

Ludwig-Maximilians-Universität München
Institut für Statistik



Simulationsstudie zur Sensitivität von Regressionsmodellen unter verschiedenen Fehlendmechanismen

Bachelorarbeit

vorgelegt von: Christoph Wiederkehr
Studiengang: Statistik
Betreuer: Prof. Dr. Christian Heumann
Prüfer: Prof. Dr. Christian Heumann
Abgabedatum: 08.09.2020

Zusammenfassung

Im Rahmen dieser Arbeit wird die Sensitivität von Regressionsmodellen unter verschiedenen Fehlendmechanismen untersucht. Zu diesem Zweck wird eine Simulationsstudie mit der Software R durchgeführt. Es wird vor allem genau auf die verwendeten Fehlendmechanismen sowie auf die Erzeugung der Datensätze eingegangen. Für die unterschiedlichen linearen Modelle und die logistische Regression werden zusätzlich noch weitere Behandlungsverfahren durchgeführt. Die Ergebnisse zeigen, dass unter den Fehlendmechanismen des Typs *missing not at random* die geschätzten Koeffizienten verzerrt sind. Die Güte hinsichtlich der geschätzten Koeffizienten ist auch stark abhängig vom Anteil der fehlenden Daten. Die verwendeten Behandlungsarten zeigen teilweise erwünschte Ansätze.

Inhaltsverzeichnis

1	Einleitung	1
1.1	Einführung	1
1.2	Betrachtungsweise und Fragestellungen	1
1.3	Aufbau der Arbeit	1
2	Theoretischer Hintergrund	2
2.1	Unterschiedliche Arten von fehlenden Daten	2
2.1.1	Missing completely at random (MCAR)	2
2.1.2	Missing at random (MAR)	3
2.1.3	Missing not at random (MNAR)	3
2.2	Behandlung von MNAR	3
2.2.1	Selektionsmodell	4
2.2.2	Pattern-Mixture Modell	4
2.3	Betrachtung der Koeffizientenschätzer und Konvergenzkriterien der Regressionsmodelle in R	4
3	Methodik	6
3.1	Datengenerierender Prozess mit multipler linearer Regression	6
3.1.1	Verwendete Fehlendmechanismen	7
3.1.2	Unterschiedliche Behandlungsarten	11
3.1.3	Verwendete Gütekriterien	12
3.2	Datengenerierender Prozess mit logistischer Regression	12
3.2.1	Verwendete Fehlendmechanismen	13
3.2.2	Unterschiedliche Behandlungsarten	13
3.2.3	Verwendete Gütekriterien	14
4	Ergebnisse	14
4.1	Ergebnisse der Modelle mit metrischer Zielgröße	14
4.1.1	Sensitivität der linearen Regression unter verschiedenen Fehlendmechanismen	15
4.1.2	Sensitivität der linearen Regression mit Ridge-Schätzung unter verschiedenen Fehlendmechanismen	24
4.1.3	Sensitivität für Brute-force-Methodiken mit linearen Modellen unter verschiedenen Fehlendmechanismen	24
4.2	Ergebnisse der Modelle mit binärer Zielgröße	25
4.2.1	Sensitivität bei logistischer Regression unter verschiedenen Fehlendmechanismen	26
4.2.2	Sensitivität bei Brute-force-Methodiken mit logistischer Regression unter verschiedenen Fehlendmechanismen	34
5	Diskussion	34
5.1	Mögliche Probleme bei den verwendeten Regressionsmodellen in R	34
5.2	Diskussion der Ergebnisse	35
5.2.1	Mögliche Erklärung der Ergebnisse	35
5.2.2	Einordnung und Bewertung der Ergebnisse	36
6	Fazit	37

A	Grafiken zu den Fehlend-Anteilen unter Generierung durch Logit-Modell	40
A.1	Datensatz mit metrischer Zielgröße	40
A.2	Datensatz mit binärer Zielgröße	44
B	Koeffizientenboxplots zu linearen Modellen	50
B.1	Normales Verfahren mit linearer Ridge Regression	50
B.2	Bruteforce1-Verfahren mit linearer Regression	58
B.3	Bruteforce2-Verfahren mit linearer Regression	66
C	Gütekriterien-Tabellen zu linearen Modellen	74
C.1	Normales Verfahren	74
C.2	Bruteforce1	92
C.3	Bruteforce2	110
D	Koeffizientenboxplots zur logistischen Regression	128
D.1	Bruteforce2 mit logistischer Regression	128
E	Gütekriterien-Tabellen zur logistischen Regression	136
E.1	Normales Verfahren	136
E.2	Bruteforce1	142
E.3	Bruteforce2	148

Tabellenverzeichnis

1	Gütekriterien-Table für MCAR mit normalen Verfahren (lm)	74
2	Gütekriterien-Table für MCAR mit normalen Verfahren (ridge, $\lambda=0$)	75
3	Gütekriterien-Table für MCAR mit normalen Verfahren (ridge)	76
4	Gütekriterien-Table für MNAR1 mit normalen Verfahren (lm)	77
5	Gütekriterien-Table für MNAR1 mit normalen Verfahren (ridge, $\lambda=0$)	78
6	Gütekriterien-Table für MNAR1 mit normalen Verfahren (ridge)	79
7	Gütekriterien-Table für MNAR2 mit normalen Verfahren (lm)	80
8	Gütekriterien-Table für MNAR2 mit normalen Verfahren (ridge, $\lambda=0$)	81
9	Gütekriterien-Table für MNAR2 mit normalen Verfahren (ridge)	82
10	Gütekriterien-Table für MNAR3 mit normalen Verfahren (lm)	83
11	Gütekriterien-Table für MNAR3 mit normalen Verfahren (ridge, $\lambda=0$)	84
12	Gütekriterien-Table für MNAR3 mit normalen Verfahren (ridge)	85
13	Gütekriterien-Table für LogitY mit normalen Verfahren (lm)	86
14	Gütekriterien-Table für LogitY mit normalen Verfahren (ridge, $\lambda=0$)	87
15	Gütekriterien-Table für LogitY mit normalen Verfahren (ridge)	88
16	Gütekriterien-Table für LogitX mit normalen Verfahren (lm)	89
17	Gütekriterien-Table für LogitX mit normalen Verfahren (ridge, $\lambda=0$)	90
18	Gütekriterien-Table für LogitX mit normalen Verfahren (ridge)	91
19	Gütekriterien-Table für MCAR mit Bruteforce1-Verfahren (lm)	92
20	Gütekriterien-Table für MCAR mit Bruteforce1-Verfahren (ridge, $\lambda=0$)	93
21	Gütekriterien-Table für MCAR mit Bruteforce1-Verfahren (ridge)	94
22	Gütekriterien-Table für MNAR1 mit Bruteforce1-Verfahren (lm)	95
23	Gütekriterien-Table für MNAR1 mit Bruteforce1-Verfahren (ridge, $\lambda=0$)	96
24	Gütekriterien-Table für MNAR1 mit Bruteforce1-Verfahren (ridge)	97

25	Gütekriterien-Tablle für MNAR2 mit Bruteforce1-Verfahren (lm)	98
26	Gütekriterien-Tablle für MNAR2 mit Bruteforce1-Verfahren (ridge, $\lambda=0$)	99
27	Gütekriterien-Tablle für MNAR2 mit Bruteforce1-Verfahren (ridge)	100
28	Gütekriterien-Tablle für MNAR3 mit Bruteforce1-Verfahren (lm)	101
29	Gütekriterien-Tablle für MNAR3 mit Bruteforce1-Verfahren (ridge, $\lambda=0$)	102
30	Gütekriterien-Tablle für MNAR3 mit Bruteforce1-Verfahren (ridge)	103
31	Gütekriterien-Tablle für LogitY mit Bruteforce1-Verfahren (lm)	104
32	Gütekriterien-Tablle für LogitY mit Bruteforce1-Verfahren (ridge, $\lambda=0$)	105
33	Gütekriterien-Tablle für LogitY mit Bruteforce1-Verfahren (ridge)	106
34	Gütekriterien-Tablle für LogitX mit Bruteforce1-Verfahren (lm)	107
35	Gütekriterien-Tablle für LogitX mit Bruteforce1-Verfahren (ridge, $\lambda=0$)	108
36	Gütekriterien-Tablle für LogitX mit Bruteforce1-Verfahren (ridge)	109
37	Gütekriterien-Tablle für MCAR mit Bruteforce2-Verfahren (lm)	110
38	Gütekriterien-Tablle für MCAR mit Bruteforce2-Verfahren (ridge, $\lambda=0$)	111
39	Gütekriterien-Tablle für MCAR mit Bruteforce2-Verfahren (ridge)	112
40	Gütekriterien-Tablle für MNAR1 mit Bruteforce2-Verfahren (lm)	113
41	Gütekriterien-Tablle für MNAR1 mit Bruteforce2-Verfahren (ridge, $\lambda=0$)	114
42	Gütekriterien-Tablle für MNAR1 mit Bruteforce2-Verfahren (ridge)	115
43	Gütekriterien-Tablle für MNAR2 mit Bruteforce2-Verfahren (lm)	116
44	Gütekriterien-Tablle für MNAR2 mit Bruteforce2-Verfahren (ridge, $\lambda=0$)	117
45	Gütekriterien-Tablle für MNAR2 mit Bruteforce2-Verfahren (ridge)	118
46	Gütekriterien-Tablle für MNAR3 mit Bruteforce2-Verfahren (lm)	119
47	Gütekriterien-Tablle für MNAR3 mit Bruteforce2-Verfahren (ridge, $\lambda=0$)	120
48	Gütekriterien-Tablle für MNAR3 mit Bruteforce2-Verfahren (ridge)	121
49	Gütekriterien-Tablle für LogitY mit Bruteforce2-Verfahren (lm)	122
50	Gütekriterien-Tablle für LogitY mit Bruteforce2-Verfahren (ridge, $\lambda=0$)	123
51	Gütekriterien-Tablle für LogitY mit Bruteforce2-Verfahren (ridge)	124
52	Gütekriterien-Tablle für LogitX mit Bruteforce2-Verfahren (lm)	125
53	Gütekriterien-Tablle für LogitX mit Bruteforce2-Verfahren (ridge, $\lambda=0$)	126
54	Gütekriterien-Tablle für LogitX mit Bruteforce2-Verfahren (ridge)	127
55	Gütekriterien-Tablle für MCAR mit normalen Verfahren (glm)	136
56	Gütekriterien-Tablle für MNAR1 mit normalen Verfahren (glm)	137
57	Gütekriterien-Tablle für MNAR2 mit normalen Verfahren (glm)	138
58	Gütekriterien-Tablle für MNAR3 mit normalen Verfahren (glm)	139
59	Gütekriterien-Tablle für LogitY mit normalen Verfahren (glm)	140
60	Gütekriterien-Tablle für LogitX mit normalen Verfahren (glm)	141
61	Gütekriterien-Tablle für MCAR mit Bruteforce1-Verfahren (glm)	142
62	Gütekriterien-Tablle für MNAR1 mit Bruteforce1-Verfahren (glm)	143
63	Gütekriterien-Tablle für MNAR2 mit Bruteforce1-Verfahren (glm)	144
64	Gütekriterien-Tablle für MNAR3 mit Bruteforce1-Verfahren (glm)	145
65	Gütekriterien-Tablle für LogitY mit Bruteforce1-Verfahren (glm)	146
66	Gütekriterien-Tablle für LogitX mit Bruteforce1-Verfahren (glm)	147
67	Gütekriterien-Tablle für MCAR mit Bruteforce2-Verfahren (glm)	148
68	Gütekriterien-Tablle für MNAR1 mit Bruteforce2-Verfahren (glm)	149
69	Gütekriterien-Tablle für MNAR2 mit Bruteforce2-Verfahren (glm)	150
70	Gütekriterien-Tablle für MNAR3 mit Bruteforce2-Verfahren (glm)	151
71	Gütekriterien-Tablle für LogitY mit Bruteforce2-Verfahren (glm)	152
72	Gütekriterien-Tablle für LogitX mit Bruteforce2-Verfahren (glm)	153

Abbildungsverzeichnis

1	Fehlend-Anteile durch LogitY auf Datensatz (n=1000, metrisches Y)	9
2	Fehlend-Anteile durch LogitX auf Datensatz (n=1000, metrisches Y)	10
3	Boxplots für β_1 mit normalen Verfahren & lm für Datensatz (n=1000, metrisches Y)	15
4	Boxplots für β_2 mit normalen Verfahren & lm für Datensatz (n=1000, metrisches Y)	16
5	Boxplots für β_3 mit normalen Verfahren & lm für Datensatz (n=1000, metrisches Y)	17
6	Boxplots für β_4 mit normalen Verfahren & lm für Datensatz (n=1000, metrisches Y)	18
7	Boxplots für β_5 mit normalen Verfahren & lm für Datensatz (n=1000, metrisches Y)	19
8	Boxplots für β_6 mit normalen Verfahren & lm für Datensatz (n=1000, metrisches Y)	20
9	Boxplots für β_7 mit normalen Verfahren & lm für Datensatz (n=1000, metrisches Y)	21
10	Boxplots für β_8 mit normalen Verfahren & lm für Datensatz (n=1000, metrisches Y)	22
11	Boxplots für β_1 mit normalen Verfahren & glm für Datensatz (n=1000, binäres Y) .	26
12	Boxplots für β_2 mit normalen Verfahren & glm für Datensatz (n=1000, binäres Y) .	27
13	Boxplots für β_3 mit normalen Verfahren & glm für Datensatz (n=1000, binäres Y) .	28
14	Boxplots für β_4 mit normalen Verfahren & glm für Datensatz (n=1000, binäres Y) .	29
15	Boxplots für β_5 mit normalen Verfahren & glm für Datensatz (n=1000, binäres Y) .	30
16	Boxplots für β_6 mit normalen Verfahren & glm für Datensatz (n=1000, binäres Y) .	31
17	Boxplots für β_7 mit normalen Verfahren & glm für Datensatz (n=1000, binäres Y) .	32
18	Boxplots für β_8 mit normalen Verfahren & glm für Datensatz (n=1000, binäres Y) .	33
19	Fehlend-Anteile durch LogitY auf Datensatz (n=50, metrisches Y)	40
20	Fehlend-Anteile durch LogitY auf Datensatz (n=200, metrisches Y)	41
21	Fehlend-Anteile durch LogitX auf Datensatz (n=50, metrisches Y)	42
22	Fehlend-Anteile durch LogitX auf Datensatz (n=200, metrisches Y)	43
23	Fehlend-Anteile durch LogitY auf Datensatz (n=50, binäres Y)	44
24	Fehlend-Anteile durch LogitY auf Datensatz (n=200, binäres Y)	45
25	Fehlend-Anteile durch LogitY auf Datensatz (n=1000, binäres Y)	46
26	Fehlend-Anteile durch LogitX auf Datensatz (n=50, binäres Y)	47
27	Fehlend-Anteile durch LogitX auf Datensatz (n=200, binäres Y)	48
28	Fehlend-Anteile durch LogitX auf Datensatz (n=1000, binäres Y)	49
29	Boxplots für β_1 mit normalen Verfahren & lm mit Ridge für Datensatz (n=1000, metrisches Y)	50
30	Boxplots für β_2 mit normalen Verfahren & lm mit Ridge für Datensatz (n=1000, metrisches Y)	51
31	Boxplots für β_3 mit normalen Verfahren & lm mit Ridge für Datensatz (n=1000, metrisches Y)	52
32	Boxplots für β_4 mit normalen Verfahren & lm mit Ridge für Datensatz (n=1000, metrisches Y)	53
33	Boxplots für β_5 mit normalen Verfahren & lm mit Ridge für Datensatz (n=1000, metrisches Y)	54
34	Boxplots für β_6 mit normalen Verfahren & lm mit Ridge für Datensatz (n=1000, metrisches Y)	55
35	Boxplots für β_7 mit normalen Verfahren & lm mit Ridge für Datensatz (n=1000, metrisches Y)	56
36	Boxplots für β_4 mit normalen Verfahren & lm mit Ridge für Datensatz (n=1000, metrisches Y)	57
37	Boxplots für β_1 mit Bruteforce1-Verfahren & lm für Datensatz (n=1000, metrisches Y)	58
38	Boxplots für β_2 mit Bruteforce1-Verfahren & lm für Datensatz (n=1000, metrisches Y)	59

39	Boxplots für β_3 mit Bruteforce1-Verfahren & lm für Datensatz (n=1000, metrisches Y)	60
40	Boxplots für β_4 mit Bruteforce1-Verfahren & lm für Datensatz (n=1000, metrisches Y)	61
41	Boxplots für β_5 mit Bruteforce1-Verfahren & lm für Datensatz (n=1000, metrisches Y)	62
42	Boxplots für β_6 mit Bruteforce1-Verfahren & lm für Datensatz (n=1000, metrisches Y)	63
43	Boxplots für β_7 mit Bruteforce1-Verfahren & lm für Datensatz (n=1000, metrisches Y)	64
44	Boxplots für β_8 mit Bruteforce1-Verfahren & lm für Datensatz (n=1000, metrisches Y)	65
45	Boxplots für β_1 mit Bruteforce2-Verfahren & lm für Datensatz (n=1000, metrisches Y)	66
46	Boxplots für β_2 mit Bruteforce2-Verfahren & lm für Datensatz (n=1000, metrisches Y)	67
47	Boxplots für β_3 mit Bruteforce2-Verfahren & lm für Datensatz (n=1000, metrisches Y)	68
48	Boxplots für β_4 mit Bruteforce2-Verfahren & lm für Datensatz (n=1000, metrisches Y)	69
49	Boxplots für β_5 mit Bruteforce2-Verfahren & lm für Datensatz (n=1000, metrisches Y)	70
50	Boxplots für β_6 mit Bruteforce2-Verfahren & lm für Datensatz (n=1000, metrisches Y)	71
51	Boxplots für β_7 mit Bruteforce2-Verfahren & lm für Datensatz (n=1000, metrisches Y)	72
52	Boxplots für β_8 mit Bruteforce2-Verfahren & lm für Datensatz (n=1000, metrisches Y)	73
53	Boxplots für β_1 mit Bruteforce2-Verfahren & glm für Datensatz (n=1000, binäres Y)	128
54	Boxplots für β_2 mit Bruteforce2-Verfahren & glm für Datensatz (n=1000, binäres Y)	129
55	Boxplots für β_3 mit Bruteforce2-Verfahren & glm für Datensatz (n=1000, binäres Y)	130
56	Boxplots für β_4 mit Bruteforce2-Verfahren & glm für Datensatz (n=1000, binäres Y)	131
57	Boxplots für β_5 mit Bruteforce2-Verfahren & glm für Datensatz (n=1000, binäres Y)	132
58	Boxplots für β_6 mit Bruteforce2-Verfahren & glm für Datensatz (n=1000, binäres Y)	133
59	Boxplots für β_7 mit Bruteforce2-Verfahren & glm für Datensatz (n=1000, binäres Y)	134
60	Boxplots für β_8 mit Bruteforce2-Verfahren & glm für Datensatz (n=1000, binäres Y)	135

1 Einleitung

1.1 Einführung

Fehlende Daten stellen in der angewandten Statistik ein häufiges Problem dar. Die Gründe für fehlende Beobachtungen können allerdings sehr vielfältig sein. Handelt es sich bspw. bei den Daten um erhobene Umfragebögen, so kann dies durch mangelndes Wissen des Befragten, durch absichtliches Nichtbeantworten, durch Übersehen einzelner Fragen oder durch mangelndes Verständnis der Fragestellung geschehen. Geht es um eine medizinische Studie für eine spezielle Therapie (Längsschnittstudie), so können Probanden ihren Wohnort wechseln oder die Probanden erscheinen nicht mehr zu den Kontrolluntersuchungen, da es ihnen entweder sehr gut oder schlecht geht.

Für den Umgang mit fehlenden Daten gibt es bereits zahlreiche Ansätze. Dabei wird meistens die Annahme getroffen, dass das Fehlen der Ausprägungen der Zielgröße ignorierbar ist. Die Rechtfertigung dafür ist im jeweiligen Zusammenhang zu überprüfen.

Die vorliegende Arbeit stellt eine kleine Simulationsstudie dar, in der in R diverse Fehlendmechanismen erstellt und angewendet werden. Auf den verbleibenden Daten werden anschließend unter unterschiedlichen Verfahren Regressionsmodelle geschätzt und ihre Sensitivität genauer betrachtet.

1.2 Betrachtungsweise und Fragestellungen

Von primärem Interesse sind die Koeffizienten der geschätzten Regressionsmodelle auf den verbliebenen Datensatz. Im Rahmen dieser Arbeit werden nur Werte für die Zielvariable Y fehlen.

- Ist beispielsweise die Methode der kleinsten Quadrate verzerrt und hängt die Stärke der Verzerrung mit dem Anteil an fehlenden Y -Werten zusammen?
- Gibt es Unterschiede bezüglich der Gütekriterien, wenn verschiedene Fehlendmechanismen verwendet werden und auch für die jeweiligen Fehlendmechanismen unterschiedlich starke Versionen auftreten?
- Wie verhalten sich die Gütekriterien der geschätzten Koeffizienten untereinander?
- Treten Unterschiede bezüglich der Gütekriterien auf, wenn diverse Vorgehensarten bei der Behandlung der fehlenden Daten benutzt werden?

1.3 Aufbau der Arbeit

Im Zweiten Kapitel werden grundlegend auf die unterschiedlichen Fehlendmechanismen beschrieben. Hierbei wird auch auf die Ignorierbarkeit fehlender Daten und auf zwei mögliche Modellierungsansätze unter *missing not at random* (MNAR) eingegangen. Desweiteren werden kurz die Konvergenzkriterien unterschiedlicher Regressionsmodelle in R, die später verwendet werden, vorgestellt.

Im dritten Kapitel wird ausführlich die Simulationsdurchführung erklärt. Diese enthält die Betrachtung der Datenerstellung, die verwendeten Fehlendmechanismen, die unterschiedlichen Behandlungsarten in Form von drei Verfahren mit diverser Modellschätzung und die dazugehörigen Gütekriterien für die Bewertung.

Im vierten Kapitel wird auf die Simulationsergebnisse eingegangen. Es wird vor allem die Sensitivität der linearen Modelle und der logistischen Regression hinsichtlich der geschätzten Koeffizienten genauer untersucht. Dabei werden zusätzlich noch die Ergebnisse der Bruteforce-Methodiken vorgestellt.

Das fünfte Kapitel schildert aufgetretene Probleme und Schwierigkeiten, die im Laufe der Entwicklung der Simulationsstudie aufgetreten sind. Außerdem werden mögliche Erklärungen für die

Ergebnisse präsentiert.

Das sechste Kapitel, das Fazit, enthält in komprimierter Form die wichtigsten Ergebnisse.

2 Theoretischer Hintergrund

2.1 Unterschiedliche Arten von fehlenden Daten

Es werden in drei Arten von fehlenden Daten unterschieden: *Missing completely at random* (MCAR), *missing at random* (MAR) und *missing not at random* (MNAR) (Kreutzmann, 2019).

Die Einteilung in die drei Kategorien soll hier in einem Regressionskontext erklärt werden. Hierfür sei

$$P(Y|R, X)$$

die Verteilung der Zielgröße Y gegeben die Kovariable X beziehungsweise

$$P(Y|R, X, \beta)$$

die parametrische Verteilung mit Parametervektor β . Fehlende Werte können dabei jedoch nur für die abhängige Zielvariable Y auftreten. Zusätzlich ist R eine Indikatorvariable, die die Realisierung 1 erhält, wenn der Wert für Y vorliegt, und 0, falls der Wert für Y nicht beobachtet wird. Desweiteren sei

$$P(R|Y, X)$$

die Verteilung von $R|Y, X$, also die Verteilung des Fehlens bedingt auf die gegebenen abhängigen und unabhängigen Variablen. Es kann ebenso eine parametrische Variante mit

$$P(R|Y, X, \phi)$$

angegeben werden (Roderick J. A. Little, 2002, S.12 ff.). Als nächstes werden die unterschiedlichen Fehlendarten unter Verwendung der parametrisierten Verteilungen von $Y|R, X$ und $R|Y, X$ betrachtet (Genbäck, 2016, S.2).

2.1.1 Missing completely at random (MCAR)

Fehlende Werte der abhängigen Zielvariable Y werden als *missing completely at random* (MCAR) bezeichnet, wenn

$$P(R|Y, X, \phi) = P(R|\phi)$$

gilt. R ist also sowohl von X als auch von Y unabhängig. Die Verteilung der Variable Y ist somit nicht von R , sondern von X abhängig, wobei

$$P(Y|R, X, \beta) = P(Y|X, \beta)$$

gilt (Lehmann, 2004, S.11). Falls ein MCAR-Mechanismus vorliegt, so können die beobachteten Daten als Zufallsstichprobe aus der Grundgesamtheit angesehen werden. Hieraus ergibt sich die sogenannte *complete case* Analyse, die unverzerrte Schätzungen ermöglicht (Genbäck, 2016, S.2).

2.1.2 Missing at random (MAR)

Fehlende Beobachtungen von Y werden als *missing at random (MAR)* bezeichnet, falls R von der beobachteten Variable X abhängt, jedoch die bedingte Dichte von $Y|X$ unabhängig von der Zielvariable Y ist, also

$$P(R|Y, X, \phi) = P(R|X, \phi)$$

gilt. Dadurch folgt unmittelbar für die Verteilung von Y , dass

$$P(Y|R, X, \beta) = P(Y|X, \beta)$$

gilt (Lehmann, 2004, S.12 ff.). Damit ist die Parameterschätzung für β bei einem Regressionsmodell im MAR- und MCAR-Fall unter *complete case* Analyse unverzerrt. Allerdings ist die Schätzung nicht so effizient, da weniger Beobachtungspaare vorliegen als im Vergleich zum vollständigen Datensatz. Außerdem gilt für Schätzer wie beispielsweise den Stichprobenmittelwert und die Stichprobenvarianz unter *complete case* Analyse und MAR, dass diese verzerrt sind (Genbäck, 2016, S.2). Ein weiteres Problem besteht darin, dass in der Praxis nicht genau bestimmt werden kann, ob der vorliegende Fehlendmechanismus vom Typ MAR ist, denn schließlich sind die fehlenden Werte gerade nicht beobachtbar. Die Entscheidung, ob ein MAR-Mechanismus vorliegt, wird im Anwendungsfall aus Plausibilitätsüberlegungen getroffen (Lehmann, 2004, S.13).

2.1.3 Missing not at random (MNAR)

Fehlende Werte der Zielvariable Y werden als *missing not at random (MNAR)* bezeichnet, wenn R , Y und X abhängig voneinander sind, also

$$P(R|Y, X, \phi) \neq P(R|X, \phi)$$

gilt. In diesem Fall hängt die Verteilung von R auch von den fehlenden Werten der Zielvariable Y selbst ab. Daraus ergibt sich wieder, dass

$$P(Y|R, X, \beta) \neq P(Y|X, \beta)$$

ist, da die Verteilung von Y sowohl von R als auch X abhängig ist (Lehmann, 2004, S.14).

2.2 Behandlung von MNAR

Der Umgang mit den fehlenden Werten für Y unter MNAR setzt im Allgemeinen die Kenntnis der bedingten Verteilung $R|Y, X, \phi$ voraus, die aufgrund der Abhängigkeit von den unbeobachteten Werten von Y jedoch in den meisten Fällen unbekannt ist. Als nächstes steht die Frage, ob die fehlenden Beobachtungen der Zielvariable Y ignoriert werden können, oder ob die Schätzer in diesem Fall verzerrt sind, im vordergründigen Interesse. Hierbei gilt innerhalb der frequentistischen Theorie, dass der Fehlendmechanismus als ignorierbar bezeichnet wird, wenn

- der Fehlendmechanismus vom Typ MCAR oder MAR ist
- der gemeinsame Parameterraum von (ϕ, β) sich aus dem kartesischen Kreuzprodukt der einzelnen Parameterräumen von ϕ und β ergibt.

Daraus lässt sich unmittelbar ableiten, dass der Fehlendmechanismus MNAR nicht ignorierbar ist und somit eine gesonderte Behandlung der fehlenden Werte nötig ist (Lehmann, 2004, S.16-21).

In der Praxis helfen folgende Überlegungen bei der Entscheidung, ob die fehlenden Daten nicht ignorierbar sind (van Buuren, 2018, Kapitel 3.8.1):

- Wichtige Variablen, die den Fehlendmechanismus bestimmen, sind nicht verfügbar.
- Die Annahme, dass sich Untersuchungsobjekte, die fehlende Werte aufweisen, von den anderen Untersuchungseinheiten unterscheiden, wirkt glaubwürdig.
- Der Datensatz ist abgeschnitten.

Im Folgenden werden zwei statistische Modellierungsansätze für nicht ignorierbare fehlende Daten vorgestellt.

2.2.1 Selektionsmodell

Das Selektionsmodell zerlegt die gemeinsame Dichte von Y und R bedingt auf X zu:

$$P(Y, R|X) = P(Y|X) \cdot P(R|Y, X).$$

Die Verteilung der Dichte von $Y|X$ wird also mit dem Prozess der fehlenden Daten $R|Y, X$ gewichtet. Das Selektionsmodell ermöglicht es, die beiden Verteilungen separat voneinander zu betrachten. Jedoch müssen diese zuerst vom Anwender spezifiziert werden, da beide Verteilungen unbekannt sind (van Buuren, 2018, Kapitel 3.8.2).

2.2.2 Pattern-Mixture Modell

Das Pattern-Mixture Modell faktorisiert die gemeinsame Dichte von Y und R bedingt auf X zu:

$$P(Y, R|X) = P(Y|R, X) \cdot P(R|X).$$

Die erhaltenen Verteilungen lassen sich weiter in

$$P(Y|R, X) \cdot P(R|X) = P(Y|R = 1, X) \cdot P(R = 1|X) + P(Y|R = 0, X) \cdot P(R = 0|X)$$

faktorisieren (van Buuren, 2018, Kapitel 3.8.3).

Die bedingten Verteilungen $P(Y|R=1, X)$, $P(R=1|X)$ und $P(R=0|X)$ lassen sich aus den beobachteten Daten ermitteln. Lediglich die Verteilung $P(Y|R=0, X)$ ist unbekannt und muss spezifiziert werden.

2.3 Betrachtung der Koeffizientenschätzer und Konvergenzkriterien der Regressionsmodelle in R

Für die lineare Regression werden die Koeffizienten mit der *Methode der kleinsten Quadrate (KQ)*

$$KQ(\beta) = (y - X\beta)^T (y - X\beta)$$

geschätzt ($X \in \mathbb{R}^{n \times p}$, $y \in \mathbb{R}^n$). Es soll die Summe der quadrierten Abweichungen bezüglich $\beta \in \mathbb{R}^p$ minimiert werden. Durch Differenzieren und Nullsetzen des Minimierungsproblems ergibt sich der KQ-Schätzer als Lösung der Normalgleichung,

$$\hat{\beta} = (X^T X)^{-1} X^T y.$$

Die Funktion `lm()` in R löst direkt die Normalgleichung

$$(X^T X)\beta = X^T y$$

(Ludwig Fahrmeir, 2009, S.91ff.).

Für die logistische Regression erfolgt die Schätzung der Koeffizienten β durch die Maximierung der Likelihood (Ludwig Fahrmeir, 2009, S.198ff.). Mithilfe der Bildung der Log-Likelihood und dem anschließenden Differenzieren ergibt sich die Scorefunktion zu

$$s(\hat{\beta}) = \sum_{i=1}^n x_i \left(y_i - \frac{\exp(x_i^T \beta)}{1 + \exp(x_i^T \beta)} \right)$$

für $i = 1, \dots, n$, wobei n die Anzahl an Beobachtungen ist. Durch Nullsetzen ergibt sich ein mehrdimensionales, nichtlineares Gleichungssystem für β . Die Funktion `glm()` in R löst das Gleichungssystem numerisch in Form einer iterativ gewichteten KQ-Schätzung (Ludwig Fahrmeir, 2009, S.222ff.):

$$\hat{\beta}^{(k+1)} = (X^T W^{(k)} X)^{-1} X^T W^{(k)} \tilde{y}^{(k)} \quad \text{mit } k = 0, 1, 2, \dots$$

und dabei ist

$$\tilde{y}_i^{(k)} = \tilde{y}_i(\hat{\beta}^{(k)}) = x_i^T \hat{\beta}^{(k)} + d_i^{-1}(\hat{\beta}^{(k)})(y_i - \hat{\pi}_i(\hat{\beta}^{(k)}))$$

und $W^{(k)}$ die Gewichtsmatrix, ausgewertet für $\beta = \hat{\beta}^{(k)}$. Nach einigen Iterationsschritten stoppt der Algorithmus, wenn das Abbruchkriterium

$$\frac{\|\hat{\beta}^{(k+1)} - \hat{\beta}^{(k)}\|}{\|\hat{\beta}^{(k)}\|} \leq \epsilon$$

erreicht wird.

Für die lineare Regression mit Ridge ergibt sich die Residuenquadratsumme (RSS)

$$RSS(\lambda) = (y - X\beta)^T (y - X\beta) + \lambda\beta^T \beta$$

mit $X \in \mathbb{R}^{n \times p}$, $y \in \mathbb{R}^n$ und $\beta \in \mathbb{R}^p$. Wenn man diese nun nach β umstellt, ergibt sich

$$\hat{\beta}_{ridge} = (X^T X + \lambda I)^{-1} X^T y.$$

Dabei ist I die Einheitsmatrix und λ der Regularisierungsparameter, der die Reduzierung/Bestrafung der Koeffizienten bestimmt. Je größer der Parameter λ ausfällt, desto größer ist die Reduzierung der Koeffizienten in Richtung null. Für $\lambda \rightarrow 0$ würde $\hat{\beta}_{ridge} \rightarrow \hat{\beta}_{KQ}$ und für $\lambda \rightarrow \infty$, $\hat{\beta}_{ridge} \rightarrow 0$ entsprechen (Trevor Hastie, 2008, S.61ff.).

Das Konvergenzkriterium für die Funktion `glmnet()` fokussiert sich auf den Einfluss der Koeffizienten auf die geschätzten Werte. Für das lineare Modell wird $\Delta_{max} = \max_j \Delta_j$ mit

$$\Delta_j = \frac{1}{n} \sum_{i=1}^n w_j (x_{ij} \hat{\beta}_j^0 - x_{ij} \hat{\beta}_j^k)^2$$

betrachtet, das den Einfluss des k-ten berechneten Koeffizienten $\hat{\beta}_j$ auf die geschätzten Werte misst. Folglich wird gestoppt, wenn der Einfluss im k-ten Iterationsschritt am größten ist (Junyang Qian, 2016, S.36).

3 Methodik

3.1 Datengenerierender Prozess mit multipler linearer Regression

Ein einzelner Datensatz enthält insgesamt acht Kovariablen und eine Zielvariable. Die Kovariablen werden durch einen vorgegebenen Zufallsmechanismus erzeugt. Zur besseren Verständlichkeit und Übersicht als auch zum späteren Vergleich werden jeweils univariate Wahrscheinlichkeitsprozesse, sodass die Kovariablen unabhängig voneinander sind, als auch multivariate Wahrscheinlichkeitsprozesse, damit diese Kovariablen nun untereinander korrelieren, angenommen.

Die Aufteilung der Kovariablen, ob diese univariat oder multivariat erzeugt wird, erfolgt im gleichen Verhältnis. Also sind in diesem Fall die ersten vier Kovariablen univariat und die letzten vier Kovariablen multivariat erzeugt.

Die Kovariable X_1 ist univariat normalverteilt mit:

$$X_1 \sim \mathcal{N}(-1, 1).$$

Die Kovariable X_2 ist ebenso univariat normalverteilt mit:

$$X_2 \sim \mathcal{N}(1, 1).$$

Die Kovariable X_3 ist binär codiert mit 0 und 1, und folgt einer univariaten Binomialverteilung:

$$X_3 \sim \mathcal{B}(1, 0.4).$$

Die Kovariable X_4 ist binär codiert mit 0 und 1, und folgt einer univariaten Binomialverteilung:

$$X_4 \sim \mathcal{B}(1, 0.85).$$

Die Kovariablen X_5 , X_6 , X_7 und X_8 werden mit einer multivariaten Verteilung generiert:

$$X \sim \mathcal{N}(\mu, \Sigma)$$

$$X = \begin{pmatrix} X_5 \\ X_6 \\ X_7 \\ X_8 \end{pmatrix} \text{ mit } \mu = \begin{pmatrix} 1 \\ 2 \\ 0 \\ 1 \end{pmatrix} \text{ und } \Sigma = \begin{pmatrix} 1 & 1 & -1 & 0 \\ 1 & 2 & 0 & 0.5 \\ -1 & 0 & 3 & -0.25 \\ 0 & 0.5 & -0.25 & 1 \end{pmatrix}.$$

Die Kovariablen X_7 und X_8 sollen selbst jedoch wieder binäre Variablen darstellen. Die Generierung mit der multivariaten Normalverteilung dient lediglich dem Zweck, dass diese nun auch untereinander korreliert sind. Die Transformierung von der metrischen zur binären Einflussgröße erfolgt durch den selbstgewählten Schwellenwert von 0. Es wird also effektiv jeder positiven reellen Zahl eingeschlossen der 0, der Wert 1 zugeordnet und jeder negativen reellen Zahl der Wert 0.

