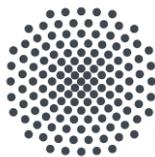


**SISS:**  
**Schriftenreihe des Instituts für Sozialwissenschaften  
der Universität Stuttgart**

**No. 50 / 2020**

***Verfahren der Multiplen Imputation bei  
Schätzung von Strukturgleichungsmodellen  
mit latenten Variablen. Ein systematischer  
Vergleich mittels Monte-Carlo-Simulationen.***

***Andreas Wahl  
Dieter Urban***



**Universität Stuttgart**  
**Abteilung für Soziologie  
und empirische Sozialforschung**



**SISS:  
Schriftenreihe  
des Instituts für Sozialwissenschaften  
der Universität Stuttgart:  
No. 50 / 2020**

Verfahren der Multiplen Imputation bei  
Schätzung von Strukturgleichungsmodellen  
mit latenten Variablen. Ein systematischer  
Vergleich mittels Monte-Carlo-Simulationen.<sup>1</sup>

Andreas Wahl  
Dieter Urban

**ISSN 2199-7780**

**Universität Stuttgart  
Institut für Sozialwissenschaften  
Abt. f. Soziologie u. empirische Sozialforschung  
70174 Stuttgart**

---

<sup>1</sup> Dieser SISS-Beitrag berichtet über Ergebnisse des SEMI-Projekts. Gefördert durch die Deutsche Forschungsgemeinschaft (DFG) – Projektnummer 370960346.

**Andreas Wahl und Dieter Urban**

**Verfahren der Multiplen Imputation bei Schätzung von Strukturgleichungsmodellen mit latenten Variablen. Ein systematischer Vergleich mittels Monte-Carlo-Simulationen.**

SISS – Schriftenreihe des Instituts für Sozialwissenschaften der Universität Stuttgart. No. 50/2020  
(<https://www.sowi.uni-stuttgart.de/dokumente/forschung/siss/SISS50.pdf>)

**Z U S A M M E N F A S S U N G:** Dieser SISS-Beitrag fasst die Ergebnisse aus dem DFG-geförderten Projekt „Monte-Carlo-Simulationen zur Evaluation der Leistungsfähigkeit moderner Missing Data Techniken bei Schätzung von Strukturgleichungsmodellen mit latenten Variablen. Eine systematische Analyse verschiedener Varianten der Multiplen Imputation“ zusammen. Im Projekt wurden mittels Monte-Carlo-Simulationstechniken (MC-Simulationstechniken) die Performanz verschiedener Varianten der Multiplen Imputation (MI) und MI-alternativer Verfahren zur Schätzung fehlender Werte im Kontext von Strukturgleichungsmodellierungen (SEM-Analyse) verglichen. Es wurden sechs Missing Data Techniken (MDTs) im Kontext von drei verschiedenen SEM-Populationsmodellen unter unterschiedlichen Simulationskonfigurationen getestet. Die variierten Konfigurationen ergaben sich aus: a) Datensätzen mit unterschiedlichen Fallzahlen, b) Datensätzen mit symmetrischen und (stark) asymmetrischen Variablenverteilungen, c) Datensätzen mit unterschiedlich hohen Anteilen an fehlenden Werten. Neben MI-Techniken mit strikten Annahmen zur Verteilung der Modellvariablen (multivariate Normalverteilung) wurden auch MI-Varianten getestet, welche dieser Annahme nicht unterliegen und kategoriale Variablen explizit im Verfahren berücksichtigen können. Zusätzlich zu den verschiedenen Varianten der Multiplen Imputation wurden zum Vergleich auch MI-alternative Verfahren eingesetzt (das „Direct Maximum Likelihood-Schätzverfahren“ sowie die „Expectation-Maximization-Methode“). Zur Bewertung der Performanz der verschiedenen MDTs wurden damit erreichbare SEM-Fit-Indices herangezogen (p-Wert der Chi<sup>2</sup>-Statistik, SRMR, RMSEA und CFI). Es wurden aber auch die Qualität der geschätzten SEM-Parameter und deren Standardfehler sowie die relative Effizienz der Parameterschätzungen untersucht.

Auf diese Weise konnten unter den sechs getesteten MDTs zwei Verfahren identifiziert werden, die bei allen konfigurierten Modell- und Datenbedingungen sehr gute Ergebnisse erbringen. Das sind das „Direct Maximum Likelihood-Schätzverfahren“ (Direct-ML-Verfahren) und eine Variante der MI, die bei der Imputation der fehlenden Werte die Modellstruktur des Analysemodells berücksichtigt: die H0-Technik. Beide erbringen neben sehr guten Ergebnissen bei den untersuchten SEM-Fit-Indices auch unverzerrte SEM-Parameterschätzungen und Standardfehler. Von den MI-Varianten kann somit allein die H0-Technik uneingeschränkt für den praktischen Einsatz empfohlen werden. Zudem kann als Nicht-MI-Variante die Direct-ML-Methode empfohlen werden. Sie hat sogar den Vorteil, dass die fehlenden Werte direkt bei der Modellschätzung berücksichtigt werden (ohne die fehlenden Werte separat imputieren zu müssen).

Alle anderen MDTs liefern zwar auch gute, unverzerrte SEM-Parameterschätzwerte und Standardfehler, aber sie generieren häufig SEM-Fit-Werte, die zur fälschlichen Ablehnung von geschätzten Strukturgleichungsmodellen führen. Zwar ist bei kleinen Missinganteilen die Modellbewertung anhand der Fit-Indices oftmals unproblematisch, aber bei höheren Anteilen (ab ca. 20 %) kann nur ein einziger Fit-Index uneingeschränkt empfohlen werden: der SRMR-Index (Standardized Root Mean-Square Residual Index).

**Andreas Wahl und Dieter Urban**

**Multiple imputation methods for estimating structural equation models with latent variables. A systematic comparison using Monte Carlo simulations.**

SISS – Schriftenreihe des Instituts für Sozialwissenschaften der Universität Stuttgart. No. 50/2020  
(<https://www.sowi.uni-stuttgart.de/dokumente/forschung/siss/SISS50.pdf>)

**A B S T R A C T:** This SISS-paper summarizes the results of the DFG-funded project "Monte Carlo simulations for evaluating the performance of modern missing data techniques in estimating structural equation models with latent variables. A systematic analysis of different variants of multiple imputation". In this project, Monte Carlo simulation (MC simulation) techniques were used to compare the performance of different variants of Multiple Imputation (MI) and other Non-Multiple Imputation methods for estimating missing values when analyzing structural equation models (SEM). In total, six missing data techniques (MDTs) applied to three different SEM-population models were investigated using various configurations for simulation. These configurations included a) data files with different numbers of cases, b) data files with symmetrical and (strong) asymmetrical value distributions, and c) data files with different proportions of missing data. Besides using MI techniques with strict assumptions of value distributions (multivariate normal distribution), we also tested MI variants which are not subject to these assumptions. For comparative reasons two Non-MI MDTs were applied (the "Direct Maximum Likelihood estimation" and the "Expectation-Maximization method"). For evaluating the performance of all six MDTs we focused on four different fit indices used most prominently in SEM analysis (p-value of the chi<sup>2</sup>-statistic, SRMR, RMSEA und CFI). We also analyzed the quality of all estimated SEM parameters and their standard errors as well as the relative efficiency of all estimated parameters.

Among the six tested missing data techniques only two techniques could be identified that deliver very good results under all model- and data configurations. These are the „Direct Maximum Likelihood estimation“ (Direct-ML-method) and a variant of MI that takes into account the model structure of the analyzed model when imputing the missing values: the H0-method. Both methods deliver high quality fit indices when applied to SEM estimation. They also deliver unbiased SEM parameter estimations and standard errors. Thus, when looking particularly at the MI variants, just the H0-method can be recommended for practical usage. In addition, the Direct ML-method (a Non-MI technique) can be recommended. The Direct ML-method integrates the process of estimating missing values into model estimation so that there is no need for an initially separate process of missing value imputation.

Although all the other MDTs deliver good and unbiased results when estimating SEM parameters and standard errors, they often generate SEM fit indices that lead to false rejections of SEMs. This is even more problematic when having data files with high proportions of missing values ( $\geq 20\%$ ). In situations like this, there is only one SEM fit index that can be fully recommended: the SRMR index (Standardized Root Mean-Square Residual index).

## Inhaltsverzeichnis

|  |    |
|--|----|
| Tabellenverzeichnis .....  | 3  |
| Abbildungsverzeichnis .....  | 3  |
| Abkürzungsverzeichnis .....  | 4  |
| 1 Einleitung .....   | 5  |
| 2 Maximum Likelihood-Schätzverfahren .....                                 | 6  |
| 2.1 Direct-ML .....  | 7  |
| 2.2 Expectation-Maximization (EM) .....                                    | 9  |
| 3 Verfahren der Multiplen Imputation .....                                 | 10 |
| 3.1 Data Augmentation (MNV) .....  | 11 |
| 3.2 Fully Conditional Specification (FCS) .....                            | 13 |
| 3.3 Zwei weitere Verfahren: EMB und H0 .....                               | 15 |
| 4 Zusammenfassung: Voraussetzungen und Annahmen der MDTs .....             | 18 |
| 5 Forschungsdesign .....   | 21 |
| 5.1 Rechentchnische Umsetzung des Forschungsdesigns .....                  | 24 |
| 6 Ergebnisse .....   | 25 |
| 6.1 Ergebnisse bzgl. Parameterschätzungen und Standardfehler .....         | 26 |
| 6.2 Ergebnisse bzgl. Fit-Indices .....                                     | 28 |
| 7 Diskussion der Ergebnisse und Handlungsempfehlungen für die Praxis ..... | 29 |
| 8 Literaturverzeichnis .....   | 31 |

## Tabellenverzeichnis

|  |    |
|--|----|
| Tabelle 1: Zusammenfassung von MDT-Annahmen (Teil 1) ..... | 20 |
| Tabelle 2: Zusammenfassung von MDT-Annahmen (Teil 2) ..... | 20 |
| Tabelle 3: Performanzkriterien .....                       | 24 |
| Tabelle 4: Handlungsempfehlungen .....                     | 30 |

## Abbildungsverzeichnis

|                                       |    |
|---------------------------------------|----|
| Abbildung 1: Populationsmodelle ..... | 22 |
| Abbildung 2: Verteilungen .....       | 23 |

## Abkürzungsverzeichnis

|               |  |
|---------------|--|
| CFI           | Comparative Fit Index (SEM-Anpassungsindex)  |
| Direct-ML     | Direct Maximum Likelihood-Verfahren (auch: Full Information Maximum Likelihood-Verfahren; kurz: FIML-Verfahren)              |
| EM            | Expectation-Maximization-Verfahren (einfaches Imputationsverfahren)  |
| EMB           | Multiple Imputationsmethode mit einem bootstrapping-basierten EM-Algorithmus   |
| FCS           | Multiple Imputationsmethode mit der Fully Conditional Specification-Technik (auch: Multiple Imputation by Chained Equations) |
| H0            | Multiple Imputationsmethode unter Berücksichtigung der Modellstruktur des Analysemodells durch Bayes-Schätzer                |
| MAR           | Missing at Random (Ausfallmechanismus bei Entstehung von Missing Values)   |
| MC            | Monte-Carlo-Simulationstechnik   |
| MDT           | Missing Data Technik   |
| MI            | Multiple Imputationsmethode  |
| ML            | Maximum Likelihood Schätzverfahren   |
| MLR           | Maximum Likelihood Schätzverfahren (Robust-Variante)   |
| MNV           | Multiple Imputationsmethode unter Annahme einer multivariaten Normalverteilung (Data Augmentation)                           |
| RMSEA         | Root Mean-Square Error of Approximation (SEM-Anpassungsindex)  |
| SEM/SE-Modell | Strukturgleichungsmodellierung/Strukturgleichungsmodell  |
| SRMR          | Standardized Root Mean-Square Residual (SEM-Anpassungsindex)   |

## 1 Einleitung<sup>2</sup>

Datensätze mit fehlenden Werten (Missing Values bzw. Missings) sind in der quantitativ-empirischen Sozialforschung ein allgegenwärtiges Problem. Insbesondere bei Befragungen in der Surveyforschung führen itemspezifische Antwortverweigerungen (item-nonresponses) zu unvollständigen Datensätzen. Je höher der Anteil an Missing Values in einem Datensatz ist, desto wahrscheinlicher wird es, dass dies Auswirkungen auf die statistischen Analysen hat: es kann dann in den Analysen zu Ergebnisverzerrungen und ggf. zu falschen inferenzstatistischen Schlüssen kommen.

Missing Values können einem *missing at random*-Ausfallmechanismus (MAR-Ausfallmechanismus) unterliegen. Unter einem MAR-Mechanismus wird verstanden, dass die fehlenden Werte einer Variablen Y abhängig von anderen, im Modell befindlichen Variablen X sind, nicht aber von Y selbst abhängig sind. Die fehlenden Werte in Y lassen sich dann auf beobachtete Werte in X zurückführen (siehe Little/Rubin 1987; vgl. Schafer 1997: 10). Weisen die Missing Values einen MAR-Ausfallmechanismus auf, dann lassen sich diese mit entsprechenden Methoden (auch: Missing Data Techniken<sup>3</sup>; MDTs) handhaben. Dafür wird in der Forschungsliteratur vor allem die Full Information Maximum Likelihood-Methode (FIML-Methode, im Weiteren: Direct-ML-Methode) und die Multiple Imputationsmethode (MI-Methode) empfohlen.

Allerdings gibt es in der Forschung nicht ‚die‘ bzw. nur eine einzige MI-Methode, sondern es werden verschiedene MI-Varianten eingesetzt, denen unterschiedliche Verfahrensalgorithmen zugrunde liegen. Vor allem im weit verbreiteten und wichtigen Forschungsbereich der Spezifikation und Schätzung von Strukturgleichungsmodellen zur Analyse von empirischen Daten (SEM-Forschung) werden oftmals diverse MI-Varianten eingesetzt, ohne über deren jeweilige Verfahrenslogik zu informieren. Jedoch ist davon auszugehen, dass jede MI-Variante für bestimmte Modellierungen und für bestimmte Datensatzstrukturen mehr oder weniger gut geeignet ist. Denn die den einzelnen MI-Varianten zugrundeliegenden Verfahrensalgorithmen haben verschiedenste Anwendungsvoraussetzungen und basieren auf unterschiedlichen Annahmen.

Aus diesem Grund wurden in dem Projekt, über dessen Ergebnisse hier berichtet wird (SEMI-Projekt), die in der empirischen Praxis wohl am häufigsten eingesetzten MI-Varianten mittels Monte-Carlo-Simulationen (MC-Simulationen) systematisch miteinander verglichen und hinsichtlich ihrer

---

<sup>2</sup> Teile dieses Beitrages wurden leicht modifiziert aber auch teilweise unverändert aus Wahl (2020) entnommen.

