi-parametrischer Regression

Erklärung SIMEX bei semiparametrischer Regression  
Probleme bei SIMEX für semiparametrische Regression  
Anwendung auf echte Daten

## Ausblick

- ▶ allgemeine Überprüfbarkeit der Wirkung von Simex über Simulationsstudie möglich
- ▶ simuliere ein wahres Modell mit allen Annahmen
- ▶ füge Messfehler hinzu, korrigiere mit SIMEX
- ▶ Vergleich von theoretischem und gefundenem Parameter

- ▶ Messfehlerbehandlung bei SIMEX funktioniert nur bei bekannter Messfehlerverteilung
- ▶ das gestaltet sich oft aber schwierig herauszufinden
- ▶ im Folgenden werden drei Möglichkeiten vorgestellt wie es trotzdem geht

## Erklärung zu SIMEX

Intuitive Einführung zu SIMEX  
der SIMEX-Algorithmus  
Theorie zu SIMEX  
Simulationsstudien zu SIMEX

## Messfehler

Kalibrierung  
Messwiederholung  
Modellbasiert

## SIMEX bei semi-parametrischer Regression

Erklärung SIMEX bei semiparametrischer Regression  
Probleme bei SIMEX für semiparametrische Regression  
Anwendung auf echte Daten

## Ausblick

## Erklärung zu SIMEX

Intuitive Einführung zu SIMEX  
der SIMEX-Algorithmus  
Theorie zu SIMEX  
Simulationsstudien zu SIMEX

## Messfehler

Kalibrierung  
Messwiederholung  
Modellbasiert

## SIMEX bei semi-parametrischer Regression

Erklärung SIMEX bei semiparametrischer Regression  
Probleme bei SIMEX für semiparametrische Regression  
Anwendung auf echte Daten

## Ausblick

## Erklärung zu SIMEX

Intuitive Einführung zu SIMEX  
der SIMEX-Algorithmus  
Theorie zu SIMEX  
Simulationsstudien zu SIMEX

## Messfehler

### Kalibrierung

Messwiederholung  
Modellbasiert

## SIMEX bei semi-parametrischer Regression

Erklärung SIMEX bei semiparametrischer Regression  
Probleme bei SIMEX für semiparametrische Regression  
Anwendung auf echte Daten

## Ausblick

- ▶ wahre Werte einer Messung bekannt
- ▶ messe diese Messung mehrfach mit der messfehlerbelasteten Methode
- ▶ daraus lässt sich dann die Verteilung der Messfehler schätzen

Erklärung zu  
SIMEX

Intuitive Einführung  
zu SIMEX  
der  
SIMEX-Algorithmus  
Theorie zu SIMEX  
Simulationsstudien zu  
SIMEX

Messfehler

**Kalibrierung**

Messwiederholung  
Modellbasiert

SIMEX bei semi-  
parametrischer  
Regression

Erklärung SIMEX bei  
semiparametrischer  
Regression  
Probleme bei SIMEX  
für semiparametrische  
Regression  
Anwendung auf echte  
Daten

Ausblick



## Erklärung zu SIMEX

Intuitive Einführung zu SIMEX  
der SIMEX-Algorithmus  
Theorie zu SIMEX  
Simulationsstudien zu SIMEX

## Messfehler

Kalibrierung  
**Messwiederholung**  
Modellbasiert

## SIMEX bei semi-parametrischer Regression

Erklärung SIMEX bei semiparametrischer Regression  
Probleme bei SIMEX für semiparametrische Regression  
Anwendung auf echte Daten

## Ausblick

### Erklärung zu SIMEX

Intuitive Einführung zu SIMEX  
der SIMEX-Algorithmus  
Theorie zu SIMEX  
Simulationsstudien zu SIMEX

### Messfehler

Kalibrierung  
**Messwiederholung**  
Modellbasiert

### SIMEX bei semi-parametrischer Regression

Erklärung SIMEX bei semiparametrischer Regression  
Probleme bei SIMEX für semiparametrische Regression  
Anwendung auf echte Daten