Die Zielvariable Y wird durch ein multiples lineares Modell mit beliebig selbstgewählten Koeffizienten erzeugt (Fahrmeier 1.Auflage S.24):

$$y = X\beta + \epsilon$$

beziehungsweise mit

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i1} + \beta_3 x_{i3} + \beta_4 x_{i4} + \beta_5 x_{i5} + \beta_6 x_{i6} + \beta_7 x_{i7} + \beta_8 x_{i8} + \epsilon_i$$

und $i = 1, \dots, n$, wobei n die Anzahl an Beobachtungen ist. Außerdem ist ϵ normalverteilt mit:

$$\epsilon \sim \mathcal{N}(0, 1)$$

Der Koeffizientenvektor β hat die folgenden Werte:

$$\beta = \begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_3 \\ \beta_4 \\ \beta_5 \\ \beta_6 \\ \beta_7 \\ \beta_8 \end{pmatrix} = \begin{pmatrix} 0 \\ -1 \\ 2 \\ 1 \\ -1 \\ 2 \\ -2 \\ 1 \\ -1 \end{pmatrix}$$

Mit diesen Vorgaben werden jeweils drei unterschiedlich große Datensätze zu je 50, 200 und 1000 Beobachtungen erzeugt. Im weiteren Verfahren werden nun insgesamt 150 solcher Datensätze mit jeweils 50, 200 und 1000 Beobachtungen angefertigt. Dabei ist jedoch zu beachten, dass die Kovariablen X nur einmal mit dem jeweiligen Zufallsprozess wie oben beschrieben und entsprechend ihrer Beobachtungsgröße, generiert werden. Somit werden die Kovariablen aus der ersten Datenerstellung praktisch gesehen festgehalten und die Zielvariable Y wird immer wieder durch das multiple lineare Modell mit den angegebenen Koeffizienten erzeugt. Dadurch sind die Realisierungen der Kovariablen in jedem dieser 150 Datensätze identisch und die Werte der Zielvariable Y variieren.

3.1.1 Verwendete Fehlendmechanismen

Für den Fehlendmechanismus MCAR werden zufällig einzelne Beobachtungen ausgewählt und dann dort die Werte der Zielvariablen Y gelöscht.

Für den Fehlendmechanismus MAR werden die Kovariablen X aus dem Datensatz als Variablen für das Logit-Modell verwendet (Fahrmeier 1.Auflage S.30ff.):

$$\pi_i = P(R_i = 1|x_i) = E(R_i|x_i) = \frac{1}{1 + \exp(-(\beta_0 + x_i^T \beta))} \in (0, 1)$$

mit $i = 1, \dots, n$ und n ist die Anzahl an Beobachtungen. Die Gleichung lässt sich wie folgt verstehen: Für eine binäre Zielgröße R_i modelliert man indirekt die Wahrscheinlichkeit π_i , dass diese 1 annimmt, gegebenen $x_i = (x_{i1}, \dots, x_{ip})^T$. Das entspricht auch dem bedingten Erwartungswert der Zielgröße R_i . Hier ist β_0 flexibel gehalten und die restlichen Koeffizienten besitzen dieselben Werte wie bei der Datenerstellung. Folglich wird einzig der Intercept dazu genutzt um die Wahrscheinlichkeit π_i zu beeinflussen und den Fehlendprozess zu steuern. Anschließend wird mit der erhaltenen Wahrscheinlichkeit ein Bernoulliprozess durchgeführt und man erhält eine binäre Variable R mit den Werten 0 und 1

$$R_i \sim \mathcal{B}(1, \pi_i).$$

Der Mechanismus bekommt die Bezeichnung *LogitX*. In der Simulation wird für $R=1$ der Y -Wert gelöscht und für $R=0$ der Y -Wert beibehalten, Wodurch relativ gesehen auch wieder eher große Y -Werte gefiltert werden.

Für den Fehlendmechanismus MNAR werden unterschiedliche Vorgehensweisen durchgeführt. Hierbei lassen sich die Methodiken in zwei Gruppen einteilen. Bei der ersten Gruppe handelt es sich um

eine zufällige Ziehung aus vorher festgelegten Quantilen der Zielvariable Y. Bei der zweiten Gruppe wurden die fehlenden Y-Werte durch ein Logit-Modell modelliert.

Bei der Methodik der Ziehung aus Quantilen wurden insgesamt drei unterschiedliche Gewichtungen der Quantile betrachtet. Der erste Mechanismus *MNAR1* zieht jeweils 50% der fehlenden Daten zufällig aus dem dritten und vierten Quantil und wählt somit viele große Werte für Y aus, sodass im verbleibenden Datensatz hauptsächlich nur noch kleinere Werte verbleiben.

Beim Mechanismus *MNAR2* werden jeweils 25% der fehlenden Daten zufällig aus dem zweiten und dritten Quantil und 50% aus dem vierten Quantil gezogen, dadurch wird der Anteil großer fehlender Y-Werte im Vergleich zu *MNAR1* etwas abgeschwächt.

Beim Mechanismus *MNAR3* werden jeweils 50% der fehlenden Daten zufällig aus dem zweiten und dritten Quantil gezogen, sodass hierbei hauptsächlich Y-Werte nahe des Medians ausgewählt werden und sowohl große als auch kleine Y-Werte im Datensatz verbleiben.

Bei der fehlenden Y-Wert Generierung durch ein Logit-Modell wird die Zielvariable Y aus dem Datensatz selbst als Variable im Logit-Modell aufgenommen

$$\pi_i = P(R_i = 1|y_i) = E(R_i|x_i) = \frac{1}{1 + \exp(-(\beta_0 + y_i\beta_1))} \in (0, 1) \text{ mit } \beta_1 = 1$$

mit $i = 1, \dots, n$. Die Wahrscheinlichkeit π_i wird wieder in Abhängigkeit von β_0 (Intercept) modelliert und der Fehlendprozess somit gesteuert. Danach wird die ermittelte Wahrscheinlichkeit für den Bernoulliprozess genutzt und man erhält eine binäre Variable R mit Realisierungen 0 und 1

$$R_i \sim \mathcal{B}(1, \pi_i).$$

Der Mechanismus bekommt die Bezeichnung *LogitY*. In der Simulation wird für R=1 der Y-Wert gelöscht und für R=0 der Y-Wert beibehalten, wodurch relativ gesehen auch wieder eher große Y-Werte gefiltert werden. Die Grundidee zur Umsetzung des Fehlend-Prozesses entstammt aus dem Buch *Induktive Statistik: Eine Einführung mit R und SPSS* (Helge Toutenburg, 2007, S.321).

Sämtliche hier vorgestellten Mechanismen werden selbstverständlich an allen jeweiligen 150 Datensätzen angewendet. Insgesamt werden drei unterschiedliche große Anteile fehlender Daten modelliert. In dieser Simulationsstudie sind die Anteile jeweils 10%, 25% und 50%. Bei den Fehlendmechanismen, die über den Prozess des Ziehens gehen, wird der Anteil der fehlenden Y-Werte über die absolute Anzahl dieser gezogenen Werte gesteuert. Also wird beispielsweise bei einer Datensatzgröße von 200 Beobachtungen und einem fehlenden Anteil von 25% die Anzahl der zu ziehenden y-Werte auf 50 festgelegt.

Bei den Fehlendmechanismen, die durch die logistische Regression bestimmt werden, erfolgt die Anpassung an die prozentuale Aufteilung durch unterschiedliche Parameter für den Intercept. Hierfür wird eine Grafik erstellt, die in Abhängigkeit des Interceptwertes den fehlenden Anteil ausgibt. Mithilfe der Grafik werden dann die Interceptwerte übernommen, die den gewünschten Anteil an fehlenden Y-Werten garantieren.

Abbildung 1 zeigt das LogitY-Modell, angewendet auf die Datensätze mit der Beobachtungsgröße 1000. Dabei werden folgende Werte für β_0 angenommen um den gewünschten Anteil fehlender Daten in der Simulation zu erzeugen: -0.2 für 50%; -2.6 für 25%; -5 für 10% (Für die kleineren Datensätze siehe Anhang A.1).

In Abbildung 2 sieht man das LogitX-Modell, angewendet auf die Datensätze mit der Beobachtungsgröße 1000. Dabei werden folgende Werte für β_0 angenommen um den gewünschten Anteil fehlender Daten in der Simulation zu erzeugen: -0.2 für 50%; -2.6 für 25%; -4.8 für 10% (Für die kleineren Datensätze siehe Anhang A.1).

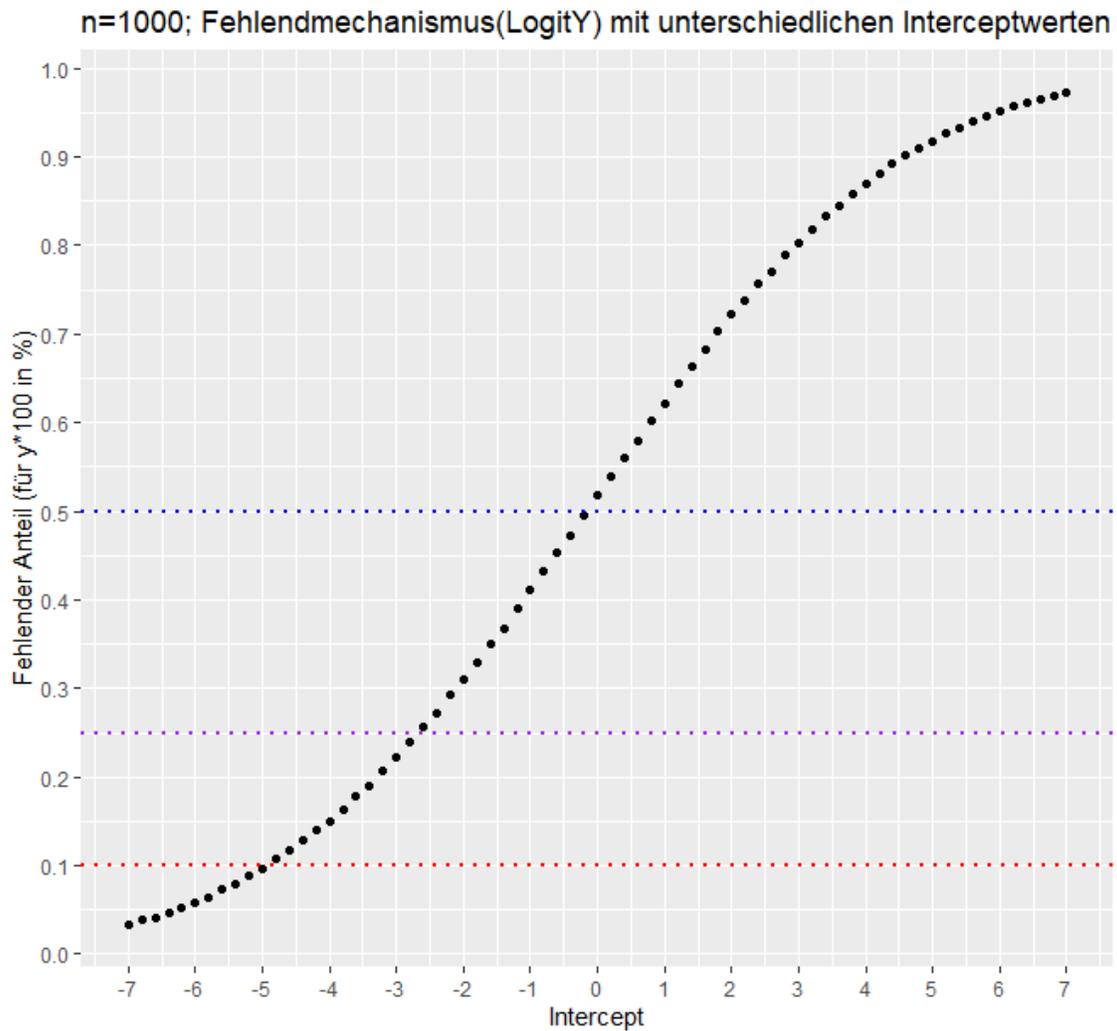


Abbildung 1: Es werden die Fehlend-Anteile, die durch LogitY mit unterschiedlichen Interceptwerten erzeugt werden, angezeigt. Zusätzlich werden durch die eingezeichneten horizontalen Linien die gewünschten Fehlend-Anteile hervorgehoben.

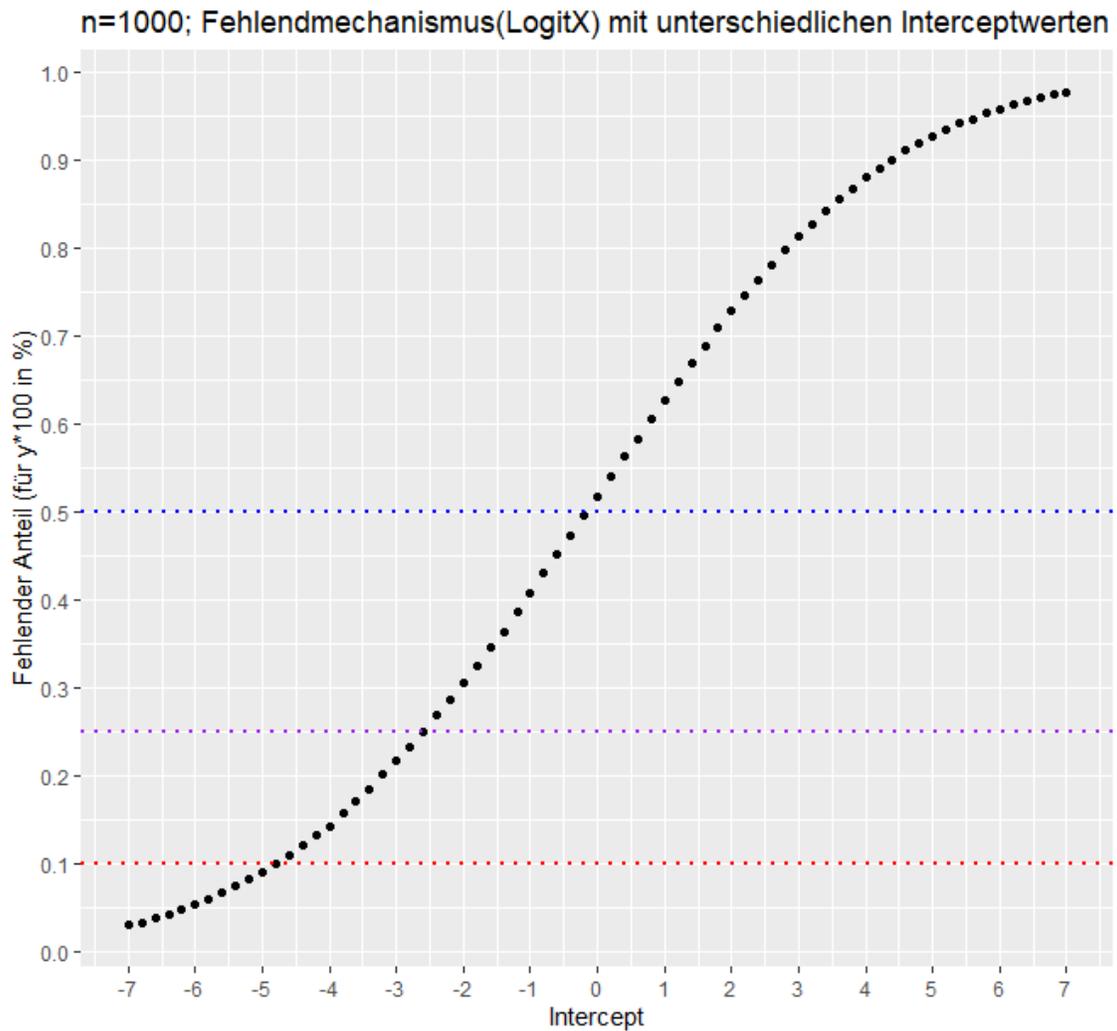


Abbildung 2: Es werden die Fehlend-Anteile, die durch LogitX mit unterschiedlichen Interceptwerten erzeugt werden, angezeigt. Zusätzlich werden durch die eingezeichneten horizontalen Linien die gewünschten Fehlend-Anteile hervorgehoben.

3.1.2 Unterschiedliche Behandlungsarten

Zuerst wird das normale Verfahren mit multipler linearer Regressionsschätzung ohne Intercept als Imputationsverfahren beschrieben: Unabhängig davon, welcher Fehlendmechanismus zur Anwendung kommt, wird zunächst auf den verbleibenden Datensatz ein lineares Regressionsmodell geschätzt (in R einfacher `lm()`-Befehl; ist in Standardpaket enthalten). Anschließend werden mit dem erhaltenen geschätzten Modell die fehlenden Y-Werte imputiert. Da der Datensatz jetzt wieder vollständig ist, wird wiederholt ein lineares Regressionsmodell (mit `lm()` in R) geschätzt. Die geschätzten Koeffizienten bleiben in beiden Modellen unverändert, da die imputierten Y-Werte dem Erwartungswert des geschätzten Modells entsprechen (Ludwig Fahrmeir, 2009, S.25). Allerdings werden jeweils die Varianzen der geschätzten Koeffizienten innerhalb des Datensatzes kleiner, da jetzt die Y-Werte in der Summe weniger streuen.

Als zweite Variante für die lineare Regressionsschätzung ohne Intercept bietet es sich an, ein lineares Regressionsmodell mit Ridge zu schätzen. Dabei ergibt sich die Residuenquadratsumme (RSS) als:

$$RSS(\lambda) = (y - X\beta)^T(y - X\beta) + \lambda\beta^T\beta$$

(Trevor Hastie, 2008, S.63). Der Penalisierungsparameter λ wird allerdings sehr klein gehalten ($\lambda=0.0001$). Dies gründet darauf, dass somit der Penalisierungsterm praktisch wegfällt und dadurch wieder die gewöhnliche Residuenquadratsumme (RSS) bzw. KQ-Schätzer betrachtet wird, wodurch von einem normalen linearen Modell ausgegangen werden kann. Das wird in R mit dem Package `glmnet` und der Funktion `glmnet()` erreicht.

Als dritte Variante wird nun ein lineares Regressionsmodell mit Ridge geschätzt; ebenfalls ohne Intercept. Dabei wird der Penalisierungsparameter λ nicht weiter als kleiner Wert festgehalten. Dies wird mit der Funktion `cv.glmnet()` durchgeführt. Die Funktion `cv.glmnet()` verwendet für die Modellschätzung, wie der Name schon verrät, eine Kreuzvalidierung. Prinzipiell wird per Standardeinstellung eine 10-fold Cross Validation verwendet. In diesem Fall wird aufgrund der teilweise geringen Beobachtungsanzahl der verbliebenen Daten eine 5-fold Cross Validation angewendet. Dies führt sogar zu einer geringeren Varianz bezüglich des geschätzten Prognosefehlers, da die Testdatensätze wesentlich größer sind und somit das Schätzergebnis weniger von möglichen Ausreißern dominiert wird. Jedoch wird im Umkehrschluss die Verzerrung hinsichtlich des wahren Prognosefehlers größer. 5-fold Cross Validation ist ein robustes Verfahren um die Genauigkeit des Modells zu schätzen (Trevor Hastie, 2008, S.241ff.). Dazu wird dann der Wert für λ verwendet, für den sich der niedrigste mean squared error (MSE) der Kreuzvalidierung ergibt. Eine andere Möglichkeit für die Ridge-Schätzung wäre die Regression mit Lasso-Schätzung. Es besteht dann jedoch die Möglichkeit, dass die Penalisierung einige Koeffizienten exakt auf 0 drückt (Trevor Hastie, 2008, S.69) und damit würde man indirekt eine Variablenselektion durchführen. Da die Zielvariable durch die Simulation tatsächlich von allen Kovariablen abhängig ist, ist das hier unerwünscht.

Das zweite Verfahren als Imputationmöglichkeit stellt die *Bruteforcemethodik* (hier als *Bruteforce1* benannt) dar. Es werden durch eine Gleichverteilung die fehlenden Y-Werte imputiert. Im ersten Schritt wird das Minimum und das Maximum der verbleibenden Y-Werte im Datensatz bestimmt und diese dienen nun als Wertebereich für die Gleichverteilung. Dieser Prozess wird in dieser Simulation 100 mal für jeden der 150 Datensätze ausgeführt. Für jeden dieser 100 imputierten Datensätze pro ursprünglichen Datensatz wird ein Modell (alle drei wie oben beschriebenen Regressionsmodelle ohne Intercept mit `lm()`, `glmnet()` und `cv.glmnet()`) geschätzt. Der nächste Schritt besteht darin den Mittelwert der Koeffizienten aus den geschätzten Modellen zu bilden und abzuspeichern. Hier erfolgte die Grundidee zur Umsetzung aus der veröffentlichten Arbeit „A GENERAL SOLUTION TO

Da im Rahmen der beschriebenen Fehlendmechanismen hauptsächlich größere Y-Werte aus den Datensätzen entfernt werden, bietet sich als drittes Verfahren an, dass zusätzlich die Annahme getroffen wird, dass nun der Wertebereich der Gleichverteilung nach rechts verschoben wird. Genauer gesagt, erstreckt sich der Wertebereich in der zweiten *Bruteforcemethodik* (genannt *Brutforce2*) vom Mittelwert der verbliebenen Y-Werte bis zum Maximum. Man zieht nun also zufällig deutlich öfter größere Y-Werte um den Fehlendenmechanismen gerecht zu werden.

3.1.3 Verwendete Gütekriterien

Als Gütekriterien gelten folgende Maßzahlen (Helge Toutenburg, 2007, S.110ff.), die zum einen vollständig in Tabellenform und zum anderen graphisch in Boxplotform präsentiert werden. Im Nachfolgenden gilt für $i = 1, \dots, N$, wobei N die Anzahl der Datensätze und $j = 1, \dots, 8$ der Index für die Koeffizienten ist.

- **Koeffizienten-Schätzung:** Es werden die Mittelwerte der geschätzten Koeffizienten über die Datensätze berechnet:

$$E(\hat{\beta}_j) = \frac{1}{N} \sum_{i=1}^N \hat{\beta}_{ij} = \overline{\hat{\beta}_j}.$$

- **Mean squared error:** Es wird die quadrierte Abweichung der geschätzten Koeffizienten mit den echten Koeffizienten berechnet:

$$MSE(\hat{\beta}_j) = Var(\hat{\beta}_j) + (Bias(\hat{\beta}_j))^2 = \frac{1}{N} \sum_{i=1}^N (\hat{\beta}_{ij} - \beta_j)^2.$$

- **Bias (Verzerrung):** Es wird die Differenz der Mittelwerte der geschätzten Koeffizienten zu den echten Koeffizienten berechnet:

$$Bias(\hat{\beta}_j) = E(\hat{\beta}_j - \beta_j) = \overline{\hat{\beta}_j} - \beta_j.$$

- **Varianz:** Es wird die Varianz der geschätzten Koeffizienten berechnet (mit den Mittelwerten der geschätzten Koeffizienten):

$$Var(\hat{\beta}_j) = \frac{1}{N} \sum_{i=1}^N (\hat{\beta}_{ij} - \overline{\hat{\beta}_j})^2.$$

3.2 Datengenerierender Prozess mit logistischer Regression

Der Datensatzaufbau ist kongruent. Es gelten demnach die gleichen Wahrscheinlichkeitsprozesse für die Kovariablen und auch deren Beobachtungsgröße ist gleich. Der Unterschied liegt hier zum einen in der Erzeugung der Zielvariable Y und zum anderen darin, dass es sich nun nicht mehr um eine metrische, sondern um eine binäre Zielvariable handelt. In diesem Fall werden zuerst Wahrscheinlichkeiten für Y mit einer logistischen Regression erzeugt (Fahrmeier 1.Auflage S.30ff.):

$$\pi_i = P(Y_i = 1|x_i) = E(Y_i|x_i) = \frac{\exp(x_i^T \beta)}{1 + \exp(x_i^T \beta)} = \frac{1}{1 + \exp(-x_i^T \beta)} \in (0, 1)$$

und $i = 1, \dots, n$. Der Koeffizientenvektor β erhält die gleichen Werte wie in Kapitel 3.1. Mit den jeweils erhaltenen Wahrscheinlichkeiten π_i wird dann nochmals ein Bernoulliprozess durchgeführt:

$$Y_i \sim \mathcal{B}(1, \pi_i).$$

Das hat den Vorteil, dass kein bestimmter Schwellenwert ausgewählt werden muss und auch kleine Wahrscheinlichkeiten eine Chance besitzen, die Realisierung 1 zu erhalten. Mit diesem Verfahren werden jeweils 250 Datensätze wie mit dem Vorgehen aus Kapitel 3.1 erzeugt.

3.2.1 Verwendete Fehlendmechanismen

Das Vorgehen entspricht hier im Wesentlichen wieder dem Vorgehen aus Kapitel 3.1.1. Für die Durchführung der MNAR-Fehlendmechanismen sind allerdings kleine Änderungen notwendig, da es sich jetzt um eine binäre statt um eine metrische Zielvariable handelt. In Kapitel 3.1.1 werden y -Werte für *MNAR1*, *MNAR2* und *MNAR3* aus bestimmten Quantilen gezogen. Statt den Quantilen werden nun die Y -Werte nach ihrer Realisierung (1 oder 0) in zwei Gruppen aufgeteilt.

Für *MNAR1* werden alle fehlenden Y -Werte aus der Gruppe mit Realisierung 1, falls genügend Werte im Datensatz vorliegen, gezogen. Ansonsten wird die Anzahl durch Y -Werte der Gruppe mit Realisierung 0 aufgefüllt.

Für *MNAR2* werden 90% aller fehlenden Werte aus der Gruppe mit $Y=1$, falls genügend Werte im Datensatz vorhanden sind (sonst aus Gruppe mit $Y=0$) und die restlichen 10% aus der anderen Gruppe gefiltert.

Für *MNAR3* werden 70% aller fehlenden Werte aus der Gruppe mit $Y=1$, wenn genügend Werte im Datensatz vorhanden sind (sonst aus Gruppe mit $Y=0$) und die restlichen 30% aus der anderen Gruppe genommen.

Der *LogitY*-Mechanismus bleibt unverändert.

Die Grafiken zu dem fehlenden Anteil für die Fehlendmechanismen *LogitY* und *LogitX* in Abhängigkeit von β_0 befinden sich in Anhang A.2. Für bspw. die Datensatzgröße 1000 werden für *LogitY* die Interceptwerte -0.6; -1.6; -2.8 und für *LogitX* die Werte -0.4; -2.8; -5 ausgewählt.

3.2.2 Unterschiedliche Behandlungsarten

Für das normale Verfahren wird auf den verbleibenden Datensatz ein logistisches Modell ohne Intercept (in R mit `glm()`-Befehl; ist in Standardpaket enthalten) geschätzt. Danach werden mit dem erhaltenen Modell die fehlenden Y -Werte geschätzt und imputiert. Auf den vollständig erhaltenen Datensatz wird erneut ein logistisches Modell trainiert. Wie in Kapitel 3.1.2 erläutert, hat das keine Auswirkungen auf die Koeffizientenschätzung (Ludwig Fahrmeir, 2009, S.190), jedoch wird die Varianz der Koeffizienten innerhalb des Datensatzes geringer.

Da es sich um eine binäre Zielvariable mit 0 oder 1 handelt, wird bei der *Bruteforcemethodik* nun eine Gleichverteilung auf den Wertebereich von $[0,1]$

$$\pi_i \sim \mathcal{U}(0, 1)$$

mit $i = 1, \dots, n$ ausgeführt. Der ermittelte Wert wird wiederum als Parameter für einen Bernoulliprozess genutzt

$$Y_i \sim \mathcal{B}(1, \pi_i)$$

und man erhält erneut eine binäre Größe als Resultat. Dieser Prozess wird für jeden fehlenden Y -Wert ausgeführt und pro Datensatz wieder 100 mal wiederholt. Anschließend wird auf alle imputierten Datensätze ein logistisches Modell trainiert und daraus der Mittelwert der jeweiligen

Koeffizienten gebildet (Verfahren wird *Bruteforce1* genannt).

Die verwendeten Fehlendmechanismen entfernen größtenteils Werte gleich 1. Damit deutlich öfter der Wert 1 imputiert wird, wird der Wertebereich der Gleichverteilung auf $[0.5,1]$ begrenzt (Verfahren wird *Bruteforce2* genannt). Der restliche Vorgehensweis entspricht der Methodik von *Bruteforce1* (Grundidee aus (Kenkel, 2013, S.9ff.)).

3.2.3 Verwendete Gütekriterien

Die verwendeten Gütekriterien sind bei den Datensätzen, bei denen die Werte für Y über die logistische Regression erzeugt werden, identisch mit jenen der Datenerstellung mit multipler linearer Regression (siehe Kapitel 3.1.3).

4 Ergebnisse

4.1 Ergebnisse der Modelle mit metrischer Zielgröße

Im Folgenden wird vor allem auf die lineare Regression ohne Intercept unter den verschiedenen Fehlendmechanismen eingegangen und ein Vergleich zu den Bruteforce-Methodiken hergestellt. Zusätzlich wird die Sensitivität der linearen Regression mit Ridge-Schätzung unter normalem Verfahren betrachtet. Dabei werden hier beispielhaft die Ergebnisse mit Datensatzgröße $n=1000$ zur genaueren Betrachtung herangezogen. Wenn von einer Unterschätzung oder Überschätzung der Beta-Koeffizienten geschrieben wird, dann ist damit inhaltlich der Einfluss, den der Koeffizient ausübt, gemeint. Also wird die Schätzung betragsmäßig betrachtet und ein Einfluss von -1 wäre demnach niedriger als der Einfluss von -2.

Im Kapitel 3.1.2 wird die Annahme beschrieben, dass die Funktion `glmnet()` mit $\lambda = 0.0001$, eine Möglichkeit neben `lm()` in R darstellt, ein lineares Regressionsmodell ohne Ridge schätzen. Für den Vergleich der linearen Regression (geschätzt mit `lm()`) zur linearen Regression mit `glmnet()` (allerdings mit $\lambda = 0.0001$) zeigt sich, dass sich ab einer Datensatzgröße von $n=200$ für alle Koeffizienten sehr ähnliche Gütekriterien ergeben (siehe Anhang C.1). Es besteht also für größere Datensätze kein Unterschied mit welcher Variante versucht wird ein lineares Modell zu schätzen (Aussage beruht auf dem Vorgehen mit normalem Verfahren).

Führt man den Vergleich der Modellschätzungen mit `lm()` und `glmnet()` mit $\lambda = 0.0001$ auch für die Bruteforce1-Methodik fort, dann lässt sich feststellen, dass sich der MSE bzw. genauer der Bias ab einer Datensatzgröße von 200 teilweise bei den einzelnen Koeffizienten unterscheidet. Für kleinere Datensätze (hier 50) ergeben sich schon größere Unterschiede hinsichtlich der Verzerrung. Unter den Fehlendmechanismen MCAR und MNAR3 decken sich aber die Gütekriterien hinsichtlich der Koeffizienten gut. Folglich gilt für beide Modellschätzmöglichkeiten für eine lineare Regression in R, dass kleinere Unterschiede bezüglich der Koeffizienten vorliegen. Tendenziell liefert die Modellschätzung mit `glmnet()` eine bessere Koeffizientenschätzung, da der Bias geringer ist.

Betrachtet man diesen Vergleich der Modellschätzungen auch für die Bruteforce2-Methodik, dann lässt sich erkennen, dass die Güte bezüglich der Koeffizienten ab einer Datensatzgröße von 200 ähnlich ist. Kleinere Unterschiede liegen bei einzelnen Koeffizienten für die Verzerrung vor. Somit liefern beide R-Befehle nicht dieselben Ergebnisse bezüglich der Koeffizienten, jedoch ist die Diskrepanz nicht allzu groß.

4.1.1 Sensitivität der linearen Regression unter verschiedenen Fehlendmechanismen

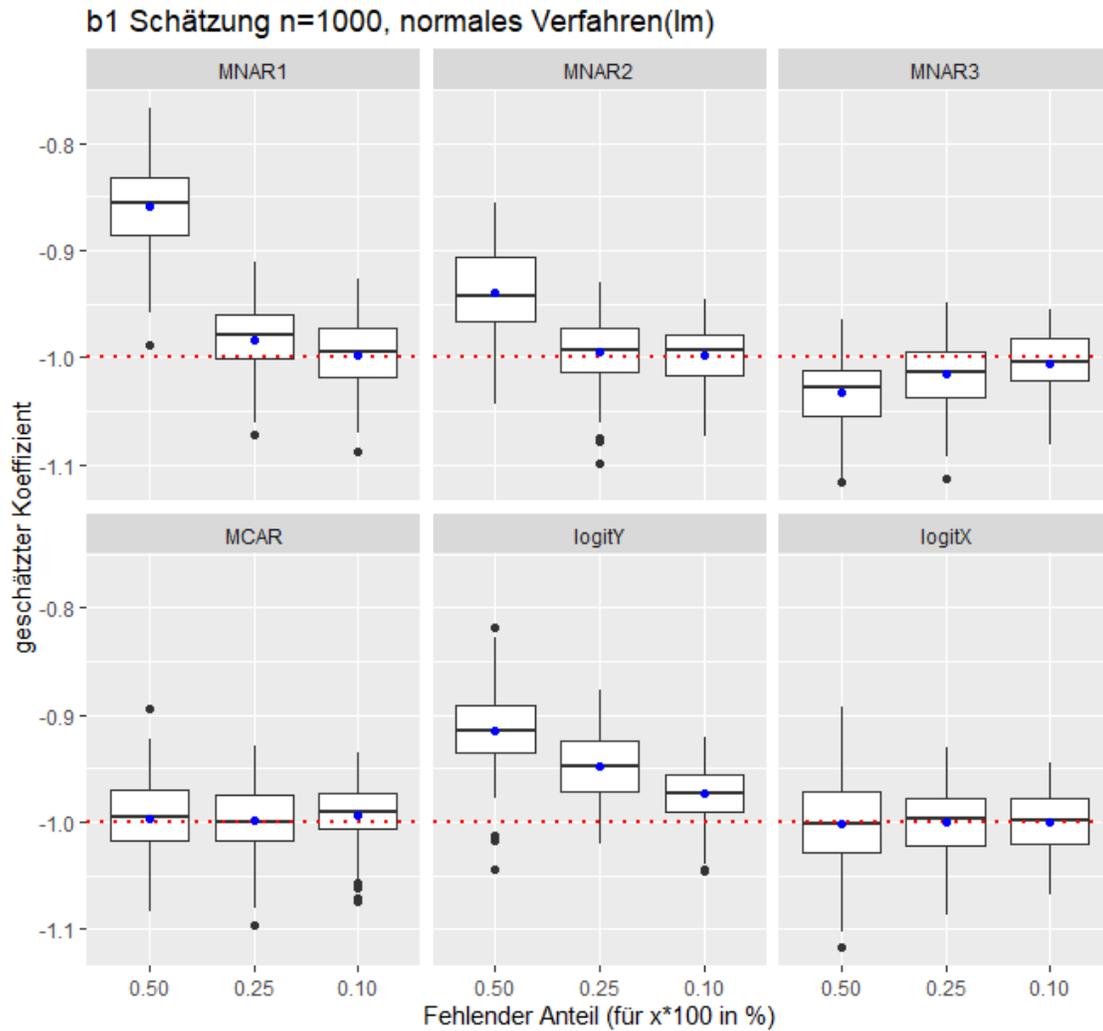


Abbildung 3: Es werden die geschätzten Werte für β_1 in Boxplotform dargestellt. Diese werden nach den Fehlendmechanismen gruppiert. Der blaue Punkt gibt den Mittelwert der geschätzten Werte an. Die rote horizontale Linie repräsentiert den echten Wert des Koeffizienten.

In Abbildung 3 ist direkt ersichtlich, dass für die lineare Regression die Koeffizientenschätzungen für β_1 unter den Fehlendmechanismen MNAR1, MNAR2, MANR3 und LogitY stark verzerrt sind. Die Verzerrung nimmt allgemein ab, wenn gleichzeitig weniger Daten verloren gehen. Augenscheinlich ist dabei, dass für 10% fehlende Y-Werte unter LogitY der Bias für β_1 noch immer ziemlich groß ist. Zusätzlich fällt auf, dass unter MNAR3 β_1 im Gegensatz zu den anderen Fehlendprozessen überschätzt wird. Unter den Mechanismen MCAR und LogitX liefert die lineare Regression keinen verzerrten Schätzer. Die Varianz ist für alle Fehlendmechanismen ähnlich groß.