<sup>3</sup> Es werden im vorliegenden SISS-Bericht folgende Bezeichnungen benutzt: Als ‚Missing Data Techniken‘ (MDTs) werden alle Methoden zum statistischen Umgang mit fehlenden Werten bezeichnet; als ‚Methode‘ werden spezifische MDTs benannt; die Direct-ML-Methode wird auch als Direct-ML-(Schätz-)Verfahren oder Direct-ML-Schätzung bezeichnet; die EM-Methode wird auch als EM-Verfahren oder EM-Algorithmus bezeichnet; Ausformungen der Multiplen Imputationsmethode werden als ‚MI-Technik/Variante/Verfahren‘ bezeichnet. Ein Verzeichnis aller benutzten Abkürzungen befindet sich auf Seite 4 dieses SISS-Beitrags.

Leistungsfähigkeit bewertet. Zu den untersuchten MI-Varianten zählten neben der FCS-Technik (Fully Conditional Specification), die MNV-Technik (MI unter Annahme einer multivariaten Normalverteilung) sowie eine bootstrapping basierte Variante mittels EM-Algorithmus (EMB). Zusätzlich wurde eine weniger bekannte Variante in die Untersuchung aufgenommen: eine modellbasierte Technik mittels Bayes-Schätzer (H0). Für all diese MI-Varianten sollte mittels MC-Simulationen ermittelt werden, ob sie unter denselben Bedingungen dieselben Ergebnisse produzieren, oder ob bereits die Wahl einer bestimmten MI-Variante das Ergebnis einer SE-Modellschätzung in entscheidender Weise beeinflussen kann.

Folglich wurden im Projekt im Wesentlichen fünf Ziele verfolgt:

1. eine Untersuchung der Performanz der oben genannten MI-Varianten im Hinblick auf die Schätzung von SE-Modellparametern (Faktorladungen, Strukturpfade und Kovarianzen);
2. eine Untersuchung der Performanz der MI-Varianten im Hinblick auf die ermittelten Standardfehler der geschätzten Modellparameter;
3. ein Vergleich der MI-Varianten mit zwei weiteren MDTs, deren Schätzwerte als Vergleichsgrößen hinzugezogen wurden (eine Einfach-Imputation mittels dem EM-Algorithmus und ein Direct-ML-Verfahren);
4. eine Untersuchung der Performanz der MI-Varianten im Hinblick auf diverse SEM-Anpassungstests (Chi<sup>2</sup>-Test, RMSEA, CFI und SRMR);
5. eine Ableitung von Handlungsempfehlungen für die Verwendung der getesteten MDTs bei Schätzung von SE-Modellen in der empirischen Praxis.

Im nächsten Kapitel folgt zunächst die Darstellung des Direct-ML- und EM-Verfahrens, bevor im dritten Kapitel die einzelnen Varianten der MI diskutiert werden.<sup>4</sup> Daran anschließend wird das Forschungsdesign des SEMI-Projekts vorgestellt und die erzielten Ergebnisse der Analysen präsentiert.

## 2 Maximum Likelihood-Schätzverfahren

Zu den ML-basierten MDTs sind das Direct-ML- und das EM-Verfahren zu zählen. Im Unterschied zum Direct-ML-Schätzverfahren, mit welchem die fehlenden Werte direkt bei der Parameterschätzung eines Modells berücksichtigt werden, ist die Behandlung der fehlenden Werte bei der EM-Methode von der eigentlichen Modellschätzung losgelöst (wie bei der MI). Hierbei werden zunächst eine Kovarianzmatrix und ein Mittelwertvektor, bestehend aus ML-Schätzwerten, generiert, in welchen die fehlenden Werte berücksichtigt werden. Diese werden dann in einem zweiten Schritt als

---

<sup>4</sup> Ergänzende Informationen zur Thematik fehlender Werte, zu anderen MDTs oder zur Darstellung der MI und ihrer Grundlagen finden sich in: Rubin 1987; Little/Rubin 2002; Schafer 1997; Allison 2000; van Buuren 2012; Marsh 1998; Pigott 2001; Urban/Mayerl 2018; Enders 2010.

Modellinput für ein Strukturgleichungsmodell o. Ä. benutzt. Ein solches Vorgehen ist äquivalent zu einer Direct-ML-Schätzung, da sich die Parameterschätzungen des Modells zwischen beiden Verfahren kaum voneinander unterscheiden. Aus diesem Grund wird das EM-Verfahren, sofern die Kovarianzmatrix und die Mittelwerte als Input benutzt werden, auch als Indirect-ML-Verfahren bezeichnet.

## 2.1 Direct-ML

„The goal of maximum likelihood estimation is to identify the population parameter values that have the highest probability of producing a particular *sample* of data” (Enders 2010: 59, Hervorhebungen im Original). Das bedeutet, dass in einer ML-Schätzung eine Kombination aus Populationsparametern wiederholt geprüft wird, bis eine Kombination gefunden wird, welche die Wahrscheinlichkeit maximiert, die beobachteten, empirischen Daten zu erhalten. Diese maximierte Wahrscheinlichkeit wird in einem Log-Likelihood-Wert ausgedrückt: der Sample-Log-Likelihood. Wenn sich diese zwischen den Wiederholungen nicht mehr signifikant unterscheidet, dann ist die ML-Schätzung konvergiert.

Die Sample-Log-Likelihood ist nichts anderes als die Summe der individuellen Log-Likelihood-Werte der einzelnen Fälle im Datensatz. Jeder einzelne Log-Likelihood-Wert gibt dabei die relative Wahrscheinlichkeit an, dass die Werte, die für jeden Fall geschätzt werden, aus einer multivariaten Normalverteilung mit einem bestimmten Mittelwertvektor und einer bestimmten Kovarianzmatrix stammen. Damit unterliegt die ML-Schätzung der Annahme einer gemeinsamen multivariaten Normalverteilung der Modellvariablen.

Das ML-Schätzverfahren ist wohl diejenige Schätzmethode, welche in der empirischen (Sozial-)Forschung am häufigsten eingesetzt wird. Das liegt daran, dass sie Eigenschaften aufweist, mit welchen Parameterschätzungen möglich sind, die a) asymptotisch konsistent sind (je größer das Sample ist, desto größer ist die Wahrscheinlichkeit, dass der geschätzte Parameter nicht vom Populationsparameter abweicht), b) asymptotisch effizient sind (in einem großen Sample ist die Varianz der geschätzten Parameter minimal) und c) asymptotisch normalverteilt sind. Dies gilt, wenn die zugrundeliegenden Annahmen der Schätzung eingehalten werden. Darunter fallen die angesprochene multivariate Normalverteilung der Modellvariablen, eine kontinuierliche Skalierung derselben, ein genügend großes Sample, ein korrekt spezifiziertes Modell und die Unabhängigkeit der Beobachtungen (vgl. Finney/DiStefano 2013: 441). Während die letzten beiden Punkte oftmals als gegeben angenommen werden, sind die ersten drei Punkte häufig Ursachen für verzerrte Schätzergebnisse der Modellschätzung.

Sowohl die Voraussetzungen der ML-Schätzung als auch ihre asymptotischen Eigenschaften liegen auch dem Direct-ML-Verfahren zugrunde. Die Direct-ML-Schätzung unterscheidet sich von der ‚normalen‘ ML-Schätzung darin, dass Fälle mit Missing Values berücksichtigt werden, wohingegen diese Fälle bei einer reinen ML-Schätzung bereits im Vorhinein ausgeschlossen werden. Das liegt daran, dass die Berechnung jeder fallweisen Log-Likelihood bei der ‚normalen‘ ML-Schätzung die gleiche Anzahl an Beobachtungen erfordert. Um also eine solche ML-Schätzung durchführen zu können, müssen alle Variablen, welche bei der interessierenden Modellschätzung benutzt werden, bei allen Fällen gültige Werte aufweisen. Wenn Missing Values vorliegen, müsste also der Datensatz erst bereinigt werden, sodass nur Fälle verbleiben, die auf den interessierenden Modellvariablen gültige Werte aufweisen.

Dieses Problem wird mit Hilfe des Direct-ML-Schätzverfahrens umgangen. Hierbei wird die fallbedingte Log-Likelihood nicht mehr mit allen Modellvariablen berechnet, sondern nur mit denjenigen, die auch tatsächlich einen gültigen Wert aufweisen. Damit stehen im Umkehrschluss auch deutlich mehr Fälle für die Modellanalyse zur Verfügung, als bei einer einfachen ML-Schätzung mit vorausgehendem Fallausschluss. Das kann sich auch im Ergebnis der Modellschätzung widerspiegeln: Mit dem Direct-ML-Verfahren sind demnach effizientere Schätzungen möglich, da sie die gesamte beobachtete Information in ihrer Schätzung berücksichtigt (siehe Arbuckle 1996; Enders 2001).

Wenn die Annahmen der ML-Schätzung eingehalten werden, dann sind die Ergebnisse der Modellschätzungen durch das Direct-ML-Schätzverfahren auch bei einem MAR-Ausfallmechanismus unverzerrt. Im Grunde besitzen die Schätzergebnisse alle gewünschten Eigenschaften, wie sie auch mit einer ‚normalen‘ ML-Schätzung einhergehen. Zusätzlich ist die Direct-ML-Methode in den meisten SEM-Paketen implementiert und einfach anzuwenden (bspw. wird in *Mplus*<sup>5</sup>, sollten Missing Values vorliegen, diese Methode per Default eingesetzt).

Sollten die Annahmen allerdings nicht zutreffen, so kann nicht mehr davon ausgegangen werden, dass das Schätzergebnis die gewünschten Eigenschaften aufweist. Stattdessen wäre dann davon auszugehen, dass es verzerrt ist und dass die daraus gezogenen Schlüsse falsch sein könnten. Widersprüchliche Ergebnisse aus MC-Studien zeigen aber an, dass es wohl auf das Ausmaß der Verletzung dieser Annahmen ankommt. Verzerrende Einflüsse auf das Ergebnis nehmen dabei vor allem kategorisierte Variablen, die Verteilung der Variablen und die Samplegröße. Es hat sich allerdings gezeigt, dass bereits eine Samplegröße von knapp 200 Fällen (je nach Komplexität des Modells) für eine gute ML-Schätzung ausreichend ist (vgl. Anderson/Gerbing 1984: 171; Curran u. a. 1996: 25).

---

<sup>5</sup> Aussagen zu *Mplus* beziehen sich auf die im Projekt verwendete Programmversion 7.31.

Auch bei Variablen mit mindestens fünf Skalenpunkten sind nur unwesentliche Verzerrungen der Schätzung zu beobachten. Sind jedoch die interessierenden Modellvariablen nicht multivariat normalverteilt, kann dies einen deutlich verzerrenden Einfluss auf das Schätzergebnis haben.

Ein weiteres Problem, das mit Abweichungen von einer multivariaten Normalverteilung einhergeht, ist die Möglichkeit, dass mehrere Maxima der Likelihood-Funktion vorliegen. Die Likelihood, zu welcher die ML-Schätzung konvergiert, kann in einem solchen Fall von den Startwerten abhängig sein. Das bedeutet, dass sich – sollten tatsächlich mehrere Maxima vorliegen – die Schätzergebnisse unterscheiden könnten, wenn unterschiedliche Startwerte verwendet werden (vgl. Dong/Peng 2013: 10).

## 2.2 Expectation-Maximization (EM)

Das EM-Verfahren basiert auf zwei Schritten: Zum einen auf einem Expectation-Step (E-Step), zum anderen auf einem Maximization-Step (M-Step). Wie beim Direct-ML-Verfahren handelt es sich auch beim EM-Algorithmus um ein Verfahren, das eine multivariate Normalverteilung der beteiligten Variablen voraussetzt (Dempster u. a. 1977). Die nachfolgenden Ausführungen zur Logik des EM-Verfahrens basieren auf den Arbeiten von Allison (2002: 19 ff.) und Little/Rubin (2002: 167 ff.).

Das Verfahren schätzt in einem ersten E-Step mittels Regressionsschätzung die fehlenden Werte im Datensatz. Dabei werden die Schätzwerte auch um zufällig ausgewählte Residuen erweitert, um die Varianz der Schätzungen zu erhöhen und Zufallsfehler zu simulieren. Nach Imputation der geschätzten Variablenwerte ist der Datensatz vollständig.

Wichtig ist nun, dass dieser vollständige Datensatz nur dazu benötigt wird, um im anschließenden M-Step neue Mittelwerte und eine neue Kovarianzmatrix berechnen zu können: Im E-Step werden demnach nur Informationen, die aufgrund von fehlenden Werten in den Mittelwerten und der Kovarianzmatrix im M-Step fehlen würden, durch erwartete (*expected*) Informationen substituiert. Wurden im M-Step neue Mittelwerte und eine neue Kovarianzmatrix berechnet, hat das Verfahren eine erste Iteration durchlaufen. Die anschließende zweite Iteration startet dann mit einem erneuten E-Step.

In dieser zweiten Iteration dienen die Mittelwerte und die Kovarianzmatrix aus dem vorherigen M-Step als Grundlage für die Regressionsgleichungen. Diese lassen nun ‚bessere Vorhersagen‘ zu, aus welchen wiederum neue Schätzwerte hervorgehen, die imputiert werden können. Nachdem die Imputationen erfolgt sind, kommt es zu einem neuen M-Step. Hierin werden wieder neue Mittelwerte und eine neue Kovarianzmatrix berechnet, welche dann dem dritten E-Step als Grundlage für die Regressionsgleichungen dienen usw.

Der Wechsel zwischen E- und M-Step dauert so lange an, bis sich die Mittelwerte und die Kovarianzmatrizen zwischen den Iterationen nicht mehr ändern, bzw. bis sich der daraus ergebende Wert für die Sample-Log-Likelihood nicht mehr signifikant ändert. Wenn sich keine Änderung mehr feststellen lässt, ist der EM-Algorithmus konvergiert und die Sample-Log-Likelihood ist maximiert. Die resultierenden Mittelwerte und die Kovarianzmatrix sind dann ML-Schätzwerte. Auf deren Basis lassen sich dann alle weiteren Datenanalysen durchführen.

Die EM-Methode ist eine gute Möglichkeit, um mit Missing Values umzugehen. Bisher zeigt sich, dass das iterative EM-Verfahren bei fehlenden Werten, die einem MAR-Ausfallmechanismus unterliegen, unverzerrte und konsistente Parameterschätzungen liefert. Gleichzeitig zeigt sich, dass das Verfahren auch dann geeignet ist, wenn die Annahme einer multivariaten Normalverteilung verletzt ist. Denn selbst wenn Daten vorliegen sollten, von denen bekannt ist, dass sie keiner Normalverteilung folgen, sind die durch den EM-Prozess generierten ML-Schätzwerte durchaus brauchbar (vgl. Allison 2002: 18). Zwar kann es mit diesem Verfahren zu Konvergenzproblemen kommen. Allerdings, so Honaker/King (2010), ist dies bei der Behandlung fehlender Werte weniger problematisch als bei einer ML-Schätzung von Modellparametern. Denn die vorhergesagten Missing Values sind unabhängig von dem jeweiligen konvergierten Maximum. In der Forschungspraxis lässt sich die Methode relativ einfach einsetzen, da sie auch in großen Statistikpaketen implementiert ist (bspw. in SPSS 25).