### Ausblick

# Messwiederholung - 1

- ▶ in (Devanarayana&Stefanski, 2002) wurde eine Methode vorgestellt, mit der Simex direkt auf Messwiederholungen angewendet wird
- ▶ es soll gelten  $Var(X_{b,i}) = \sigma_x^2 * (1 + \lambda)$
- ▶ dies soll durch Bildung einer Linearkombination der bekannten Punkte erreicht werden

$$X_{b,i}(\lambda) = \sum_{j=1}^{m_i} c_{bi} X_{ij}$$

Erklärung zu  
SIMEX

Intuitive Einführung  
zu SIMEX  
der  
SIMEX-Algorithmus  
Theorie zu SIMEX  
Simulationsstudien zu  
SIMEX

Messfehler

Kalibrierung  
**Messwiederholung**  
Modellbasiert

SIMEX bei semi-  
parametrischer  
Regression

Erklärung SIMEX bei  
semiparametrischer  
Regression  
Probleme bei SIMEX  
für semiparametrische  
Regression  
Anwendung auf echte  
Daten

Ausblick

- ▶ nach (Devanarayana&Stefanski, 2002) geht das wenn

$$\sum_1^m c_i = 1$$

und

$$\sum_1^m c_i^2 = \frac{1 + \lambda}{m}$$

der Fall ist

- ▶ es ist auch möglich die  $c_i$  computational effizient zu ziehen

Erklärung zu  
SIMEX

Intuitive Einführung  
zu SIMEX  
der  
SIMEX-Algorithmus  
Theorie zu SIMEX  
Simulationsstudien zu  
SIMEX

Messfehler

Kalibrierung  
**Messwiederholung**  
Modellbasiert

SIMEX bei semi-  
parametrischer  
Regression

Erklärung SIMEX bei  
semiparametrischer  
Regression  
Probleme bei SIMEX  
für semiparametrische  
Regression  
Anwendung auf echte  
Daten

Ausblick

## Erklärung zu SIMEX

Intuitive Einführung zu SIMEX  
der SIMEX-Algorithmus  
Theorie zu SIMEX  
Simulationsstudien zu SIMEX

## Messfehler

Kalibrierung  
Messwiederholung  
**Modellbasiert**

## SIMEX bei semi-parametrischer Regression

Erklärung SIMEX bei semiparametrischer Regression  
Probleme bei SIMEX für semiparametrische Regression  
Anwendung auf echte Daten

## Ausblick

### Erklärung zu SIMEX

Intuitive Einführung zu SIMEX  
der SIMEX-Algorithmus  
Theorie zu SIMEX  
Simulationsstudien zu SIMEX

### Messfehler

Kalibrierung  
Messwiederholung  
**Modellbasiert**

### SIMEX bei semi-parametrischer Regression

Erklärung SIMEX bei semiparametrischer Regression  
Probleme bei SIMEX für semiparametrische Regression  
Anwendung auf echte Daten

### Ausblick

- ▶ in der Psychologie werden häufig latente Variablen wie beispielsweise Angst über mehrere Fragen eines Fragebogens gemessen
- ▶ jeder dieser Fragen bezieht sich auf eine andere Facette dieser gleichen latenten Variable
- ▶ jede Frage wird zwischen 1 (stimme gar nicht zu) und 5 (stimme ganz zu) beantwortet

## Erklärung zu SIMEX

Intuitive Einführung zu SIMEX  
der SIMEX-Algorithmus  
Theorie zu SIMEX  
Simulationsstudien zu SIMEX

## Messfehler

Kalibrierung  
Messwiederholung  
**Modellbasiert**

## SIMEX bei semi-parametrischer Regression

Erklärung SIMEX bei semiparametrischer Regression  
Probleme bei SIMEX für semiparametrische Regression  
Anwendung auf echte Daten

## Ausblick

- ▶ nach dem Modell der einfaktoriellen Faktoranalyse gilt dann

$$x_{ij} = \mu_j + L_j * f_i + u_{ij}$$

- ▶ Schätzung von  $L_j$  und  $f_i$  beispielsweise über ML bei Annahme einer Verteilung der  $u_{ij}$
- ▶  $x_{ij} = x_{ij}^{theor} + u_{ij}$  bei bekanntem L und  $f_i$  ist ein klassisches Messfehlerproblem und kann über SIMEX bestimmt werden
- ▶ die  $u_{ij}$  werden bei diesem Modell mitgeschätzt und somit ist auch die Messfehlerverteilung bekannt