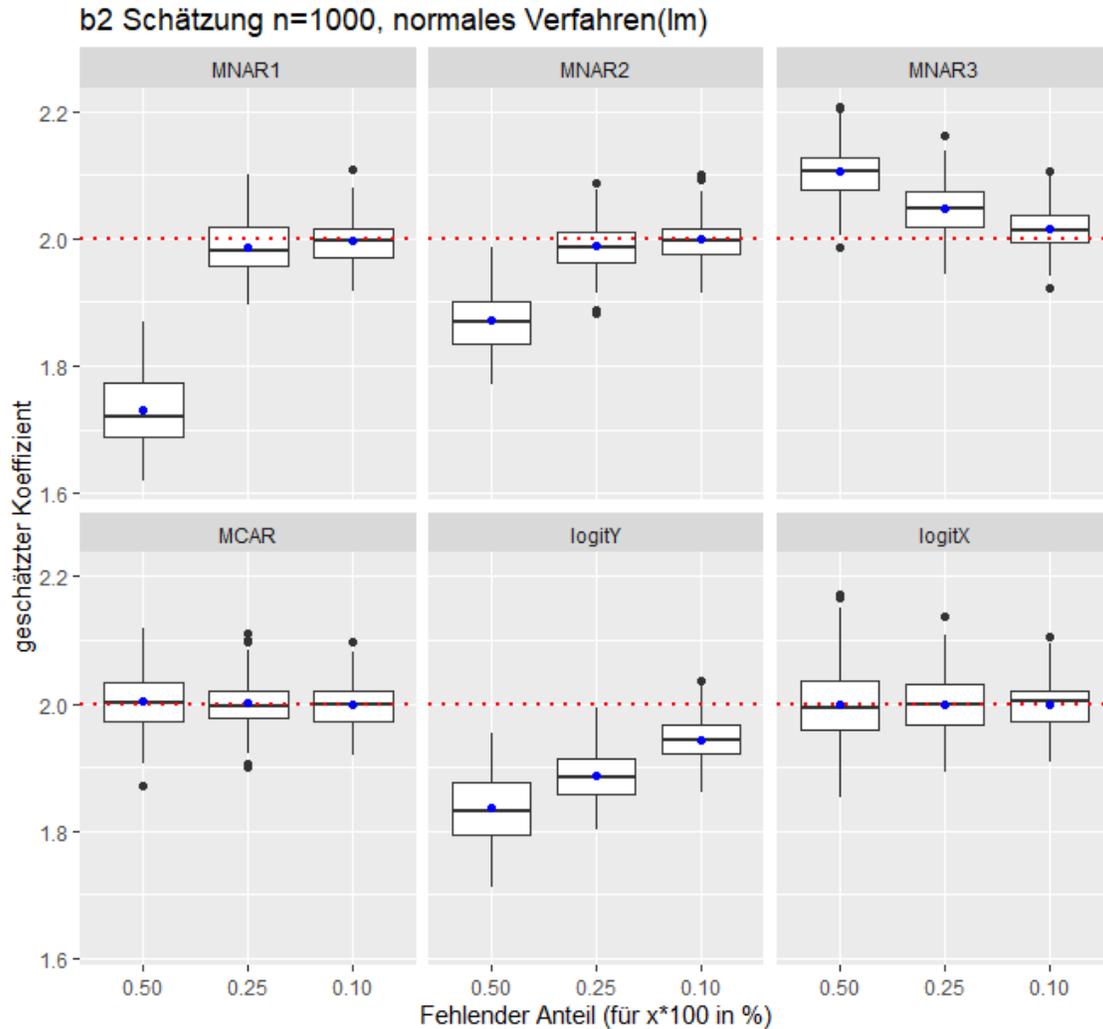


Abbildung 4: Es werden die geschätzten Werte für β_2 in Boxplotform dargestellt. Diese werden nach den Fehlendmechanismen gruppiert. Der blaue Punkt gibt den Mittelwert der geschätzten Werte an. Die rote horizontale Linie repräsentiert den echten Wert des Koeffizienten.

Aus Abbildung 4 ist festzuhalten, dass unter MNAR1, MNAR2 und LogitY β_2 unterschätzt wird, wohingegen unter MNAR3 das Gegenteil der Fall ist. Im Vergleich zur Abbildung 3 ist der Bias zahlenmäßig ausgedrückt wesentlich größer. Dabei nimmt die Verzerrung und die Varianz mit dem fehlenden Anteil ab. Dennoch wird unter LogitY mit 10% fehlenden Daten der Koeffizient deutlich verzerrt unterschätzt.

Für Abbildung 5 lässt sich feststellen, dass der KQ-Schätzer unter MNAR1, MNAR2 und LogitY für β_3 deutlich verzerrt ist. Für MNAR3 ist der Bias jedoch sehr gering. Das ist ebenfalls unter MCAR und LogitX der Fall.

In Abbildung 6 ist zu erkennen, dass im Gegensatz zu β_3 der Bias unter MNAR3 am größten ist und MNAR1, MNAR2 sowie LogitY eher gering ausfällt. Betrachtet man die Varianz, dann sind keine erkennbaren Unterschiede unter den Mechanismen ersichtlich.

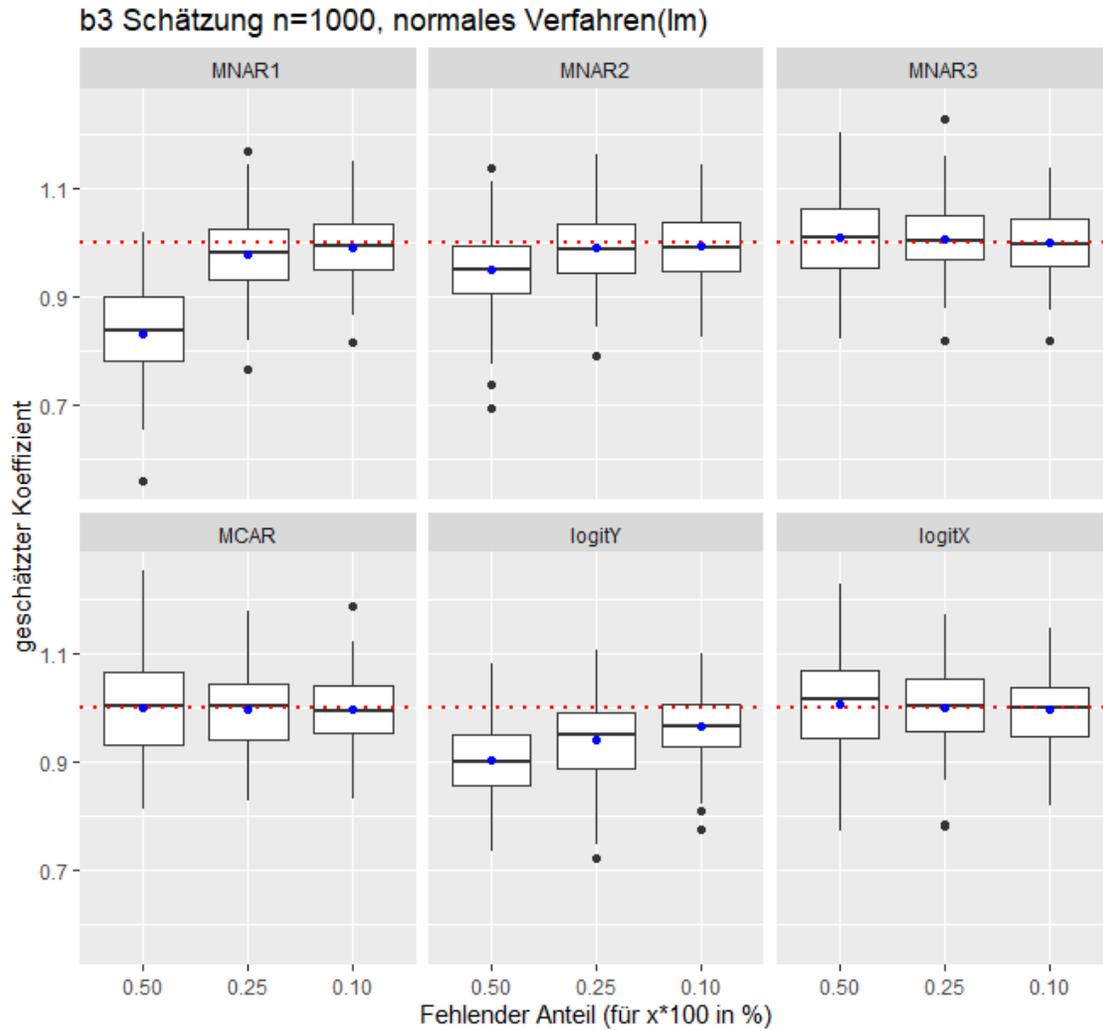


Abbildung 5: Es werden die geschätzten Werte für β_3 in Boxplotform dargestellt. Diese werden nach den Fehlendmechanismen gruppiert. Der blaue Punkt gibt den Mittelwert der geschätzten Werte an. Die rote horizontale Linie repräsentiert den echten Wert des Koeffizienten.

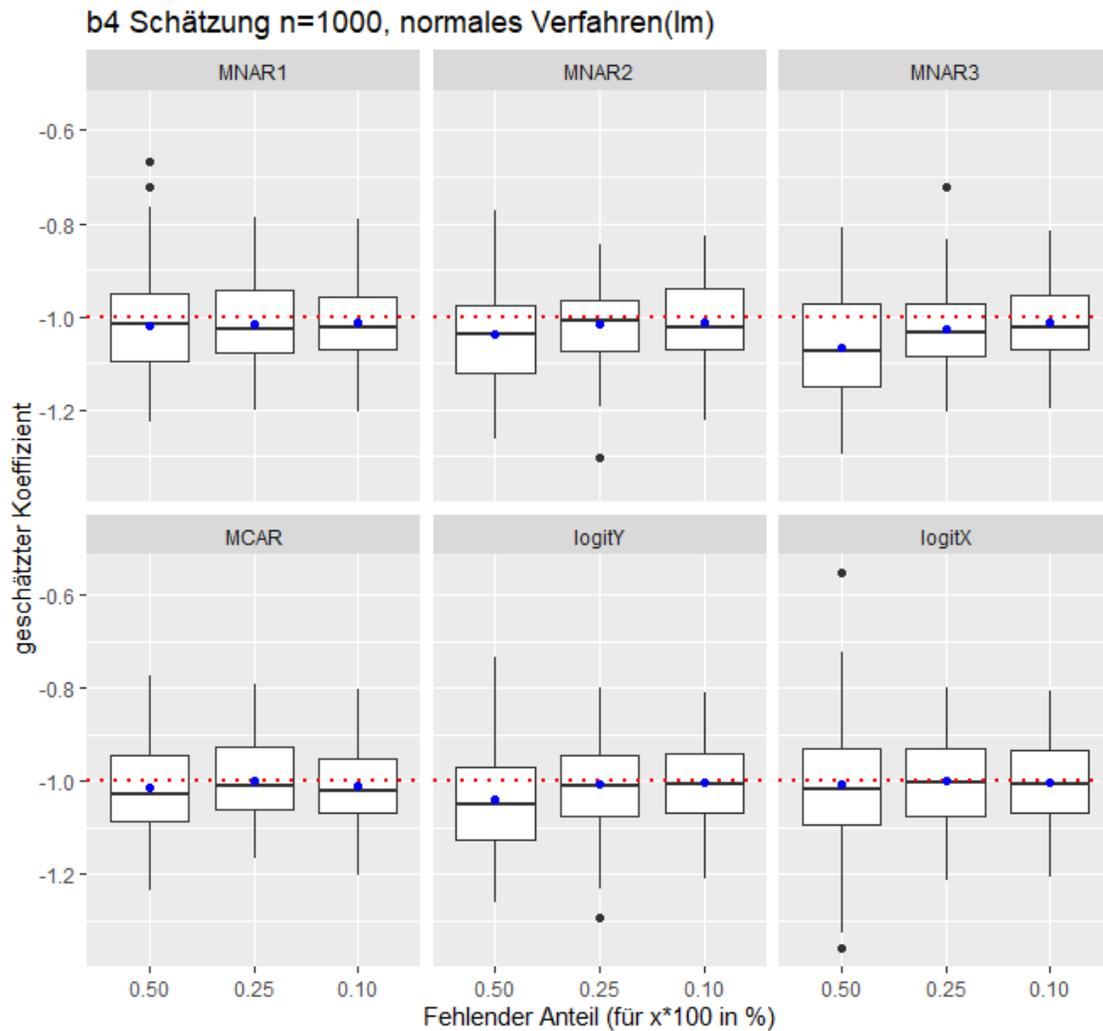


Abbildung 6: Es werden die geschätzten Werte für β_4 in Boxplotform dargestellt. Diese werden nach den Fehlendmechanismen gruppiert. Der blaue Punkt gibt den Mittelwert der geschätzten Werte an. Die rote horizontale Linie repräsentiert den echten Wert des Koeffizienten.

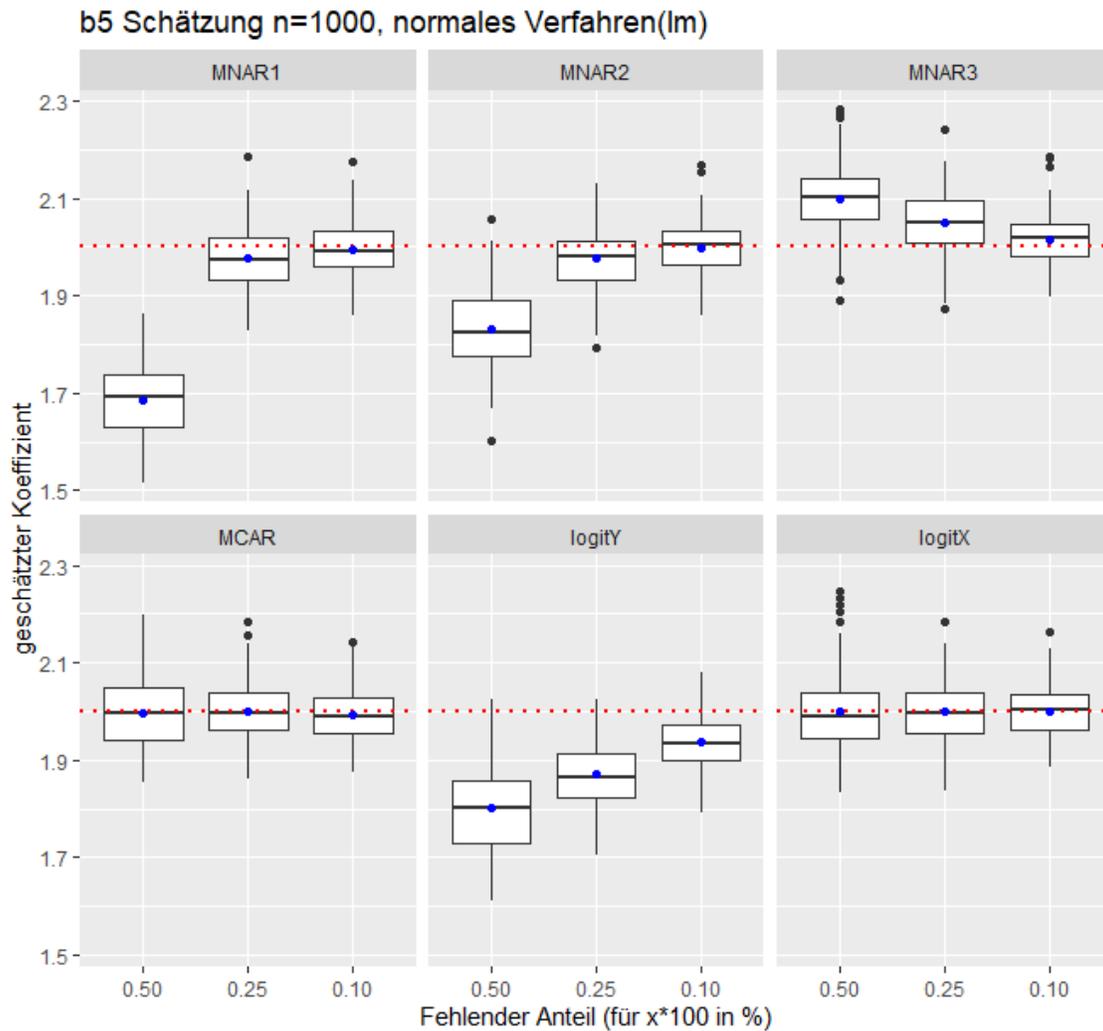


Abbildung 7: Es werden die geschätzten Werte für β_5 in Boxplotform dargestellt. Diese werden nach den Fehlendmechanismen gruppiert. Der blaue Punkt gibt den Mittelwert der geschätzten Werte an. Die rote horizontale Linie repräsentiert den echten Wert des Koeffizienten.

Aus Abbildung 7 ist ersichtlich, dass unter MNAR1, MNAR2 und LogitY β_5 deutlich unterschätzt wird. Unter MNAR3 wird β_5 überschätzt. Außerdem nimmt sowohl die Verzerrung als auch die Varianz mit dem fehlenden Anteil ab. Für MCAR und LogitX erhält man unverzerrte Koeffizientenschätzungen.

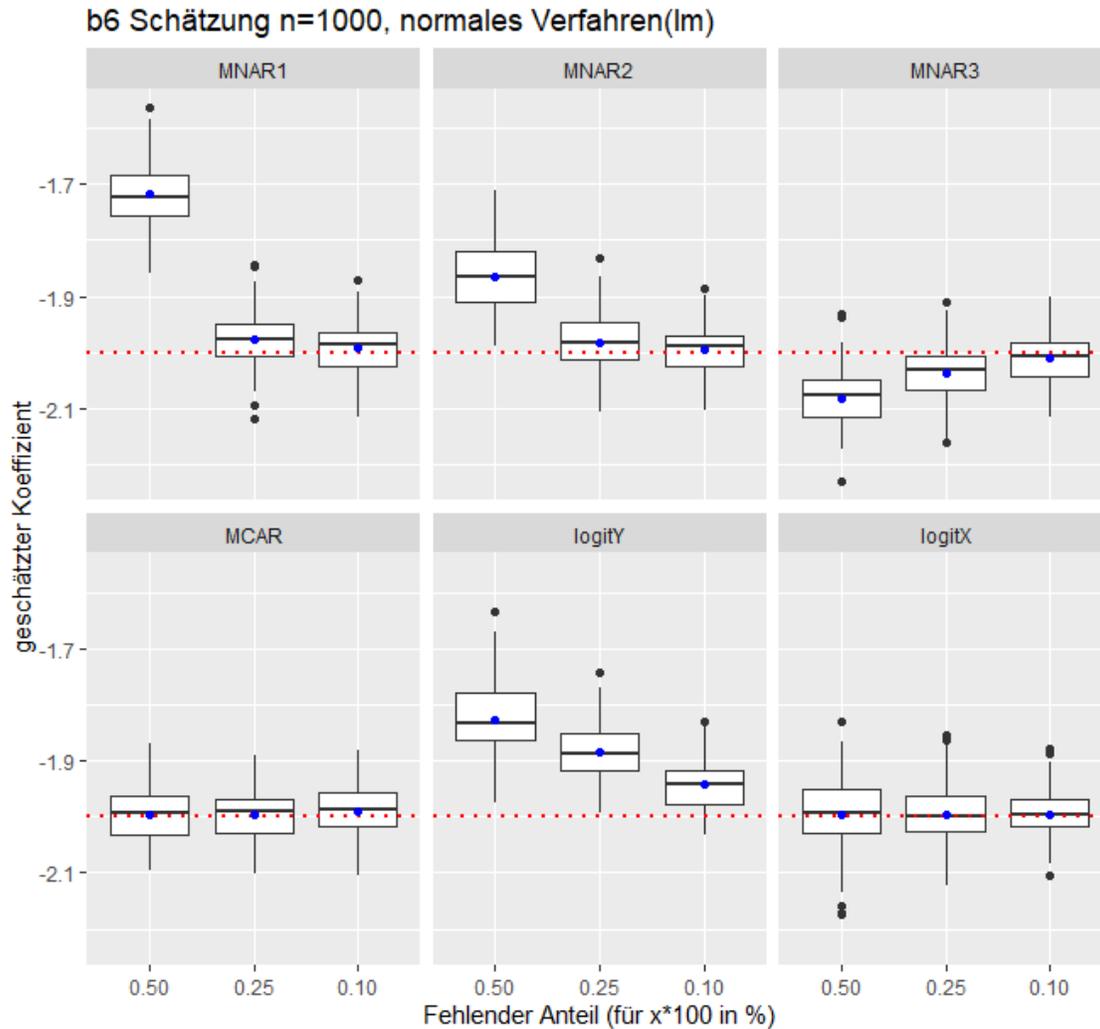


Abbildung 8: Es werden die geschätzten Werte für β_6 in Boxplotform dargestellt. Diese werden nach den Fehlendmechanismen gruppiert. Der blaue Punkt gibt den Mittelwert der geschätzten Werte an. Die rote horizontale Linie repräsentiert den echten Wert des Koeffizienten.

Bei Abbildung 8 lassen sich die gleichen Beobachtungen wie bei Abbildung 7 machen. Außerdem zeigt sich wie auch in den anderen Abbildungen, dass im Fall MNAR1 und MNAR2 der Bias für 50% fehlende Daten ziemlich groß ist und dann schon sehr stark für 25% abflacht.

Für Abbildung 9 lässt sich festhalten, dass die Güte hinsichtlich der Koeffizientenschätzung für β_7 ähnlich wie in Abbildung 5 ist. Ein wesentlicher Unterschied ist jedoch, dass der Bias unter MNAR1, MNAR2, MNAR3 und LogitY größer ist und auch die Varianz unter LogitX sichtlich zugelegt hat.

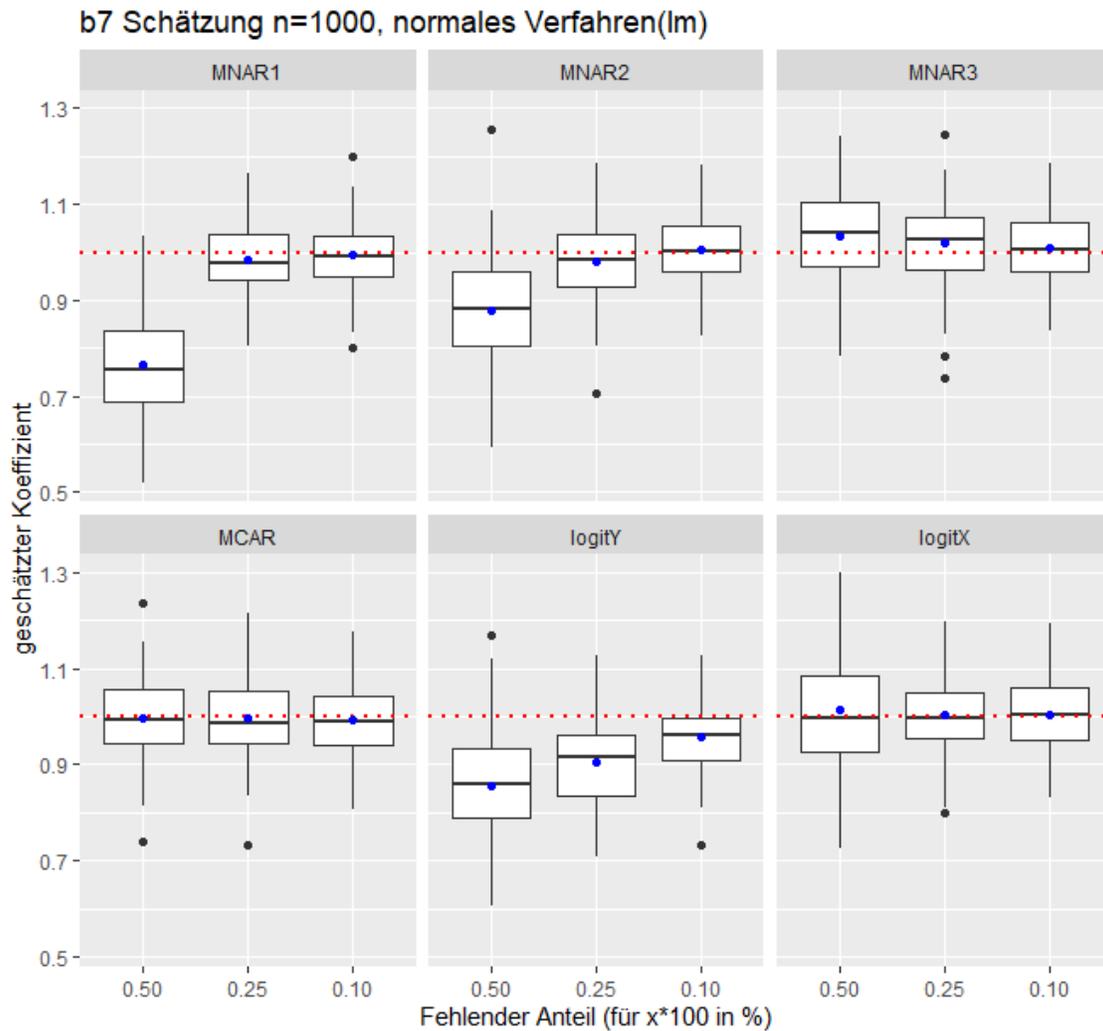


Abbildung 9: Es werden die geschätzten Werte für β_7 in Boxplotform dargestellt. Diese werden nach den Fehlendmechanismen gruppiert. Der blaue Punkt gibt den Mittelwert der geschätzten Werte an. Die rote horizontale Linie repräsentiert den echten Wert des Koeffizienten.

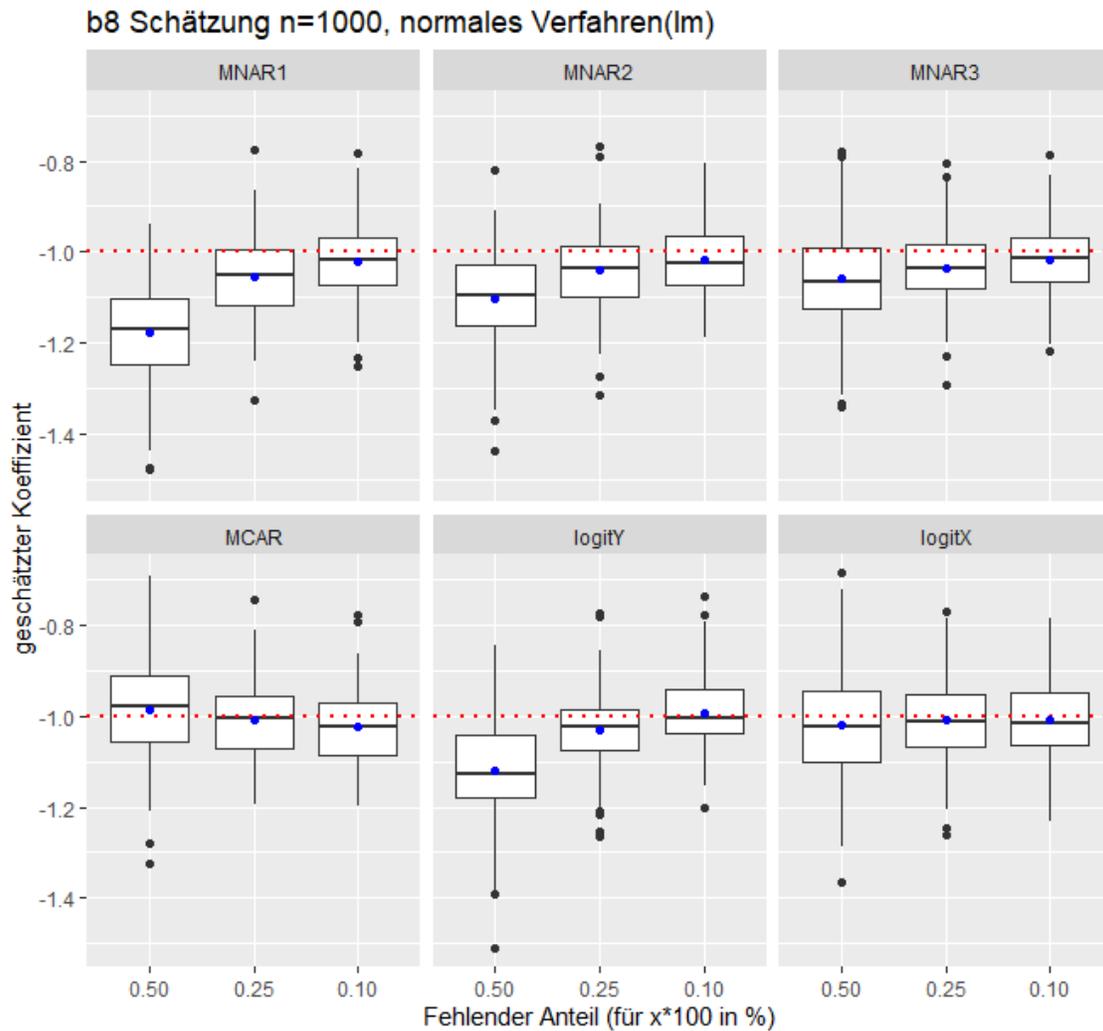


Abbildung 10: Es werden die geschätzten Werte für β_8 in Boxplotform dargestellt. Diese werden nach den Fehlendmechanismen gruppiert. Der blaue Punkt gibt den Mittelwert der geschätzten Werte an. Die rote horizontale Linie repräsentiert den echten Wert des Koeffizienten.

In Abbildung 10 sieht man, dass der KQ-Schätzer unter MNAR1, MNAR2, MNAR3 und LogitY für β_8 deutlich verzerrt ist. Die Schätzungen für β_7 aus der vorherigen Abbildung haben allerdings einen größeren Bias. Auffällig ist, dass für MNAR3 die Richtung der Verzerrung mit den anderen Fehlendmechanismen für MNAR übereinstimmt. Außerdem ist noch zu erwähnen, dass die Schätzungen für β_8 im linearen Regressionmodell unter MCAR und LogitX geringfügig verzerrt sind.

Insgesamt kann man anhand der Abbildungen und der Tabellen aus Anhang C.1 feststellen, dass die allgemeine Güte (gemessen am MSE) der geschätzten Koeffizienten zunimmt, wenn der Anteil der fehlenden Daten geringer wird. Zusätzlich lässt sich beobachten, dass der MSE mit steigender Datensatzgröße deutlich sinkt und als Konsequenz die Schätzungen für die Beta-Koeffizienten genauer werden.

Zudem ist auch zu erkennen, dass der MSE sowie der Bias für die Koeffizienten unterschiedlich groß ist. Beispielsweise ist der MSE von β_2 immer deutlich größer als der von β_1 (Ausnahme bei MCAR und LogitX). Allgemein betrachtet, besitzen die Koeffizienten für die metrischen Kovariablen X_1 , X_2 , X_5 und X_6 die größere Verzerrung und damit auch einen höheren MSE als für die binären Kovariablen. Dies resultiert hauptsächlich aus der Verzerrung. Denn wenn die Varianz über alle Fehlendmechanismen isoliert betrachtet wird, so ergibt sich, dass die Schätzungen der Koeffizienten für binäre Kovariablen mehr streuen als die für metrische Kovariablen. Wenn man sich nur auf die Verzerrung fokussiert, dann fällt außerdem auf, dass bei den Fehlendenmechanismen MNAR1, MNAR2, MNAR3 und LogitY die geschätzten Koeffizienten β_2 , β_5 und β_6 annähernd doppelt so stark verzerrt sind als β_1 . Dies könnte daran liegen, dass die gerade genannten Fehlendenmechanismen dazu neigen große Y-Werte auszuwählen und die echten Koeffizienten mit $\beta_2 = 2$, $\beta_5 = 2$ und $\beta_6 = -2$, betragsmäßig einen größeren Einfluss auf eine hohe Y-Realisierung haben als $\beta_1 = 1$. Diese Beobachtung tritt bei den Fehlendenmechanismen MCAR und LogitY nicht ein, denn dort unterscheiden sich der Bias und der MSE für die vergleichsweise Betrachtung von β_1 zu β_2 , β_5 und β_6 im Wesentlichen nicht. Außerdem passt sich das Verhältnis der Verzerrung von β_1 zu β_2 , β_5 und β_6 dem Anteil fehlender Daten an. Je geringer also der Anteil der fehlenden Daten ist, desto kleiner sind die Unterschiede von β_1 zu β_2 , β_5 und β_6 unabhängig vom Fehlendenmechanismus.

Wenn die Fehlendenmechanismen MCAR und LogitX mit denen für MNAR verglichen werden, so fällt auf, dass unter MCAR und LogitX die Varianzen der Koeffizienten größer sind als die Verzerrung. Unter MNAR ist das Verhältnis genau umgekehrt, man erhält also stärker verzerrte als gestreute geschätzte Koeffizienten. Allerdings ist der MSE unter MNAR meistens höher als unter MCAR und LogitX. Ein weiterer interessanter Aspekt ist, die Erzeugung MNAR durch ein Logit-Modell zu betrachten, um zu sehen welche Auswirkung die Variablenauswahl für das Logit-Modell auf die Güte der geschätzten Koeffizienten hat. Wählt man nämlich die Variante mit Y als Variable, so erhält man deutlich verzerrte Koeffizienten und geringere Varianzen. Wählt man die ursprünglichen Kovariablen aus, so bekommt man wieder eine geringere Verzerrung und größere Varianzen. Für alle Datensatzgrößen zeigt sich zudem, dass der MSE und vor allem der Bias für das Logit-Modell mit Y als Variable größer ist als mit den Kovariablen im Modell. Unter diesen Aspekten erscheint die Annahme, dass die Methode mit der ursprünglichen Zielgröße Y als Variable im Logit-Modell als MNAR angesehen werden kann, als sinnvoll.

Zuletzt hat sich noch bemerkbar gemacht, dass unterschiedliche MNAR-Mechanismen auch unterschiedlich starke Verzerrungen herbeiführen. MNAR1 ist die stärkere Version von MNAR2 und man erkennt auch aus den Abbildungen, dass die Verzerrung der Koeffizienten unter MNAR2 deutlich schwächer und dementsprechend auch der MSE niedriger ist. LogitY stellt durch die Modellierung durch das logistische Modell eine Art fließenden Übergang hinsichtlich der Intensität, mit der große Y-Werte aus dem Datensatz entfernt werden, von MNAR2 zu MNAR1 dar. Für MNAR3 erhält man auch verzerrte Koeffizienten, jedoch ist die Richtung der Verzerrung für die meisten Koeffizi-

enten gegenläufig zu MNAR1, MNAR2 und LogitY. Unter MNAR1, MNAR2 und MNAR3 nimmt der Bias gleichzeitig mit dem Anteil fehlender Daten ab. Das gilt auch unter LogitY, allerdings geschieht dies deutlich langsamer bzw. die Abnahme der Verzerrung scheint linear zu sein. Folglich werden mit 10% fehlende Y-Werte die Koeffizienten erheblich verzerrt geschätzt.

4.1.2 Sensitivität der linearen Regression mit Ridge-Schätzung unter verschiedenen Fehlendmechanismen

Wird anstelle der linearen Regression ein lineares Modell mit Ridge auf die gleichen Datensätze unter Anwendung der verschiedenen Fehlendmechanismen geschätzt, dann lässt sich direkt aus den Grafiken im Anhang B.1 erkennen, dass sämtliche geschätzten Koeffizienten unter den Prozessen MCAR und LogitX verzerrt sind (im Gegensatz zur linearen Regression). Die geschätzten Koeffizienten weisen allerdings eine ähnlich große Varianz auf. Damit wird der resultierende höhere MSE für die Regression mit Ridge durch den Bias und nicht durch die Varianz erklärt.

Allgemein zeigt sich hier, dass unter allen Fehlendmechanismen der Bias wesentlich größer für das lineare Modell mit Ridge-Schätzung als für die lineare Regression ist. Es ist klar zu erkennen, dass für 10% fehlende Daten die Koeffizienten immer noch stark verzerrt sind. Für die lineare Regression hingegen sind die Koeffizienten relativ gesehen deutlich näher am echten Wert. Das lässt sich hauptsächlich mit dem Shrinkage-Parameter λ erklären, dessen Hauptaufgabe gerade die Reduzierung hinsichtlich der Koeffizienten darstellt.

Vielmehr lässt sich auch beobachten, dass die unterschiedlichen Fehlendmechanismen keinen großen Einfluss auf den Bias haben, sondern dieser vor allem vom Anteil der fehlenden Y-Werte und dem Parameter λ abhängig ist.

Interessanterweise ist für die Abbildungen (Anhang B.1) und Tabellen (Anhang C.1) zu erkennen, dass der MSE für die geschätzten Koeffizienten der univariat erzeugten Kovariablen unter MNAR3 am niedrigsten ist. Da unter MNAR3 die Koeffizienten dazu neigen überschätzt zu werden, wird durch die Penalisierung die Verzerrung in die „richtige“ Richtung reduziert. Besonders der Fall für 50% fehlende Daten ist hervorzuheben, da dort für die univariaten Variablen der jeweilige Koeffizient eine niedrigere Verzerrung aufweist, als wenn weniger Y-Werte verloren gehen. Also ist hier die Tendenz genau umgekehrt und der MSE wird größer, falls der fehlenden Anteil ansteigt.

Für die multivariaten Kovariablen, welche untereinander korreliert sind, erkennt man, dass die Koeffizienten stärker bestraft werden und damit im Endeffekt die Verzerrung der geschätzten Koeffizienten größer wird, als es für die univariaten Variablen der Fall war. Als Resultat lässt sich für die Regression mit Ridge festhalten, dass sich die Schätzungen für die Koeffizienten erheblich weniger vom vorliegenden Fehlendmechanismus beeinflussen lassen, da ein großer Anteil der Verzerrung vom Penalisierungsparameter λ selbst abhängt. Der MSE für die Koeffizienten wird wie im normalen linearen Modell mit steigender Datensatzgröße und steigender Anzahl beobachteter Y-Werte kleiner. Allerdings zeigt sich für die Fehlendmechanismen, für die eher große Y-Werte ausgewählt werden, dass die zusätzliche Reduzierung der Koeffizienten einen noch größeren Bias zur Folge haben.

4.1.3 Sensitivität für Brute-force-Methodiken mit linearen Modellen unter verschiedenen Fehlendmechanismen

Für die lineare Regression mit BruteForce1 ergibt sich aus den Abbildungen (Anhang B.2) und Tabellen (Anhang B.3), dass der MSE für die Koeffizienten im Vergleich zum linearen Modell mit normalem Verfahren für alle Verfahren um ein mehrfaches größer ist. Das resultiert zum einen aus einem extrem größeren Bias und zum anderen auch aus größeren Varianzen.