Neben diesen sicherlich vorteilhaften Eigenschaften des Verfahrens, ist noch auf eine Problematik hinzuweisen, die mit diesem Verfahren verbunden ist. Auf Kovarianzmatrizen basierende Modellanalysen sind darauf angewiesen, dass bei der Modellspezifikation eine Fallzahl angegeben wird. Weil es für die maximierte Kovarianzmatrix aus einem EM-Prozess aber keine einheitliche Fallzahl gibt und es deshalb nicht klar ist, welche Fallzahl für die Modellanalysen verwendet werden sollte, kann deren willkürliche Wahl dazu führen, dass die Schätzung der Standardfehler und die darauf aufbauenden inferenzstatistischen Schlüsse verzerrt werden (vgl. Graham/Schafer 1999: 3).

### **3 Verfahren der Multiplen Imputation**

Grundsätzlich lässt sich das Vorgehen der MI in drei Phasen unterteilen, welche als Imputations-, als Analyse- und als Poolingphase bezeichnet werden. In der ersten Phase wird der Datensatz mit fehlenden Werten vervielfältigt, sodass  $m$  Datensätze vorliegen. In den  $m$  Datensätzen werden die fehlenden Werte durch geschätzte Werte ersetzt. Nach der Imputationsphase liegen dann  $m$  Datensätze vor, die sich in den imputierten Werten voneinander unterscheiden (die bereits vorhandenen Werte bleiben von diesem Vorgehen unberührt).

Nachdem die fehlenden Werte in der Imputationsphase ersetzt wurden, können die Datensätze so behandelt werden, als ob keine Missing Values vorgelegen hätten. Es folgt die Analysephase. Hierin wird für jeden  $m$ -ten Datensatz das interessierende Modell (bspw. ein Strukturgleichungsmodell) separat geschätzt. Es liegen dann  $m$  verschiedene Ergebnisse vor. Das Zusammenführen der geschätzten Parameter und Standardfehler geschieht in der Poolingphase: Aus mehreren Ergebnissen ergibt sich durch die Anwendung der Regeln nach Rubin (*Rubin's rules*) das letztliche Ergebnis für die Modellschätzung (siehe dazu u. a. Enders 2010 und Rubin 1987).

Damit die Ergebnisse der statistischen Analyse mit MI-imputierten Datensätzen unverzerrt sind, müssen die folgenden drei Bedingungen eingehalten werden: a) es muss ein MAR-Ausfallmechanismus vorliegen, b) es muss ein genügend großes  $m$  an Datensätzen gewählt werden,<sup>6</sup> c) es muss für den MI-Einsatz sichergestellt werden, dass das Imputationsmodell mit dem Analysemodell übereinstimmt. Hierbei geht es nicht um die Struktur des späteren Analysemodells, sondern nur darum, die im Analysemodell benötigten Variablen zu berücksichtigen (vgl. Schafer/Graham 2002: 167).

### 3.1 Data Augmentation (MNV)

Im Folgenden soll zunächst diejenige Technik vorgestellt werden, die mit am häufigsten in der empirischen Forschung eingesetzt wird: die Technik der „Data Augmentation“ (im Weiteren: „MNV-Technik“ genannt<sup>7</sup>). Die MNV-Technik ist ein zweistufiger, iterativer Prozess, der aus einem Imputations-Step (I-Step) und einem Posterior-Step<sup>8</sup> (P-Step) besteht. Das MNV-Verfahren beruht auf der Annahme, dass die vorliegenden Daten  $Y$  (sowohl die beobachteten als auch die unbeobachteten) einer multivariaten Normalverteilung folgen, wobei die Parameter  $\theta$ , die diese Verteilung beschreiben, durch eine Kovarianzmatrix und einen Mittelwertvektor bestimmt sind.

Für den ersten I-Step werden ein anfänglicher Mittelwertvektor und eine anfängliche Kovarianzmatrix benötigt. Aus diesem Grund wird mit den vorhandenen Daten und dem EM-Verfahren (Kapitel 2.2) diese Kovarianzmatrix mit dem dazugehörigen Mittelwertvektor geschätzt. Liegen beide vor, beginnt das MNV-Verfahren mit der ersten Iteration.

---

<sup>6</sup> Empfohlen wird, dass  $m$  ungefähr dem Anteil an Missing Values entsprechen sollte (vgl. Graham u. a. 2007: 212).

<sup>7</sup> MNV = Multiple Imputationsmethode unter Annahme einer multivariaten Normalverteilung. Das MNV-Verfahren geht auf Tanner/Wong (1987) und Schafer (1997) zurück (dort finden sich auch detaillierte Informationen zu diesem Verfahren). Die nachfolgenden Ausführungen zum Ablauf der MNV-Technik sind von den Ausführungen in Allison (2003: 551 f.) und Enders (2010: 187-254) inspiriert.

<sup>8</sup> Die MI ist eine Methode, die auf dem Theorem von Bayes fußt. Danach ist eine posterior-Verteilung eine Verteilung der Parameter, welche die Wahrscheinlichkeit maximieren die beobachteten Daten zu erhalten, wenn die Vorstellungen bzgl. dieser Parameter, die vor der Beobachtung der Daten vorgelegen haben, miteinbezogen werden. Da die posterior-Verteilung bei der MI aber im Normalfall keine vorherigen Informationen miteinbezieht, wird diese zum Großteil von den vorliegenden Daten bestimmt. Damit liefert die posterior-Verteilung der Parameter nichts anderes als eine Parameterverteilung, welche die Wahrscheinlichkeit maximiert, die beobachteten Daten zu erhalten (vgl. Allison 2002: 35).

Mit Hilfe der Kovarianzmatrix und des Mittelwertvektors lassen sich im ersten I-Step die Koeffizienten für die zur Imputation benötigten, linearen Regressionen berechnen. Mit den Regressionen werden Werte für die Missing Values geschätzt, um diese (nachdem den geschätzten Werten zufällig Residuen zugeordnet wurden) zu ersetzen. Somit liegt dann ein Datensatz vor, der vollständig ist und ersetzte Werte enthält. Mit diesem werden erneut eine Kovarianzmatrix und ein Mittelwertvektor berechnet. Es folgt der P-Step.

Im P-Step werden die aufgefüllten Daten aus dem vorherigen I-Step herangezogen, um eine posterior-Verteilung der Kovarianzmatrix als auch des Mittelwertvektors zu spezifizieren. Die MNV-Technik nutzt hierbei Monte Carlo (MC)-Simulationen, um aus den posterior-Verteilungen einen neuen Mittelwertvektor und eine neue Kovarianzmatrix zufällig zu ziehen. Durch diese Zufallsziehungen aus den posterior-Verteilungen wird sichergestellt, dass die einzelnen imputierten Werte innerhalb eines Datensatzes zufällig und voneinander unabhängig sind.

Nachdem der erste Iterationsschritt abgeschlossen ist, erfolgt die zweite Iteration, beginnend mit einem erneuten I-Step. In diesem werden die neu gezogene Kovarianzmatrix und der neue Mittelwertvektor als Grundlage für die Regressionsgleichungen benutzt, die dann wiederum die fehlenden Werte vorhersagen. Nachdem erneut Residuen zugeordnet wurden, werden neue Mittelwerte und Kovarianzen berechnet. Danach kommt es zu einem neuen P-Step. Auch hierbei werden zunächst die posterior-Verteilungen der Kovarianzmatrix und des Mittelwertvektors spezifiziert, woraus durch Zufallsziehungen wieder eine neue Kovarianzmatrix und ein neuer Mittelwertvektor hervorgehen. Beide dienen sodann den Regressionsgleichungen im folgenden dritten Iterationsschritt als Grundlage usw.

Der Wechsel zwischen I- und P-Step wird vielfach wiederholt. Nachdem die vorgegebene Anzahl an Iterationen durchlaufen ist, liegt eine Kette mit verschiedenen Datensätzen vor (nach jedem I-Step liegt ein vollständiger Datensatz vor). Alle imputierten Werte innerhalb einer jeden Iteration sind zufällig und voneinander unabhängig. Allerdings sind die imputierten Werte zwischen zwei aufeinanderfolgenden Iterationen nicht unabhängig voneinander. Um sicherzustellen, dass die  $m$  ausgewählten Datensätze keine Abhängigkeiten untereinander aufweisen, muss das Verfahren konvergiert sein.

Nachdem die Kette konvergiert ist, können die  $m$  Datensätze ausgewählt werden. Zu beachten ist dabei, dass zwischen den ausgewählten  $m$  Datensätzen genügend Iterationsschritte liegen, um etwaige Abhängigkeiten zwischen den Iterationen auszuschließen. Bspw. könnte nach Konvergenz der Kette jeder 150ste Datensatz ausgewählt werden, bis das Ziel von  $m$  Datensätzen erreicht ist. Eine andere Möglichkeit wäre, dass anstatt die  $m$  Datensätze aus einer Kette auszuwählen, insgesamt  $m$  Ketten

gerechnet werden und jeweils der Datensatz der letzten Iteration gespeichert wird. Die Berechnung von  $m$  Ketten hat den Vorteil, dass zwischen den  $m$  Datensätzen keine Abhängigkeiten bestehen können.

Folgendes ist zur kritischen Einschätzung der MNV-Technik anzuführen. Die MNV-Technik ist eine recht unflexible MI-Variante. Denn es sind im Prinzip nur Imputationen mit metrischen Daten möglich, die annähernd normalverteilt sind. Werden auch kategorial-geordnete Variablen in den Imputationsprozess mit aufgenommen, ist dies problematisch. Zwar zeigt sich, dass mit der MNV-Technik auch bei Verletzung der Normalverteilungsannahme und mit kategorialen Variablen durchaus stabile Schätzungen möglich sind. Allerdings ist das Ausmaß von Verzerrungen davon abhängig, wie stark die einzelnen Abweichungen von der Normalverteilung sind und wie viele Kategorien die einzelnen kategorialen Variablen tatsächlich aufweisen. Eine Verletzung der Normalverteilungsannahme muss somit nicht unbedingt problematisch sein, kann aber durchaus einen Einfluss auf die imputierten Werte und das spätere Schätzergebnis nehmen.

Für alle MI-Verfahren gilt zudem, dass die Rechenzeit bis zur Konvergenz davon abhängt, wie groß der Datensatz ist, also wie viele Fälle und Variablen im Datensatz vorhanden sind (a), wie viele Missing Values sich darin befinden (b) und wie viele Datensätze ( $m$ ) erstellt werden sollen (c). Zudem ist zu berücksichtigen, ob verschieden skalierte Variablen vorliegen oder nicht (d). Wenn der Datensatz dementsprechend komplex ist, kann es sehr lange dauern, bis die einzelnen MI-Varianten konvergieren. Es kann auch sein, dass überhaupt keine Konvergenz erzielt wird. Denn je komplexer der Datensatz ist, desto unwahrscheinlicher ist es, dass eine gemeinsame Verteilung existiert, welche die vorliegenden Daten beschreiben kann.

### **3.2 Fully Conditional Specification (FCS)**

Die FCS-Variante der MI beruht nicht auf der Annahme einer gemeinsamen multivariaten Normalverteilung aller Modellvariablen, sondern legt für jede Variable mit fehlenden Werten ein eigenes Schätzmodell an, das den fehlenden Wert mit Hilfe aller anderen im Datensatz befindlichen Variablen vorhersagt. Die Idee dahinter ist, dass bei einer binären Variablen ein logistisches, bei einer ordinalen ein multinominal logistisches und bei einer metrischen ein lineares Schätzmodell spezifiziert wird.

Während also die MNV-Technik eine gemeinsame multivariate Normalverteilung aller Variablen annimmt und nur die Parameter sampelt (die Kovarianzmatrix und die Mittelwertvektoren), welche diese beschreiben, trifft das FCS-Verfahren keine solche Restriktion. Stattdessen geht es davon aus, dass eine gemeinsame Verteilung – wie immer diese aussehen mag – aller Variablen implizit vorhanden ist. Die gemeinsame Verteilung der beobachteten und unbeobachteten vorliegenden Daten  $Y$ , die

durch die Parameter  $\theta$  beschrieben werden kann, wird mit der FCS-Variante also erst ‚gesucht‘ und nicht wie bei der MNV-Technik bereits vorgegeben. Wie dies vonstattengeht, zeigen die nachfolgenden Ausführungen. Diese basieren auf den Arbeiten von van Buuren (2015) und Raghunathan u. a. (2001).

Das FCS-Verfahren benötigt für jede unvollständige Variable in einem Datensatz ein eigenes Schätzmodell, welches von der Skalierung der jeweiligen Variablen mit fehlenden Werten abhängig ist. Damit mit diesen Schätzmodellen die Imputationen getätigt werden können, muss der Datensatz mit fehlenden Werten zunächst vervollständigt werden. Zuerst werden aus den univariaten Verteilungen der beobachteten Variablenwerte rein zufällig solche Werte gezogen, die die fehlenden Werte im Datensatz ersetzen können. Die fehlenden Werte sind also nach dieser Prozedur durch Platzhalter ersetzt und der Datensatz ist vollständig.

Im Anschluss daran werden für die Variable mit den wenigsten Missing Values diese Platzhalter wieder entfernt (hier:  $Y_1$ ). Es liegt damit ein Datensatz vor, der nur in einer Variablen fehlende Werte aufweist. Für diese Variable wird ein Modell spezifiziert, das die Skalierung der Variablen berücksichtigt (bspw. ein lineares Regressionsmodell für eine metrische Variable) und auf alle Informationen im Datensatz zurückgreift (also auf alle Variablen einschließlich ihrer Platzhalter). Mit Hilfe dieses Modells werden dann die fehlenden Werte in der ersten Variablen geschätzt und imputiert.

Nachdem in einem ersten Schritt die erste Variable imputiert wurde, folgt nun in einem zweiten Schritt die Imputation für die zweite Variable. Bei der zweiten Variablen handelt es sich um diejenige Variable, die nach der ersten Variablen, die wenigsten Missing Values aufweist. Nachdem also die fehlenden Werte in der Variablen  $Y_1$  ersetzt wurden, dient diese bei der nächsten Variablen ( $Y_2$ ) (nachdem die Platzhalter wieder gelöscht wurden) mit allen anderen Variablen als unabhängige Variable. Für  $Y_2$  wird wiederum das für das jeweilige Skalenniveau angebrachte Schätzmodell gewählt. Dieser Vorgang erfolgt im weiteren Verlauf auch mit allen anderen Variablen, die fehlende Werte aufweisen. Die Ersetzung der fehlenden Werte endet dann, wenn für die letzte Variable die Missing Values imputiert wurden.