Erklärung zu  
SIMEX

Intuitive Einführung  
zu SIMEX  
der  
SIMEX-Algorithmus  
Theorie zu SIMEX  
Simulationsstudien zu  
SIMEX

Messfehler

Kalibrierung  
Messwiederholung  
Modellbasiert

SIMEX bei semi-  
parametrischer  
Regression

Erklärung SIMEX bei  
semiparametrischer  
Regression  
Probleme bei SIMEX  
für semiparametrische  
Regression  
Anwendung auf echte  
Daten

Ausblick

## Erklärung zu SIMEX

- Intuitive Einführung zu SIMEX
- der SIMEX-Algorithmus
- Theorie zu SIMEX
- Simulationsstudien zu SIMEX

## Messfehler

- Kalibrierung
- Messwiederholung
- Modellbasiert

## SIMEX bei semi-parametrischer Regression

- Erklärung SIMEX bei semiparametrischer Regression
- Probleme bei SIMEX für semiparametrische Regression
- Anwendung auf echte Daten

## Ausblick

## Erklärung zu SIMEX

- Intuitive Einführung zu SIMEX
- der SIMEX-Algorithmus
- Theorie zu SIMEX
- Simulationsstudien zu SIMEX

## Messfehler

- Kalibrierung
- Messwiederholung
- Modellbasiert

## SIMEX bei semi-parametrischer Regression

- Erklärung SIMEX bei semiparametrischer Regression
- Probleme bei SIMEX für semiparametrische Regression
- Anwendung auf echte Daten

## Ausblick

- ▶ Splines sind Funktionen der Art  $f(x) = \sum_1^n \kappa_j B(X)$
- ▶ Grundlage sind Basisfunktionen  $B(X)$  mit verschiedenen Knotenpunkten, aus der die Designmatrix  $Z_{ij} = B_j(x_i)$  gebildet wird

## Erklärung zu SIMEX

Intuitive Einführung zu SIMEX  
der SIMEX-Algorithmus  
Theorie zu SIMEX  
Simulationsstudien zu SIMEX

## Messfehler

Kalibrierung  
Messwiederholung  
Modellbasiert

## SIMEX bei semiparametrischer Regression

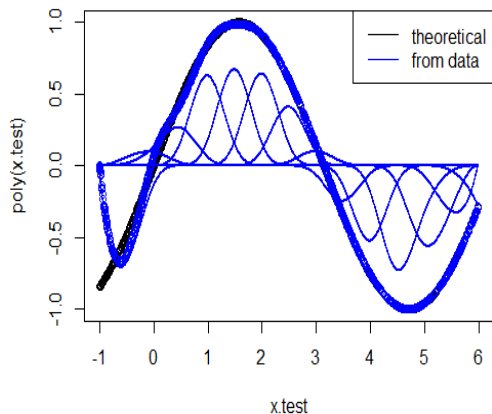
**Erklärung SIMEX bei semiparametrischer Regression**

Probleme bei SIMEX für semiparametrische Regression  
Anwendung auf echte Daten

## Ausblick



# Visualisierung semiparametrische Regression



## Erklärung zu SIMEX

Intuitive Einführung zu SIMEX  
der SIMEX-Algorithmus  
Theorie zu SIMEX  
Simulationsstudien zu SIMEX

## Messfehler

Kalibrierung  
Messwiederholung  
Modellbasiert

## SIMEX bei semiparametrischer Regression

**Erklärung SIMEX bei semiparametrischer Regression**

Probleme bei SIMEX für semiparametrische Regression  
Anwendung auf echte Daten

## Ausblick

## Visualisierung von Splines



- ▶ minimiere  $(y - Z\gamma)(y - Z\gamma)$
- ▶ um Overfitting zu vermeiden gibt es einen Penalisierungparameter  $\lambda$
- ▶ die Minimierung von  $(y - Z\gamma)(y - Z\gamma) + \kappa\gamma K\gamma$  führt dann zur Parametervektor  $\gamma = (Z^T Z + \kappa K)^{-1} Z^T y$