Ursprünglich zeigte sich, dass vor allem die Mechanismen unter MNAR1, MNAR2, MNAR3 und

LogitY im normalen Verfahren verzerrte Schätzungen hervorrufen. Unter der Bruteforce1-Methodik zeigt sich jetzt, dass sich auch unter den Fehlendmechanismen MCAR und LogitX stark verzerrte Schätzer bezüglich der Koeffizienten ergeben. Die allgemeine Güte für die Koeffizienten unter LogitX unterscheidet sich nur sehr geringfügig vom Prozess LogitY unabhängig davon wie viele Y-Werte fehlen.

Verwendet man die Bruteforce1-Methodik, dann verfügt die lineare Regression unter dem Fehlendmechanismus MNAR3 für alle Koeffizienten über den geringsten MSE. Dies liegt hauptsächlich an der geringeren Verzerrung im Vergleich zu den anderen Fehlend-Prozessen. Jedoch sind die Koeffizientenschätzungen nicht besser, als wenn im Vergleich dazu das normale Verfahren angewendet wird.

Grundsätzlich lässt sich noch beobachten, dass für die Kovariablen, die multivariat erzeugt wurden, die Varianz und der Bias bezüglich der Koeffizienten höher sind als für die univariat erzeugten Kovariablen.

Darüber hinaus werden die Koeffizienten für metrische Kovariablen unabhängig des Fehlendmechanismus unterschätzt.

Insgesamt lässt sich für die Bruteforce2-Verfahren mit der linearen Regression für die Abbildungen (im Anhang B.3) und die Tabellen (Anhang C.2) feststellen, dass der Bias im Vergleich zu Bruteforce1 niedriger ist. Allerdings werden ebenfalls die Koeffizienten für alle Fehlendmechanismen verzerrt geschätzt.

Betrachtet man die Schätzungen unter Fehlendmechanismus MNAR1, dann wird sichtbar, dass für 50% fehlende Daten die Koeffizienten extrem verzerrt sind. Für die Fälle 25% und 10% fehlende Daten sind die Koeffizienten allerdings erheblich geringer verzerrt. Im Vergleich zum normalen Verfahren sind diese im geringem Maße schlechter. Betrachtet man als nächstes die Varianz, so verhält sich deren Güte umgekehrt. Das bedeutet die Varianz ist für 25% und 10% fehlenden Datenanteil wesentlich größer als für den Fall 50% fehlende Daten.

Unter dem Fehlendmechanismus MNAR2 lassen sich gleichartige Beobachtung machen. Insgesamt ist die Verzerrung für 50% fehlende Daten deutlich geringer als unter MNAR1. Für jeweils 25% und 10% fehlende Y-Werte nimmt der Bias hingegen nicht so stark ab wie im Fall MNAR1.

Unter der Bruteforce2-Methodik ergibt sich wie für Bruteforce1, dass zusätzlich unter den Fehlendmechanismen MCAR und LogitX stark verzerrte Schätzer bezüglich der Koeffizienten resultieren. Für den Fehlend-Prozess MCAR erhält man deutlich stärker verzerrte Koeffizienten als für MNAR1, MNAR2 und MNAR3. Erstaunlich ist, dass sich die allgemeine Güte für die Koeffizienten unter LogitX einzig in einer geringen Verzerrung zum Mechanismus LogitY unterscheidet. Die Varianz und der Bias verändern sich im gleichen Maße mit den unterschiedlichen fehlend-Anteilen.

Für Bruteforce2 werden die Koeffizienten für metrische Kovariablen nicht mehr wie für Bruteforce1 unabhängig des Fehlendmechanismus unterschätzt, sondern in einigen Fällen auch überschätzt.

Die Ergebnisse für die Gütekriterien bezüglich der Beta-Koeffizienten für das lineare Modell mit Ridge-Schätzung befinden sich als Tabellen im Anhang (C.2 für Bruteforce1, C.3 für Bruteforce2). Kurz zusammengefasst, ergibt sich in beiden Verfahren jeweils, dass die Koeffizienten unabhängig vom Fehlendmechanismus stark verzerrt sind und daraus ein höherer MSE resultiert. Die Verzerrung wird besonders durch den Penalisierungsparameter λ beeinflusst.

4.2 Ergebnisse der Modelle mit binärer Zielgröße

Im Nachfolgenden wird auf die Sensitivität hinsichtlich der Koeffizienten der logistischen Regression unter verschiedenen Fehlendmechanismen eingegangen.

4.2.1 Sensitivität bei logistischer Regression unter verschiedenen Fehlendmechanismen

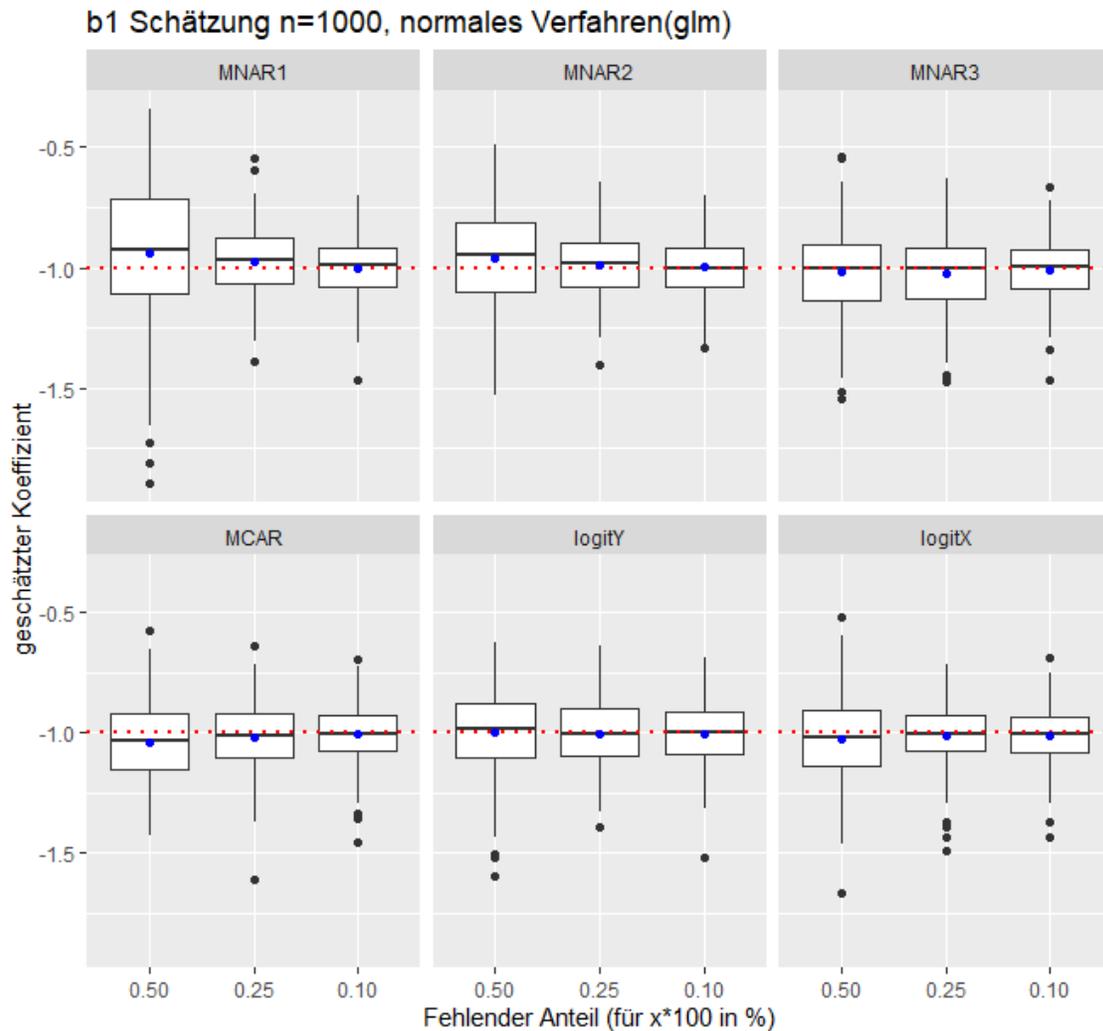


Abbildung 11: Es werden die geschätzten Werte für β_1 in Boxplotform dargestellt. Diese werden nach den Fehlendmechanismen gruppiert. Der blaue Punkt gibt den Mittelwert der geschätzten Werte an. Die rote horizontale Linie repräsentiert den echten Wert des Koeffizienten.

Aus Abbildung 11 ist ersichtlich, dass unter den Fehlendmechanismen MNAR1 und MNAR2 die Koeffizienten für β_1 im Vergleich zu den anderen Mechanismen deutlich verzerrt sind und auch die Varianz für 50% und 25% fehlende Daten erheblich größer ist. Allgemein betrachtet nimmt die Varianz und die Verzerrung für alle Fehlendmechanismen ab, wenn der fehlend-Anteil sinkt. Die Mechanismen MNAR3, MCAR, LogitY und LogitX weisen eine sehr geringe Varianz auf.

Für die Abbildung 12 lässt sich feststellen, dass bei fallendem Fehlend-Anteil die Varianz abnimmt. Die Varianz unter den Mechanismen MCAR, LogitY, LogitX und MNAR3 ist unabhängig vom Fehlend-Anteil gleich groß. Des Weiteren sind die Koeffizienten für β_2 unabhängig vom Fehlendmechanismus kaum verzerrt.

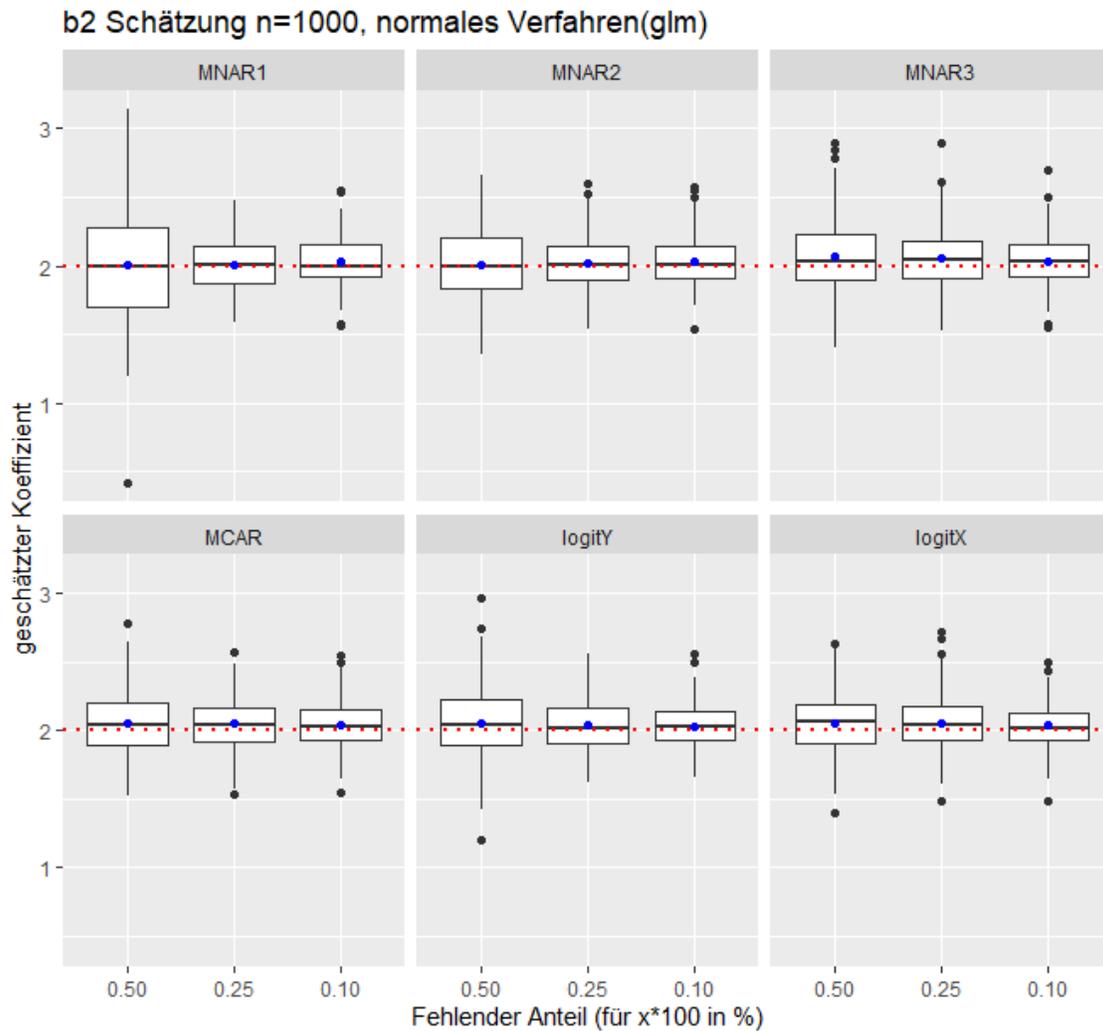


Abbildung 12: Es werden die geschätzten Werte für den Koeffizienten β_2 in Boxplotform dargestellt. Diese werden nach den Fehlendmechanismen gruppiert. Der blaue Punkt gibt den Mittelwert der geschätzten Werte an. Die rote horizontale Linie repräsentiert den echten Wert des Koeffizienten.

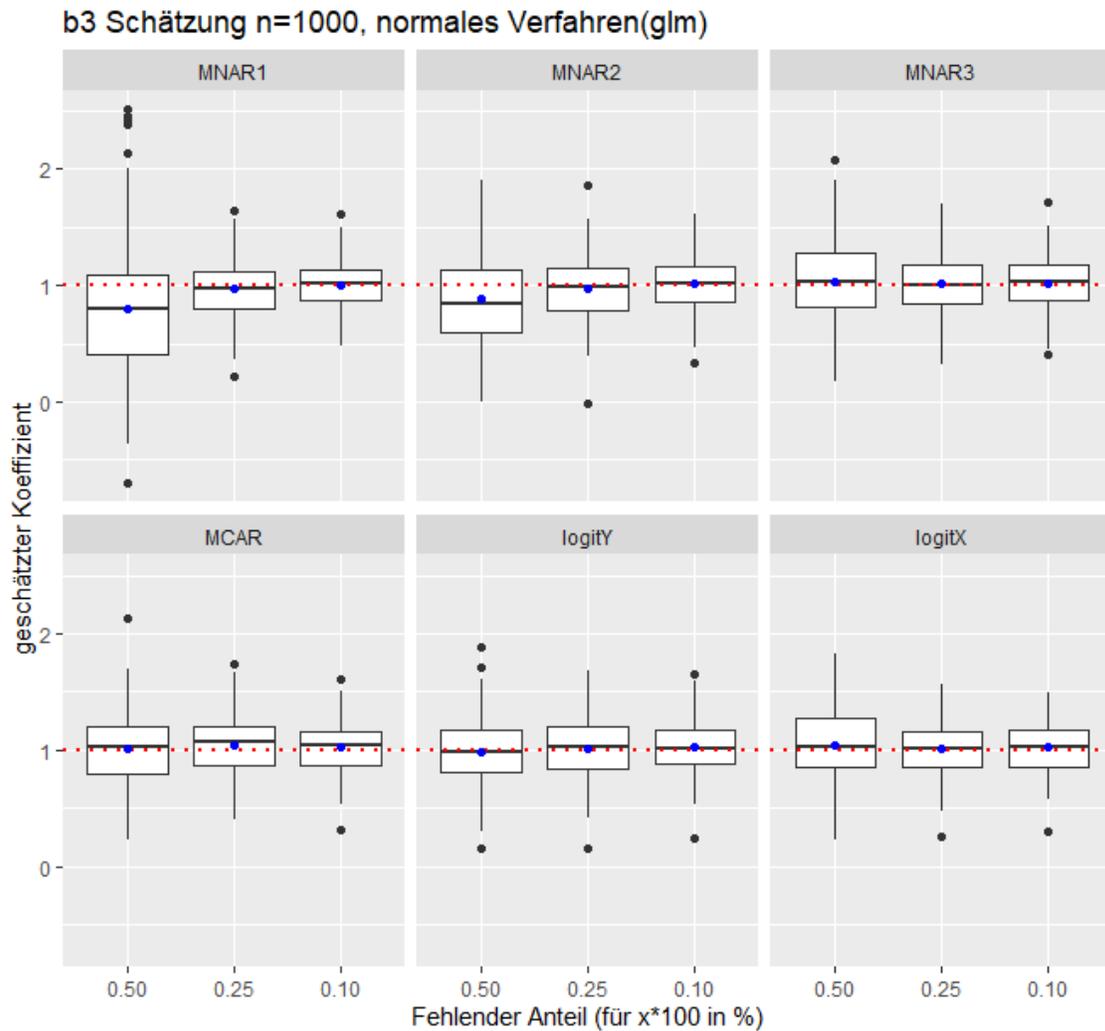


Abbildung 13: Es werden die geschätzten Werte für den Koeffizienten β_3 in Boxplotform dargestellt. Diese werden nach den Fehlendmechanismen gruppiert. Der blaue Punkt gibt den Mittelwert der geschätzten Werte an. Die rote horizontale Linie repräsentiert den echten Wert des Koeffizienten.

Aus Abbildung 13 ist zu erkennen, dass für die Fehlendmechanismen MNAR1 und MNAR2 $\hat{\beta}_3$ verzerrt ist und eine höhere Varianz aufweist als unter den Mechanismen MCAR, LogitY und LogitX. Sowohl die Verzerrung als auch die Varianz nehmen ab, wenn mehr Y-Werte beobachtet werden. Für die anderen Mechanismen ist die Verzerrung der Beta-Schätzung eher gering. Es ist dort auch zu sehen, dass die Verzerrung und die Varianz geringer sind, falls weniger Beobachtungen verloren gehen.

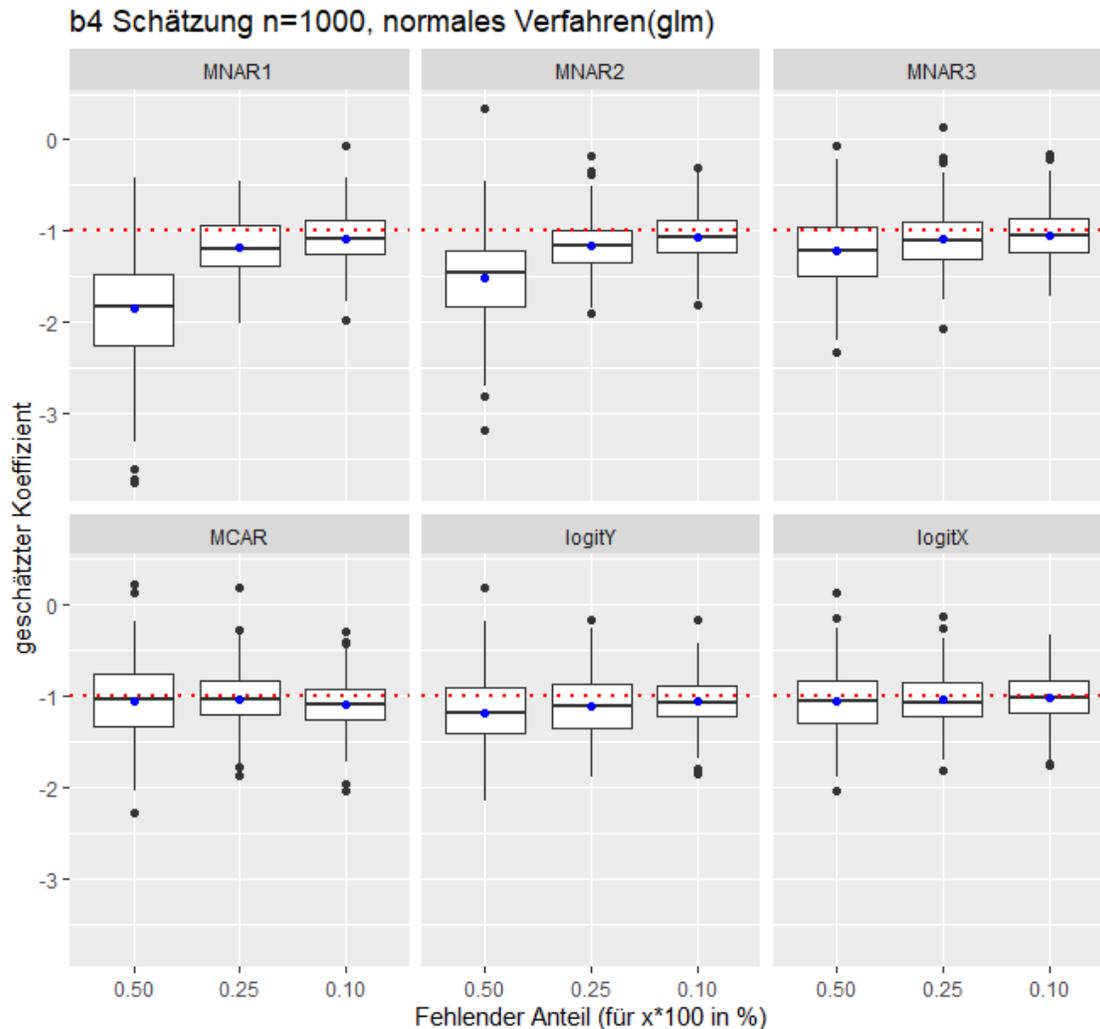


Abbildung 14: Es werden die geschätzten Werte für den Koeffizienten β_4 in Boxplotform dargestellt. Diese werden nach den Fehlendmechanismen gruppiert. Der blaue Punkt gibt den Mittelwert der geschätzten Werte an. Die rote horizontale Linie repräsentiert den echten Wert des Koeffizienten.

Aus Abbildung 14 wird ersichtlich, dass für alle Fehlendmechanismen, die als MNAR einzuordnen sind, wesentlich verzerrte Schätzungen für $\hat{\beta}_4$ geliefert werden. Hierarchisch lässt sich dabei MNAR1 mit der stärksten Verzerrung und LogitY mit einer schwächeren Verzerrung festhalten. Als besonders einflussreich stellt sich der Fehlend-Anteil heraus. So wird bspw. für den Fall von 50% fehlenden Y-Werten $\hat{\beta}_4$ unter MNAR1 im Mittel auf 1,85 geschätzt, obwohl der wahre Wert -1 ist.

In Abbildung 15 lässt sich erkennen, dass unter MNAR1, MNAR2 die Koeffizientenschätzungen für

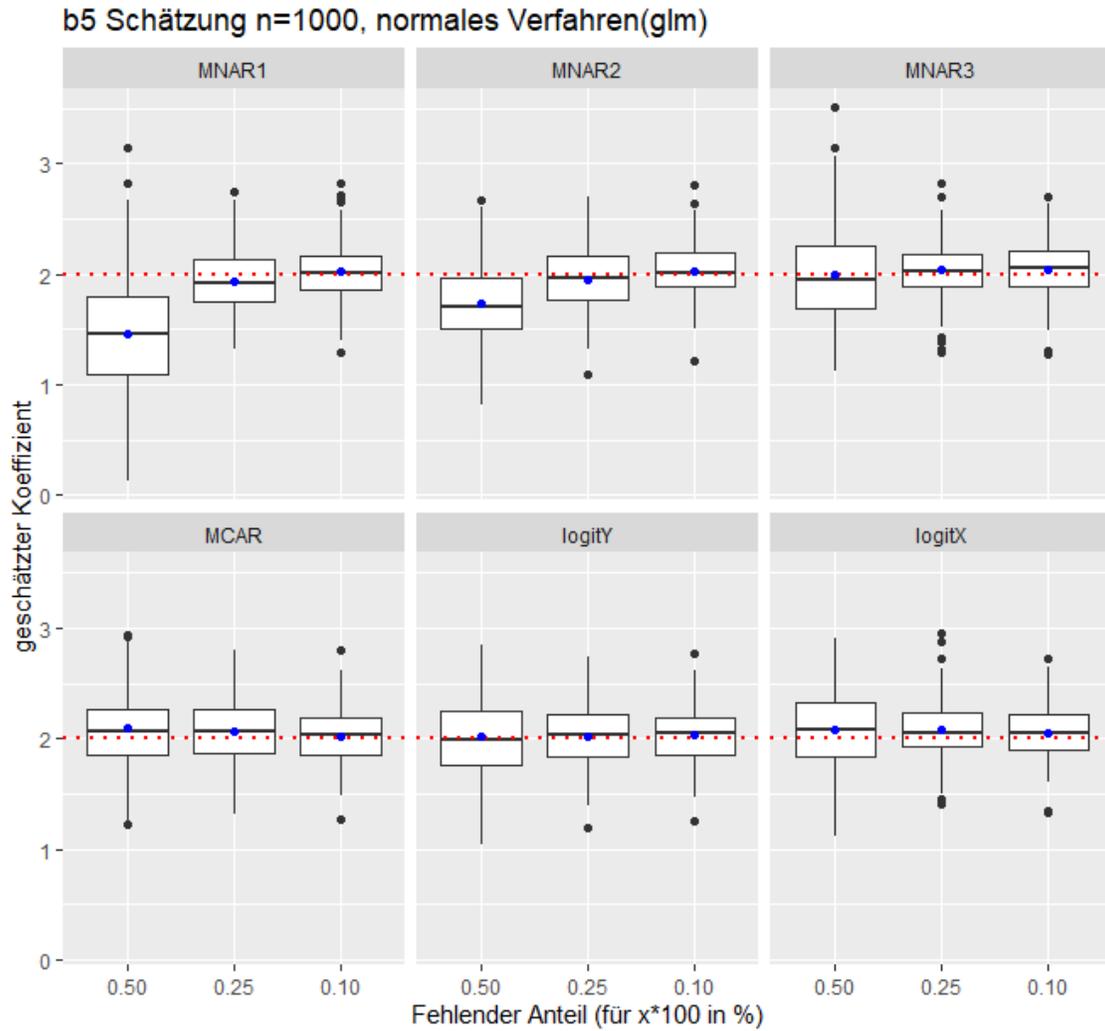


Abbildung 15: Es werden die geschätzten Werte für den Koeffizienten β_5 in Boxplotform dargestellt. Diese werden nach den Fehlendmechanismen gruppiert. Der blaue Punkt gibt den Mittelwert der geschätzten Werte an. Die rote horizontale Linie repräsentiert den echten Wert des Koeffizienten.

β_5 besonders für 50% fehlenden Daten deutlich verzerrt und gestreut sind. Für MNAR3 und MCAR, LogitY und LogitX sind die Koeffizienten unabhängig vom fehlenden Anteil kaum verzerrt.

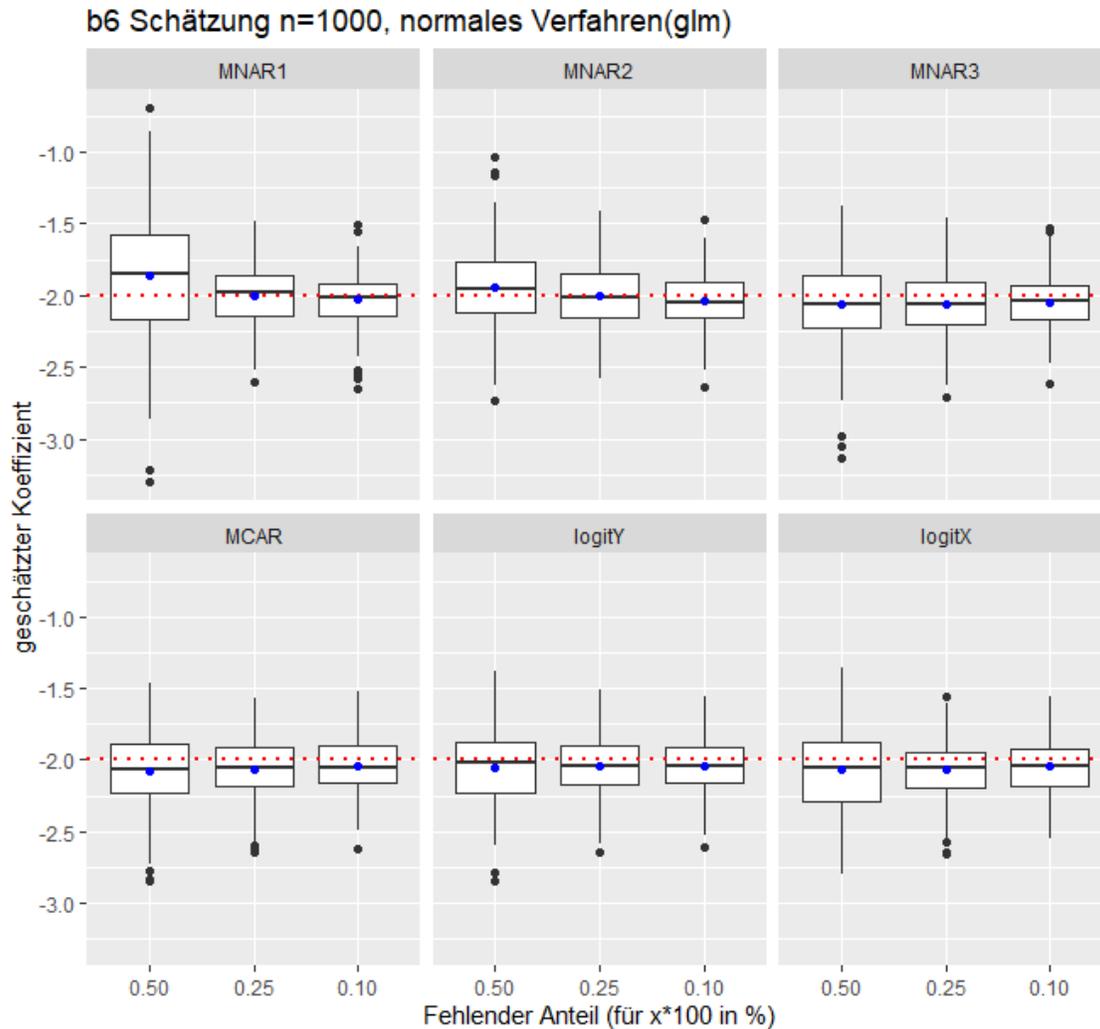


Abbildung 16: Es werden die geschätzten Werte für den Koeffizienten β_6 in Boxplotform dargestellt. Diese werden nach den Fehlendmechanismen gruppiert. Der blaue Punkt gibt den Mittelwert der geschätzten Werte an. Die rote horizontale Linie repräsentiert den echten Wert des Koeffizienten.

Aus der Abbildung 16 lässt sich ablesen, dass unter den Mechanismen MNAR1, MNAR2 und MNAR3 die Schätzungen für β_6 für 50% fehlende Y-werte deutlich verzerrt sind. Aus der Abbildung hervorzuheben sind MCAR, LogitY, LogitX und MNAR3, die unabhängig vom Anteil der fehlenden Y-Werte verzerrte Koeffizienten liefern.

In Abbildung 17 ist ersichtlich, dass unter den Mechanismen MNAR1 und MNAR2 stark verzerrte Schätzungen für β_7 herauskommen.

Aus Abbildung 18 lässt sich ableiten, dass die Koeffizienten für β_8 unter den Fehlendmechanismen MNAR1, MNAR2 und MNAR3 extrem verzerrt sind und die Varianz allgemein unter allen Fehlendmechanismen hoch ist.

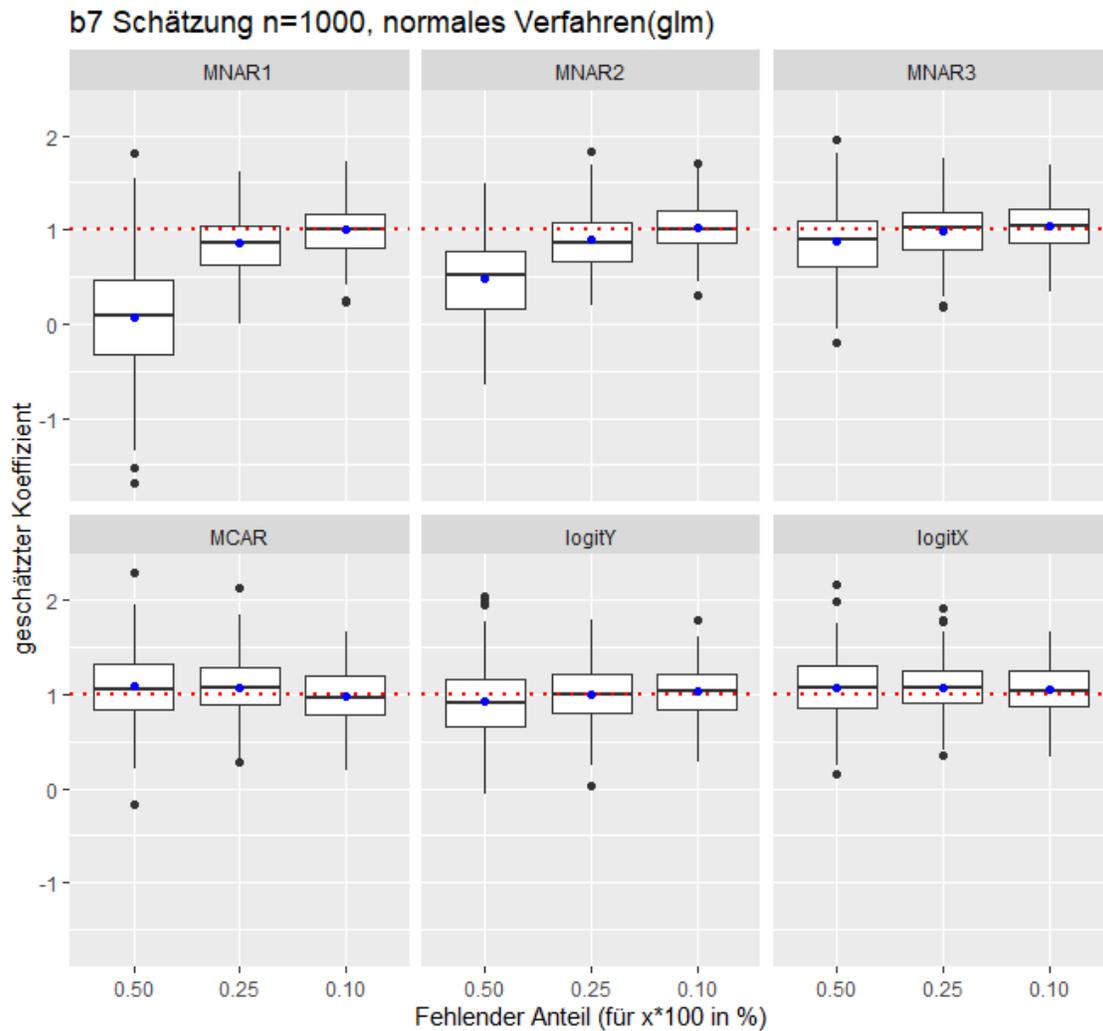


Abbildung 17: Es werden die geschätzten Werte für den Koeffizienten β_7 in Boxplotform dargestellt. Diese werden nach den Fehlendmechanismen gruppiert. Der blaue Punkt gibt den Mittelwert der geschätzten Werte an. Die rote horizontale Linie repräsentiert den echten Wert des Koeffizienten.

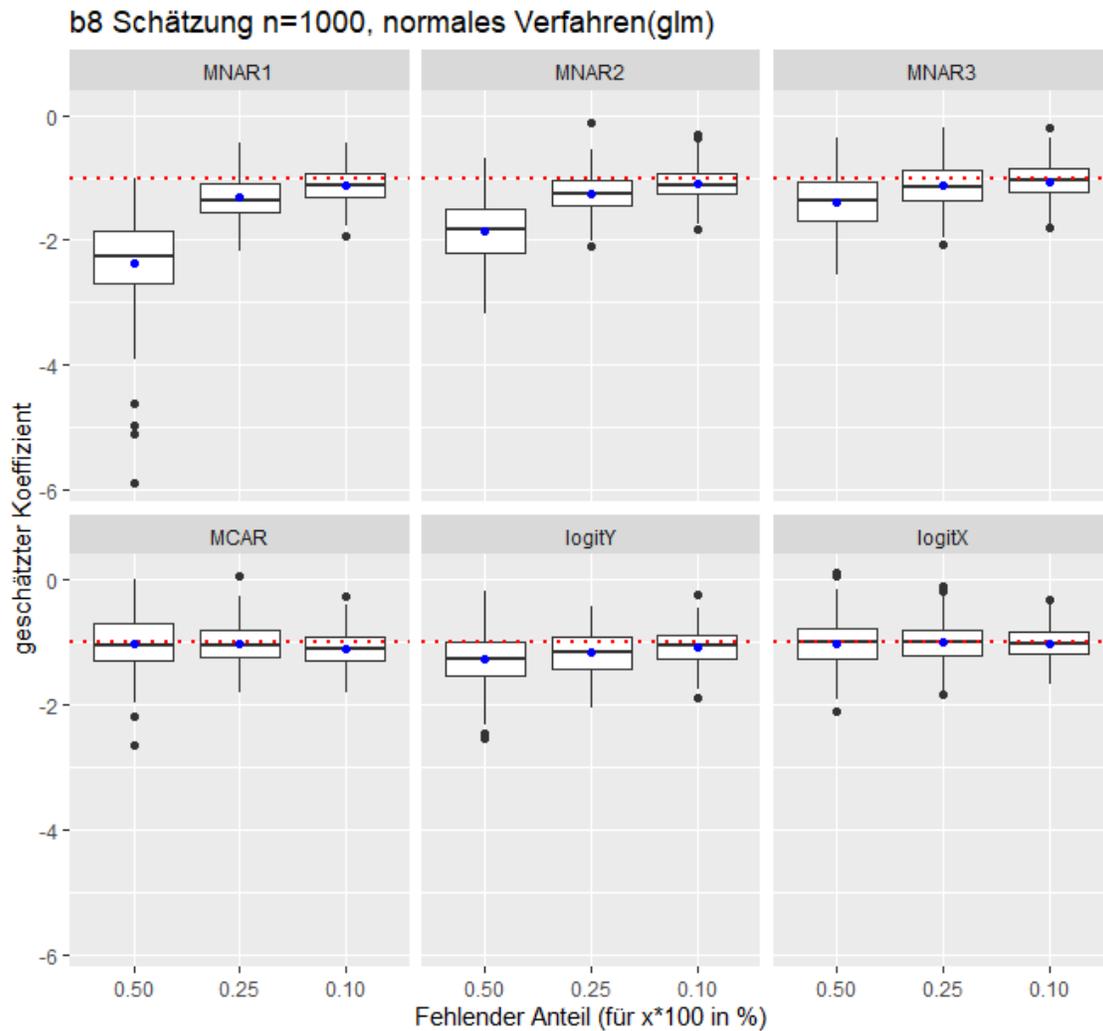


Abbildung 18: Es werden die geschätzten Werte für den Koeffizienten β_8 in Boxplotform dargestellt. Diese werden nach den Fehlendmechanismen gruppiert. Der blaue Punkt gibt den Mittelwert der geschätzten Werte an. Die rote horizontale Linie repräsentiert den echten Wert des Koeffizienten.