An dieser Stelle hat das Verfahren dann eine erste Iteration durchlaufen. Danach startet es mit der zweiten Iteration. Dieser liegt nun der ersetzte Datensatz aus der ersten Iteration zugrunde. Der ganze Vorgang wird so lange wiederholt, bis die angegebene Anzahl an Iterationen durchlaufen wurde. Zum Schluss liegt dann ein vollständiger Datensatz vor.

Mit diesen Schritten wird angenommen, dass die gemeinsame multivariate Verteilung, die durch die Parameter  $\theta$  beschrieben werden kann, durch die einzelnen, auf die Variablenskalierungen zugeschnittenen Schätzmodelle, iterativ approximiert werden kann (wobei diese aber unbekannt bleibt). Somit unterscheidet sich die FCS- zur MNV-Technik dadurch, dass bei letzterer die gemeinsame Verteilung explizit definiert wird.

Auch die Iterationskette im FCS-Verfahren muss sich stabilisieren. Denn erst nach der Stabilisation der Kette, sind die Imputationen für die fehlenden Werte innerhalb eines Datensatzes zufällig und unabhängig voneinander. Wiederum gilt, dass aufeinanderfolgende Iterationen voneinander abhängig sind. Aus diesem Grund müssen entweder zwischen den Datensätzen verschieden viele Iterationsschritte liegen, oder es müssen  $m$  Ketten gerechnet werden, aus welchen dann der letzte Datensatz ausgewählt wird. Dieses Vorgehen ist Standard bei der FCS-Technik (als Variante der MI-Methodik). Dabei reicht es meist aus, je nachdem wie groß und komplex der Datensatz ist, oder wie viele Missing Values vorliegen, zwischen 5 und 20 Iterationen zu benutzen (vgl. van Buuren 2012: 113).

Während das MNV-Verfahren eine gemeinsame Verteilung der Variablen vorgibt, ist für die FCS-Technik auch nach der Stabilisation der Iterationsketten nicht bekannt, ob für die Variablen überhaupt eine gemeinsame Verteilung existiert. Es kann also sein, dass die FCS- Technik nicht zu einer stabilen Verteilung konvergiert. Allerdings funktioniert sie auch dann, wenn die einzelnen Verteilungen nicht kompatibel sind (vgl. Horton/Kleinman 2007: 83).

Mithin sind im Gegensatz zur MNV-Variante die Annahmen, die über die Daten getroffen werden, bei der FCS-Variante weniger restriktiv: Je nach Skalenniveau wird das ‚richtige‘ Modell gewählt. Das hat zur Folge, dass sich die FCS-Technik für verschiedenste Skalenniveaus eignet. Ihre Flexibilität und Robustheit gegenüber den einzelnen Datensets verschafft dieser Technik einen nicht zu unterschätzenden Vorteil.

### **3.3 Zwei weitere Verfahren: EMB und H0**

Im Folgenden sollen noch zwei weitere Varianten der MI vorgestellt werden, die sich etwas von den bereits vorgestellten MI-Varianten unterscheiden. Das ist zum einen eine bootstrapping basierte Variante, die mithilfe des EM-Verfahrens die fehlenden Werte ersetzt (EMB) und zum anderen die H0-Technik. Im Gegensatz zur MNV-, FCS- oder EMB-Variante ist die H0-Technik eine MI-Variante, die explizit die Modellstruktur des Analysemodells berücksichtigt. Damit ähnelt sie dem Direct-ML-Verfahren, beruht aber auf dem Ansatz der MI.

Wie bei den bereits dargestellten MI-Techniken, lässt sich auch die EMB-Variante in drei Schritte aufteilen. Zunächst werden die Datensätze vervielfältigt, danach werden die fehlenden Werte imputiert, bevor dann die  $m$  Schätzergebnisse zu einem Endergebnis zusammengeführt werden. Während sich die MNV- und FCS-Technik nur in der Ersetzung der fehlenden Werte unterscheiden, unterscheidet sich das EMB-Verfahren bereits im ersten Schritt von diesen. Denn zur Herstellung der  $m$  verschiedenen Datensätze benutzt es ein Bootstrapping-Verfahren. Dabei werden aus dem vorliegenden Datensatz mit fehlenden Werten  $m$  verschiedene Datensätze zufällig gezogen. Mit Hilfe des mehrfachen und zufälligen Resamplings (mit Zurücklegen) schafft es die EMB-Technik, Variation zwischen den Datensätzen herzustellen (vgl. Honaker/King 2010: 564 f.; Honaker u. a. 2011: 4).

Nachdem die  $m$  Bootstrapping-Datensätze vorliegen, werden mit dem EM-Verfahren (Kapitel 2.2) für jeden dieser Datensätze die Punktschätzungen der Kovarianzmatrix und des Mittelwertvektors generiert. Schlussendlich liegen dann  $m$  verschiedene Mittelwertvektoren und Kovarianzmatrizen vor. Diese werden im Anschluss dazu verwendet, im *ursprünglichen* Datensatz, also demjenigen Datensatz mit den fehlenden Werten, die Missings zu imputieren. Damit liegen  $m$  ursprüngliche Datensätze vor, die sich in den ersetzten Missing Values unterscheiden. Infolgedessen kann jeder dieser Datensätze analysiert werden. Die einzelnen Analyseergebnisse werden dann zu einem gesamten Ergebnis gepoolt.

Mit der EMB-Variante der Multiplen Imputation gehen alle Vorteile der MI-Methode einher und unter der MAR-Annahme sind unverzerrte Modellschätzungen zu erwarten. Im Gegensatz zum reinen EM-Verfahren findet mit der EMB-Variante tatsächlich eine Imputation statt, sodass die Problematik der Fallzahl nicht vorliegt. Zudem sind mehrere Maxima der Likelihood-Funktion eher unproblematisch, wie dies auch bei der reinen EM-Methode der Fall ist. Auch das EMB-Verfahren unterliegt (wie die MNV-Technik oder die Direct-ML- oder die EM-Methode) der Annahme einer gemeinsamen multivariaten Normalverteilung aller Modellvariablen. Das führt dazu, dass sie auf metrische Skalierungen und nur geringe Abweichungen von der Normalverteilung der Daten angewiesen ist.

Die bereits vorgestellten Techniken sind in der empirischen Forschung, die wohl am häufigsten verwendeten. Ihnen ist gemein, dass die Imputationen anhand eines ‚unrestricted models‘ erfolgen, was bedeutet, dass das Imputationsmodell nicht der Modellstruktur des Analysemodells entspricht. Eine weitaus weniger bekannte Variante ist die eigens für das Datenanalysepaket „*Mplus*“ implementierte H0-Technik. Das Besondere an dieser ist, dass sie auf Grundlage der spezifizierten Modellstruktur die fehlenden Werte ersetzt. Es handelt sich bei den Imputationen durch das H0-Verfahren also um Ersetzungen anhand eines ‚restricted models‘. Zudem ist es möglich, sowohl kategoriale als auch

metrische Variablen zu berücksichtigen. Wie bereits die FCS-Technik ist auch das H0-Verfahren recht flexibel gegenüber unterschiedlich skalierten Modellvariablen.

Im Gegensatz zu den anderen Varianten der MI erfolgt bei der H0-Variante die Behandlung der fehlenden Werte mit Hilfe eines Bayes-Schätzers direkt im Rahmen der Parameterschätzungen. Im Grunde wird iterativ das spezifizierte Modell wiederholt geschätzt, bis sich die Iterationen stabilisiert haben. Ist das eingetreten, dann liegen die Parameter  $\theta$  vor, die die gemeinsame Verteilung der Daten  $Y$  beschreiben können. Damit ähnelt dieses Verfahren der Direct-ML-Schätzung, da auch diese die Missing Values direkt bei der Parameterschätzung berücksichtigt (vgl. Asparouhov/Muthén 2010a: 16 f.).

Das Vorgehen der H0-Methode soll an dieser Stelle ein wenig ausführlicher vorgestellt werden. Zunächst wird das spezifizierte Modell auf den vorliegenden Daten geschätzt (die Missing Values wurden zuvor durch Platzhalter ersetzt, sodass ein vollständiger Datensatz gegeben ist). Basierend auf den Parameterschätzwerten des spezifizierten Modells werden dann für die metrischen Variablen<sup>9</sup> mit fehlenden Werten im Datensatz lineare Regressionsgleichungen spezifiziert. Dieser Prozess ist äquivalent zum Vorgehen bei der FCS-Technik. Die Regressionsschätzungen liefern die Information, die notwendig ist, um die posterior-Verteilungen der Parameter für jede Variable mit Missings zu spezifizieren. Mittels MC-Simulationen werden aus diesen Verteilungen die Parameter für eine weitere lineare Regression gezogen. Mit dieser werden dann die fehlenden Werte vorhergesagt und imputiert (nachdem Residuen zugeordnet wurden). Es liegen dann Imputationen für die erste metrische Variable vor und der Algorithmus fährt mit der zweiten metrischen Variablen fort, usw. Nachdem alle Variablen ersetzt wurden, liegt ein neuer vollständiger Datensatz vor. Auf diesen wird das Modell erneut geschätzt, um danach wiederum die fehlenden Werte zu ersetzen. Diese Prozedur wird solange wiederholt, bis sich die Iterationskette stabilisiert hat. Ab diesem Punkt können dann die einzelnen Datensätze gespeichert werden.

Um die Konvergenz des Verfahrens zu erfassen, werden mehrere Iterationsketten benötigt. *Mplus* benutzt per Default zwei Iterationsketten. Die nach der Konvergenz der Ketten vorliegende Modellschätzung berücksichtigt unter einem MAR-Ausfallmechanismus bereits adäquat die fehlenden Werte. Damit liegt nach der Konvergenz der Ketten ein Modell vor, das bereits für die Analyse von inhaltlichen Fragestellungen verwendet werden kann. Mit dem Bayes-Schätzer ist es also im Grunde

---

<sup>9</sup> Wird im Modell eine kategoriale Variable verwendet (hier:  $X$ ) dann wird diese zunächst in eine metrische Variable überführt ( $X^*$ ). Damit geht ein Set an Schwellenwerten einher. Die metrische Variable  $X^*$  bildet dann zusammen mit den entsprechenden Schwellenwerten die betreffende kategoriale Variable  $X$  ab (vgl. Asparouhov/Muthén 2010a: 9 f.). Wenn  $X$  Missing Values aufweist, dann werden diese bei der Generierung der metrischen Variablen ( $X^*$ ) direkt berücksichtigt (vgl. ebd.: 16).

genommen nicht nötig, dass imputierte Datensätze erzeugt werden. Wird dies dennoch gewünscht, dann werden nach jeder 100sten Iteration (um die Unabhängigkeit der Iterationen zu gewährleisten) die imputierten Datensätze ausgespeichert, sodass am Ende die  $m$  gewünschten Datensätze vorliegen, die dann analysiert werden können (vgl. Muthén/Muthén 2017: 443 ff.).

Grundsätzlich sollten sich MI-Techniken, ob nun eine Modellstruktur vorgegeben wird oder nicht, in ihren Ergebnissen nur gering voneinander unterscheiden. Was die H0-Variante jedoch etwas benachteiligt ist, dass das spezifizierte Modell im Imputationsprozess korrekt sein und mit dem Populationsmodell übereinstimmen muss. Sollte das Imputationsmodell fehlspezifiziert sein, dürfte sich das auch in den ausgeführten Imputationen widerspiegeln. Asparouhov/Muthén (2010b) relativieren dies jedoch etwas, indem sie zeigen, dass leichte Fehlspezifikationen im Imputationsprozess keine Auswirkungen auf die Modellanalyse haben (vgl. ebd.: 22). Was genau die Autoren unter ‚leichten Fehlspezifikationen‘ verstehen, wird aber leider nicht ausgeführt. Es ist deshalb sinnvoll, anzunehmen, dass etwaige Fehlspezifikationen einen Einfluss nehmen können. Damit impliziert die H0-Technik, im Gegensatz zu den anderen drei MI-Techniken, die Annahme, dass das Schätzmodell dem Populationsmodell entspricht.

Ein Vorteil des H0-Verfahrens ist die flexible Ersetzung der fehlenden Werte auf Grundlage der Variablenskalierung. So ist die Verteilung der Variablen, also ob eine gemeinsame multivariate Normalverteilung vorliegt oder nicht, irrelevant. Aufgrund der Berücksichtigung von kategorialen Skalenniveaus innerhalb der Modellschätzung bietet die H0-Variante eine ähnliche Flexibilität wie die weitaus bekanntere FCS-Variante. Gleichzeitig ist die Bayes-Schätzung auch gegenüber kleineren Samples robust. Das wiederum bevorteilt sie gegenüber dem Direct-ML-Schätzverfahren, das bspw. nur unter größeren Samples gute Schätzergebnisse liefert (vgl. Schmitt 2011: 308). Zusätzlich kann das H0-Verfahren die Komplexität des Imputationsprozesses reduzieren, wenn viele unvollständige Variablen in einem großen Datensatz vorliegen. Das liegt daran, dass die vorliegende Modellstruktur die zu schätzenden Parameter reduziert, die unter einem nicht restringierten Modell vorliegen würden (vgl. Asparouhov/Muthén 2010b: 22). Damit dürften vor allem etwaige Konvergenzprobleme reduziert werden.

#### **4 Zusammenfassung: Voraussetzungen und Annahmen der MDTs**

Wie im vorangegangenen Kapitel 3 aufgezeigt, ist der Einsatz der MI variantenunabhängig an folgende Bedingungen geknüpft:

1. Es muss zumindest ein MAR-Ausfallmechanismus angenommen werden können.
2. Die Anzahl der Datensätze ( $m$ ) muss in etwa dem Anteil der fehlenden Werte entsprechen.

3. Alle inhaltlich relevanten Variablen und inhaltlich relevanten Variablenbeziehungen müssen bei der Imputation der fehlenden Werte berücksichtigt werden.
4. Es ist sicherzustellen, dass die jeweiligen Iterationsketten konvergiert sind. Erst dann entsprechen die Imputationen zufälligen Werten, die unabhängig voneinander sind.
5. Aufeinanderfolgende Datensätze in den Iterationsketten sollten nicht für die Analyse ausgewählt werden, da die Imputationen zwischen diesen Datensätzen unter Umständen sehr stark miteinander korrelieren können (gilt nicht für die EMB-Variante).

Wird all dies sichergestellt, entscheidet die ausgewählte MI-Variante darüber, ob nach der Imputationsphase plausible Imputationen vorliegen und aus der Analyse- und Poolingphase unverzerrte Parameter und Standardfehler hervorgehen. Problematisch ist jedoch, dass jede MI-Variante zusätzliche Annahmen über die vorliegenden Daten mit in die Imputationsphase einbringt. Diese zusätzlichen Annahmen einer jeden MI-Technik lassen sich den nachfolgenden Tabellen 1 und 2 entnehmen.