## Erklärung zu SIMEX

Intuitive Einführung zu SIMEX  
der SIMEX-Algorithmus  
Theorie zu SIMEX  
Simulationsstudien zu SIMEX

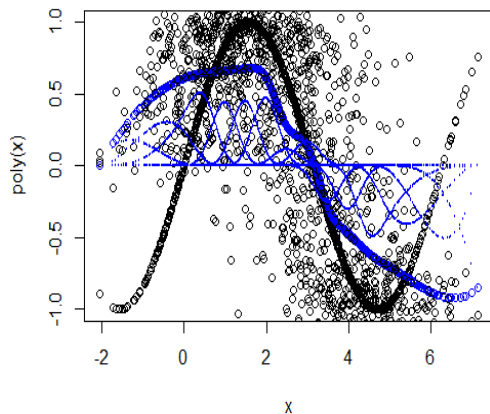
## Messfehler

Kalibrierung  
Messwiederholung  
Modellbasiert

## SIMEX bei semiparametrischer Regression

Erklärung SIMEX bei semiparametrischer Regression  
Probleme bei SIMEX für semiparametrische Regression  
Anwendung auf echte Daten

## Ausblick



Splines bei messfehlerbehafteter abhängiger Variable

Erklärung zu  
SIMEX

Intuitive Einführung  
zu SIMEX

der  
SIMEX-Algorithmus

Theorie zu SIMEX

Simulationsstudien zu  
SIMEX

Messfehler

Kalibrierung

Messwiederholung

Modellbasiert

SIMEX bei semi-  
parametrischer  
Regression

Erklärung SIMEX bei  
semiparametrischer  
Regression

Probleme bei SIMEX  
für semiparametrische  
Regression

Anwendung auf echte  
Daten

Ausblick

# Idee für Simex bei semiparametrischer Regression

- ▶ simuliere Pseudo-daten für verschiedene  $\lambda$ :  
 $X_{b,i}(\lambda) = X_i^* + \sqrt{\lambda_b} U_{b,i}$  mit  $U_{b,i}$  aus angenommener Messfehlerverteilung

- ▶ berechne

$$\gamma_b = (Z^T Z + \kappa K)^{-1} Z^T y$$

und middle für jedes einzelne  $\lambda$  die  $\gamma$  über die B Ziehungen

- ▶ erstelle für jedes einzelne  $\gamma$  ein Modell des Zusammenhangs zwischen  $\lambda$  und  $\gamma$  und extrapoliere zu  $\lambda = -1$

Erklärung zu  
SIMEX

Intuitive Einführung  
zu SIMEX  
der  
SIMEX-Algorithmus  
Theorie zu SIMEX  
Simulationsstudien zu  
SIMEX

Messfehler

Kalibrierung  
Messwiederholung  
Modellbasiert

SIMEX bei semi-  
parametrischer  
Regression

Erklärung SIMEX bei  
semiparametrischer  
Regression

Probleme bei SIMEX  
für semiparametrische  
Regression  
Anwendung auf echte  
Daten

Ausblick

## Erklärung zu SIMEX

Intuitive Einführung zu SIMEX  
der SIMEX-Algorithmus  
Theorie zu SIMEX  
Simulationsstudien zu SIMEX

## Messfehler

Kalibrierung  
Messwiederholung  
Modellbasiert

## SIMEX bei semi-parametrischer Regression

Erklärung SIMEX bei semiparametrischer Regression  
Probleme bei SIMEX für semiparametrische Regression  
Anwendung auf echte Daten

## Ausblick

## Erklärung zu SIMEX

Intuitive Einführung zu SIMEX  
der SIMEX-Algorithmus  
Theorie zu SIMEX  
Simulationsstudien zu SIMEX

## Messfehler

Kalibrierung  
Messwiederholung  
Modellbasiert

## SIMEX bei semi-parametrischer Regression

Erklärung SIMEX bei semiparametrischer Regression  
**Probleme bei SIMEX für semiparametrische Regression**  
Anwendung auf echte Daten