Allgemein auffällig ist, dass für die binären Kovariablen X_3 , X_4 , X_7 und X_8 die Koeffizienten unter den Fehlendmechanismen des Typs MNAR ausschließlich unterschätzt werden. Außerdem sind die Verzerrung und die Varianz bei den korrelierten Kovariablen größer als bei den unkorrelierten Variablen. Dies ist deutlich zu erkennen, wenn man die Schätzungen für β_1 und β_2 mit β_5 und β_6 vergleicht. Ebenfalls ist der beschriebene Unterschied vor allem unter MNAR1, MNAR2 und MNAR3 für β_3 und β_4 zu β_7 und β_8 ersichtlich. Interessant ist, dass sich LogitY und LogitX hinsichtlich ihrer Gütekriterien für β ähnlich sind, obwohl LogitY vom Typ MNAR und LogitX vom Typ MCAR ist. Zusätzlich ist festzuhalten, dass unter MNAR1 und MNAR2 erheblich verzerrte Koeffizientenschätzungen ermittelt werden. Allgemein betrachtet ist die Varianz der Koeffizienten im logistischen Modell für alle Mechanismen hoch verglichen mit den Varianzen aus den linearen Modellen.

4.2.2 Sensitivität bei Brute-force-Methodiken mit logistischer Regression unter verschiedenen Fehlendmechanismen

Für das Brute-force2-Verfahren mit logistischer Regression (Anhang D.1) lässt sich allgemein festhalten, dass alle Beta-Koeffizienten unabhängig vom Fehlendmechanismus unterschätzt werden. Zieht man ebenfalls die Tabellen aus dem Anhang E.2 heran, dann sieht man, dass die Verzerrung der Beta-Koeffizienten mit Brute-force2 im Vergleich zu Brute-force1 kleiner ist. Interessanterweise steigen die Varianzen der geschätzten Koeffizienten mit sinkendem fehlend-Anteil. Der Bias hingegen nimmt leicht ab, wenn mehr beobachtete Daten vorliegen. Der MSE ist für die geschätzten Koeffizienten für das logistische Modell unter dem Fehlendmechanismus LogitX allgemein im Vergleich zu den anderen Mechanismen am größten. Anhand der Grafiken (aus Anhang D.1) sieht man, dass sich die Gütekriterien für die Koeffizienten tendenziell unter allen Fehlendmechanismen gleich entwickeln.

5 Diskussion

5.1 Mögliche Probleme bei den verwendeten Regressionsmodellen in R

Bei der Modellierung der linearen Regression auf die Datensätze, für die der Fehlendmechanismus bereits angewendet wurde, können in R gegebenenfalls Warnmeldungen ausgegeben werden. Im Rahmen dieser Arbeit tritt für die lineare Regressionsschätzung nur eine spezielle Warnmeldung auf: *In:prediction from a rank-deficient fit may be misleading.*

Diese Warnmeldung erscheint, wenn mithilfe des geschätzten linearen Modells die fehlenden Y-Werte imputiert (bzw. geschätzt) werden sollen (mit Funktion `predict.lm()`). Bei näherer Beschäftigung mit der Warnmeldung ergeben sich vor allem zwei mögliche Gründe. Zum einen wird mit der Funktion überprüft, ob der Rang der Datenmatrix mindestens der Anzahl der Parameter entspricht. Da das Modell jedoch nur acht Kovariablen enthält und auch die Kolinearität der Kovariablen für zu gering erscheint, ist diese Variante eher auszuschließen. Zum anderen besteht die Möglichkeit, dass die Datenmenge als zu gering angesehen wird. Dieser Fall ist als wesentlich wahrscheinlicher anzusehen, da die Warnmeldung ausschließlich auftritt, wenn mit dem Datensätzen mit jeweils 50 Beobachtungen gearbeitet wird. Die Warnmeldung wird jedoch hauptsächlich nur ausgelöst, falls von den 50 Beobachtungen unter den Mechanismen MNAR 50% der Y-Werte verloren gehen.

Ein Ansatz um diese Problematik zu umgehen oder um zu überprüfen wie stark dies die Ergebnisse beeinflusst, ist dass man eine Ridge Regression mit `glmnet()` mit einem λ nahe an der Null schätzt. Die groben Hintergründe und die Umsetzung werden in Kapitel 3.1.2 erläutert. Die Ergebnisse des

Vergleichs zwischen `lm()` und `glmnet()` mit $\lambda = 0.0001$ werden in Kapitel 4.1 vorgestellt.

Grundsätzlich kann sich die Dauer der Laufzeit, die benötigt wird um die Simulation für viele Durchgänge laufen zu lassen als problematisch erweisen. Einerseits ist es erforderlich möglichst effizient die gewünschten Schritte zu programmieren, andererseits benötigen komplexere Modelle (besonders `glmnet()` und `cv.glmnet()`), die geschätzt werden sollen, einiges mehr an Zeit als einfache Funktionen. In dieser Simulationsstudie wurde die Anzahl der Modellschätzungen durch die Bruteforce-Methodik äußerst stark und rapide erhöht. So werden bspw. für die Bruteforce-Methodik, wenn eine spezielle Datensatzgröße mit einem bestimmten fehlend-Anteil unter einem Fehlendmechanismus betrachtet werden soll, $150 \cdot 100 = 15000$ Modellschätzungen nötig. Da in dieser Simulation unter anderem der beste Wert für λ über 5-fold Cross Validation und anschließend noch die Ridge Regression mit dem ermittelten λ -Wert geschätzt wird, werden die 15000 Modellschätzungen nochmals vervielfacht. Außerdem erscheint es als durchaus sinnvoll, die abschließende Modellschätzung wegzulassen, wenn mithilfe eines zuvor geschätzten Regressionsmodells die fehlenden Werte imputiert wurden, und einfach direkt die Koeffizienten vom ersten Modell zu nehmen.

Für die Modellierung einer logistischen Regression auf Datensätze mit Größe 25, für die der Fehlendmechanismus bereits angewendet wurde, treten in R deutlich mehr Warnmeldungen auf als im linearen Modell (teilweise auch für Beobachtungsgröße 200). Im Rahmen der logistischen Regression erscheinen die Warnmeldungen *fitted probabilities numerically 0 or 1 occurred* und *algorithm did not converge*.

Die erste Warnmeldung bedeutet, dass die geschätzten Wahrscheinlichkeiten so hohe/niedrige Werte annehmen, dass sich im Prinzip numerisch 1/0 als Resultat ergibt. Dieser Fall tritt vor allem für separierende Variablen auf, z.B. wenn für eine Kovariable ein Schwellenwert existiert, für den sich die Realisierungen 0 und 1 der Zielvariable aufteilen lassen. Im mehrdimensionalen Raum tritt dieses Problem häufig auf, wenn eine hohe Multikolarität der Kovariablen vorliegt.

Die andere Warnmeldung bringt zum Ausdruck, dass die `glm()`-Schätzung nicht konvergiert, also der Algorithmus nicht das Abbruchkriterium erfüllt. Die Probleme, die mit den Warnmeldungen einhergehen, werden im Anhang E.1 für die Tabelle 56, Tabelle 57 und Tabelle 58 für die Datensatzgröße 200 ersichtlich. Eine einfache Möglichkeit zur Lösung besteht darin, die Anzahl der Iterationsschritte zu erhöhen. Die andere Option, die beide Warnmeldungen behandelt, ist, dass auch hier ein logistisches Modell mit Ridge (mit $\lambda = 0.0001$) geschätzt wird. Dies kann ebenfalls mit der `glmnet()`-Funktion in R durchgeführt werden. In der Simulation kommt allerdings zusätzlich das Problem auf, dass unter den Fehlendmechanismen MNAR ausschließlich die Zielvariable $Y = 1$ gefiltert wird. Da in den Datensätzen kein großes Ungleichgewicht bezüglich der Verteilung der Zielvariable vorliegt, treten vereinzelt Fälle auf, in denen im verbleibenden Datensatz für eine Gruppe (0 oder 1) keine Beobachtungen vorliegen.

5.2 Diskussion der Ergebnisse

5.2.1 Mögliche Erklärung der Ergebnisse

Für die lineare Modellschätzung mit normalem Vorgehen ergibt sich, dass die Varianz der geschätzten Koeffizienten eher klein ausfällt und die Verzerrung eine dominantere Rolle spielt, wenn der MSE herangezogen wird. Das begründet sich daraus, dass jeder der 150 Datensätze mit metrischer Zielgröße dieselben Realisierungen bezüglich der Kovariablen hat. Da die Y-Werte über eine lineare Regression generiert werden, wird die einzige Variation bzw. Unsicherheit durch ϵ , wie in Kapitel 3.1 beschrieben, erzeugt. Folglich ist die Abhängigkeit bzw. der Zusammenhang von den Kovariablen

zu Y sehr groß und wird bei den Modellschätzungen deutlich.

Die geringfügig höheren Varianzen für die Koeffizienten unter den Bruteforce-Methodiken wirken im ersten Moment leicht verwunderlich. Allerdings lassen sich die in diesem Kontext verhältnismäßig geringen Varianzen damit erklären, dass im ersten Schritt für die fehlenden Y -Werte über eine Gleichverteilung zufällig Werte zugeteilt werden und dadurch der direkte Zusammenhang der Kovariablen, der zu der Zielgröße besteht, abnimmt und damit die gestiegene Unsicherheit zu größeren Varianzen der Koeffizienten resultieren. Dieser Vorgang wird für jeden einzelnen Datensatz 100 mal wiederholt. Für die 100 imputierten Datensätze werden anschließend die linearen Modelle geschätzt. Für diese kann der geschätzte Zusammenhang durchaus verstärkt schwanken. Allerdings werden für die 100 vollständigen Datensätze die geschätzten Koeffizienten jeweils gemittelt. Betrachtet man die geschätzten Koeffizienten als unabhängige, identisch verteilte Zufallsvariable, dann wird die ursprüngliche Varianz mit dem Faktor $1/100$ gewichtet. Dies lässt sich direkt für die Varianz des arithmetischen Mittels herleiten. Dadurch ergibt sich im Prinzip für die geschätzten Koeffizienten für die 150 Datensätzen keine allzu große Streuung.

Das lineare Modell mit Ridge-Schätzung liefert erheblich verzerrte Koeffizienten. Diese werden hauptsächlich durch den Penalierungsparameter λ hervorgerufen. Allerdings besteht gerade das Hauptziel der Ridge Regression darin, dass die geschätzten Koeffizienten abgeschwächt werden um *Overfitting* auf dem Trainingsdatensatz zu vermeiden und anschließend eine bessere Prognosegüte zu erzielen. Der Parameter λ wird in diesem Fall durch eine 5-fold Cross Validation ermittelt. Damit ist offensichtlich, dass die Koeffizienten für das lineare Modell mit Ridge im Vergleich zur linearen Regression einen höheren Bias und in diesem Zusammenhang eine niedrigere Güte aufweisen. Außerdem hat die höhere Verzerrung bezüglich der Koeffizienten für die multivariat erzeugten Kovariablen gezeigt, dass die Funktion `cv.glmnet()` die Kolinearität der abhängigen Variablen erfolgreich identifiziert hat. Dies war vor allem für die Koeffizienten β_5 und β_6 zu erwarten, wie aus der Kovarianzmatrix aus Kapitel 3.1 ersichtlich ist, dass diese über eine moderate bis hohe positive Korrelation (ca. 0.7) verfügen.

5.2.2 Einordnung und Bewertung der Ergebnisse

Im Falle des Datensatzes mit metrischer Zielgröße lässt sich feststellen, dass das normale Verfahren, das direkt auf den verbliebenen Datensatz ein lineares Modell ohne Ridge schätzt, den besten MSE für die geschätzten Koeffizienten liefert. Dennoch lässt sich unter den Fehlendmechanismen MNAR1, MNAR2, MNAR3 und LogitY eine systematische Verzerrung feststellen. Vor allem die Größe des fehlenden Anteils ist für die Verzerrung ausschlaggebend.

Für das lineare Modell mit Ridge ist aufgefallen, dass der MSE für die geschätzten Koeffizienten der univariat erzeugten Kovariablen unter dem Prozess MNAR3 am niedrigsten ist. Da im linearen Modell ohne Ridge unter MNAR3 die Koeffizienten dazu neigen überschätzt zu werden, wird durch die Penalisierung die Verzerrung in die „richtige“ Richtung reduziert. Besonders hervorzuheben ist der Fall für 50% fehlende Daten, da dort für die univariaten Variablen der jeweiligen Koeffizienten eine niedrigere Verzerrung vorliegt. Verglichen mit der normalen linearen Regression ist der Bias sogar niedriger. Sollten folglich die Koeffizienten überschätzt werden, da beispielsweise durch einen Fehlendmechanismus primär kleinere Y -Werte verloren gehen, dann besteht durch die Penalisierung mit λ die Möglichkeit, dass die Verzerrung im richtigen Maße reduziert werden kann.

Der Bias unter den Fehlendmechanismen vom Typ MNAR bezüglich der geschätzten Koeffizienten im linearen Modell zeigt, dass die Fehlendmechanismen MNAR1, MNAR2, MNAR3 und LogitY sinngemäß erstellt worden sind.

Der Versuch, mit Bruteforce 2 absichtlich höhere Y -Werte für die fehlenden Werte zufällig zu-

zuteilen und damit die verzerrten Koeffizientenschätzungen abzuschwächen, ist im Wesentlichen fehlgeschlagen. Mit Bruteforce2 ist es offensichtlich, dass das Verfahren unter den Fehlendmechanismen MCAR und MAR (LogitX) für die Koeffizienten zu erheblichen Verzerrungen führt. Zusätzlich dazu wurden noch die Fehlendmechanismen LogitY und MNAR3 deutlich verzerrt. Hauptmotivation des Bruteforce2-Verfahrens sind die Mechanismen MNAR1 und MNAR2, da bei diesen gezielt über Quantile große Werte für Y ausgewählt werden. Wie in Kapitel 4.1.3 gezeigt wurde, sind die Koeffizienten, wenn 50% der Y-Werte fehlen unter MNAR1 und MNAR2 extrem stark verzerrt. Für 25% und 10% fehlende Daten wird die Verzerrung unter beiden Mechanismen mit dem Bruteforce2-Verfahren deutlich reduziert und erfüllt somit teilweise seinen Zweck. Allerdings ist der Bias im Vergleich zum normalen Verfahren mit linearer Regression immer noch höher und damit verliert die Methodik seine Berechtigung. Für den Fehlendprozess LogitY konnte der Bias nicht reduziert werden, sondern wurde sogar noch größer, verglichen zum normalen Verfahren mit linearer Regression. Das Bruteforce1-Verfahren teilt für die fehlenden Y-Werte zufällig Werte, die mit einer Gleichverteilung über den Wertebereich von Minimum bis Maximum der verbliebenen Y-Werte generiert werden, zu. Damit wird hauptsächlich versucht, der Verzerrung der Koeffizienten im linearen Modell unter MNAR3, das tendenziell mittelgroße Y-Werte aus dem Datensatz filtert, entgegenzuwirken. Im Vergleich zu den anderen Fehlendmechanismen ist der Bias der Schätzungen deutlich kleiner und dies trifft auch für 50% fehlende Daten zu. Vergleicht man die Verzerrung jedoch mit dem normalen Verfahren mit linearer Regression, dann ist die Verzerrung der geschätzten Koeffizienten größer und das Bruteforce1-Verfahren kann seinen Zweck nicht erfüllen.

6 Fazit

Sowohl die Verzerrung als auch die Varianz hinsichtlich der geschätzten Koeffizienten in den linearen Modellen hängen von unterschiedlichen Faktoren ab. Die Güte der Beta-Koeffizienten wird demnach vorrangig durch die Datensatzgröße, den Anteil fehlender Daten und der Art des Fehlendmechanismus beeinflusst. Als weiterer Einflussfaktor kann ein eigens gewollter Prozess vorgegeben werden, der bspw. die fehlenden Daten durch gezielte Werte auffüllt.

Aus dem Ergebnisteil lässt sich insgesamt für den Datensatz mit metrischer Zielvariable feststellen, dass die Verzerrung hinsichtlich der Koeffizienten im linearen Modell zusammen mit dem fehlend-Anteil sinkt. Es zeigt sich ebenso, dass die Varianz und der Bias der Koeffizientenschätzungen in größeren Datensätzen kleiner werden.

Unter den Fehlendmechanismen MNAR1, MNAR2 und LogitY ergeben sich stark verzerrte Koeffizientenschätzer, die Varianz bleibt dabei aber weitestgehend unberührt. Zusätzlich hat sich herausgestellt, dass unterschiedlich starke Versionen der MNAR-Mechanismen auch verschiedenen starke Verzerrungen herbeiführen. So sind die Beta-Koeffizienten unter MNAR1 stärker verzerrt als unter MNAR2. Für die logistische Regression zeigen sich diese Tendenzen hauptsächlich für die multivariat erzeugten Variablen und die binären Kovariablen.

Das normale Verfahren, das direkt auf den verbliebenen Datensatz ein lineares Regressionsmodell ohne Ridge schätzt, hat für die geschätzten Koeffizienten unter allen Fehlendmechanismen den niedrigsten MSE. Dennoch lässt sich unter den Fehlendmechanismen MNAR1, MNAR2, MNAR3 und LogitY eine deutliche Verzerrung feststellen.

Mit dem Bruteforce1-Verfahren wird versucht, die Verzerrung der Koeffizienten, die sich mit dem linearen Regressionsmodell ergibt, zu reduzieren. Mit dem Verfahren lässt sich zwar zeigen, dass der MSE und vor allem der Bias mit dem Verfahren unter MNAR3 niedriger als unter den anderen Mechanismen ist, allerdings ist der Bias nicht niedriger als im normalen Verfahren.

Das Ziel für Bruteforce2 ist, dass es gelingt, für die Fehlendmechanismen MNAR1 und MNAR2, ein Verfahren im Zusammenhang unterschiedlicher linearer Modelle bereitzustellen, das nahezu unverzerrte Schätzungen der Beta-Koeffizienten liefert. Dies ist für 25% und 10% fehlende Daten nur mäßig gelungen, denn für die lineare Regression mit normalem Verfahren ergibt sich ein geringerer Bias.

Literatur

- [van Buuren 2018] BUUREN, Stef van: *Flexible Imputation of Missing Data*. Chapman Hall, 2018
- [Genbäck 2016] GENBÄCK, Minna: Uncertainty Intervals and Sensitivity Analysis for Missing Data. (2016). <https://www.diva-portal.org/smash/get/diva2:1043969/FULLTEXT01.pdf>
- [Helge Toutenburg 2007] HELGE TOUTENBURG, Christian H.: *Induktive Statistik: Eine Einführung mit R und SPSS*. Springer, 2007
- [Junyang Qian 2016] JUNYANG QIAN, Trevor H.: An Introduction to glmnet. (2016). <https://cran.r-project.org/web/packages/glmnet/vignettes/glmnet.pdf>
- [Kenkel 2013] KENKEL, Brenton: A GENERAL SOLUTION TO NONIGNORABLE MISSING OUTCOMES IN BINARY CHOICE DATA. (2013). <https://bkenkel.com/data/idlogit.pdf>
- [Kreutzmann 2019] KREUTZMANN, Ann-Kristin: Vom Umgang mit fehlenden Werten. (2019). <https://wikis.fu-berlin.de/display/fustat/Vom+Umgang+mit+fehlenden+Werten+VomUmgangmitfehlendenWerten-TraditionelleMethoden>
- [Lehmann 2004] LEHMANN, Thomas: Behandlung von fehlenden Werten bei nicht ignorierbaren Ausfallmechanismen. (2004). https://www.db-thueringen.de/servlets/MCRFileNodeServlet/dbt_derivate_00004775/Promo.pdf
- [Ludwig Fahrmeir 2009] LUDWIG FAHRMEIR, Stefan L. Thomas Kneib K. Thomas Kneib: *Regression: Modelle, Methoden und Anwendungen*. Springer, 2009
- [Roderick J. A. Little 2002] RODERICK J. A. LITTLE, Donald B. R.: *Statistical Analysis with Missing Data*. John Wiley Sons, 2002
- [Trevor Hastie 2008] TREVOR HASTIE, Jerome F. Robert Tibshirani T. Robert Tibshirani: *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Springer, 2008

A Grafiken zu den Fehlend-Anteilen unter Generierung durch Logit-Modell

A.1 Datensatz mit metrischer Zielgröße

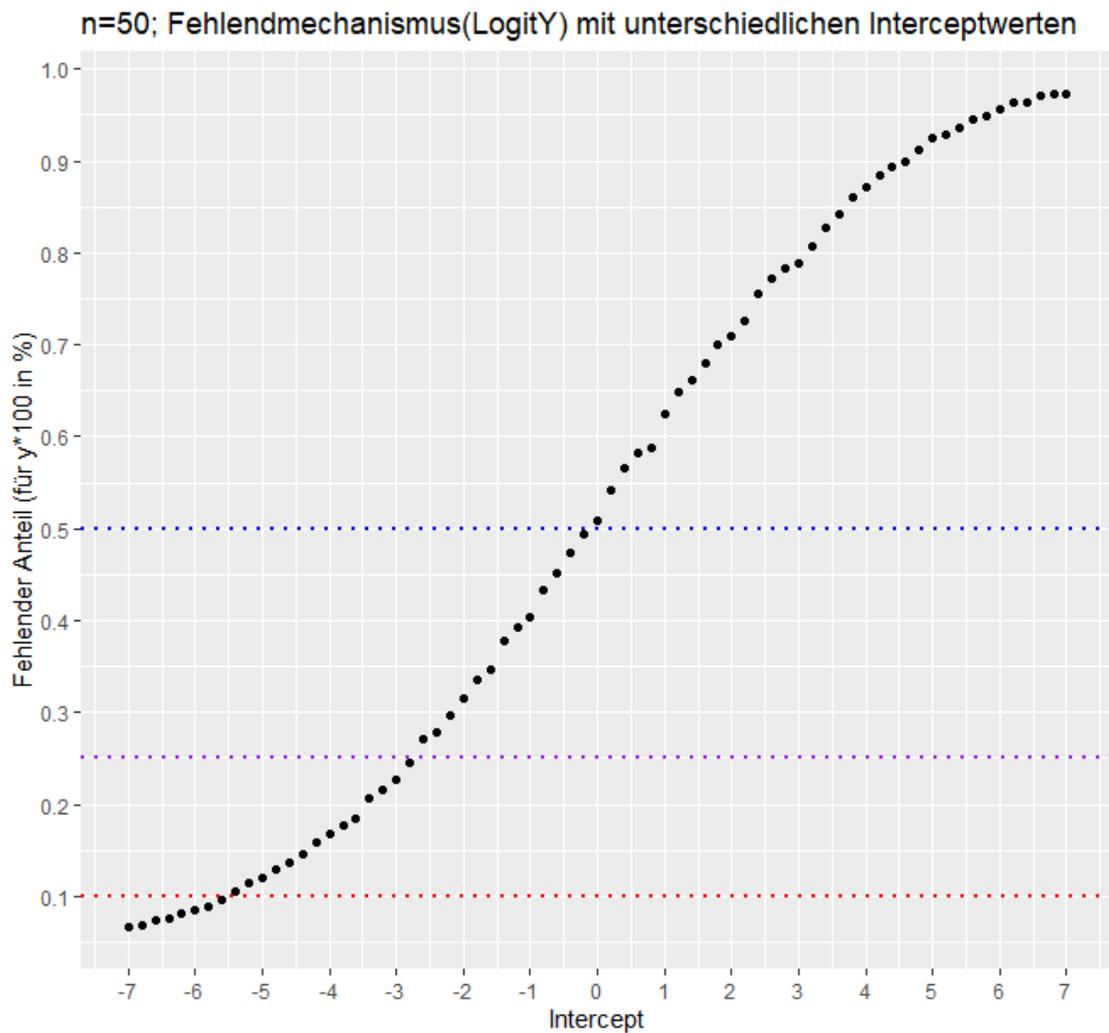


Abbildung 19: Es werden die Fehlend-Anteile, die durch LogitY mit unterschiedlichen Interceptwerten erzeugt werden, angezeigt. Zusätzlich werden durch die eingezeichneten horizontalen Linien die gewünschten Fehlend-Anteile hervorgehoben.

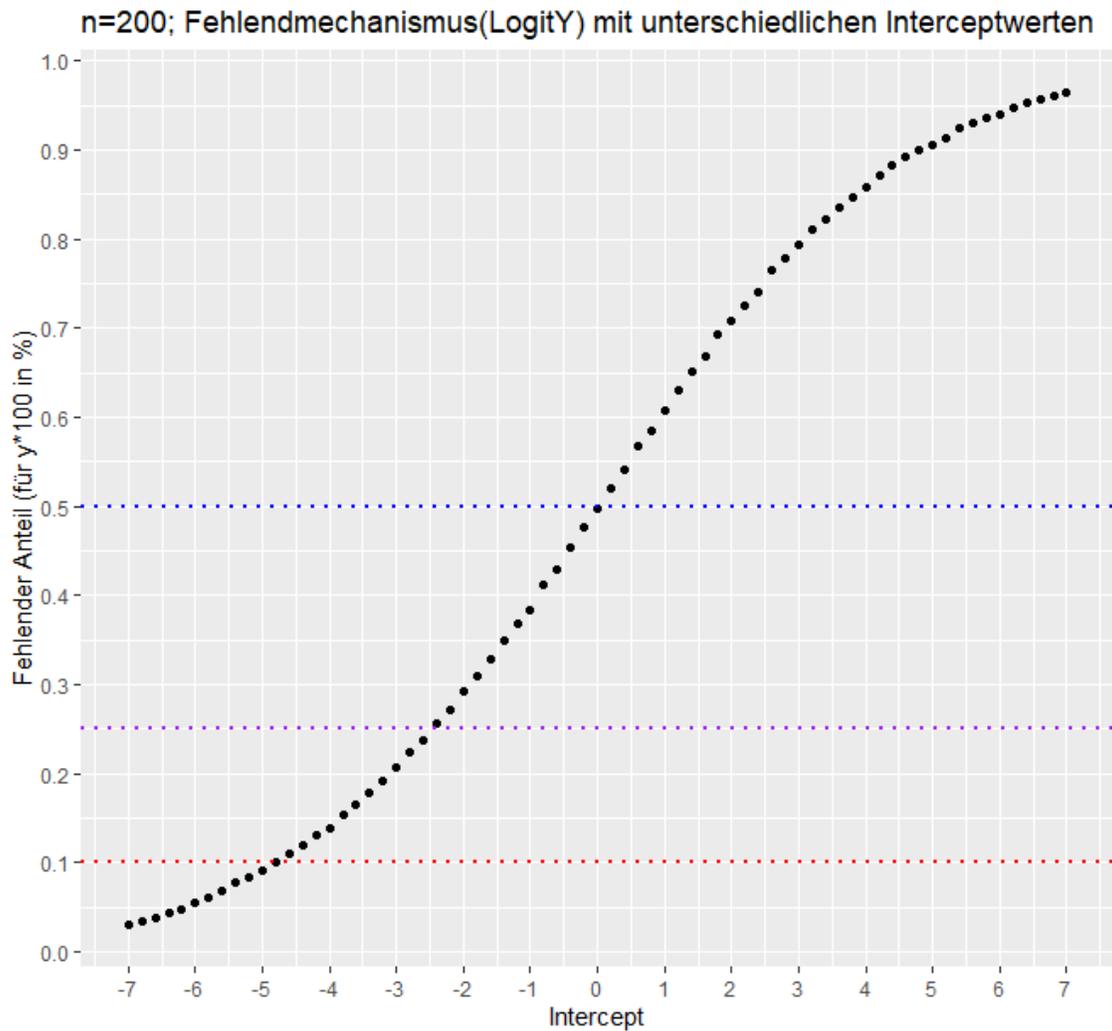


Abbildung 20: Es werden die Fehlend-Anteile, die durch LogitY mit unterschiedlichen Interceptwerten erzeugt werden, angezeigt. Zusätzlich werden durch die eingezeichneten horizontalen Linien die gewünschten Fehlend-Anteile hervorgehoben.

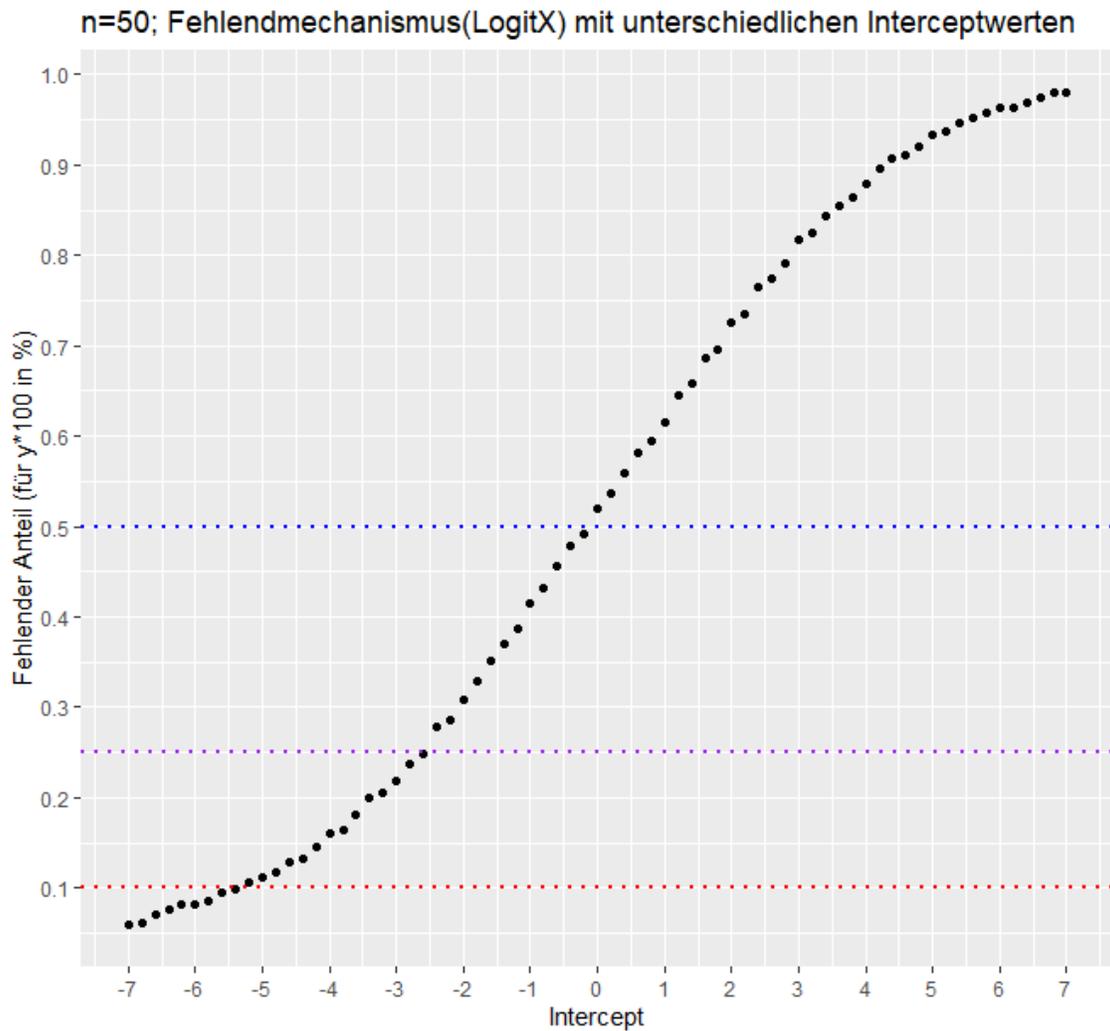


Abbildung 21: Es werden die Fehlend-Anteile, die durch LogitX mit unterschiedlichen Interceptwerten erzeugt werden, angezeigt. Zusätzlich werden durch die eingezeichneten horizontalen Linien die gewünschten Fehlend-Anteile hervorgehoben.

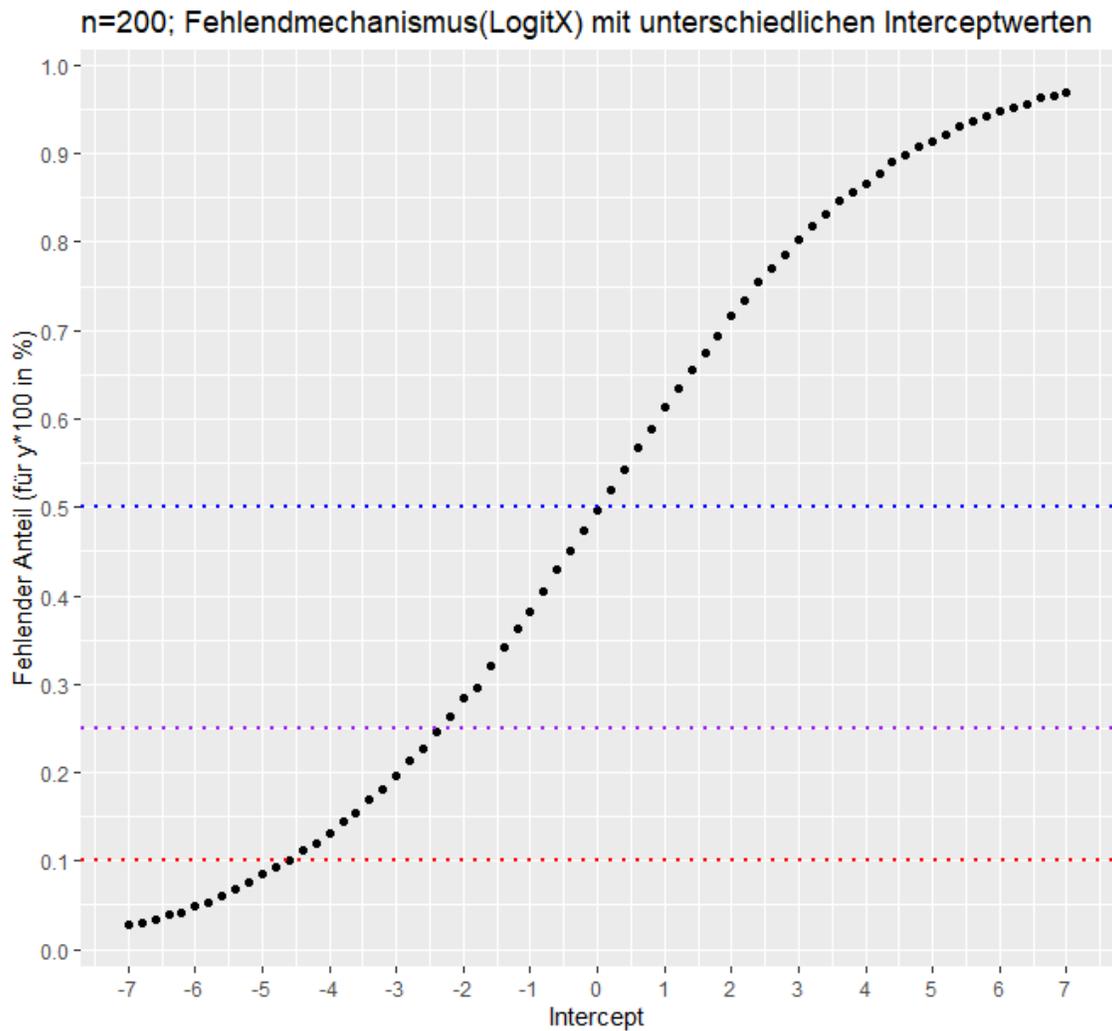


Abbildung 22: Es werden die Fehlend-Anteile, die durch LogitX mit unterschiedlichen Interceptwerten erzeugt werden, angezeigt. Zusätzlich werden durch die eingezeichneten horizontalen Linien die gewünschten Fehlend-Anteile hervorgehoben.