Für das Direct-ML- und EM-Verfahren gilt, wie für die MI auch, dass zumindest ein MAR-Mechanismus vorliegen muss, dass alle relevanten Variablen und Variablenbeziehungen im Imputationsmodell (EM) oder im Analysemodell (Direct-ML) berücksichtigt werden und dass die Schätzverfahren konvergieren. Eine Besonderheit besteht für das EM-Verfahren darin, dass nicht klar ist, welche Fallzahl den statistischen Analysen mitgeteilt werden soll (sofern die Kovarianzmatrix als Analysegrundlage dient). Allerdings könnte der imputierte Datensatz verwendet werden, um der Fallzahlproblematik zu entgehen, was aber möglicherweise die Schätzung der Standardfehler beeinflussen kann, sodass mit einer Unterschätzung der Schätzwerte zu rechnen. Gleichzeitig bringen Direct-ML- und EM-Verfahren zusätzliche Annahmen über die zugrundeliegenden Daten mit ein. Auch diese sind den folgenden Tabellen 1 und 2 zu entnehmen.

Unabhängig davon, welche der hier vorgestellten MDTs in der empirischen Modellanalyse verwendet wird, dürfte es klar sein, dass empirische Datensätze nicht immer alle Annahmen einer jeden Methode erfüllen können, was bedeutet, dass trotz sorgfältiger Anwendung einzelner MDTs die geschätzten Parameter und Standardfehler mehr oder weniger stark verzerrt sein können. Das dürfte vor allem dann der Fall sein, wenn die empirischen Daten die Annahmen hinsichtlich der Variablenskalierung und -verteilung verletzen (zusätzlich haben auch die Größe des Datensatzes bzw. die Komplexität und der Anteil an Missing Values einen Einfluss). Aus diesem Grund sollte im hier beschriebenen SEMI-Projekt vor allem geprüft werden, welche Verletzungen von Annahmen der einzelnen MDTs so wirkmächtig sind, dass sie zu verzerrten Modellschätzungen führen können.

Tabelle 1: Zusammenfassung von MDT-Annahmen (Teil 1)

| MDTs                             | Zusatzannahmen   | Folgen bei Verletzung der Zusatzannahmen  |
|----------------------------------|--|---|
| Direct-ML,<br>EM,<br>EMB,<br>MNV | Multivariate Normalverteilung<br>aller Modellvariablen         | <b>Verletzung durch:</b> <ul style="list-style-type: none"> <li>Abweichungen von der Normalverteilung</li> <li>Kategoriale Variablen (mit weniger als fünf Ausprägungen)</li> </ul> <b>Folgen:</b> <ul style="list-style-type: none"> <li>Konvergenz zu lokalen Maxima (Direct-ML)</li> <li>Konvergenzprobleme; keine Konvergenz (EMB, MNV)</li> <li>Verzerrte Modellschätzungen</li> </ul> |
| Direct-ML,<br>H0                 | korrekt spezifiziertes Modell                                  | <b>Verletzung durch:</b> <ul style="list-style-type: none"> <li>Inkorrekt spezifiziertes Modell</li> </ul> <b>Folgen:</b> <ul style="list-style-type: none"> <li>Konvergenzprobleme</li> <li>Verzerrte Modellschätzungen</li> </ul>   |
| Direct-ML                        | ausreichend großes Sample                                      | <b>Verletzung durch:</b> <ul style="list-style-type: none"> <li>Kleine Fallzahl</li> </ul> <b>Folgen:</b> <ul style="list-style-type: none"> <li>Verzerrte Modellschätzungen</li> </ul>   |
| FCS,<br>H0,<br>MNV               | <b>Umfang/Komplexität des<br/>Datensatzes oder<br/>Modells</b> | <b>Folgen:</b><br>Dauer bis zur Konvergenz<br>Keine Konvergenz  |
| EM,<br>EMB                       |  | <b>Folgen:</b><br>Dauer bis zur Konvergenz  |

Tabelle 2: Zusammenfassung von MDT-Annahmen (Teil 2)

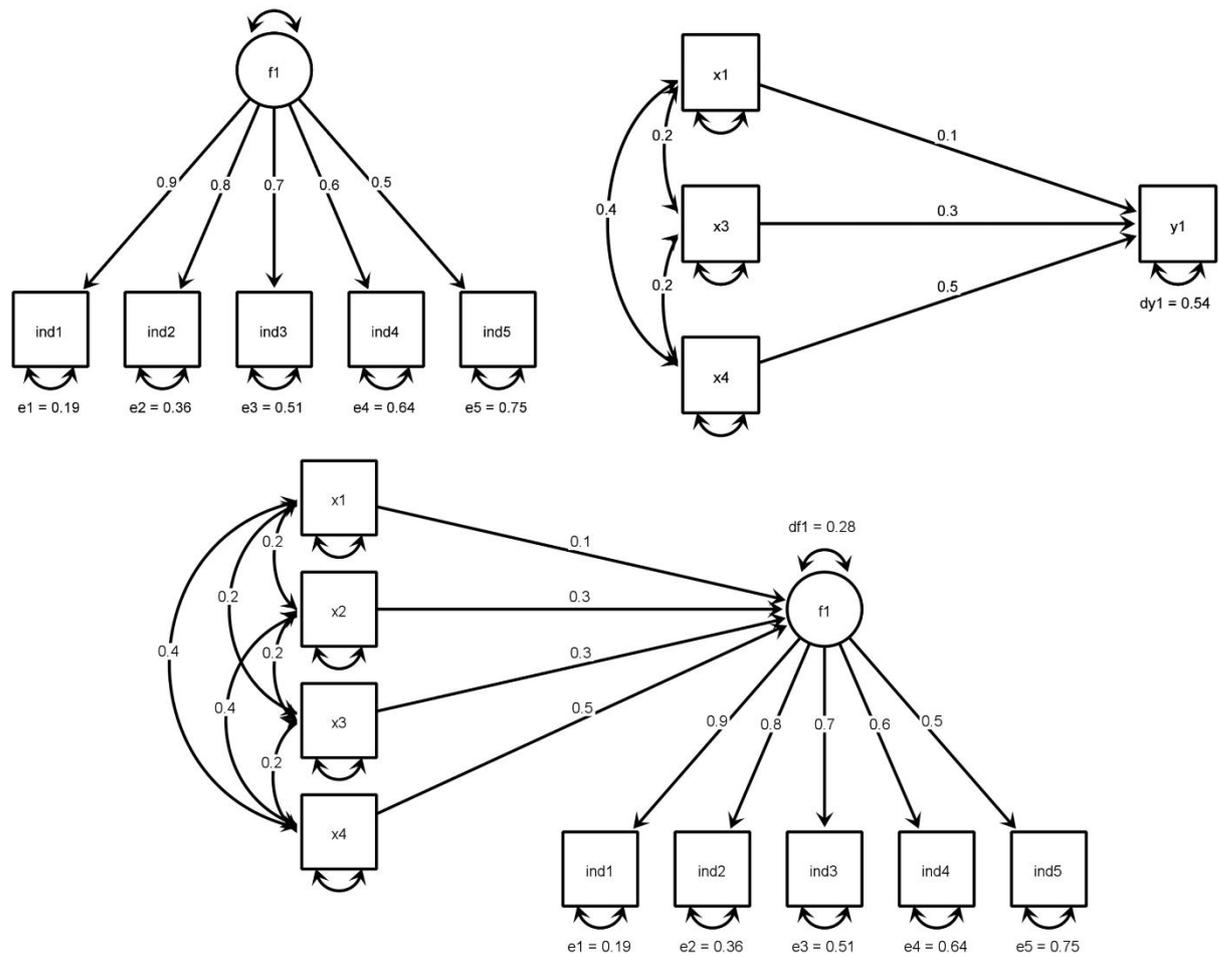
|  | MNV | EMB | FCS | H0 | EM | Direct-ML |  |
|--|-----|-----|-----|----|----|-----------|--|
| <b>Multivariate Normalverteilung</b>   |     |     |     |    |    |           |  |
| Unverzerrtheit bei Vorliegen   | ja  | ja  | ja  | ja | ja | ja        |  |
| Verzerrungen bei Verletzungen  |     | ja  | ja  | ja | ja | ja        |  |
| • Verzerrungen vernachlässigbar  | ja  |     |     |    |    |           |  |
| <b>Kategoriale Variablen</b>   |     |     |     |    |    |           |  |
| Unverzerrtheit bei nicht metrisch definierten Modellvariablen  | ja  | ja  | ja  | ja | ja | ja        |  |
| Verzerrungen bei asymmetrischen, kategorialen Modellvariablen  | ja  | ja  |     |    | ja | ja        |  |
| • Verzerrungen vernachlässigbar  |     |     | ja  | ja |    |           |  |
| <b>Samplegröße</b>   |     |     |     |    |    |           |  |
| Verzerrungen bei kleinen Samples   |     |     |     |    |    | ja        |  |
| • unerheblich für  | ja  | ja  | ja  | ja | ja |           |  |
| <b>Anteil an Missing Values</b>  |     |     |     |    |    |           |  |
| Anzahl nicht konvergierter Lösungen steigt mit Missinganteil   | ja  |     | ja  | ja |    | ja        |  |
| • unerheblich für  |     | ja  |     |    | ja |           |  |
| vermehrt falsche inhaltliche Schlussfolgerungen (zu kleine Standardfehler)   |     |     |     |    | ja |           |  |
| <b>Imputationsmodell</b>   |     |     |     |    |    |           |  |
| Modellstruktur im Imputationsmodell unerheblich  | ja  | ja  | ja  |    | ja |           |  |
| <b>Fit-Indices</b>   |     |     |     |    |    |           |  |
| Offene Forschungsfrage: Beeinflusst die Behandlung der fehlenden Werte mit den verschiedenen MDTs und unter den dargelegten Bedingungen die Fit-Indices? |     |     |     |    |    |           |  |

## 5 Forschungsdesign

Im Projekt wurden zur Überprüfung der Leistungsfähigkeit der hier vorgestellten MDTs diejenigen Performanzkriterien ausgewählt, nach welchen die MDTs mittels MC-Simulationen bewertet werden sollten. Probleme entstanden dabei insbesondere hinsichtlich der Bestimmung von modellfit-bezogenen Performanzkriterien. Denn da die Performanz einzelner SEM-Fit-Indices bei Modellschätzungen mit Missing Values bislang noch nicht systematisch untersucht worden ist, konnten für die Festlegung dieser Performanzkriterien allein die Resultate der SEM-Methodenforschung bezüglich von Fit-Indices bei Modellschätzungen mit vollständigen Daten benutzt werden.

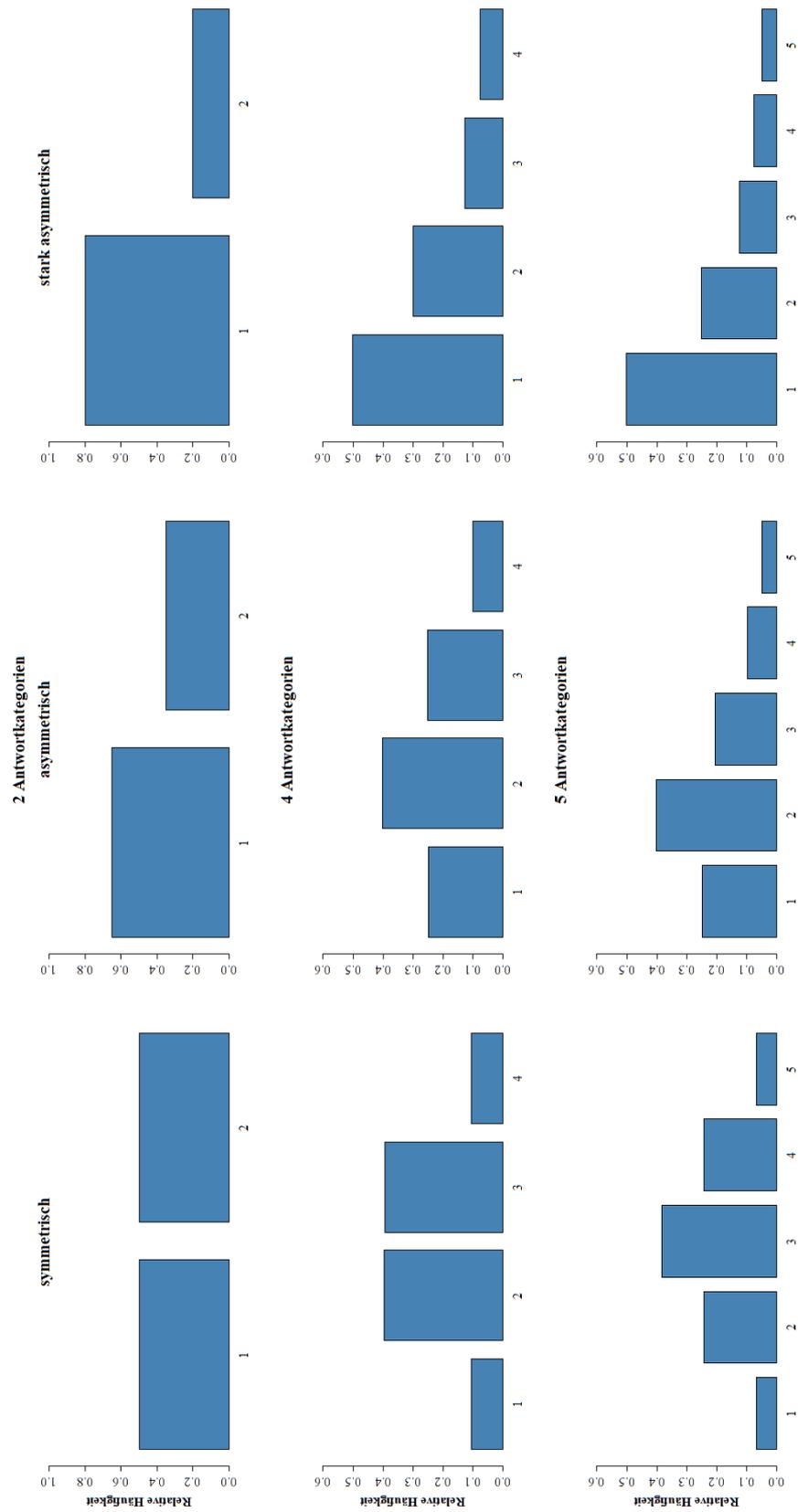
Für die Untersuchungen wurden drei unterschiedliche SE-Modelle geschätzt (vgl. die nachfolgende Abbildung 1), deren Modellstrukturen und definierte Parameterwerte einigen typischen und häufig anzutreffenden Modellen in der gegenwärtigen SEM-Forschung entsprechen. Bei den Schätzungen wurden mehrere Schätzbedingungen variiert. So wurden für die jeweiligen MC-basierten Modellschätzungen verschiedene Fallzahlen (250 und 750) sowie verschieden hohe Anteile an Missing Values (unter Annahme eines MAR-Ausfallmechanismus) vorgegeben (5 %, 20 %, 35 %). Zudem wurden insgesamt drei verschiedene Datenverteilungen simuliert (symmetrisch, asymmetrisch, stark asymmetrisch; vgl. Abbildung 2). Damit wurde versucht, möglichst häufig anzutreffende empirische Datenverteilungen zu implementieren und so die Übertragbarkeit der erwarteten Ergebnisse auf die praktische Erforschung empirisch relevanter Modellkonstruktionen zu erleichtern.