## Ausblick

# Probleme bei Simex bei semiparametrischer Regression - 1

- ▶ für Simex müssen Knotenpunkte bestimmt werden, diese müssen jedesmal gleich sein
- ▶ dann aber müsste  $\gamma_b = (Z^T Z + \kappa K)^{-1} Z^T y$  eine Schätzgleichung sein und der Schätzer asymptotisch richtig sein
- ▶ aber: Problem der Schätzung des Glättungsparameters  $\kappa$ . Normalerweise geht dies über Kreuzvalidierung, hier liegen aber verschiedene Daten je nach hinzuaddierten Messfehler vor

Erklärung zu  
SIMEX

Intuitive Einführung  
zu SIMEX  
der  
SIMEX-Algorithmus  
Theorie zu SIMEX  
Simulationsstudien zu  
SIMEX

Messfehler

Kalibrierung  
Messwiederholung  
Modellbasiert

SIMEX bei semi-  
parametrischer  
Regression

Erklärung SIMEX bei  
semiparametrischer  
Regression

Probleme bei SIMEX  
für semiparametrische  
Regression

Anwendung auf echte  
Daten

Ausblick

# Probleme bei Simex bei semiparametrischer Regression -2

- ▶ Bestimmung der  $\kappa$  für jeden Messfehler. In (Carroll&Küchenhoff et al. 1996) gezeigten asymptotischen Eigenschaften gelten nicht.
- ▶ Bestimmung der  $\kappa$  am gegebenen Datensatz. Dann ist  $\kappa$  nicht angebracht für andere Datensätze

## Erklärung zu SIMEX

Intuitive Einführung zu SIMEX  
der SIMEX-Algorithmus  
Theorie zu SIMEX  
Simulationsstudien zu SIMEX

## Messfehler

Kalibrierung  
Messwiederholung  
Modellbasiert

## SIMEX bei semiparametrischer Regression

Erklärung SIMEX bei semiparametrischer Regression

## Probleme bei SIMEX für semiparametrische Regression

Anwendung auf echte Daten

## Ausblick

# Probleme bei Simex bei semiparametrischer Regression -3

- ▶ In (Carroll&Ruppert, 2004) werden beide Möglichkeiten vorgestellt aber keine Vergleichsstudie
- ▶ Hier wurde  $\lambda$  jedes mal über Kreuzvalidierung neu bestimmt
- ▶ aber nicht bekannt ob dies die bessere Vorgehensweise ist

## Erklärung zu SIMEX

Intuitive Einführung zu SIMEX  
der SIMEX-Algorithmus  
Theorie zu SIMEX  
Simulationsstudien zu SIMEX

## Messfehler

Kalibrierung  
Messwiederholung  
Modellbasiert

## SIMEX bei semiparametrischer Regression

Erklärung SIMEX bei semiparametrischer Regression

## Probleme bei SIMEX für semiparametrische Regression

Anwendung auf echte Daten

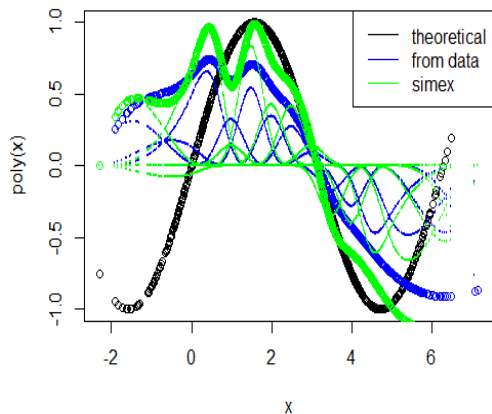
## Ausblick



# Visualisierung von semiparametrischem SIMEX

SIMEX

Andreas Hölzl



Erklärung zu  
SIMEX

Intuitive Einführung  
zu SIMEX  
der  
SIMEX-Algorithmus  
Theorie zu SIMEX  
Simulationsstudien zu  
SIMEX