A.2 Datensatz mit binärer Zielgröße

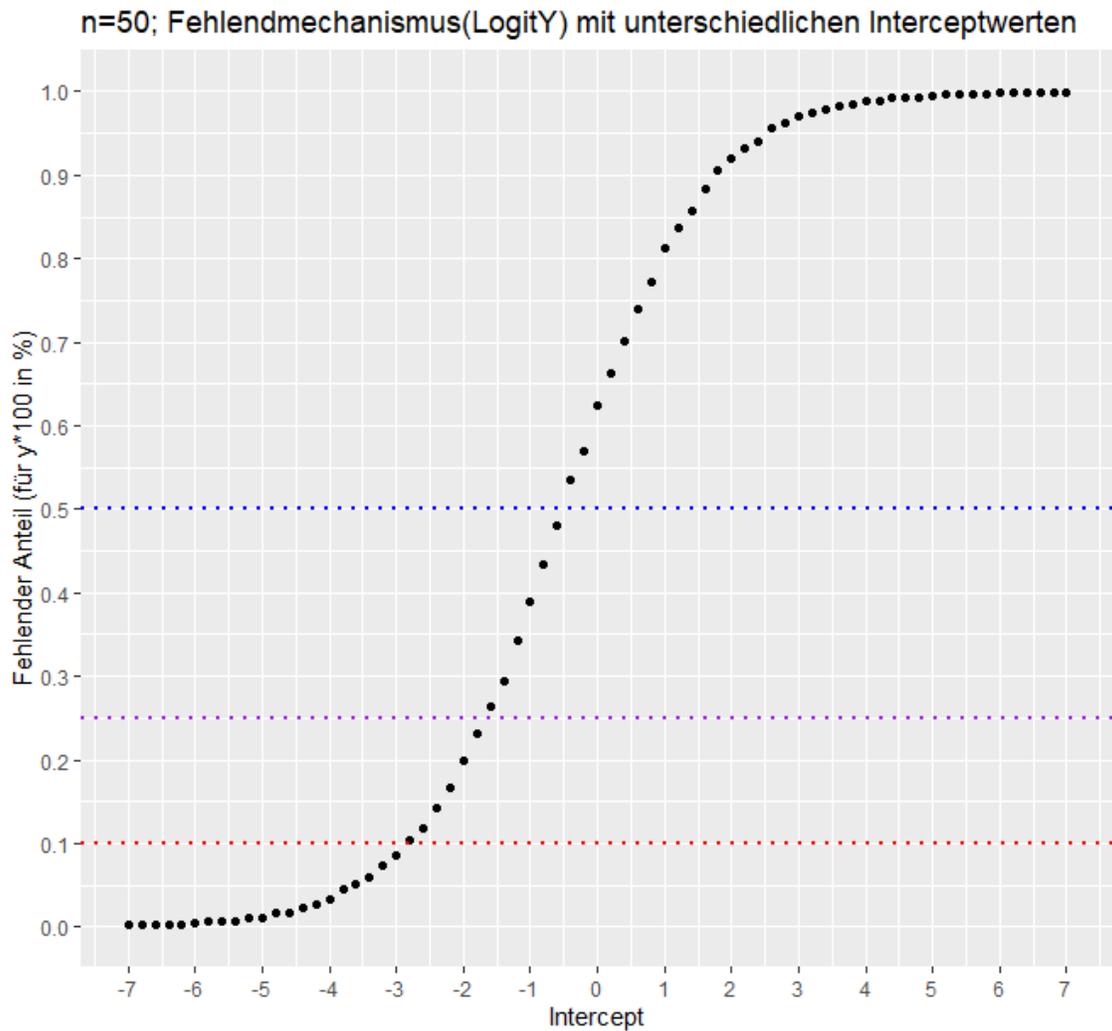


Abbildung 23: Es werden die Fehlend-Anteile, die durch LogitY mit unterschiedlichen Interceptwerten erzeugt werden, angezeigt. Zusätzlich werden durch die eingezeichneten horizontalen Linien die gewünschten Fehlend-Anteile hervorgehoben.

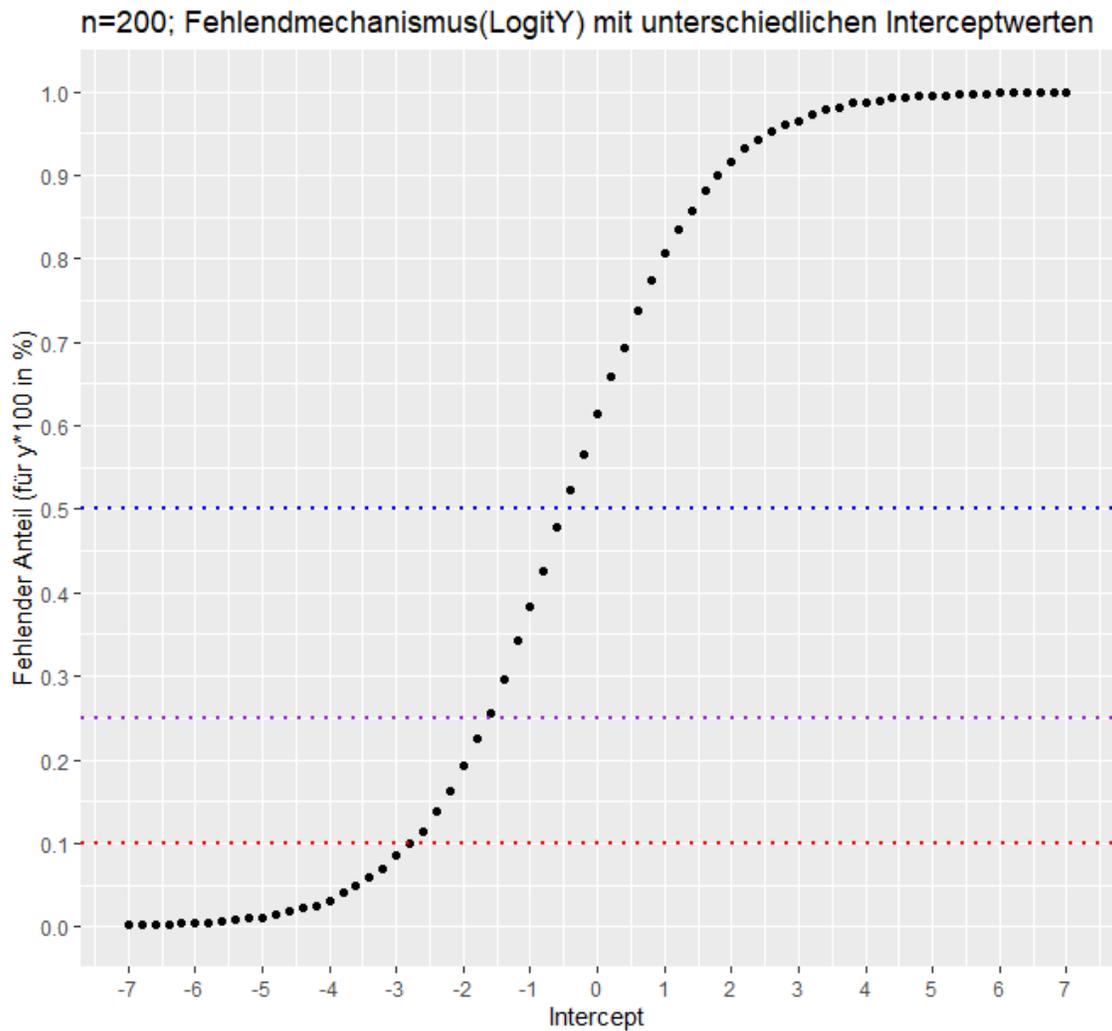


Abbildung 24: Es werden die Fehlend-Anteile, die durch LogitY mit unterschiedlichen Interceptwerten erzeugt werden, angezeigt. Zusätzlich werden durch die eingezeichneten horizontalen Linien die gewünschten Fehlend-Anteile hervorgehoben.

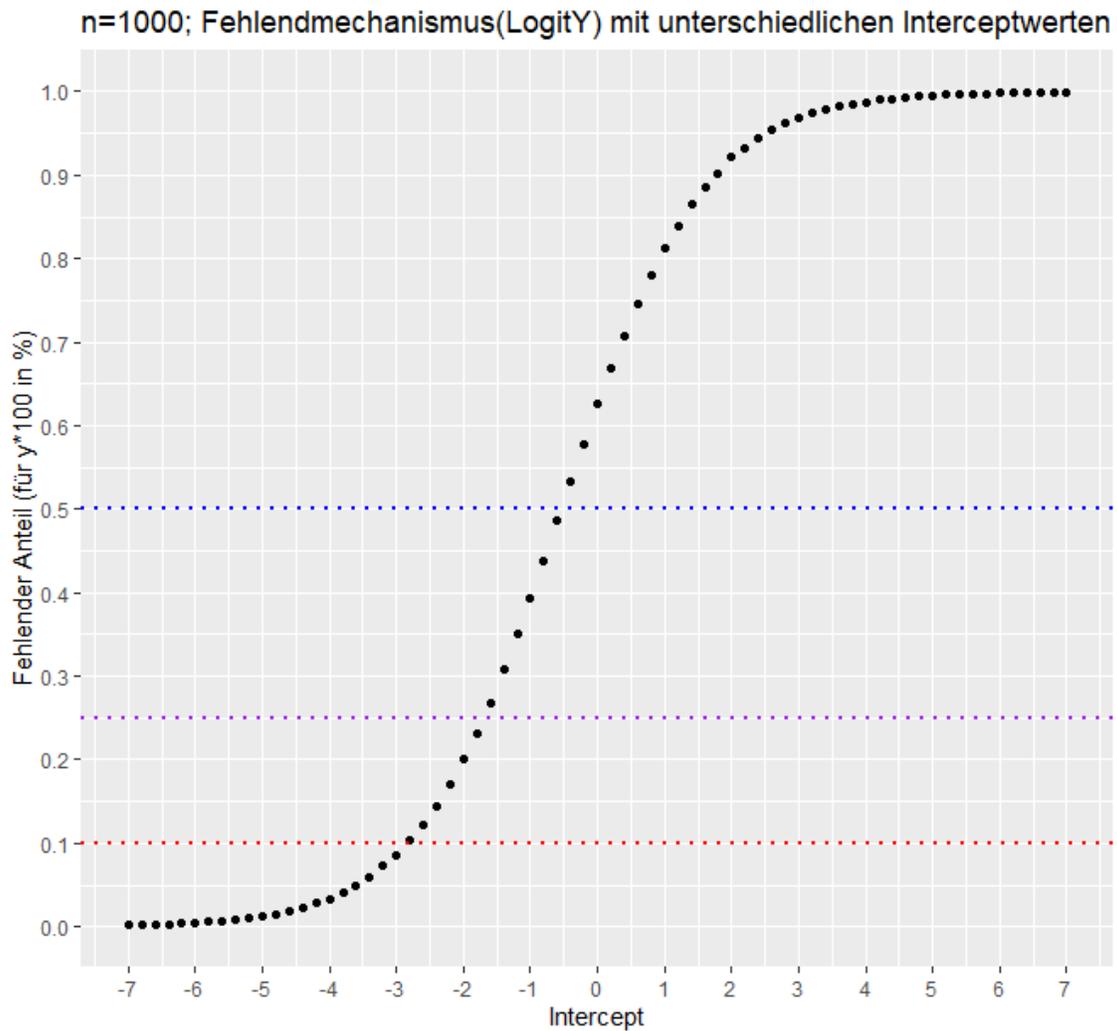


Abbildung 25: Es werden die Fehlend-Anteile, die durch LogitY mit unterschiedlichen Interceptwerten erzeugt werden, angezeigt. Zusätzlich werden durch die eingezeichneten horizontalen Linien die gewünschten Fehlend-Anteile hervorgehoben.

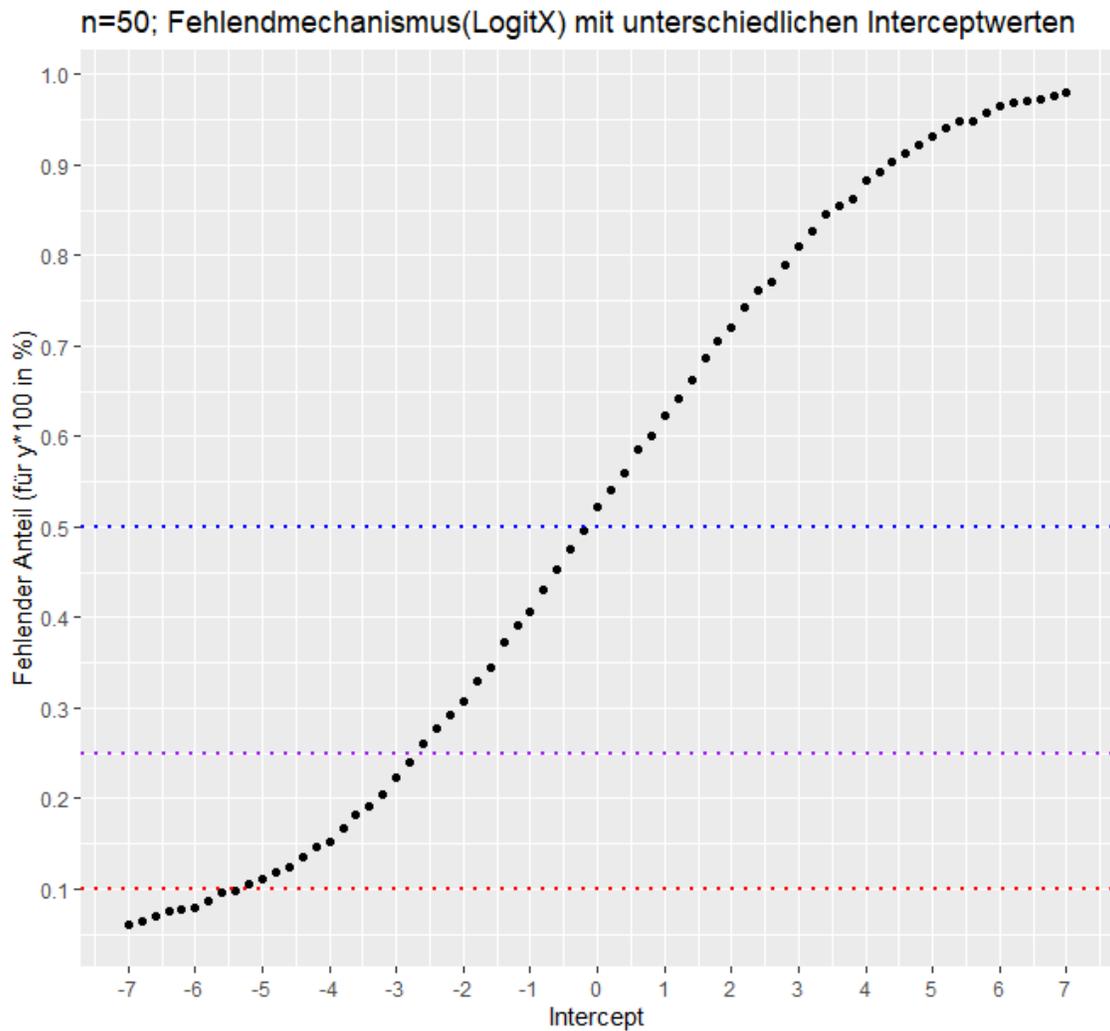


Abbildung 26: Es werden die Fehlend-Anteile, die durch LogitX mit unterschiedlichen Interceptwerten erzeugt werden, angezeigt. Zusätzlich werden durch die eingezeichneten horizontalen Linien die gewünschten Fehlend-Anteile hervorgehoben.

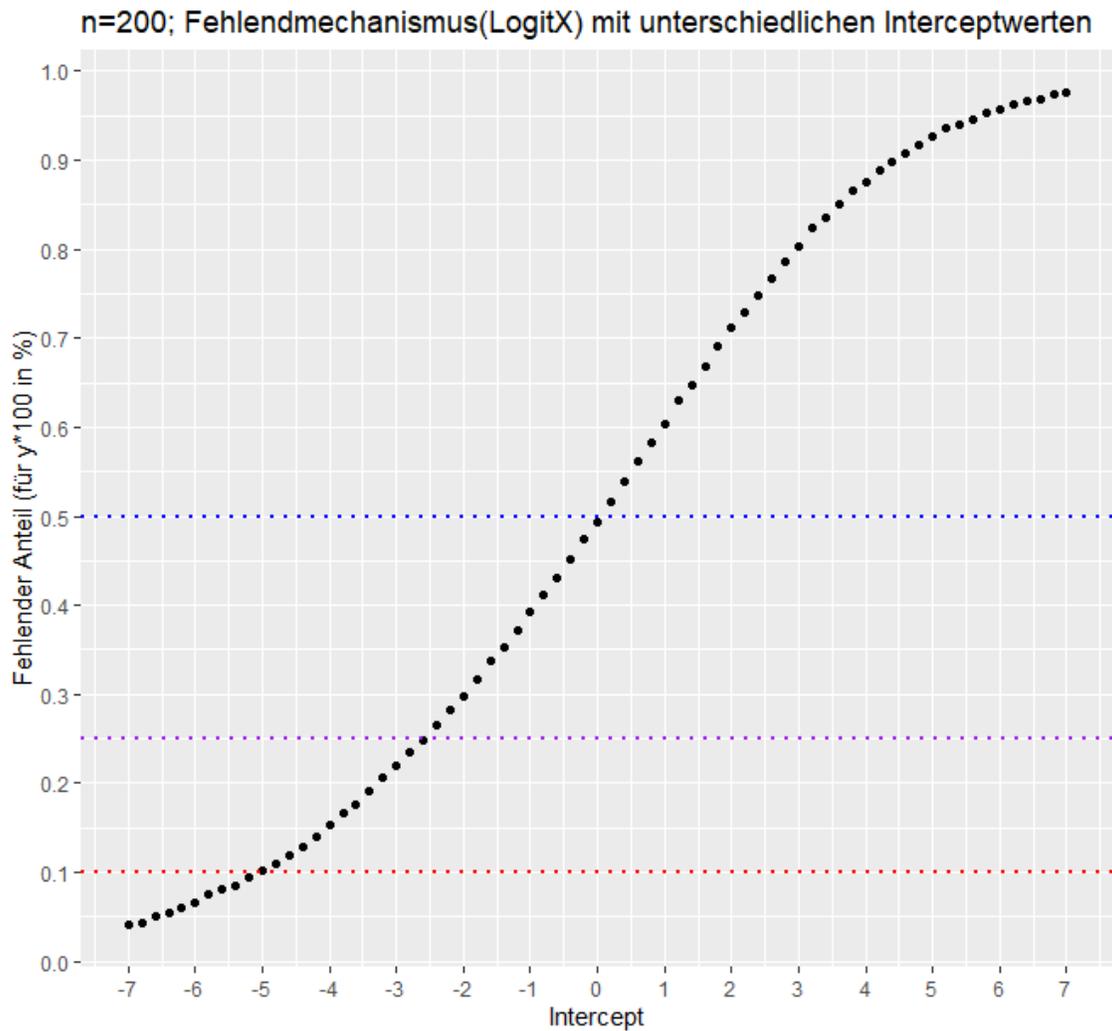


Abbildung 27: Es werden die Fehlend-Anteile, die durch LogitX mit unterschiedlichen Interceptwerten erzeugt werden, angezeigt. Zusätzlich werden durch die eingezeichneten horizontalen Linien die gewünschten Fehlend-Anteile hervorgehoben.

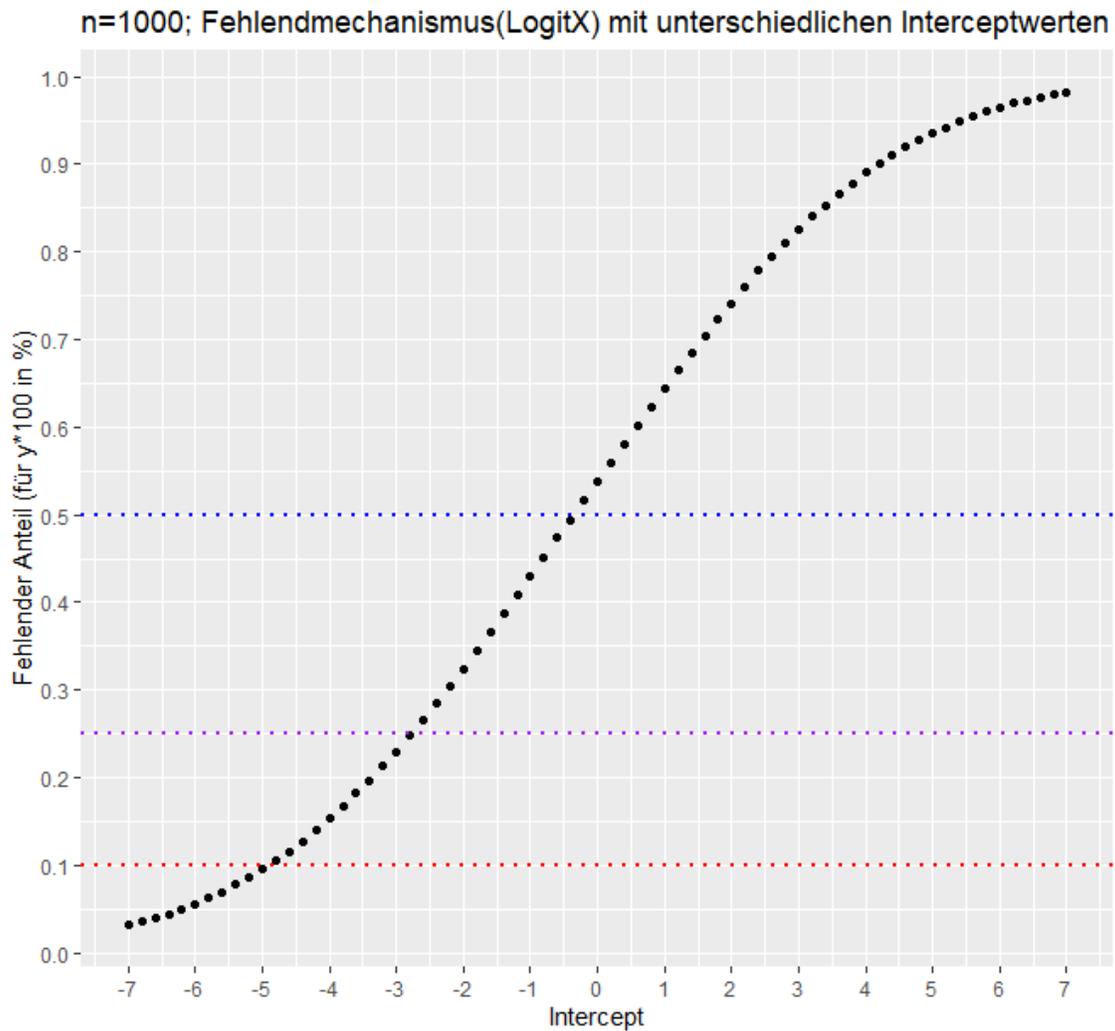


Abbildung 28: Es werden die Fehlend-Anteile, die durch LogitX mit unterschiedlichen Interceptwerten erzeugt werden, angezeigt. Zusätzlich werden durch die eingezeichneten horizontalen Linien die gewünschten Fehlend-Anteile hervorgehoben.

B Koeffizientenboxplots zu linearen Modellen

B.1 Normales Verfahren mit linearer Ridge Regression

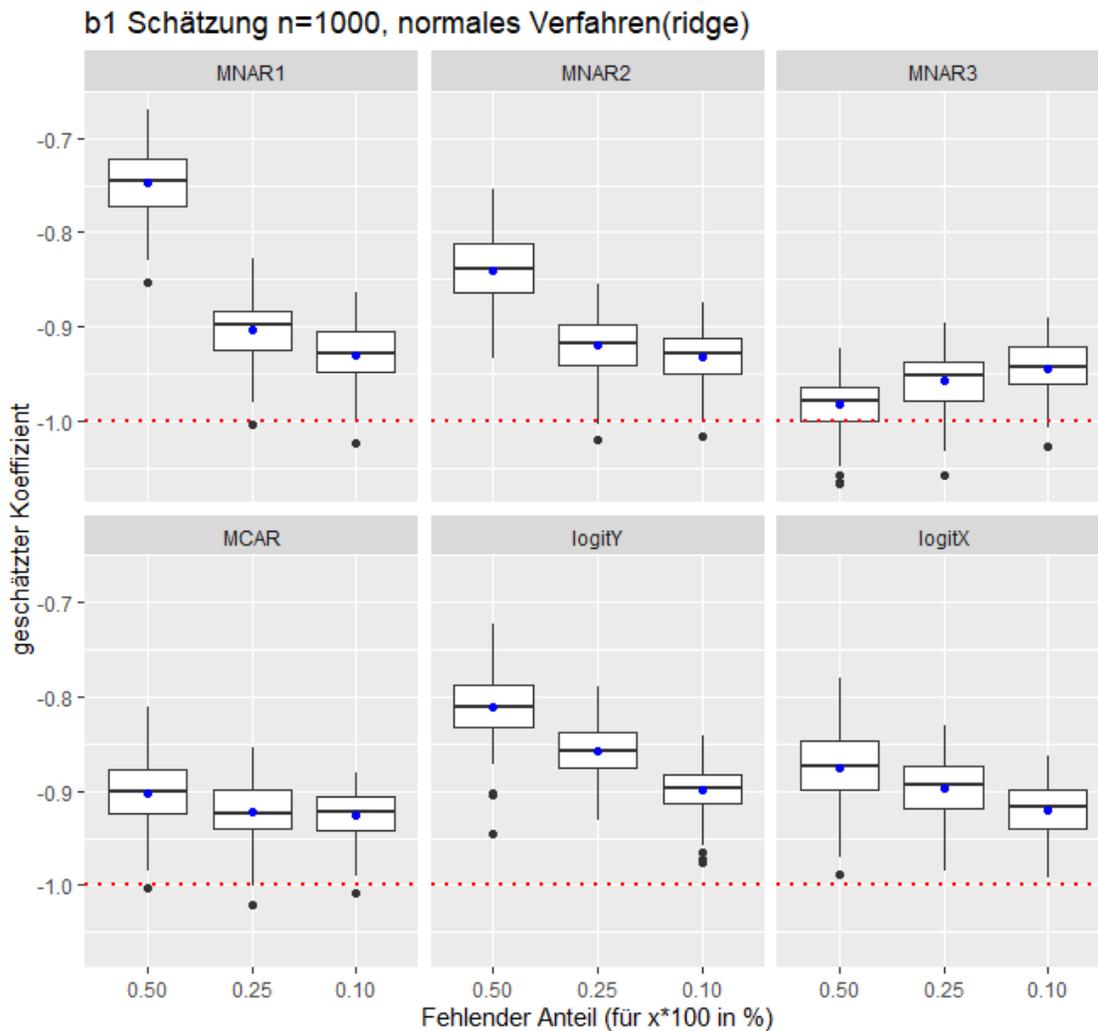


Abbildung 29: Es werden die geschätzten Werte für β_1 in Boxplotform dargestellt. Diese werden nach den Fehlendmechanismen gruppiert. Der blaue Punkt gibt den Mittelwert der geschätzten Werte an. Die rote horizontale Linie repräsentiert den echten Wert des Koeffizienten.

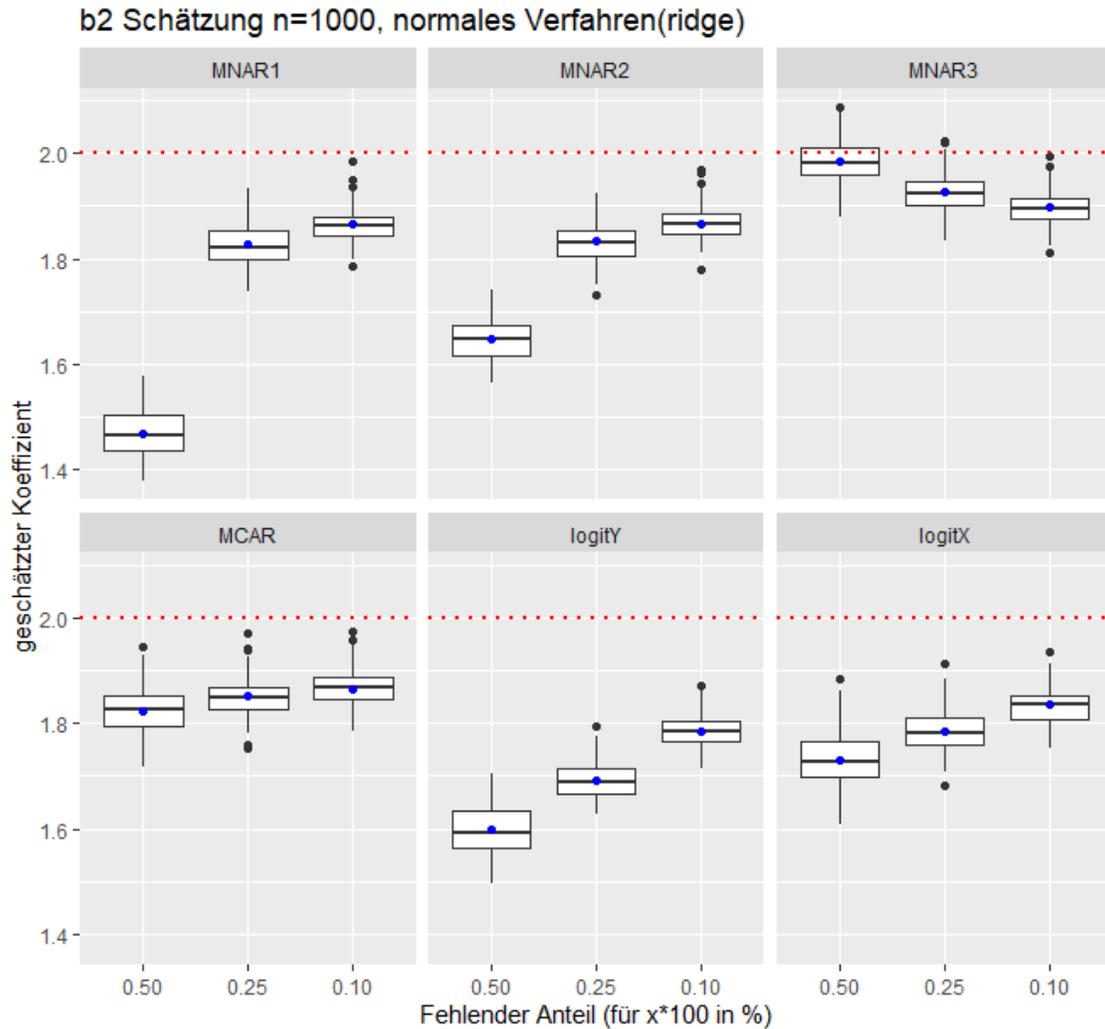


Abbildung 30: Es werden die geschätzten Werte für β_2 in Boxplotform dargestellt. Diese werden nach den Fehlendmechanismen gruppiert. Der blaue Punkt gibt den Mittelwert der geschätzten Werte an. Die rote horizontale Linie repräsentiert den echten Wert des Koeffizienten.

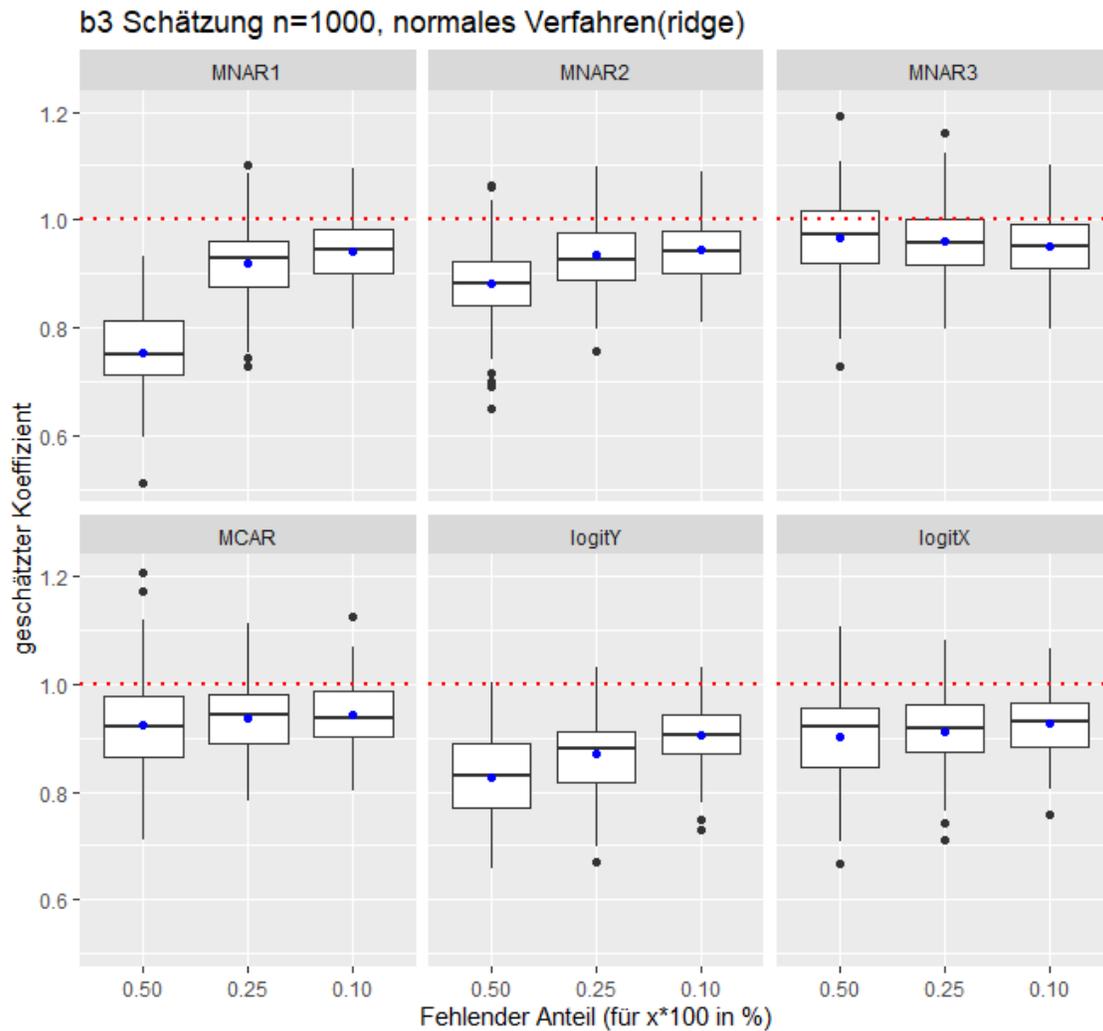


Abbildung 31: Es werden die geschätzten Werte für β_3 in Boxplotform dargestellt. Diese werden nach den Fehlendmechanismen gruppiert. Der blaue Punkt gibt den Mittelwert der geschätzten Werte an. Die rote horizontale Linie repräsentiert den echten Wert des Koeffizienten.

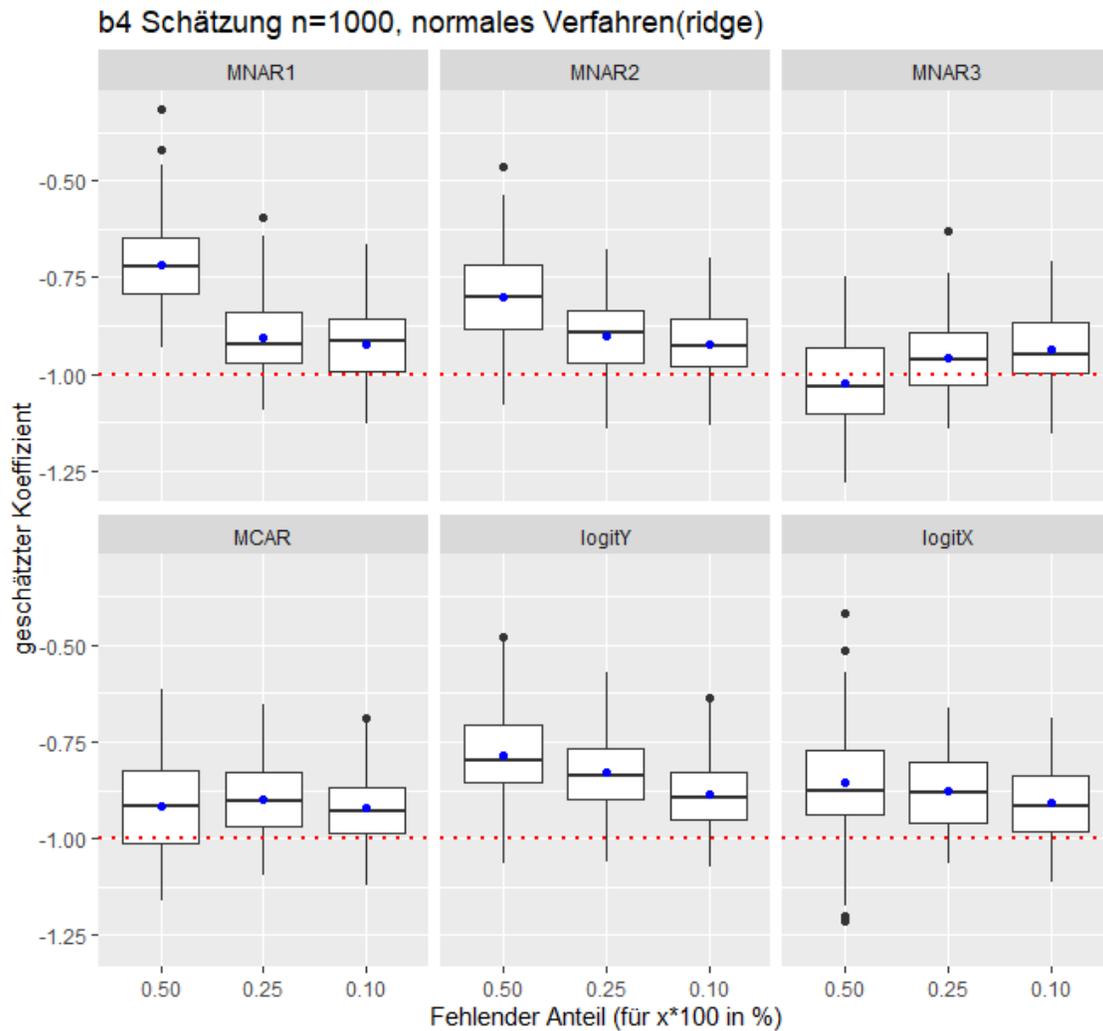


Abbildung 32: Es werden die geschätzten Werte für β_4 in Boxplotform dargestellt. Diese werden nach den Fehlendmechanismen gruppiert. Der blaue Punkt gibt den Mittelwert der geschätzten Werte an. Die rote horizontale Linie repräsentiert den echten Wert des Koeffizienten.