Abbildung 1: Populationsmodelle



Anmerkungen: oben links: Modell 1; oben rechts: Modell 2; unten: Modell 3. Ein doppelseitiger Pfeil, der nur auf eine jeweilige Variable/ein jeweiliges latentes Konstrukt zeigt, stellt die Varianzen/Residualvarianzen dar. Varianzen der unabhängigen Variablen sind auf 1.0 festgesetzt (Wert nicht dargestellt).

Abbildung 2: Verteilungen



Als Performanzkriterien wurden die Durchschnittswerte der Fit-Indices und die prozentuale Modell-Ablehnungsrate benutzt. Zur Bewertung der Modellparameterschätzungen und ihrer Standardfehler wurde ein Biasmaß (relativ und absolut) eingesetzt. Die folgende Tabelle 3 informiert über die Bewertungskriterien sowie die jeweils gewählten Analysemethoden.

Tabelle 3: Performanzkriterien

| Bewertungskriterium    |  | Schwellenwerte | Analysen                                |
|------------------------|--|----------------|---|
| <b>Modellebene:</b>    |  |                |   |
| Konvergenz             | Anteil erfolgreicher Schätzungen pro Konfig. | --             | deskriptiv                              |
| <i>rejection rate</i>  | Fit-Indices pro Replikation:                 |                |   |
|                        | 1. p-Wert Chi <sup>2</sup>                   | 1. $\geq .05$  | deskriptiv +<br>log. Regressionsmodelle |
|                        | 2. RMSEA                                     | 2. $\leq .05$  |   |
|                        | 3. SRMR                                      | 3. $\leq .08$  |   |
| 4. CFI                 | 4. $\geq .95$                                |                |   |
| Durchschnitt           | $\emptyset$ Fit-Werte pro Konfiguration      | --             | deskriptiv                              |
| <b>Parameterebene:</b> |  |                |   |
| Parameterbias          | absoluter Bias pro Replikation               | --             | lin. Regressionsmodelle                 |
|                        | relativer Bias pro Konfiguration             | 10 % bzw. 15 % | deskriptiv                              |
| Relative Effizienz     | Vergleich MDT zu BL                          | --             | deskriptiv                              |
| Standardfehlerbias     | absoluter SE-Bias pro Replikation            | --             | lin. Regressionsmodelle                 |
|                        | relativer SE-Bias pro Konfiguration          | 10 % bzw. 15 % | Deskriptiv                              |

## 5.1 Rechentechnische Umsetzung des Forschungsdesigns

Für die oben angekündigten Simulationen wurden zunächst R-Skripte verfasst, mit deren Hilfe die Prozesse der Datengenerierung, der Imputation und der Modellschätzung sowie der Auswertung im Hinblick auf die Bewertungskriterien automatisiert<sup>10</sup> werden konnten. Die Datengenerierung sowie die späteren Modellschätzungen erfolgten mittels der SEM-Software *Mplus* (Version 7.31; Muthén/Muthén 2012) und dem darin implementierten, standardmäßigen MLR-Schätzer. Auch die Behandlung der fehlenden Werte mit der H0-Technik und der Direct-ML-Schätzung erfolgte in *Mplus*, wohingegen für alle anderen MDTs verschiedene Pakete in R eingesetzt wurden (R in den Versionen 3.2.2<sup>11</sup> und 3.4.3; R Core Team 2015; 2017). Für die FCS-Variante der MI wurde das R-Paket *mice* in der Version 2.46.0 benutzt. Die MNV-Variante wurde mithilfe des *norm2*-R-Pakets umgesetzt (Version 2.0.1) und die EM- sowie die EMB-Analysen wurden mit dem R-Paket *Amelia* (Version 1.7.4) durchgeführt. Gemäß der Verfahrenslogik der MI wurden für jede ihrer Varianten insgesamt 50 Datensätze imputiert (*m*) und die Konvergenz der Verfahren sichergestellt.<sup>12</sup>

<sup>10</sup> Durch das R-Paket *MplusAutomation* (Hallquist/Wiley 2018; Hallquist 2018).

<sup>11</sup> Die R-Version 3.4.3 diente allen notwendigen Berechnungen als Grundlage. Eine Ausnahme stellte die Implementation der MNV-Technik dar. Hierbei wurde die Version 3.2.2 verwendet, da das benutzte R-Paket (*norm2*) zum Zeitpunkt der Simulationsdurchführungen nicht für eine neuere Version verfügbar war.

<sup>12</sup> Folgender Link führt zum Online-Anhang mit allen notwendigen Templates und Programmcodes, um die Ergebnisse zu replizieren: [https://osf.io/pufkj/?view\\_only=411b2ed2437b4bb4a01595ffd836299a](https://osf.io/pufkj/?view_only=411b2ed2437b4bb4a01595ffd836299a). In dem zur Verfügung stehenden Ordner finden sich die Templates, die zur Datengenerierung benutzt wurden (B1.1). Mithilfe dieser wird es möglich,

Insgesamt lagen pro MDT 54 mögliche Konfigurationen vor. Das wiederum entspricht im Gesamten einem Umfang von 324<sup>13</sup> verschiedenen Simulationskonfigurationen, die mit dieser Studie geprüft wurden. Weil die Performanz einzelner MDTs im Fokus des Projektes stand, wurden für jede der MDTs identische Datensätze verwendet. Es wurden jeweils 500 verschiedene Datensätze (Replikationen) für die 54 Konfigurationen erstellt, die dann versechsfacht wurden.

Zusätzlich wurden auch Datensätze generiert, die keine Missing Values enthielten. Die Ergebnisse der Modellschätzungen, die mit diesen Daten durchgeführt wurden, konnten im Laufe der weiteren Projektarbeit mit den Ergebnissen der Modellschätzungen bei Verwendung unvollständiger Datensätze und bei Einsatz von MDTs verglichen werden.<sup>14</sup>

Da die MI-Varianten iterative Verfahren sind, konnten die Simulationen nicht auf handelsüblichen Desktop-Rechnern durchgeführt werden. Denn damit hätte allein die Imputation der fehlenden Werte (ohne die Daten zu generieren und ohne die Modellschätzungen durchzuführen und deren Ergebnisse für die Analyse aufzubereiten) insgesamt 175 Rechentage in Anspruch genommen. Erst durch den Einsatz des bwUniClusters (ein Parallelrechner der im Rahmen des baden-württembergischen Umsetzungskonzepts für Hochleistungsrechner den Universitäten des Landes zur Verfügung steht) konnte die Dauer der Simulationen deutlich reduziert werden.<sup>15</sup>

## 6 Ergebnisse

Die Resultate der Modell- und Simulationsschätzungen wurden unter Verwendung der ausgewählten Performanzkriterien analysiert. Es wurden uni- und bivariate Auswertungen aber auch multivariate logistische und lineare Regressionsanalysen durchgeführt. Während die uni- und bivariaten Analysen dazu benutzt wurden, potentiell problematische Kombinationen von Modellparametern zu identifi-

---

dieselben Daten zu generieren. In B1.2 findet sich die *Mplus*-Syntax für die Implementation der H0-Technik. B1.3 beinhaltet die Konfiguration der MDTs, die in R vorgenommen wurden: Das ist neben dem EM-Verfahren auch die EMB-, FCS- und MNV-Technik. Mithilfe der zur Verfügung gestellten Programmcodes lässt sich deren Implementation und Konfiguration nachvollziehen. In B1.4 findet sich die Analysesyntax für die aufbereiteten Daten. Da sich die Analyse der Daten für die MDTs voneinander unterscheidet (für das Direct-ML-Verfahren beruht die Analyse bspw. auf einem Datensatz, während für die MI-Techniken eine Liste als Grundlage dient, mit welcher die einzelnen *m* Datensätze angesteuert werden), muss die Analysesyntax der Modellschätzung angepasst werden. Die Anpassung betrifft nicht den Modellschätzer oder das Modell, sondern bezieht sich auf die Art der vorliegenden Daten (ob Liste oder einfacher Datensatz) oder auf die Reihenfolge der Variablen in den Datensätzen.

<sup>13</sup> 3 Modelle\*2 Samplegrößen\*3 Verteilungen\*3 Missingquoten\* 6 MDTs = 324 Simulationskonfigurationen.

<sup>14</sup> Zu den bereits vorliegenden 324 Simulationskonfigurationen kamen in diesem Fall noch einmal zusätzliche 18 Konfigurationen (3 Modelle\*2 Samplegrößen\*3 Verteilungen\*1 Missingquote) hinzu. Damit erhöhte sich die gesamte Zahl an Simulationskonfigurationen auf 342.

<sup>15</sup> In B1.5 findet sich ein Beispiel für den Imputationsprozess auf dem bwUniCluster. Mit dem ersten File wird das zweite File generiert, das die Informationen enthält, welche Software das Cluster benötigt und wo das eigentliche Skript liegt, das im Cluster durchlaufen werden soll. Das dritte File stellt eines dieser Prozess-Skripte dar. Das vierte File beinhaltet den Auftrag das dritte File an das Cluster zu senden und den Prozess zu starten. Das Skript läuft dann automatisiert durch und speichert die Ergebnisse an dem gewünschten Ort ab.

zieren, wurden die logistischen und linearen Regressionen eingesetzt, um diejenigen Modellparameter zu identifizieren, die den größten/stärksten Einfluss auf die Performanz der MDTs haben. Dabei zeigte sich, dass es hinsichtlich der MDT-Performanz nur sehr geringe Unterschiede zwischen den drei getesteten SE-Modellen gibt, sodass die Ergebnisse im Folgenden nicht separat für jedes einzelne Modell dargestellt werden müssen.

### **6.1 Ergebnisse bzgl. Parameterschätzungen und Standardfehler**

1) Es zeigte sich, dass alle MDTs unverzerrte Parameterschätzwerte und Standardfehler (mit Ausnahme der Standardfehler im EM-Verfahren, s.u.) erbringen. In keiner Simulationskonfiguration konnten Bias-Werte beobachtet werden, welche die gesetzten Grenzen überschreiten.

2) Auch nicht-normalverteilte Variablen erzeugten in unseren Schätzungen keine gravierenden Verzerrungen (mit Ausnahme der Standardfehler im EM-Verfahren, s.u.). Zwar führten Verteilungseinflüsse zu kleineren Verzerrungen von Modellparametern und Standardfehlern, jedoch waren diese Einflüsse nur wenig bedeutsam. Dementsprechend sind nicht-normalverteilte Variablen beim Einsatz von MDTs eher als unproblematisch einzustufen und führen nicht zu verzerrten Ergebnissen.

3) Werden in SE-Modellen kategoriale Variablen berücksichtigt, die nicht asymmetrisch verteilt sind (hier: Modell 2 und Modell 3 in der symmetrischen Verteilungskonfiguration), ergeben sich bei allen MDTs nur unwesentlich verzerrte Parameter und nur unwesentlich verzerrte Standardfehler (mit Ausnahme der Standardfehler im EM-Verfahren, s.u.). In den meisten Fällen wurden dabei in unseren Tests keine Grenzwerte überschritten.

4) Auch mit zunehmend asymmetrisch verteilten Variablen konnten keine unverhältnismäßigen Verzerrungen beobachtet werden (mit Ausnahme der Standardfehler im EM-Verfahren, s.u.). Es ließen sich zwar Verteilungseinflüsse auf das Ausmaß der Verzerrungen der Parameter und der Standardfehler erkennen, allerdings sind diese Einflüsse wenig problematisch: Asymmetrisch verteilte Variablen sind bei Verwendung von MDTs für SEM-Schätzungen unproblematisch und führen nicht zu verzerrten Ergebnissen.

5) Bei den MI-Techniken (bzgl. Modellparameter und Standardfehler) und bei der Direct-ML-Schätzung (bzgl. der Standardfehler) konnte die Samplegröße als entscheidend für das Ausmaß an Verzerrung identifiziert werden. Dementsprechend ließ sich bei kleinen Fallzahlen ein größerer Bias für Parameterschätzungen mit den MI-Techniken identifizieren als bei Schätzungen nach dem EM- und Direct-ML-Verfahren. Dennoch wurden auch dabei die gesetzten Grenzwerte nicht überschritten. Im Ganzen bedeutet dies, dass zwar mit Verzerrungen unter kleinen Fallzahlen zu rechnen ist, diese

aber bis zu Samplegrößen von ca. 250 Fällen unproblematisch bleiben. Es ist aber anzunehmen, dass mit noch kleineren Fallzahlen größere Verzerrungen als die hier angesprochenen auftreten können.

6) Insgesamt betrachtet gab es in den Modellanalysen keine nicht-konvergierenden Schätzungen. Demnach ist die Annahme, wonach der Anteil an Missing Values einen Einfluss auf die MNV- oder FCS-Technik bzw. auf die Direct-ML- und H0-Methode haben, nicht haltbar. Für die EM-Methode und für die EMB-Technik konnte auch eine Konvergenzrate von 100 % erzielt werden. Allerdings wurden die Standardfehler beim EM-Einsatz tlw. deutlich unterschätzt. Der Missinganteil war bisweilen für das EM-Verfahren auch entscheidend für den Bias der Standardfehler. Mit der EM-Methode konnte ab 20 % an Missing Values kein Bias mehr beobachtet werden, der als unverzerrt einzuordnen wäre. Wird das EM-Verfahren demnach eingesetzt, könnten signifikante Effekte aufgrund der ersetzten Werte zustande kommen. Im Ergebnis zeigte sich, wie erwartet, dass mit der EM-Technik die Standardfehler unterschätzt werden, was vermehrt zu falschen inhaltlichen Schlüssen führen kann. Abgesehen davon, konnte für das Direct-ML- und das EM-Verfahren der Missinganteil als einflussreichste Größe für die Parameterschätzung identifiziert werden. Je höher der Anteil an Missing Values wird, desto größer wird auch der Bias der Parameterschätzung. Zwar konnten (wie für die MI-Techniken) auch für die ML-Verfahren keine Grenzwertüberschreitungen festgestellt werden, allerdings ist anzunehmen, dass mit noch höheren Anteilen an Missing Values noch größere Verzerrungen in den Parameterschätzwerten einhergehen.

7) Die Performanz der einzelnen MDTs unterscheidet sich nicht zwischen den drei SE-Modellen. Die MDTs erzielten in jedem Modell ähnlich gute Ergebnisse.