Messfehler

Kalibrierung  
Messwiederholung  
Modellbasiert

SIMEX bei semi-  
parametrischer  
Regression

Erklärung SIMEX bei  
semiparametrischer  
Regression

Probleme bei SIMEX  
für semiparametrische  
Regression

Anwendung auf echte  
Daten

Ausblick

Splines bei messfehlerbehafteter abhängiger Variable



# Fazit Simex bei semiparametrischer Regression

- ▶ Testen von SIMEX auf semiparametrischer Regression in Simulationsstudie
- ▶ erste Ergebnisse weisen darauf hin, dass es zu einer Verbesserung des Modells führt
- ▶ keine asymptotische Theorie

Erklärung zu  
SIMEX

Intuitive Einführung  
zu SIMEX  
der  
SIMEX-Algorithmus  
Theorie zu SIMEX  
Simulationsstudien zu  
SIMEX

Messfehler

Kalibrierung  
Messwiederholung  
Modellbasiert

SIMEX bei semi-  
parametrischer  
Regression

Erklärung SIMEX bei  
semiparametrischer  
Regression

**Probleme bei SIMEX  
für semiparametrische  
Regression**

Anwendung auf echte  
Daten

Ausblick

## Erklärung zu SIMEX

- Intuitive Einführung zu SIMEX
- der SIMEX-Algorithmus
- Theorie zu SIMEX
- Simulationsstudien zu SIMEX

## Messfehler

- Kalibrierung
- Messwiederholung
- Modellbasiert

## SIMEX bei semi-parametrischer Regression

- Erklärung SIMEX bei semiparametrischer Regression
- Probleme bei SIMEX für semiparametrische Regression
- Anwendung auf echte Daten

## Ausblick

## Erklärung zu SIMEX

- Intuitive Einführung zu SIMEX
- der SIMEX-Algorithmus
- Theorie zu SIMEX
- Simulationsstudien zu SIMEX

## Messfehler

- Kalibrierung
- Messwiederholung
- Modellbasiert

## SIMEX bei semi-parametrischer Regression

- Erklärung SIMEX bei semiparametrischer Regression
- Probleme bei SIMEX für semiparametrische Regression
- Anwendung auf echte Daten**

## Ausblick

# Anwendung SIMEX bei semiparametrischer Regression

- ▶ Anwendung des Modells auf psychologischen Datensatz
- ▶ Untersuchung des Zusammenhangs von Angst vor Mathematik und der Mathematiknote
- ▶ die abhängigen Variablen liegen als Fragebogendaten vor
- ▶ per einfaktorieller Faktoranalyse wurde die latente Variable und die Messfehlerabschätzung bestimmt

## Erklärung zu SIMEX

Intuitive Einführung zu SIMEX  
der SIMEX-Algorithmus  
Theorie zu SIMEX  
Simulationsstudien zu SIMEX

## Messfehler

Kalibrierung  
Messwiederholung  
Modellbasiert

## SIMEX bei semiparametrischer Regression

Erklärung SIMEX bei semiparametrischer Regression  
Probleme bei SIMEX für semiparametrische Regression

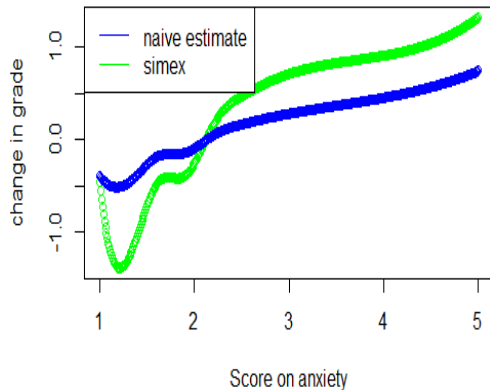
**Anwendung auf echte Daten**

## Ausblick

# Anwendung von semiparametrischen Simex

SIMEX

Andreas Hölzl



Erklärung zu  
SIMEX

Intuitive Einführung  
zu SIMEX

der  
SIMEX-Algorithmus

Theorie zu SIMEX

Simulationsstudien zu  
SIMEX

Messfehler

Kalibrierung

Messwiederholung

Modellbasiert

SIMEX bei semi-  
parametrischer  
Regression

Erklärung SIMEX bei  
semiparametrischer  
Regression

Probleme bei SIMEX  
für semiparametrische  
Regression

Anwendung auf echte  
Daten

Ausblick

Anwendung von semiparametrischen Simex auf echte Daten



- ▶ SEM Modelle in der Psychologie üblich
- ▶ meint Annahme linearen Zusammenhangs zwischen latenter Variable und Zielvariable
- ▶ In (Fahrmeir L.; Raach A., 2006) auch nichtlineare Modell vorgestellt
- ▶ Vergleich mit SIMEX-Verfahren sollte gemacht werden