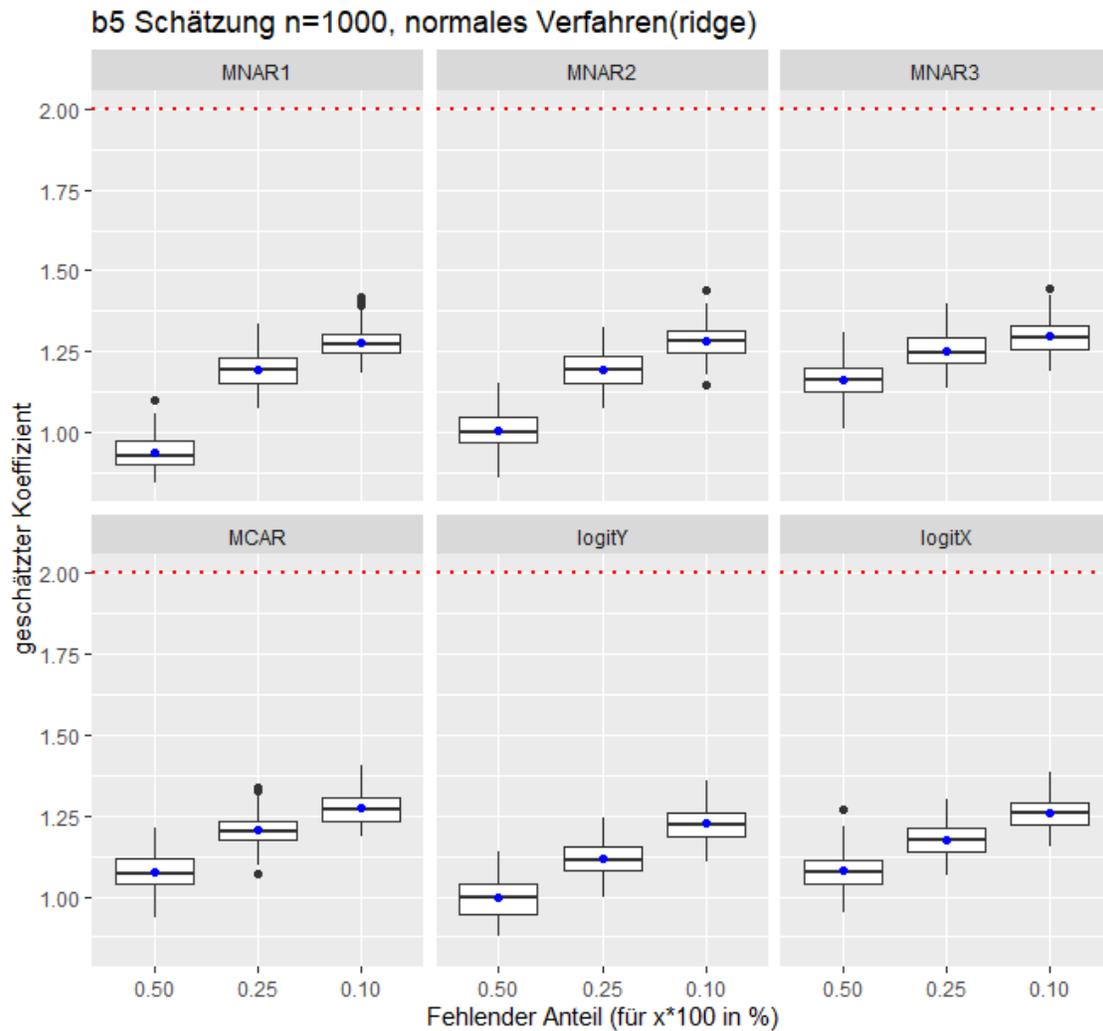


Abbildung 33: Es werden die geschätzten Werte für β_5 in Boxplotform dargestellt. Diese werden nach den Fehlendmechanismen gruppiert. Der blaue Punkt gibt den Mittelwert der geschätzten Werte an. Die rote horizontale Linie repräsentiert den echten Wert des Koeffizienten.

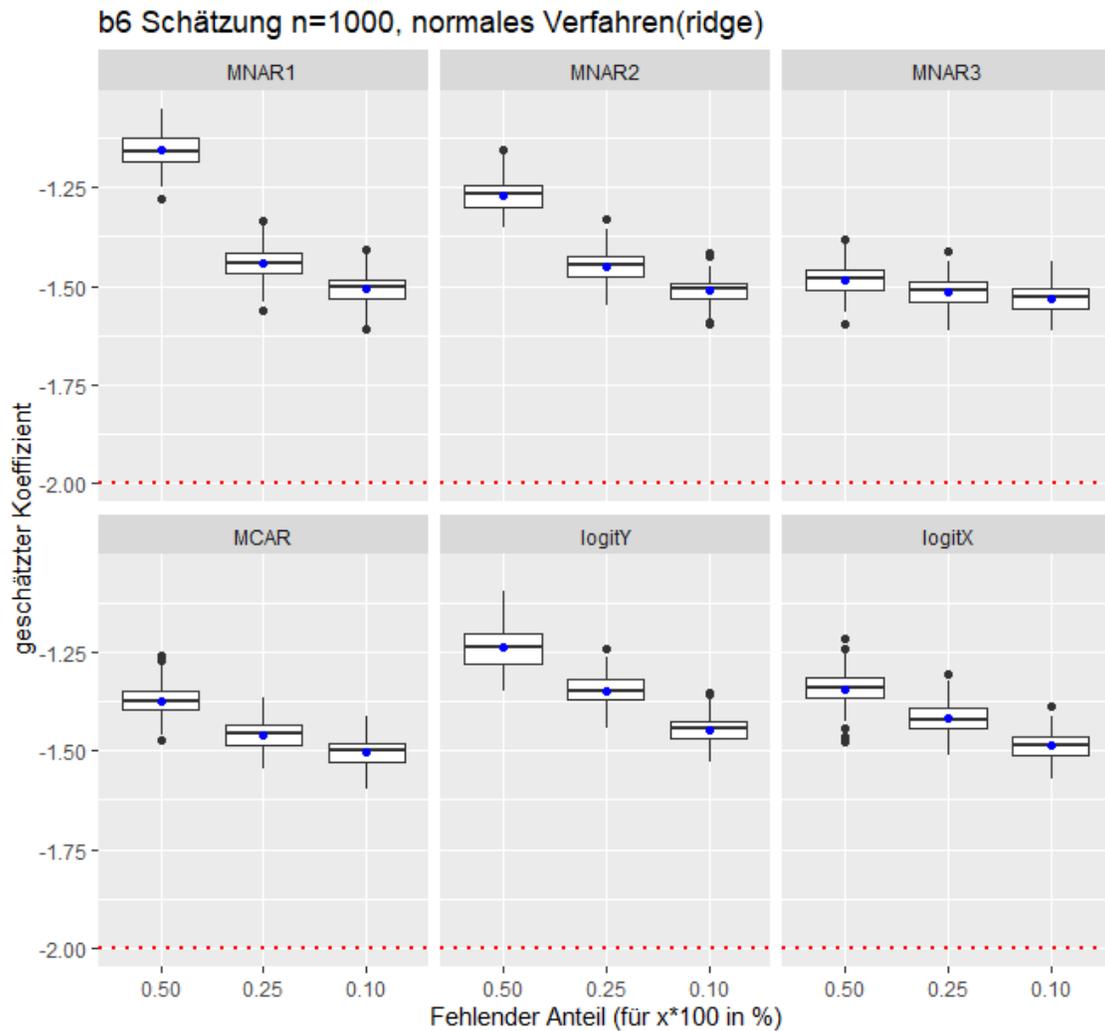


Abbildung 34: Es werden die geschätzten Werte für β_6 in Boxplotform dargestellt. Diese werden nach den Fehlendmechanismen gruppiert. Der blaue Punkt gibt den Mittelwert der geschätzten Werte an. Die rote horizontale Linie repräsentiert den echten Wert des Koeffizienten.

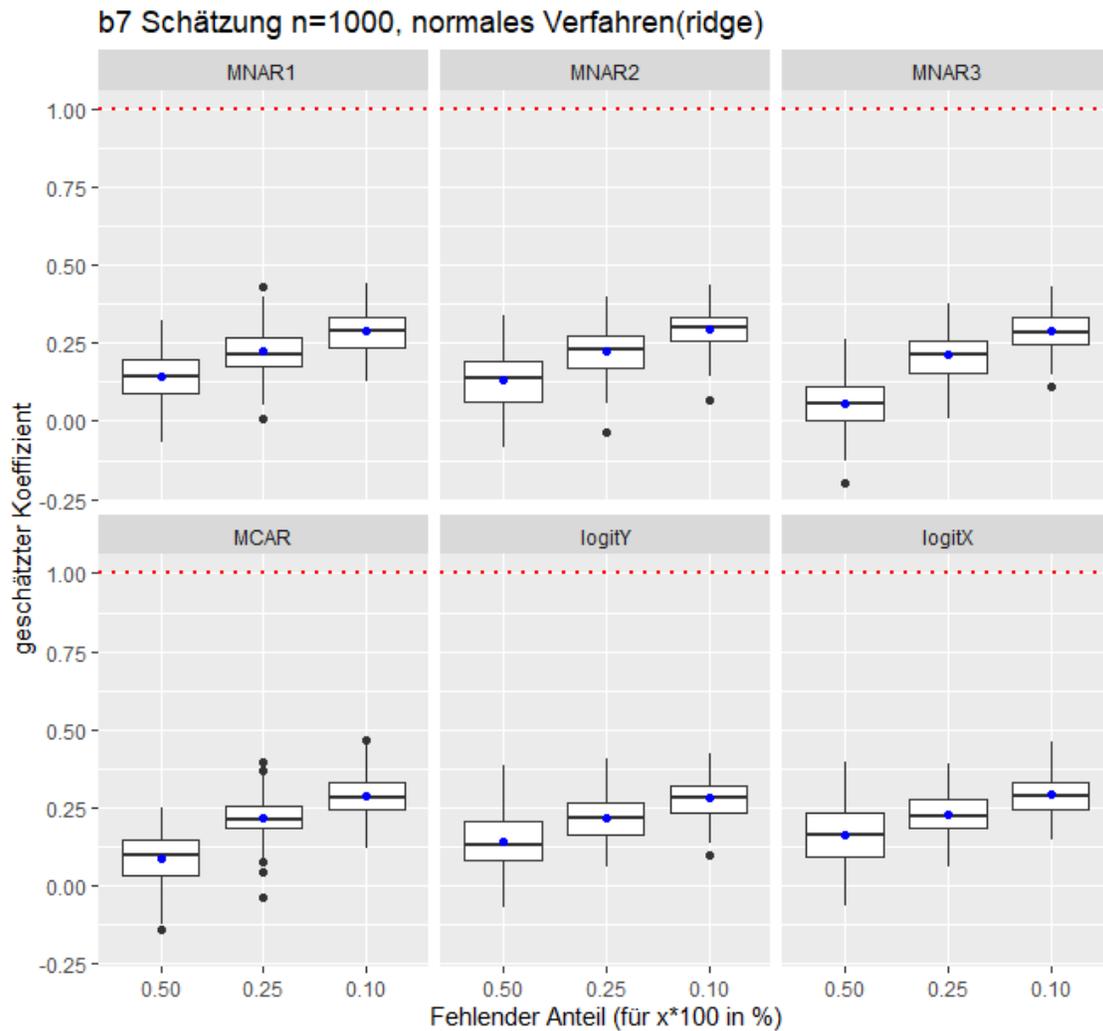


Abbildung 35: Es werden die geschätzten Werte für β_7 in Boxplotform dargestellt. Diese werden nach den Fehlendmechanismen gruppiert. Der blaue Punkt gibt den Mittelwert der geschätzten Werte an. Die rote horizontale Linie repräsentiert den echten Wert des Koeffizienten.

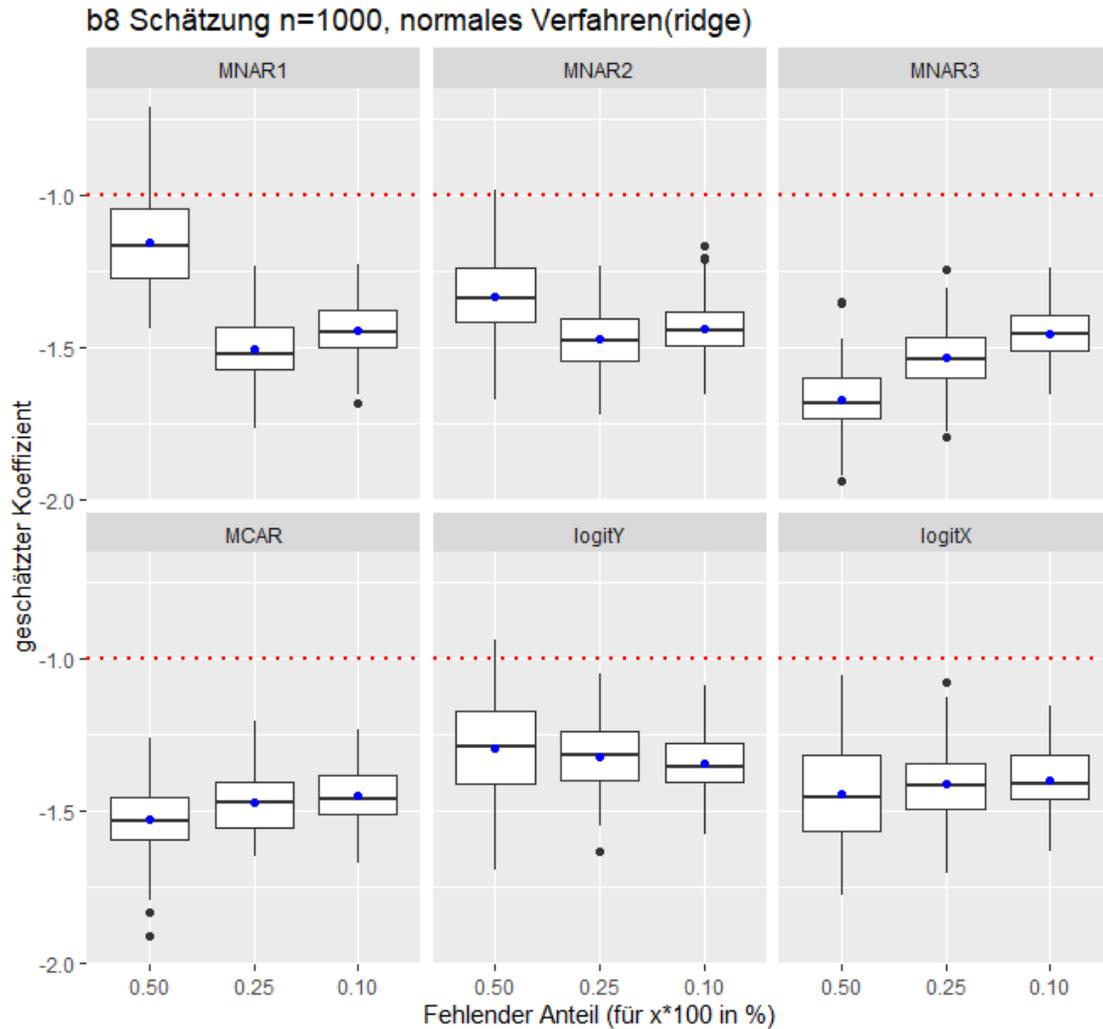


Abbildung 36: Es werden die geschätzten Werte für β_8 in Boxplotform dargestellt. Diese werden nach den Fehlendmechanismen gruppiert. Der blaue Punkt gibt den Mittelwert der geschätzten Werte an. Die rote horizontale Linie repräsentiert den echten Wert des Koeffizienten.

B.2 Bruteforce1-Verfahren mit linearer Regression

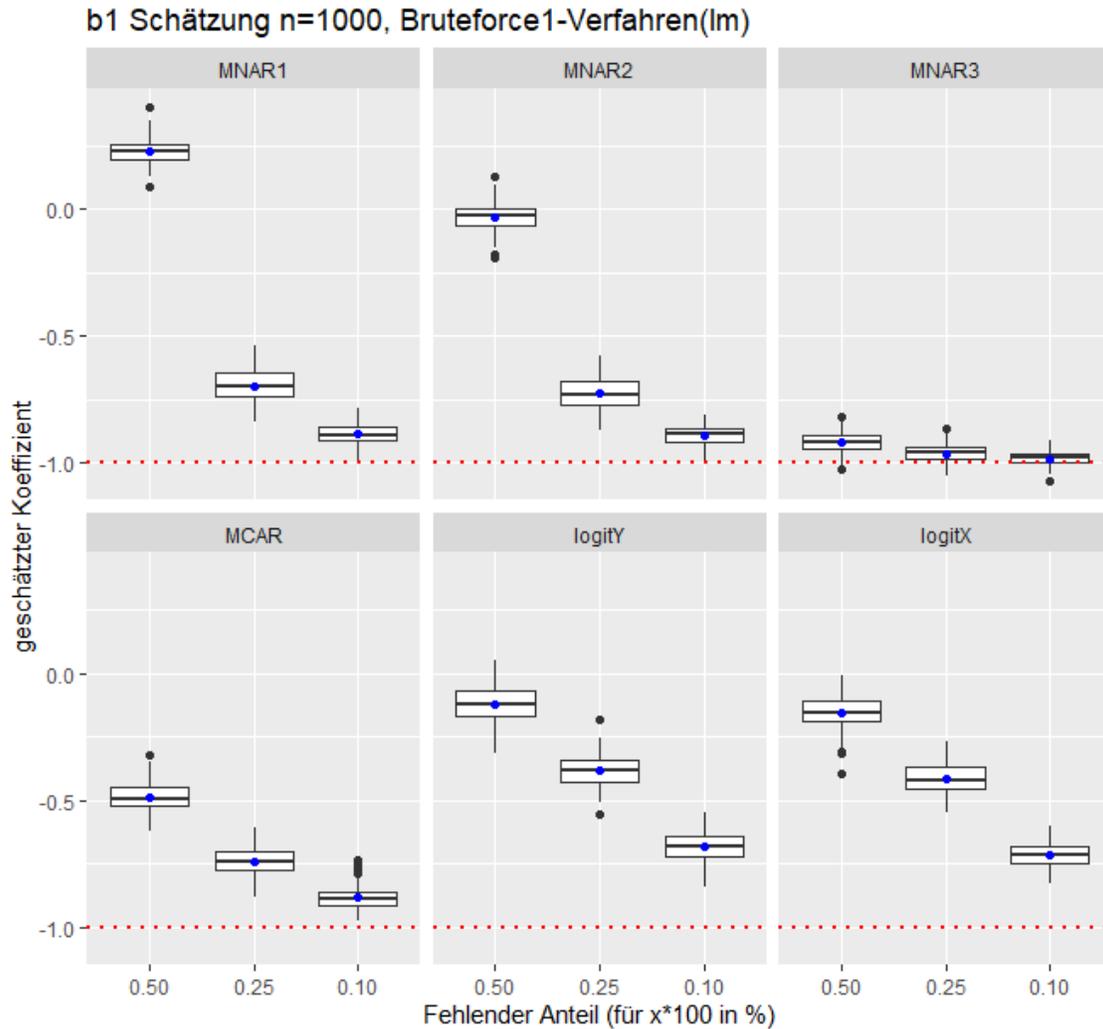


Abbildung 37: Es werden die geschätzten Werte für β_1 in Boxplotform dargestellt. Diese werden nach den Fehlendmechanismen gruppiert. Der blaue Punkt gibt den Mittelwert der geschätzten Werte an. Die rote horizontale Linie repräsentiert den echten Wert des Koeffizienten.

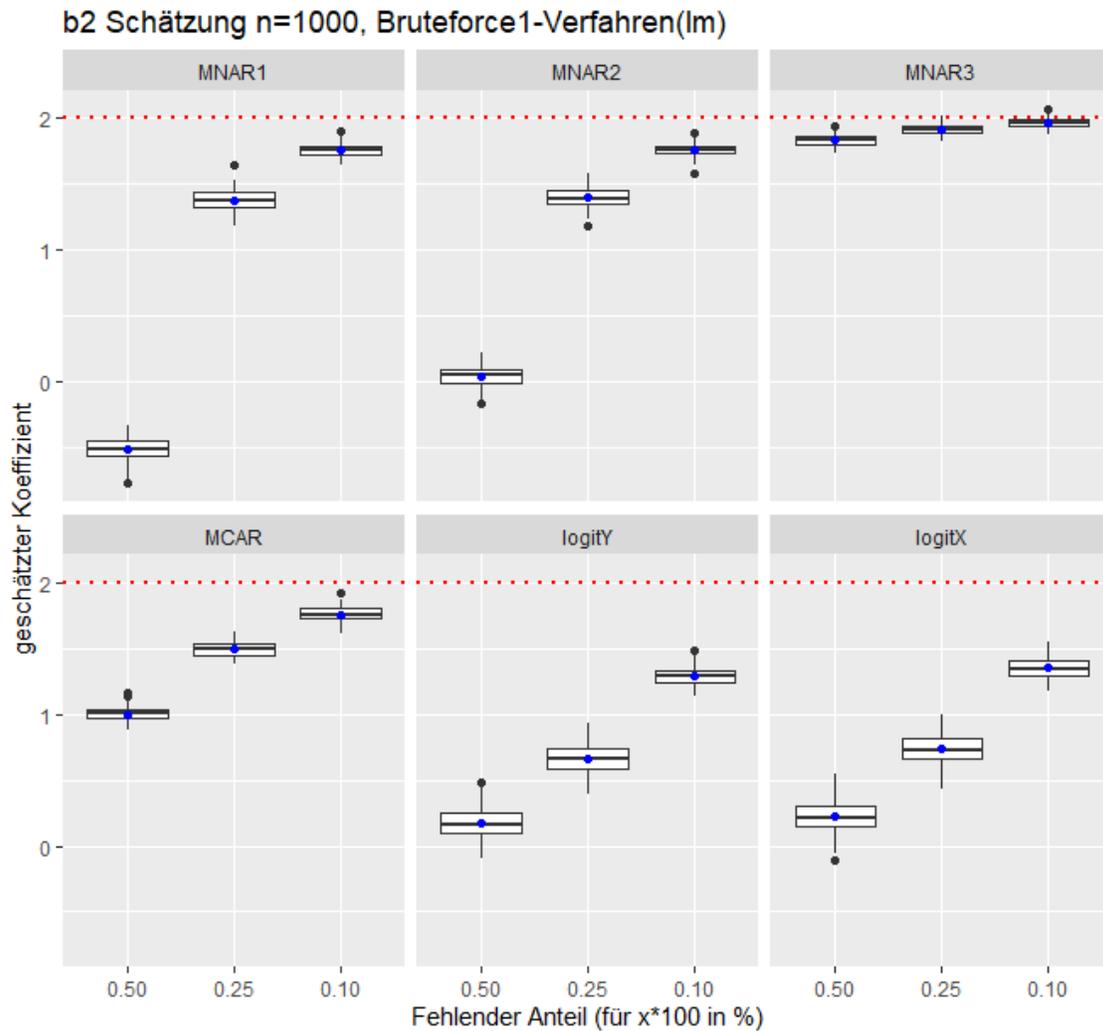


Abbildung 38: Es werden die geschätzten Werte für β_2 in Boxplotform dargestellt. Diese werden nach den Fehlendmechanismen gruppiert. Der blaue Punkt gibt den Mittelwert der geschätzten Werte an. Die rote horizontale Linie repräsentiert den echten Wert des Koeffizienten.

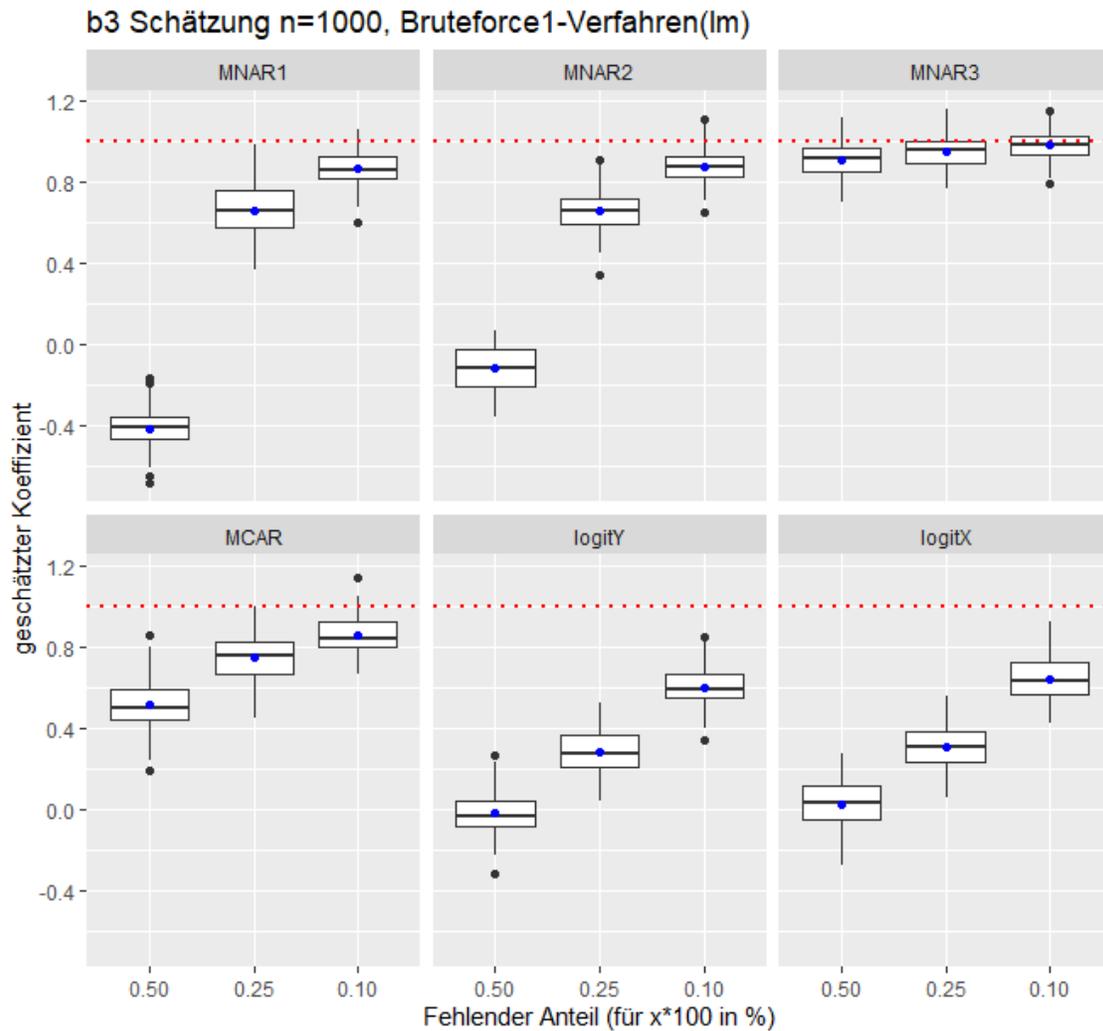


Abbildung 39: Es werden die geschätzten Werte für β_3 in Boxplotform dargestellt. Diese werden nach den Fehlendmechanismen gruppiert. Der blaue Punkt gibt den Mittelwert der geschätzten Werte an. Die rote horizontale Linie repräsentiert den echten Wert des Koeffizienten.

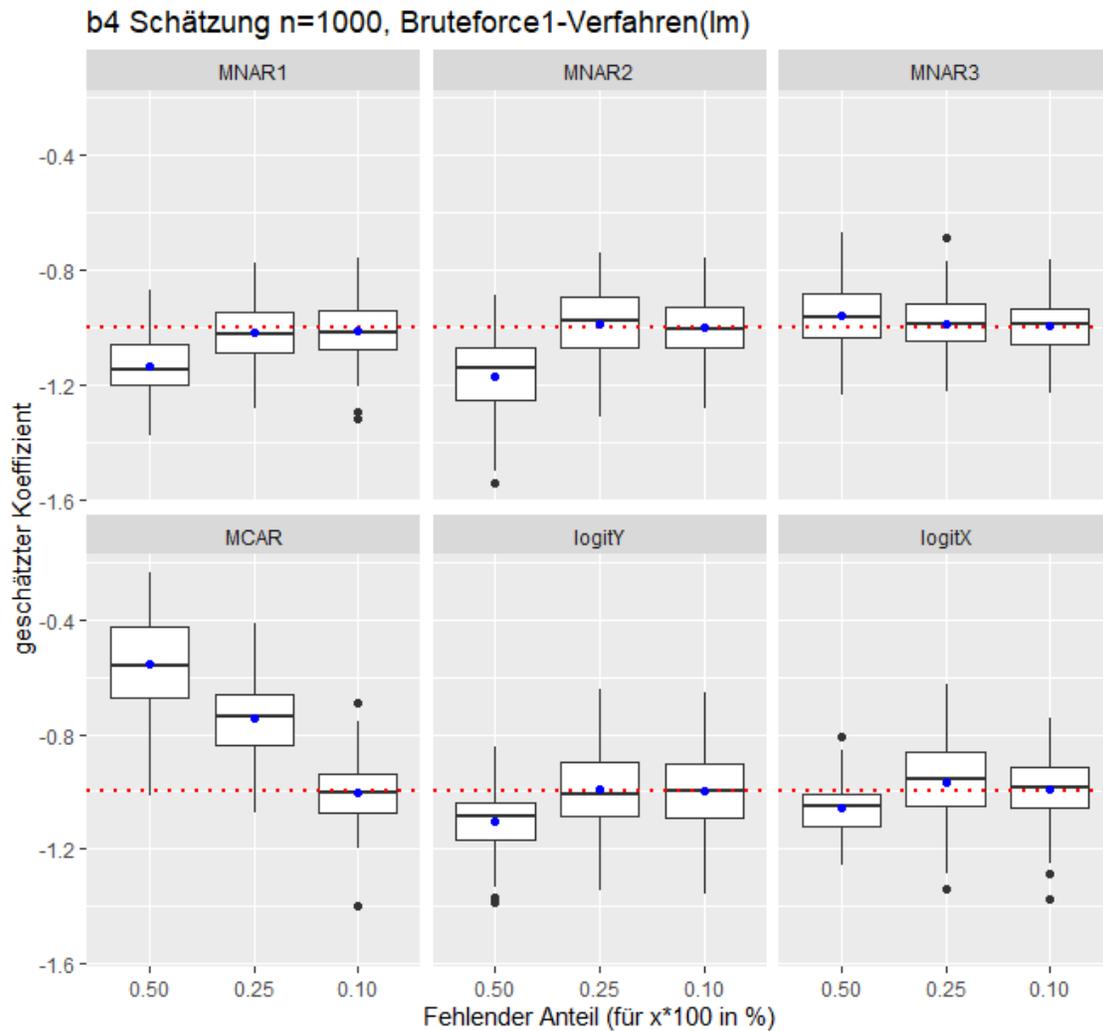


Abbildung 40: Es werden die geschätzten Werte für β_4 in Boxplotform dargestellt. Diese werden nach den Fehlendmechanismen gruppiert. Der blaue Punkt gibt den Mittelwert der geschätzten Werte an. Die rote horizontale Linie repräsentiert den echten Wert des Koeffizienten.

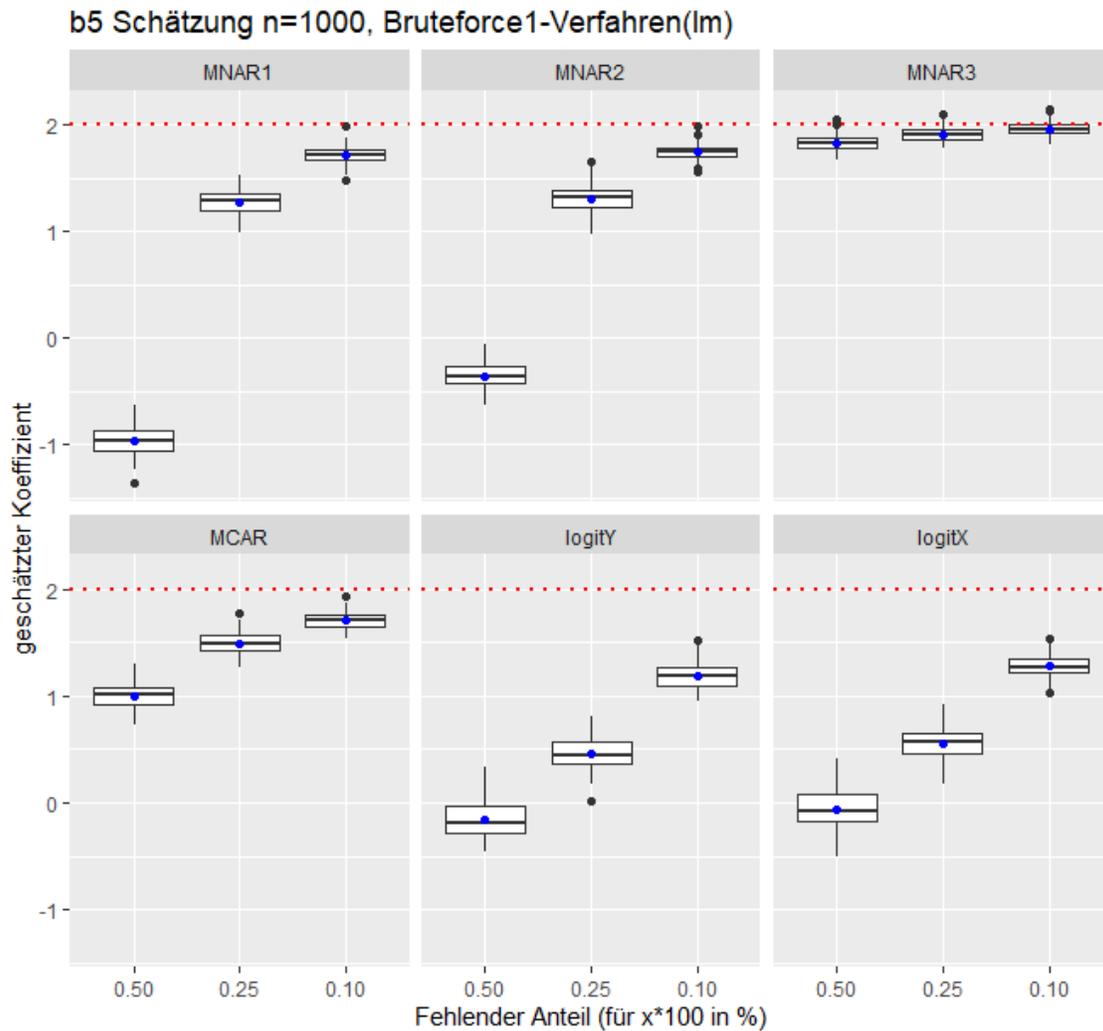


Abbildung 41: Es werden die geschätzten Werte für β_5 in Boxplotform dargestellt. Diese werden nach den Fehlendmechanismen gruppiert. Der blaue Punkt gibt den Mittelwert der geschätzten Werte an. Die rote horizontale Linie repräsentiert den echten Wert des Koeffizienten.