Fassen wir zusammen: Alle getesteten MDTs lassen sich bzgl. der Parameterschätzungen uneingeschränkt empfehlen; sie liefern nahezu identische, zufriedenstellende Ergebnisse. Lediglich die EM-Methode ist im Vergleich mit den anderen MDTs etwas weniger genau und liefert eine etwas größere Variabilität der Schätzwerte, was allerdings nicht zu inakzeptablen Parameterschätzungen führt. Allerdings konnten die linearen Regressionsanalysen zeigen, dass für den Bias in den Parameterschätzwerten vor allem die Samplegrößen und der Missinganteil verantwortlich sind, nicht aber die Werteverteilungen der beteiligten Variablen. Somit ist nicht auszuschließen, dass mit noch kleineren Fallzahlen und/oder mit noch höheren Missinganteilen auch der Bias in den Parameterschätzungen noch problematischer wird.

Auch in Bezug auf die geschätzten Standardfehler konnten zwar in aller Regel (ausgenommen: EM-Methode) keine problematischen Verzerrungen beobachtet werden. Dennoch zeigten auch hier die Regressionsanalysen, dass nicht die Verteilungen, sondern die Samplegrößen und der Missinganteil

für die Verzerrungen entscheidend sind. Auch hier könnten sich somit mit noch kleineren Fallzahlen und/oder noch höheren Missinganteilen letztlich problematische Verzerrungen ergeben.

## 6.2 Ergebnisse bzgl. Fit-Indices

Zufriedenstellende Ergebnisse konnten unter allen Bedingungen mit der Direct-ML-Schätzung und der H0-Variante der MI erzielt werden, denn bei beiden wird die jeweilige Modellstruktur bereits bei der Behandlung der fehlenden Werte berücksichtigt. So konnten sowohl akzeptable Durchschnittswerte als auch akzeptable Ablehnungsraten ermittelt werden. Allerdings schneidet das H0-Verfahren in Bezug auf die ermittelten Anpassungswerte immer dann etwas besser ab als das Direct-ML-Verfahren, wenn neben einem hohen Missinganteil (35 %) auch noch ein kleines Sample vorliegt. Interessanterweise ‚verbessern‘ beide die Durchschnittswerte der Fit-Indices bzw. reduzieren die Ablehnungsraten gegenüber den Referenzmodellen. Bezüglich beider MDTs konnte keine Kombination von Datenstrukturen ermittelt werden, die in Bezug auf die Fit-Indices problematisch wäre. Wenn demnach Missing Values vorliegen, dann kann nach deren Behandlung mit dem Direct-ML- oder dem H0-Verfahren davon ausgegangen werden, dass die geschätzten SE-Modelle mit den hier getesteten Fit-Indices korrekt beurteilt werden können (sofern die Modelle korrekt spezifiziert wurden).

Weitaus problematischer sind die Ergebnisse für die EM-Methode und die EMB-, FCS- und MNV-Technik. Hier konnten bei allen Simulationskonfigurationen nur für den SRMR-Fit-Index zufriedenstellende Ergebnisse beobachtet werden.

Generell konnten als Haupteinflussgrößen für die Performanz der Fit-Indices der Missinganteil und die Samplegröße identifiziert werden. Beide haben einen derart starken Einfluss auf die Fit-Performanz, dass der Einfluss der Variablenverteilungen kaum mehr ins Gewicht fällt und eher zu vernachlässigen ist. So konnten bei einem Anteil von 5 % an Missing Values pro Variable noch mit allen Fit-Indices korrekte Schlussfolgerungen für alle Simulationskonfigurationen erzielt werden. Aber ab einem Missinganteil von 20 % wurden die Ergebnisse deutlich heterogener. Liegt der Anteil der Missing Values bei 20 % (oder darüber) und ist die Fallzahl gering, sollte bei Verwendung der EM-Methode oder der EMB-, FCS- und MNV-Technik die Modellbewertung nach Imputation der fehlenden Werte nur noch mit dem SRMR-Index erfolgen. Denn alle anderen Fit-Indices erweisen sich dann als eher unzuverlässig (siehe dazu auch Tabelle 4). Bei einer größeren Fallzahl können dagegen auch mit dem CFI und RMSEA-Index zufriedenstellende Ergebnisse erzielt werden. Allerdings ist ab einer Missingquote von 35 % auch bei größeren Fallzahlen der RMSEA-Index nicht mehr und der CFI nur noch bedingt empfehlenswert (auf die Analyse des  $\chi^2$ -Werts sollte generell ab 20 % Missing Values verzichtet werden).

## 7 Diskussion der Ergebnisse und Handlungsempfehlungen für die Praxis

Mittels Einsatz des bwUniClusters konnten alle MDTs im Projektzeitraum hinsichtlich ihrer Performanz untersucht werden. Es zeigte sich, dass die verschiedenen MI-Varianten unter denselben Bedingungen und unabhängig von den zugrundeliegenden Populationsmodellen und deren einzelnen Komponenten (Faktorladungen, Strukturpfade und Kovarianzen) zu ähnlichen Ergebnissen führen, obwohl sie teilweise unterschiedliche Annahmen treffen und differente Verfahrensalgorithmen benutzen. Mit allen getesteten MI-Varianten ließen sich im angewandten Simulationstest korrekte Modellschätzungen erzielen. Dies gilt auch für das Direct-ML-Schätzverfahren.

Einzig für die EM-Methode konnte nachgewiesen werden, dass diese bei höheren Missinganteilen nicht mehr als MDT bei der Schätzung von SEMs eingesetzt werden sollte. Mit ihr werden die Standardfehler der geschätzten Modellparameter sehr stark nach unten verzerrt, was die Gefahr von inferenzstatistischen Fehlschlüssen deutlich erhöht.

Während die MDTs größtenteils ähnliche Schätzergebnisse für die Modellparameter der untersuchten SE-Modelle liefern, sind die untersuchten Fit-Indices doch unterschiedlich gut für die Beschreibung der Modellanpassung bei Einsatz von MDTs geeignet. Während für das Direct-ML- und H0-Verfahren die Modellbewertung mittels der getesteten Fit-Indices zufriedenstellend verläuft, trifft dies bei allen anderen getesteten imputationsbasierten Techniken nicht mehr zu. Bei Einsatz dieser Verfahren würde in vielen Fällen aufgrund eines fälschlicherweise zu geringen/hohen Fit-Indices ein korrekt spezifiziertes Modell, das zudem korrekte Parameterschätzungen und Standardfehler liefert, abgelehnt werden. Einzig der SRMR-Index würde in allen getesteten Fällen korrekte Modellbewertungen erlauben.

Die Ergebnisse der MC-Simulationen ermöglichen Handlungsempfehlungen für die empirisch-statistische Forschungspraxis. Diese lassen sich der folgenden Tabelle 4 entnehmen. Darin finden sich die Empfehlungen für den Einsatz bestimmter MDTs und bestimmter Modell-Fit-Indices unter verschiedensten SEM-Forschungskonstellationen (anzumerken ist, dass in der Tabelle die Variablenverteilungen keine Berücksichtigung finden, da sich diese als eher unbedeutend in den Modellanalysen herausgestellt haben).

Folgende drei Beispiele mögen die Aussagekraft der Tabelle 4 verdeutlichen: Wenn in einem Forschungsprojekt zur Schätzung eines SE-Modells ein Datensatz mit ca. 750 Fällen vorliegt, der eine mittlere Quote an Missing Values von ca. 20 % aufweist, dann sollte

1. bei Interesse am Chi<sup>2</sup> p-Wert die Direct-ML-Methode oder die H0-Technik eingesetzt werden;

2. bei Interesse am SRMR-Index sowie an unverzerrten Parameterschätzungen und deren Standardfehlern das Direct-ML-, EMB-, MNV-, H0- oder FCS-Verfahren ausgewählt werden;
3. bei Interesse an allen Fit-Indices, unverzerrten Parametern und deren Standardfehlern die Direct-ML-Methode oder die H0-Technik angewendet werden.

Tabelle 4: Handlungsempfehlungen

|                     |                                    | bei ca. 250 Fällen                   |                                       |                                       | bei ca. 750 Fällen                   |                                       |                                       |                  |
|---------------------|------------------------------------|--------------------------------------|---------------------------------------|---------------------------------------|--------------------------------------|---------------------------------------|---------------------------------------|------------------|
|                     |                                    | bei ca.<br>5 %<br>Missing-<br>anteil | bei ca.<br>20 %<br>Missing-<br>anteil | bei ca.<br>35 %<br>Missing-<br>anteil | bei ca.<br>5 %<br>Missing-<br>anteil | bei ca.<br>20 %<br>Missing-<br>anteil | bei ca.<br>35 %<br>Missing-<br>anteil |                  |
| Modell-<br>ebene    | Chi <sup>2</sup> p-Wert            | alle MDTs                            | Direct-ML,<br>H0                      | Direct-ML,<br>H0                      | alle MDTs                            | Direct-ML,<br>H0                      | Direct-ML,<br>H0                      |                  |
|                     | SRMR                               |                                      | alle MDTs                             | alle MDTs                             |                                      | alle MDTs                             | alle MDTs                             |                  |
|                     | RMSEA                              |                                      | Direct-ML,<br>H0                      | Direct-ML,<br>H0                      |                                      | alle MDTs                             | alle MDTs                             | Direct-ML,<br>H0 |
|                     | CFI                                |                                      | Direct-ML,<br>H0                      | Direct-ML,<br>H0                      |                                      | alle MDTs                             | alle MDTs                             | alle MDTs        |
| Parameter-<br>ebene | unverzerrte<br>Parameter           | alle MDTs                            |                                       |                                       |                                      |                                       |                                       |                  |
|                     | unverzerrte<br>Standard-<br>fehler | alle MDTs                            | alle außer EM                         |                                       | alle MDTs                            | alle außer EM                         |                                       |                  |

## 8 Literaturverzeichnis

- Allison, Paul D. 2000: Multiple Imputation for Missing Data. A Cautionary Tale. In: *Sociological Methods & Research*, Jg. 28, Nr. 3, 301–309.
- Allison, Paul D. 2002: *Missing Data*. Thousand Oaks, CA: SAGE.
- Allison, Paul D. 2003: Missing Data Techniques for Structural Equation Modeling. In: *Journal of Abnormal Psychology*, Jg. 112, Nr. 4, 545–557.
- Anderson, James C./Gerbing, David W. 1984: The Effect Of Sampling Error On Convergence, Improper Solutions, And Goodness-Of-Fit Indices For Maximum Likelihood Confirmatory Factor Analysis. In: *Psychometrika*, Jg. 49, Nr. 2, 155–173.
- Arbuckle, James L. 1996: Full Information Estimation in the Presence of Incomplete Data. In: Marcoulides, George A./Schumacker, Randall E. (Hrsg.): *Advanced Structural Equation Modeling. Issues and Techniques*, Mahwah, NJ: Lawrence Erlbaum Associates, 243–278.
- Asparouhov, Tihomir/Muthén, Bengt O. 2010a: Bayesian Analysis Using Mplus: Technical Implementation. In: <https://www.statmodel.com/download/Bayes3.pdf>, zugegriffen am 06.08.2019.
- Asparouhov, Tihomir/Muthén, Bengt O. 2010b: Multiple Imputation with Mplus. In: <https://www.statmodel.com/download/Imputations7.pdf>, zugegriffen am 06.08.2019.
- Curran, Patrick J./West, Stephen G./Finch, John F. 1996: The Robustness of Test Statistics to Nonnormality and Specification Error in Confirmatory Factor Analysis. In: *Psychological Methods*, Jg. 1, Nr. 1, 16–29.
- Dempster, A. P./Laird, Nan M./Rubin, Donald B. 1977: Maximum Likelihood from Incomplete Data via the EM Algorithm. In: *Journal of the Royal Statistical Society: Series B (Methodological)*, Jg. 39, Nr. 1, 1–38.
- Dong, Yiran/Peng, Chao-Ying 2013: Principled Missing Data Methods for Researchers. In: *SpringerPlus*, Jg. 2, Nr. 1, 1–17.
- Enders, Craig K. 2001: A Primer on Maximum Likelihood Algorithms Available for Use with Missing Data. In: *Structural Equation Modeling: A Multidisciplinary Journal*, Jg. 8, Nr. 1, 128–141.
- Enders, Craig K. 2010: *Applied Missing Data Analysis*. New York, NY: The Guilford Press.
- Finney, Sara J./DiStefano, Christine 2013: Nonnormal And Categorical Data In Structural Equation Modeling. In: Hancock, Gregory R./Mueller, Ralph O. (Hrsg.): *Structural Equation Modeling. A Second Course*, 2nd ed., Charlotte, NC: Information Age Publishing, 439–492.
- Graham, John W./Olchowski, Allison E./Gilreath, Tamika D. 2007: How Many Imputations are Really Needed? Some Practical Clarifications of Multiple Imputation Theory. In: *Prevention Science*, Jg. 8, Nr. 3, 206–213.

- Graham, John W./Schafer, Joseph L. 1999: On the Performance of Multiple Imputation for Multivariate Data With Small Sample Size. In: Hoyle, Rick H. (Hrsg.): *Statistical Strategies for Small Sample Research*, Thousand Oaks, CA: SAGE, 1–29.
- Hallquist, Michael N. 2018: Package ‘MplusAutomation’. In: <https://cran.r-project.org/web/packages/MplusAutomation/MplusAutomation.pdf>, zugegriffen am 24.10.2018.
- Hallquist, Michael N./Wiley, Joshua F. 2018: MplusAutomation: An R Package for Facilitating Large-Scale Latent Variable Analyses in Mplus. In: *Structural Equation Modeling: A Multidisciplinary Journal*, Jg. 25, Nr. 4, 621–638.
- Honaker, James/King, Gary 2010: What to Do about Missing Values in Time-Series Cross-Section Data. In: *American Journal of Political Science*, Jg. 54, Nr. 2, 561–581.
- Honaker, James/King, Gary/Blackwell, Matthew 2011: Amelia II: A Program for Missing Data. In: *Journal of Statistical Software*, Jg. 45, Nr. 7, 1–47.
- Horton, Nicholas J./Kleinman, Ken P. 2007: Much Ado about Nothing: A Comparison of Missing Data Methods and Software to Fit Incomplete Data Regression Models. In: *The American Statistician*, Jg. 61, Nr. 1, 79–90.
- Little, Roderick J. A./Rubin, Donald B. 1987: *Statistical Analysis with Missing Data*. New York, NY: John Wiley & Sons.
- Little, Roderick J. A./Rubin, Donald B. 2002: *Statistical Analysis with Missing Data*. 2nd ed. Hoboken, NJ: John Wiley & Sons.
- Marsh, Herbert W. 1998: Pairwise Deletion for Missing Data in Structural Equation Models: Nonpositive Definite Matrices, Parameter Estimates, Goodness of Fit, and Adjusted Sample Sizes. In: *Structural Equation Modeling: A Multidisciplinary Journal*, Jg. 5, Nr. 1, 22–36.
- Muthén, Bengt O./Asparouhov, Tihomir 2012: Bayesian Structural Equation Modeling: A More Flexible Representation of Substantive Theory. In: *Psychological Methods*, Jg. 17, Nr. 3, 313–335.
- Muthén, Linda K./Muthén, Bengt O. 2012: *Mplus User’s Guide*. 7th ed. Los Angeles, CA: Muthén & Muthén.
- Muthén, Linda K./Muthén, Bengt O. 2017: *Mplus User’s Guide*. 8th ed. Los Angeles, CA: Muthén & Muthén.
- Peters, Cara Lee Okleshen/Enders, Craig K. 2002: A primer for the estimation of structural equation models in the presence of missing data: Maximum likelihood algorithms. In: *Journal of Targeting, Measurement and Analysis for Marketing*, Jg. 11, Nr. 1, 81–95.
- Pigott, Therese D. 2001: A Review of Methods for Missing Data. In: *Educational Research and Evaluation*, Jg. 7, Nr. 4, 353–383.