## Erklärung zu SIMEX

Intuitive Einführung zu SIMEX  
der SIMEX-Algorithmus  
Theorie zu SIMEX  
Simulationsstudien zu SIMEX

## Messfehler

Kalibrierung  
Messwiederholung  
Modellbasiert

## SIMEX bei semiparametrischer Regression

Erklärung SIMEX bei semiparametrischer Regression  
Probleme bei SIMEX für semiparametrische Regression  
Anwendung auf echte Daten

## Ausblick

- ▶ mögliche Alternative zu semiparametrischer Regression wäre Kernel-Regression



$$E(Y|x) = \int y * \frac{f(x, y)}{f(x)} dy$$

- ▶ Problem auf messfehlerbehaftete Kerndichtebestimmung verschoben, als kernel deconvolution bekanntes Problem
- ▶ In Delaigle A. & Hall P., 2008 werden dazu mehrere Methoden miteinander verglichen
- ▶ Idee: Warum nicht nur  $f(x, y)$  mit einer dieser Methoden schätzen, daraus ziehen und dann Parameter ziehen?

Erklärung zu  
SIMEX

Intuitive Einführung  
zu SIMEX  
der  
SIMEX-Algorithmus  
Theorie zu SIMEX  
Simulationsstudien zu  
SIMEX

Messfehler




Kalibrierung  
Messwiederholung  
Modellbasiert

SIMEX bei semi-  
parametrischer  
Regression

Erklärung SIMEX bei  
semiparametrischer  
Regression  
Probleme bei SIMEX  
für semiparametrische  
Regression  
Anwendung auf echte  
Daten




Ausblick

# Bibliographie I




-  Cook, J.R. and Stefanski, L.A. (1994) Simulation-extrapolation estimation in parametric measurement error models. *Journal of American Statistical Association*, 89, 1314 - 1328
-  Carroll R., Küchenhoff, H. et al. Asymptotics for the SIMEX estimator in Nonlinear Measurement Error Model, *Journal of the American Statistical Association*; Mar 1996; 91, 433; *ABI/INFORM Global* pg. 242
-  Shang Y. (2012), Measurement Error Adjustment Using the SIMEX Method: An Application to Student Growth Percentiles, *Journal of Educational Measurement* Winter 2012, Vol. 49, No. 4, pp. 446–465




# Bibliographie II

-  Hu B, Liang L. (2012), Nonparametric Multi-state Representations of Survival and Longitudinal Data with Measurement Error, Stat Med. 2012 September 20; 31(21)
-  Mallick R, Fung K (2012) Adjusting for measurement error in the Cox proportional hazards regression model., J Cancer Epidemiol Prev. 2002;7(4):155-64
-  Devanarayana V.; Stefanski L.(2002) Empirical simulation extrapolation for measurement error models with replicate measurements, Statistics and Probability Letters 59 (2002) 219–225

# Bibliographie III

-  Carroll R.; Ruppert D.; Crainiceanu C..(2004), Nonlinear and Nonparametric Regression and Instrumental Variables, Journal of the American Statistical Association September 2004, Vol. 99, No. 467
-  Carroll R.; Maca C.; Ruppert D. (1999), Nonparametric regression in the presence of measurement error , Biometrika 86,3 pp. 541 - 554
-  Fahrmeir L.; Raach A., 2006, A Bayesian semiparametric latent variable model for mixed responses , Discussion paper // Sonderforschungsbereich 386 der Ludwig-Maximilians-Universität München, No. 471

-  Delaigle A. & Hall P., 2008, Using SIMEX for Smoothing-Parameter Choice in Errors-in-Variables Problems , Journal of the American Statistical Association March 2008, Vol. 103, No. 481, Theory and Methods DOI 10.1198/016214507000001355