- R Core Team 2015: R: A language and environment for statistical computing. Wien: R Foundation for Statistical Computing.
- R Core Team 2017: R: A language and environment for statistical computing. Wien: R Foundation for Statistical Computing.
- Raghunathan, Trivellore E./Lepkowski, James M./van Hoewyk, John/Solenberger, Peter 2001: A Multivariate Technique for Multiply Imputing Missing Values Using a Sequence of Regression Models. In: *Survey Methodology*, Jg. 27, Nr. 1, 85–95.
- Rubin, Donald B. 1987: Multiple Imputation for Nonresponse in Surveys. New York, NY: John Wiley & Sons.
- Schafer, Joseph L. 1997: Analysis of Incomplete Multivariate Data. 1st ed. Boca Raton, FL: CRC Press.
- Schafer, Joseph L./Graham, John W. 2002: Missing Data: Our View of the State of the Art. In: *Psychological Methods*, Jg. 7, Nr. 2, 147–177.
- Schmitt, Thomas A. 2011: Current Methodological Considerations in Exploratory and Confirmatory Factor Analysis. In: *Journal of Psychoeducational Assessment*, Jg. 29, Nr. 4, 304–321.
- Tanner, Martin A./Wong, Wing Hung 1987: The Calculation of Posterior Distributions by Data Augmentation. In: *Journal of the American Statistical Association*, Jg. 82, Nr. 398, 528–540.
- Urban, Dieter/Mayerl, Jochen 2018: Angewandte Regressionsanalyse: Theorie, Technik und Praxis. 5., überarb. Aufl. Wiesbaden: Springer.
- van Buuren, Stef 2012: Flexible Imputation of Missing Data. Boca Raton, FL: CRC Press.
- van Buuren, Stef 2015: Fully Conditional Specification. In: Molenberghs, Geert (Hrsg.): *Handbook of Missing Data Methodology*, Boca Raton, FL: CRC Press, 267–294.
- Wahl, Andreas 2020: Multiple Imputation by Chained Equations: Eine Leistungsevaluation bei Schätzung von Strukturgleichungsmodellen mittels Monte-Carlo-Simulationen. Diss., Universität Stuttgart.

## **SISS:**

### **Schriftenreihe des Instituts für Sozialwissenschaften der Universität Stuttgart**

<https://www.sowi.uni-stuttgart.de/forschung/schriftenreihe-des-instituts-fuer-sozialwissenschaften-siss/>

- No.50/2020 Verfahren der Multiplen Imputation bei Schätzung von Strukturgleichungsmodellen mit latenten Variablen. Ein systematischer Vergleich mittels Monte-Carlo-Simulationen.  
Andreas Wahl, Dieter Urban
- No. 49/2020 Psychische Gesundheit: Eine heikle Thematik in der empirischen Umfrageforschung?  
Zur Validierung des Crosswise Modells.  
Anna Götze, Andreas Wahl
- No. 48/2020 Analyse und Schätzung von Mehrgruppen-Strukturgleichungsmodellen mittels SPSS und EQS. Eine praxisnahe Anleitung.  
Andreas Wahl, Dieter Urban
- No. 47/2020 Machine Learning basierte Response Style Identifikation:  
eine simulations-statistische Pilotstudie.  
Thomas Krause
- No. 46/2020 Informationskriterien zur latenten Klassenidentifikation für Multiple Indicator Growth Mixture Models.  
Thomas Krause
- No. 45/2017 Die Angst vor dem Fremden. Ein Vignettenexperiment zur Bedrohungswahrnehmung. Ergebnisse eines studentischen Projektseminars.  
Alexander Geisler, Alexandra Möck, Rebecca Nell, Svenja Wiechmann, Thomas Krause
- No. 44/2016 Regressionsanalyse bei fehlenden Variablenwerten (missing values): Imputation oder Nicht-Imputation? Eine Anleitung für die Regressionspraxis mit SPSS.  
Dieter Urban, Jochen Mayerl, Andreas Wahl
- No. 43/2014 Meta-Analyse mit Mehrebenenmodellen unter Verwendung von HLM.  
Joachim Fiebig, Dieter Urban
- No. 42/2013 Panelanalyse mit Mehrebenenmodellen. Eine anwendungsorientierte Einführung.  
Thomas Krause, Dieter Urban
- No. 41/2012 Meta-Analyse und "publication bias". Eine empirisch-statistische Fallstudie.  
Joachim Fiebig, Gerrit Rüdebusch, Dieter Urban
- No. 40/2011 Pädosexuelle Viktimisierung und pädosexuelle Straffälligkeit. Ein „harter“ empirischer Hypothesentest.  
Dieter Urban, Joachim Fiebig
- No. 39/2010 Binär-logistische Regressionsanalyse.  
Grundlagen und Anwendung für Sozialwissenschaftler.  
Jochen Mayerl, Dieter Urban
- No. 38/2010 Das Working-Poor-Problem in Deutschland.  
Empirische Analysen zu den Ursachen von Armut trotz Erwerbstätigkeit.  
Leonie Hellmuth, Dieter Urban
- No. 37/2010 Der Bystander-Effekt in alltäglichen Hilfsituationen: Ein nicht-reaktives Feldexperiment.  
Katrin Alle, Jochen Mayerl
- No. 36/2008 Berufseintritt und Berufssituation von Soziologieabsolventen der Universität Stuttgart.  
Deskriptive Ergebnisse einer Absolventenbefragung aus dem Jahr 2007.  
Jochen Mayerl, Dieter Urban

## **SISS:**

### **Schriftenreihe des Instituts für Sozialwissenschaften der Universität Stuttgart**

<https://www.sowi.uni-stuttgart.de/forschung/schriftenreihe-des-instituts-fuer-sozialwissenschaften-siss/>

- No. 35/2007 Antwortlatenzzeiten in TRA-Modellen. Zur statistischen Erklärung von (Geld)-Spendenverhalten.  
Dieter Urban, Jochen Mayerl
- No. 34/2006 Zur soziologischen Erklärung individuellen Geldspendens. Eine Anwendung und Erweiterung der Theory of Reasoned Action unter Verwendung von Antwortlatenzzeiten in einem Mediator-Moderator-Design.  
Jochen Mayerl
- No. 33/2006 Ist Bildung gleich Bildung? Der Einfluss von Schulbildung auf ausländerablehnende Einstellungen in verschiedenen Alterskohorten.  
Klaus Hadwiger
- No. 32/2005 Analyzing cognitive processes in CATI-Surveys with response latencies: An empirical evaluation of the consequences of using different baseline speed measures.  
Jochen Mayerl, Piet Sellke, Dieter Urban
- No. 31/2005 Persönliche Netzwerke in der Systemtheorie.  
Jan A. Fuhse
- No. 30/2004 Erhöht ein Opfer-Täter-Zyklus das Risiko, Sexualstraftaten als pädosexuelle Straftaten zu begehen? Ergebnisse einer ereignisanalytischen Pilotstudie.  
Dieter Urban, Heiko Lindhorst
- No. 29/2003 Können Nonattitudes durch die Messung von Antwortreaktionszeiten ermittelt werden? Eine empirische Analyse computergestützter Telefoninterviews.  
Jochen Mayerl
- No. 28/2003 Systeme, Netzwerke, Identitäten. Die Konstitution sozialer Grenzziehungen am Beispiel amerikanischer Straßengangs.  
Jan A. Fuhse
- No. 27/2002 Probleme bei der Messung individueller Veränderungsraten. 13 empirisch und methodisch induzierte Effekte, die es schwierig machen, Veränderungen von generalisierten Bewertungen zu ermitteln.  
Dieter Urban
- No. 26/2002 Vertrauen und Risikoakzeptanz. Zur Relevanz von Vertrauen bei der Bewertung neuer Technologien.  
Martin Slaby, Dieter Urban
- No. 25/2002 Risikoakzeptanz als individuelle Entscheidung. Zur Integration der Risikoanalyse in die nutzentheoretische Entscheidungs- und Einstellungsforschung.  
Martin Slaby, Dieter Urban
- No. 24/2001 Ist Glück ein affektiver Sozialindikator subjektiven Wohlbefindens? Dimensionen des subjektiven Wohlbefindens und die Differenz zwischen Glück und Zufriedenheit.  
Jochen Mayerl
- No. 23/2001 Religiosität und Profession. Longitudinale Analysen zur Entwicklung des religiösen Selbstbildes bei Erzieherinnen.  
Heiko Lindhorst

## **SISS:**

### **Schriftenreihe des Instituts für Sozialwissenschaften der Universität Stuttgart**

<https://www.sowi.uni-stuttgart.de/forschung/schriftenreihe-des-instituts-fuer-sozialwissenschaften-siss/>

- No. 22/2001    Differentielle Technikakzeptanz, oder: Nicht immer führt die Ablehnung einer Technik auch zur Ablehnung ihrer Anwendungen. Eine nutzentheoretische und modell-statistische Analyse.  
Martin Slaby, Dieter Urban
- No. 21/2001    Unser „wir“ - ein systemtheoretisches Modell von Gruppenidentitäten.  
Jan A. Fuhse
- No. 20/2000    Längsschnittanalysen mit latenten Wachstumskurvenmodellen in der politischen Sozialisationsforschung.  
Dieter Urban
- No. 19/2000    Die Zufriedenheit von Stuttgarter Studierenden mit ihrer Lebens- und Wohnsituation. Erste deskriptive Ergebnisse einer sozialwissenschaftlichen Studie zu allgemeinen und bereichsspezifischen Zufriedenheiten der Studierenden des Campus Vaihingen und des Campus Hohenheim.  
Projektgruppe Campus: Slaby, M.; Grund, R.; Mayerl, J.; Noak, T.; Payk, B.; Sellke, P.; Urban, D.; Zudrell, I.
- No. 18/1999    Role Models and Trust in Socio-Political Institutions: A Case Study in Eastern Germany, 1992-96.  
Joachim Singelmann, Toby A. Ten Ayck, Dieter Urban
- No. 17/1998    Zur Interaktion zwischen Befragten und Erhebungsinstrument. Eine Untersuchung zur Konstanz des Meinungsurteils von Befragten im Interviewverlauf.  
Martin Slaby
- No. 16/1998    Technikeinstellungen: gibt es die überhaupt? Ergebnisse einer Längsschnittanalyse von Bewertungen der Gentechnik.  
Dieter Urban
- No. 15/1998    Bewertende Einstellungen zur Gentechnik: ihre Form, ihre Inhalte und ihre Dynamik. Kurzbericht zu Ergebnissen des Forschungsprojektes „Einstellungen zur Gentechnik“.  
Dieter Urban, Uwe Pfenning, Joachim Allhoff
- No. 14/1997    Politische Sozialisation im Transformationsprozeß: Die Entwicklung demokratiebezogener Einstellungen von ostdeutschen Jugendlichen und deren Eltern zwischen 1992 und 1996.  
Barbara Schmidt, Dieter Urban, Joachim Singelmann
- No. 13/1997    Die soziale Vererbung von „Ausländerfeindlichkeit“. Eine empirische Längsschnittanalyse der intra- und intergenerativen Transmission von sozialen Einstellungen.  
Dieter Urban und Joachim Singelmann
- No. 12/1997    Gentechnik: „Fluch oder Segen“ versus „Fluch und Segen“. Bilanzierende und differenzierende Bewertungen der Gentechnik in der öffentlichen Meinung.  
Dieter Urban und Uwe Pfenning
- No. 11/1996    Einstellungsmessung oder Einstellungsgenerierung? Die Bedeutung der informationellen Basis bei Befragten für die empirische Rekonstruktion von Einstellungen zu gentechnischen Anwendungen.  
Martin Slaby

## **SISS:**

### **Schriftenreihe des Instituts für Sozialwissenschaften der Universität Stuttgart**

<https://www.sowi.uni-stuttgart.de/forschung/schriftenreihe-des-instituts-fuer-sozialwissenschaften-siss/>

- No. 10/1996 Jugend und Politik im Transformationsprozeß. Eine Fallstudie zur Stabilität und Veränderung von politischen Einstellungen bei ostdeutschen Jugendlichen zwischen 1992 und 1995.  
Dieter Urban, Joachim Singelmann, Helmut Schröder
- No. 09/1995 Social Indicators in a Nonmetropolitan County: Testing the Representativeness of a Regional Nonrandom Survey in Eastern Germany.  
Dieter Urban, Joachim Singelmann
- No. 08/1995 Handbuch zur empirischen Erhebung von Einstellungen/Kognitionen zur Bio- und Gentechnologie (inklusive Diskette)  
(zweite, überarbeitete und erweiterte Auflage)  
Uwe Pfenning, Dieter Urban, Volker Weiss
- No. 07/1995 Latent Attitude Structures Directing the Perception of New Technologies.  
An Application of SEM-Methodology to the Construction of Attitude Measurement Models Related to Technologies of Prenatal Genetic Engineering and Testing.  
Dieter Urban
- No. 06/1995 Regionale Strukturprobleme.  
Sammelband der Beiträge zum Symposium vom 13. und 14. Oktober 1994.  
Frank C. Englmann (Hrsg.)
- No. 05/1995 Die Bedeutung Neuronaler Netze in der Ökonomie.  
Hermann Schnabl
- No. 04/1994 Rechtstatsachen zum Unternehmens- und Gesellschaftsrecht aus den neuen Bundesländern.  
Udo Kornblum
- No. 03/1994 Neue württembergische Rechtstatsachen zum Unternehmens- und Gesellschaftsrecht.  
Udo Kornblum
- No. 02/1994 Report on the German Machine Tool Industry.  
Frank C. Englmann, Christian Heyd, Daniel Köstler, Peter Paustian  
with the assistance of Susanne Baur and Peter Bergmann
- No. 01/1994 "Vertrauen" - soziologisch betrachtet. Ein Beitrag zur Analyse binärer Interaktionssysteme.  
Peter Antfang, Dieter Urban