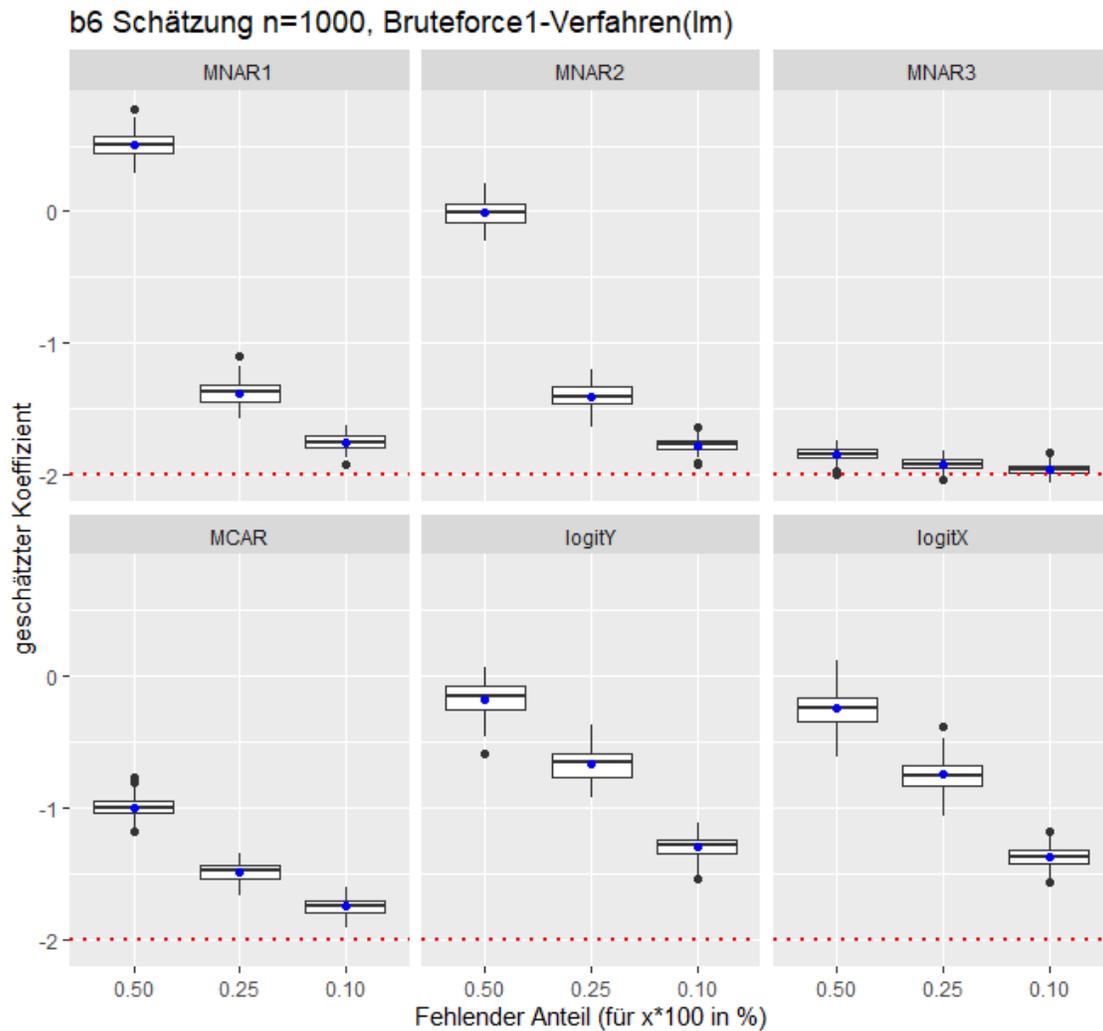


Abbildung 42: Es werden die geschätzten Werte für β_6 in Boxplotform dargestellt. Diese werden nach den Fehlendmechanismen gruppiert. Der blaue Punkt gibt den Mittelwert der geschätzten Werte an. Die rote horizontale Linie repräsentiert den echten Wert des Koeffizienten.

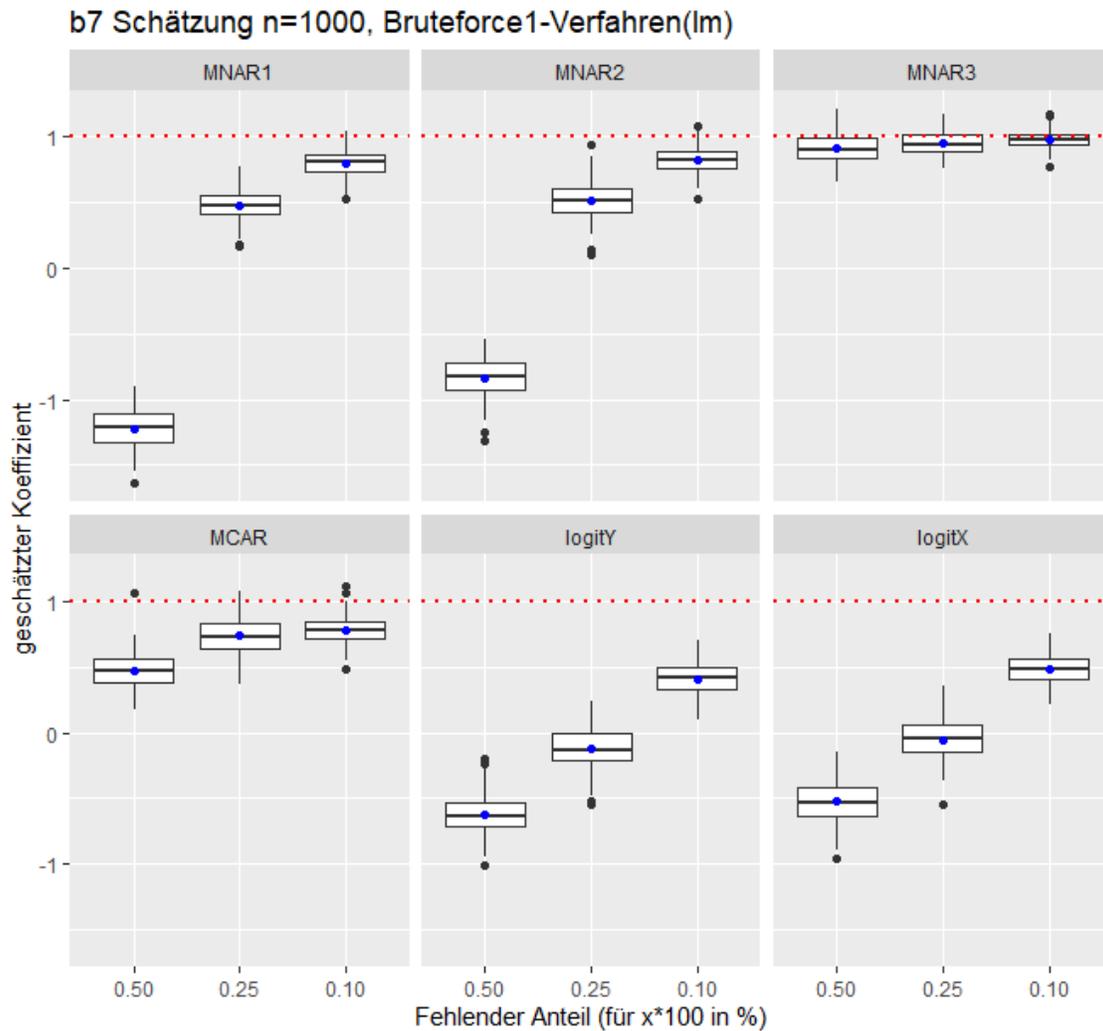


Abbildung 43: Es werden die geschätzten Werte für β_7 in Boxplotform dargestellt. Diese werden nach den Fehlendmechanismen gruppiert. Der blaue Punkt gibt den Mittelwert der geschätzten Werte an. Die rote horizontale Linie repräsentiert den echten Wert des Koeffizienten.

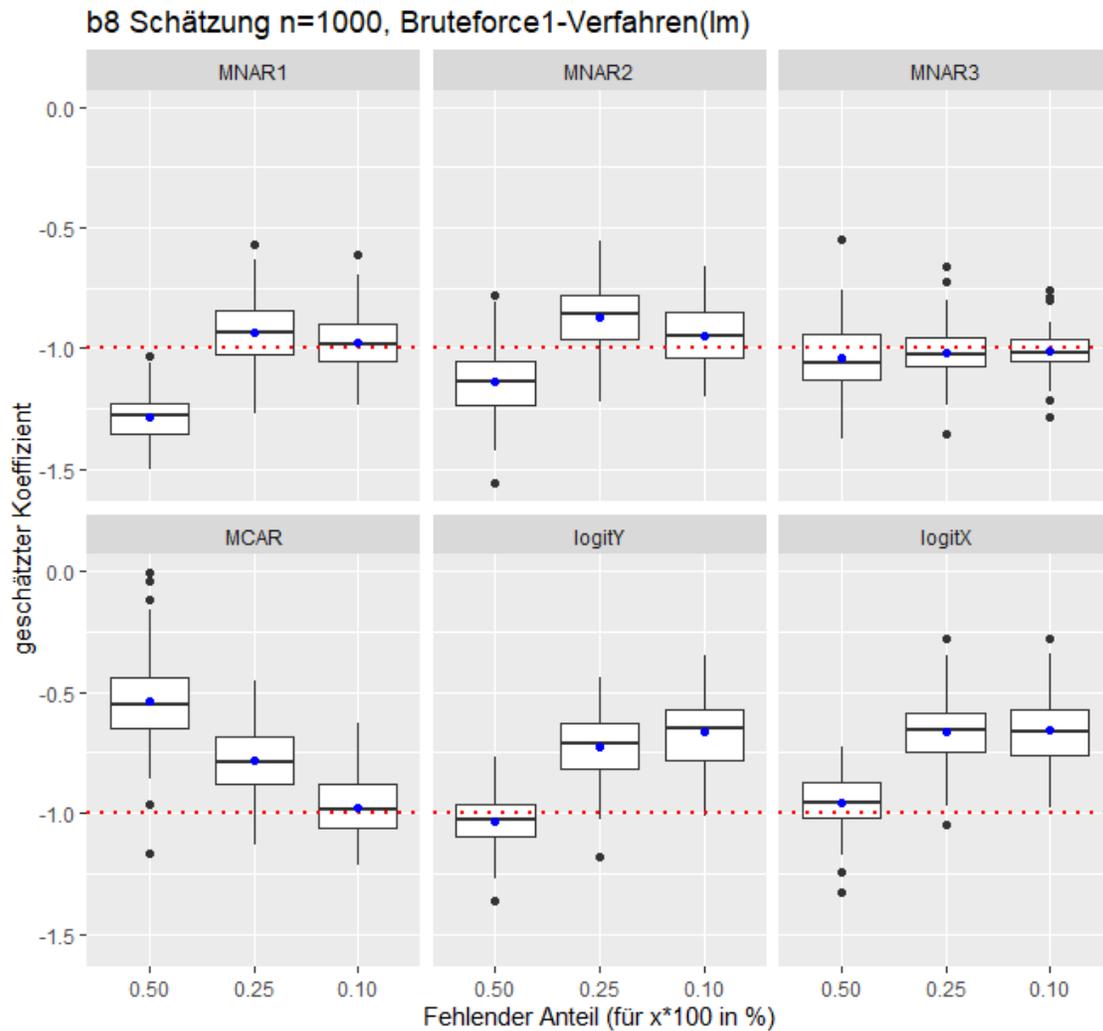


Abbildung 44: Es werden die geschätzten Werte für β_8 in Boxplotform dargestellt. Diese werden nach den Fehlendmechanismen gruppiert. Der blaue Punkt gibt den Mittelwert der geschätzten Werte an. Die rote horizontale Linie repräsentiert den echten Wert des Koeffizienten.

B.3 Bruteforce2-Verfahren mit linearer Regression

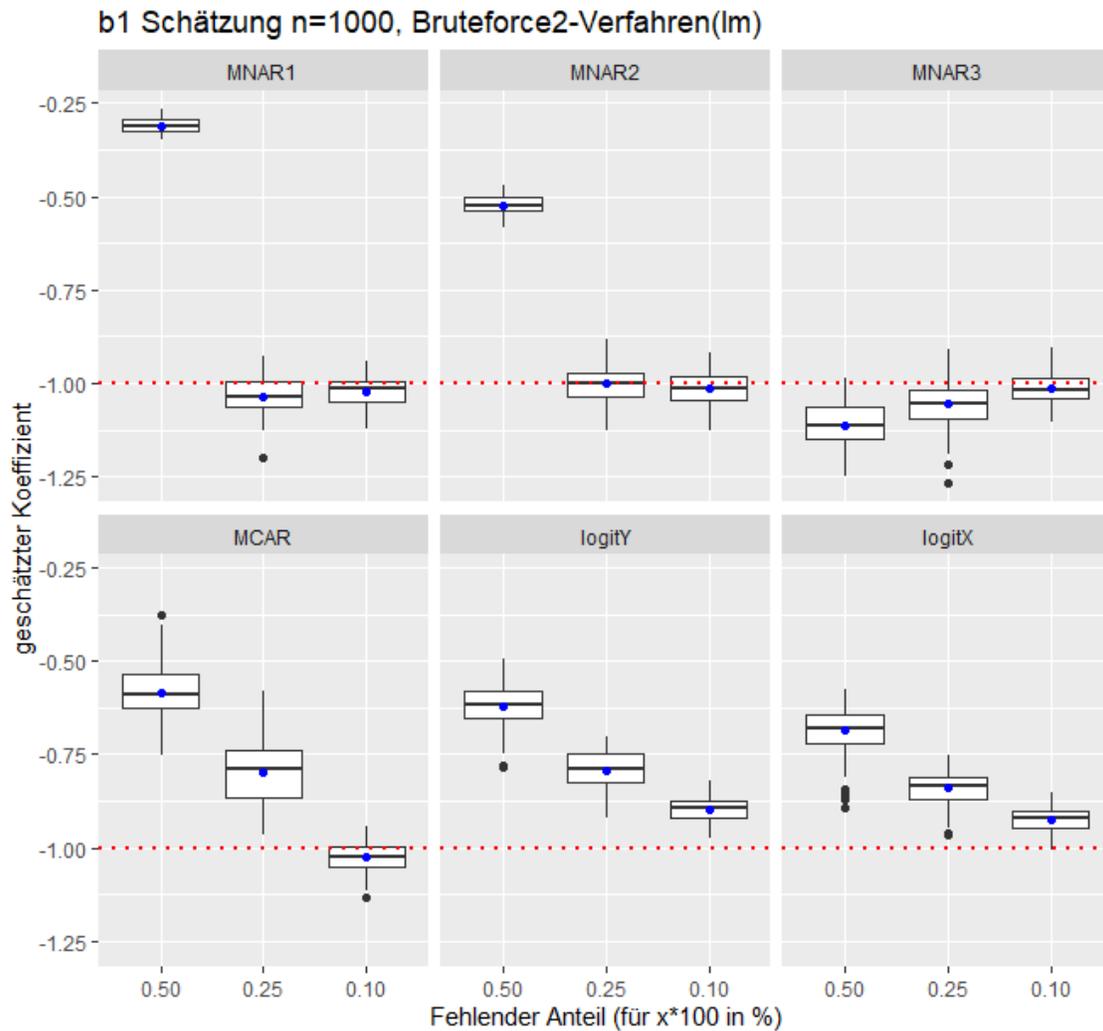


Abbildung 45: Es werden die geschätzten Werte für β_1 in Boxplotform dargestellt. Diese werden nach den Fehlendmechanismen gruppiert. Der blaue Punkt gibt den Mittelwert der geschätzten Werte an. Die rote horizontale Linie repräsentiert den echten Wert des Koeffizienten.

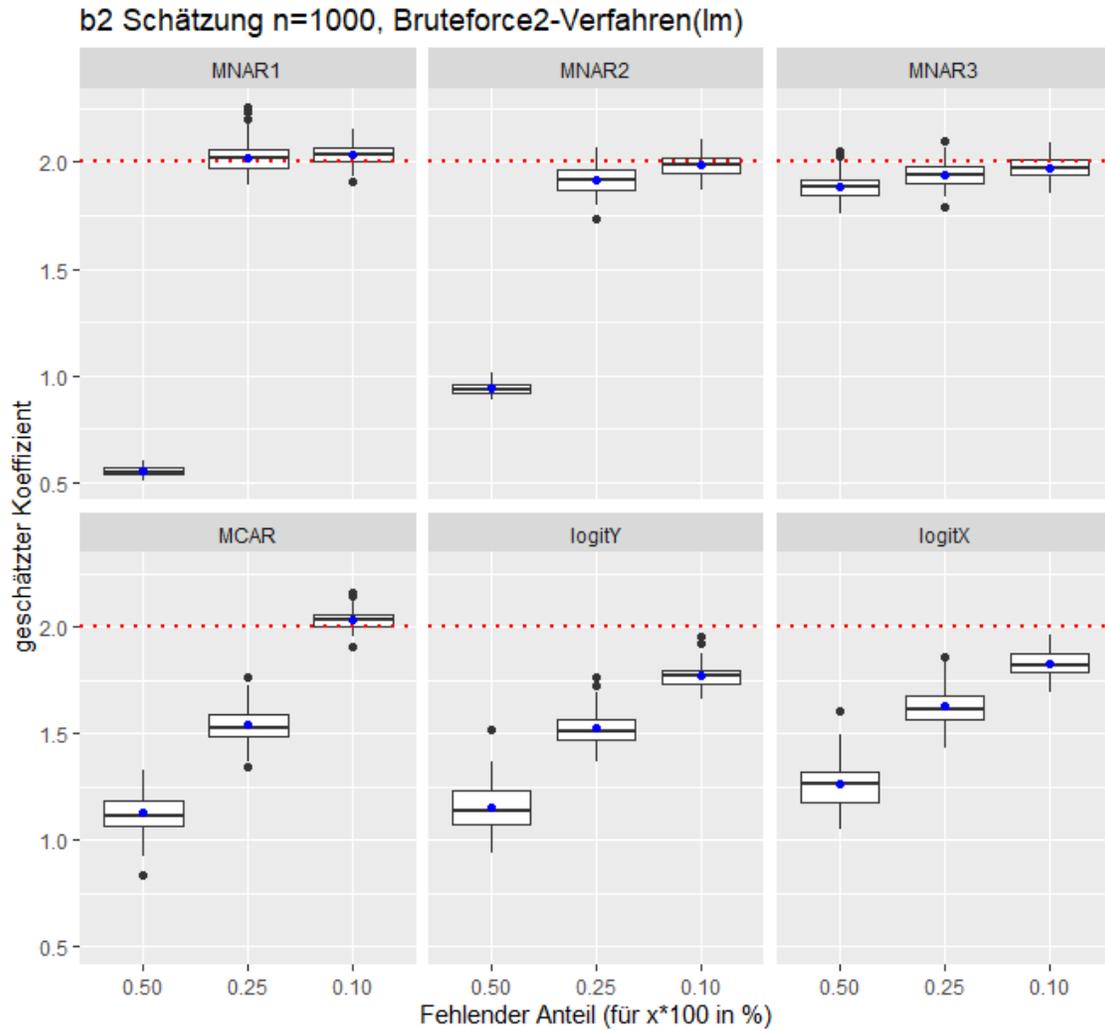


Abbildung 46: Es werden die geschätzten Werte für β_2 in Boxplotform dargestellt. Diese werden nach den Fehlendmechanismen gruppiert. Der blaue Punkt gibt den Mittelwert der geschätzten Werte an. Die rote horizontale Linie repräsentiert den echten Wert des Koeffizienten.

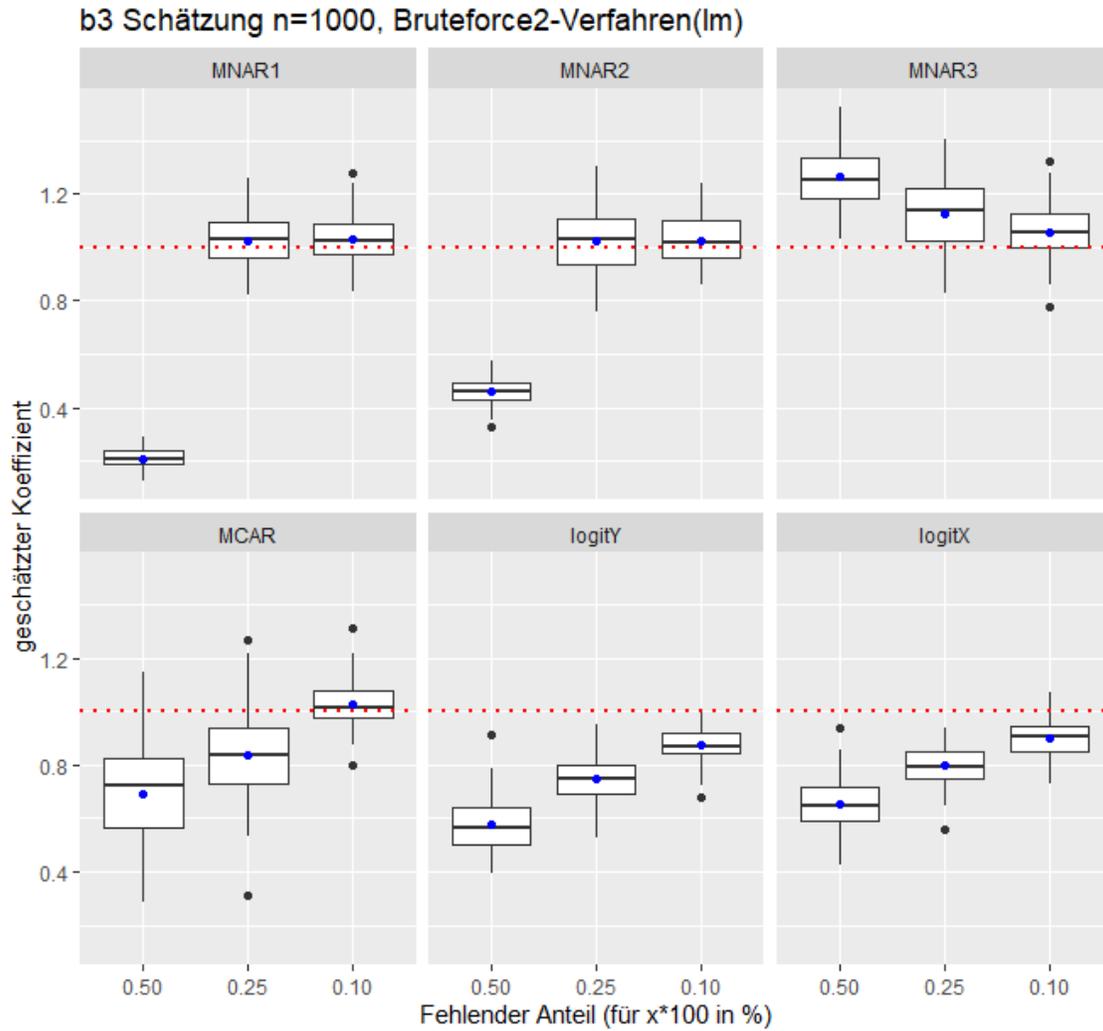


Abbildung 47: Es werden die geschätzten Werte für β_3 in Boxplotform dargestellt. Diese werden nach den Fehlendmechanismen gruppiert. Der blaue Punkt gibt den Mittelwert der geschätzten Werte an. Die rote horizontale Linie repräsentiert den echten Wert des Koeffizienten.

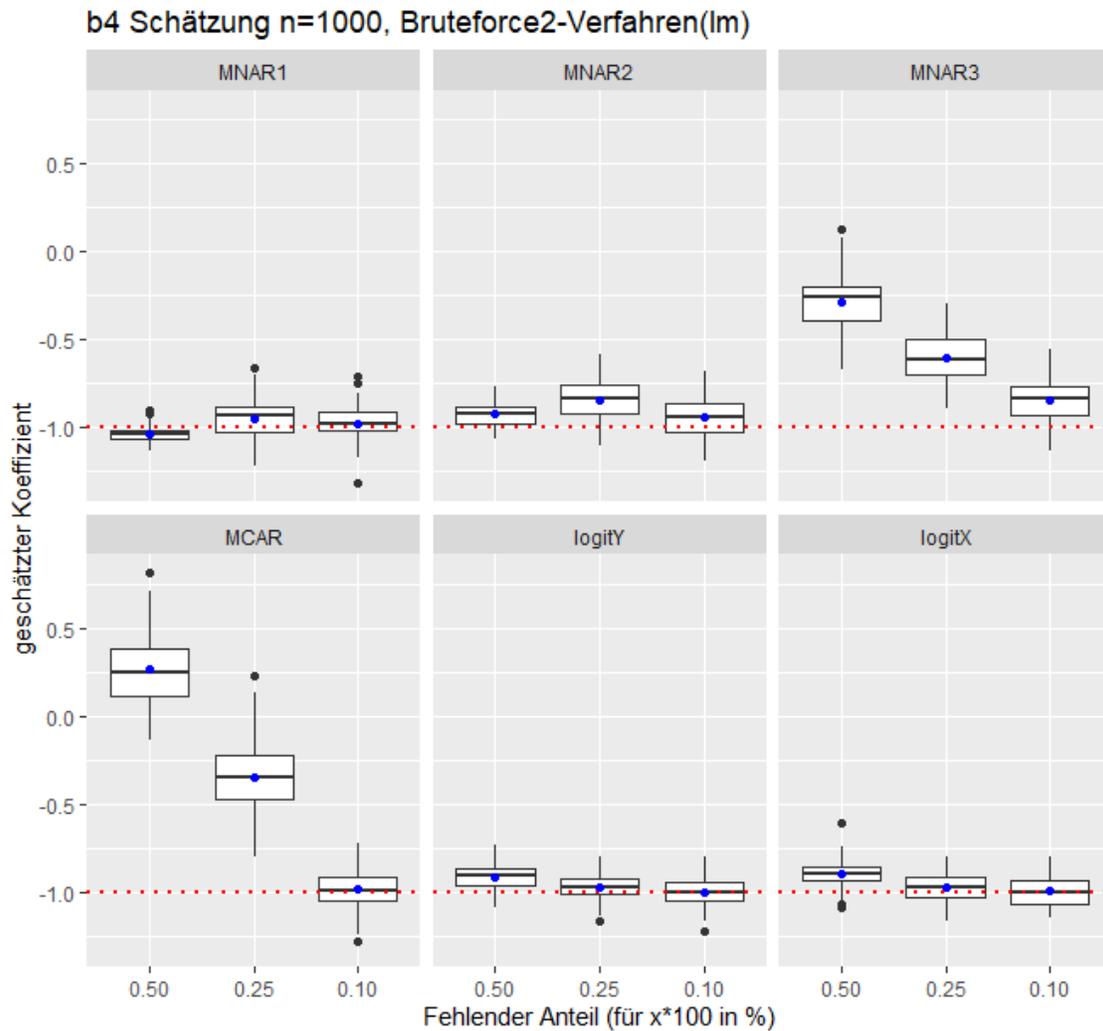


Abbildung 48: Es werden die geschätzten Werte für β_4 in Boxplotform dargestellt. Diese werden nach den Fehlendmechanismen gruppiert. Der blaue Punkt gibt den Mittelwert der geschätzten Werte an. Die rote horizontale Linie repräsentiert den echten Wert des Koeffizienten.

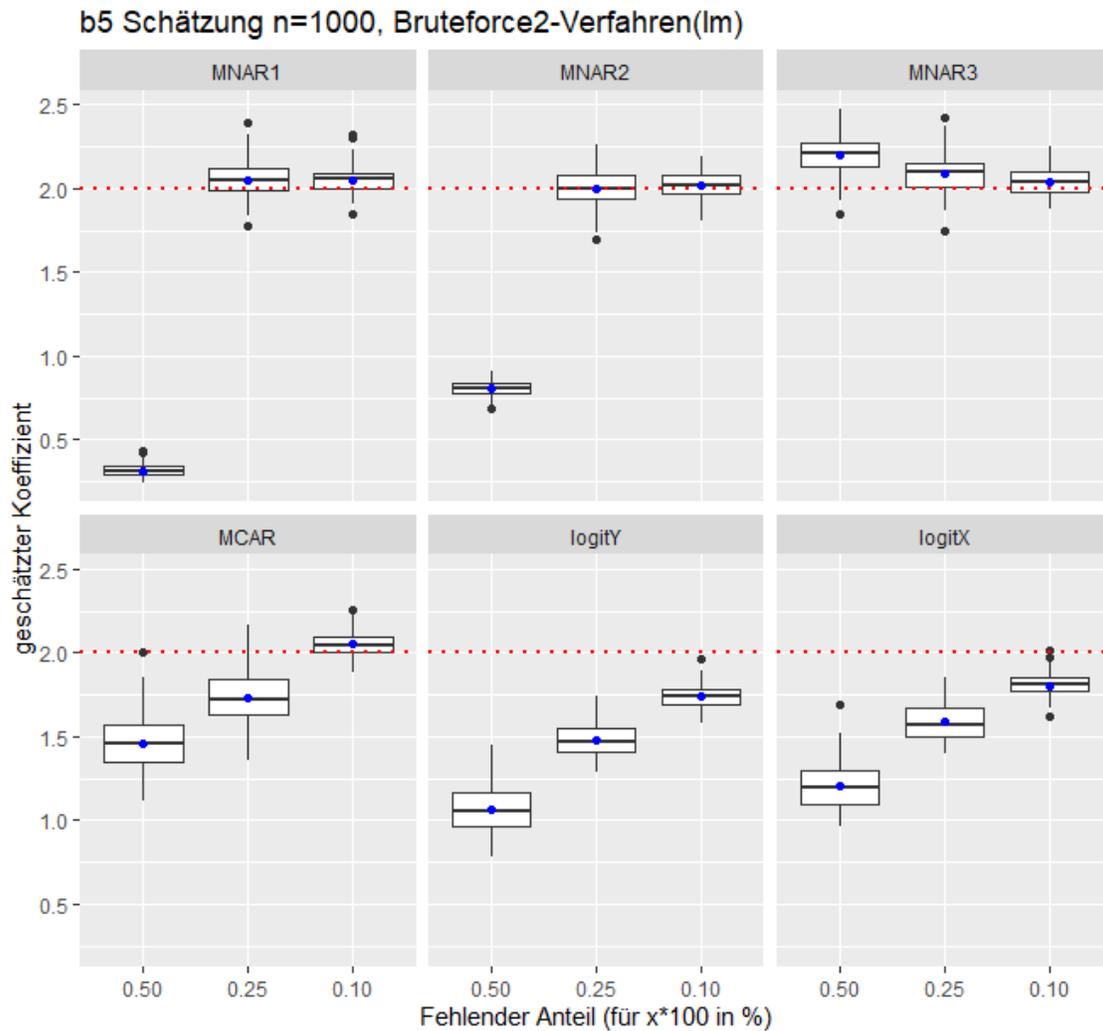


Abbildung 49: Es werden die geschätzten Werte für β_5 in Boxplotform dargestellt. Diese werden nach den Fehlendmechanismen gruppiert. Der blaue Punkt gibt den Mittelwert der geschätzten Werte an. Die rote horizontale Linie repräsentiert den echten Wert des Koeffizienten.

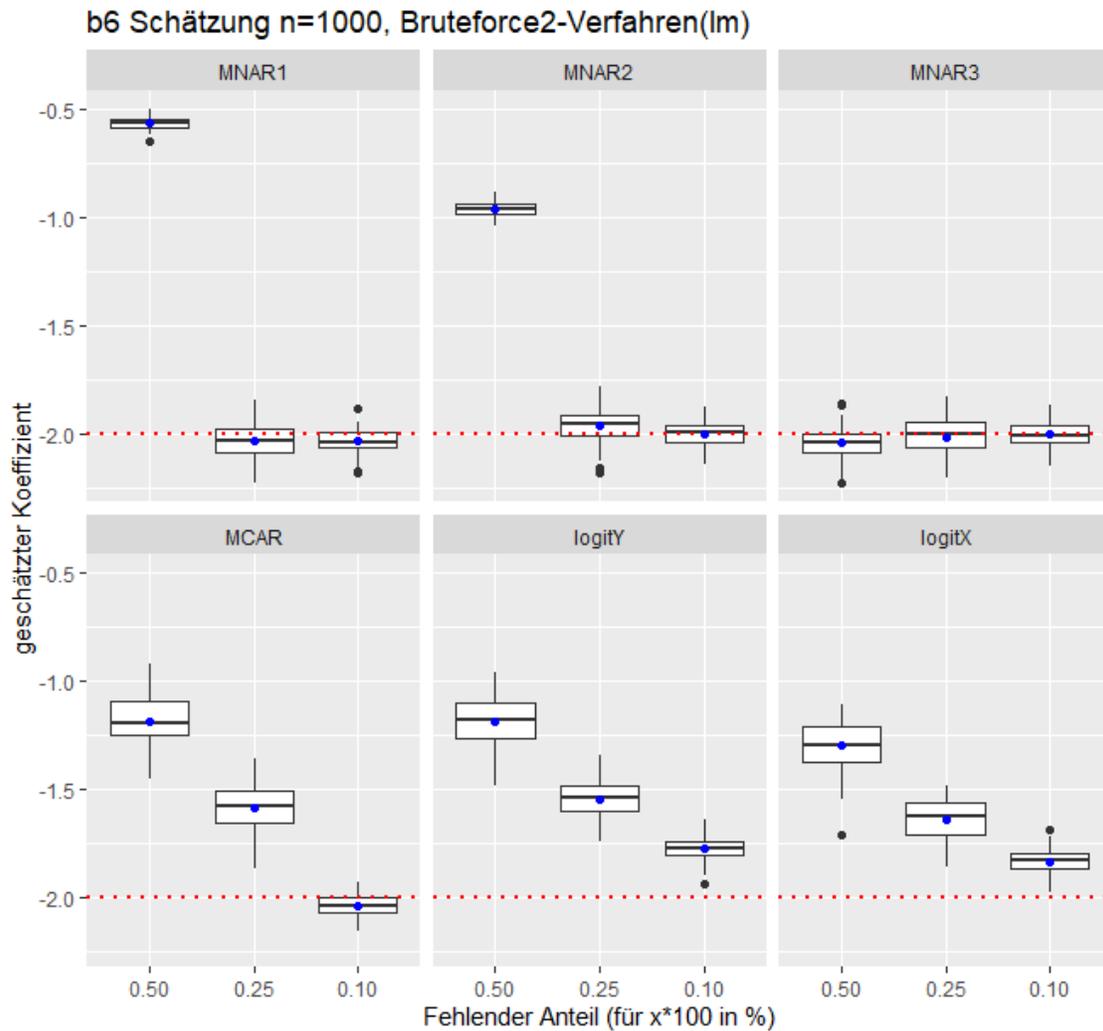


Abbildung 50: Es werden die geschätzten Werte für β_6 in Boxplotform dargestellt. Diese werden nach den Fehlendmechanismen gruppiert. Der blaue Punkt gibt den Mittelwert der geschätzten Werte an. Die rote horizontale Linie repräsentiert den echten Wert des Koeffizienten.

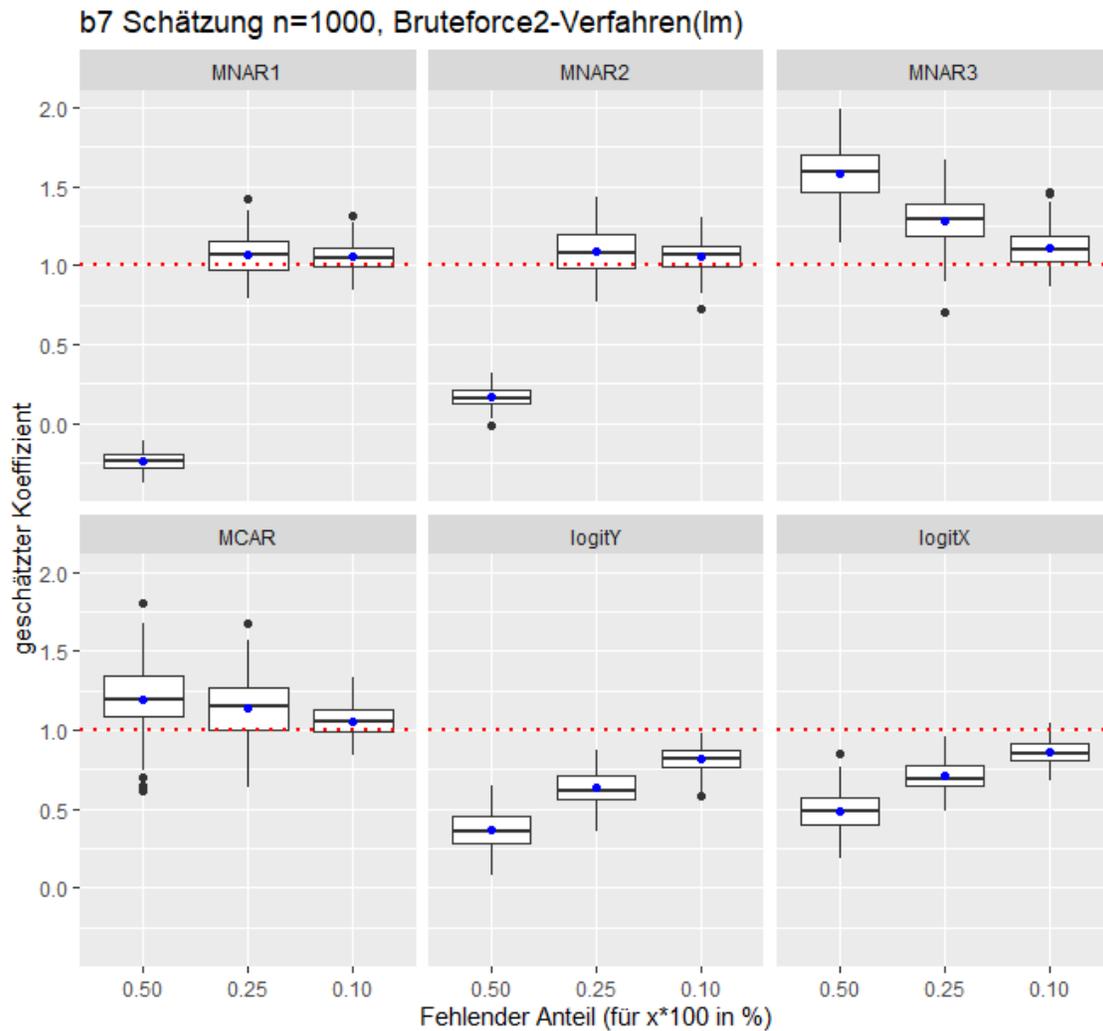


Abbildung 51: Es werden die geschätzten Werte für β_7 in Boxplotform dargestellt. Diese werden nach den Fehlendmechanismen gruppiert. Der blaue Punkt gibt den Mittelwert der geschätzten Werte an. Die rote horizontale Linie repräsentiert den echten Wert des Koeffizienten.

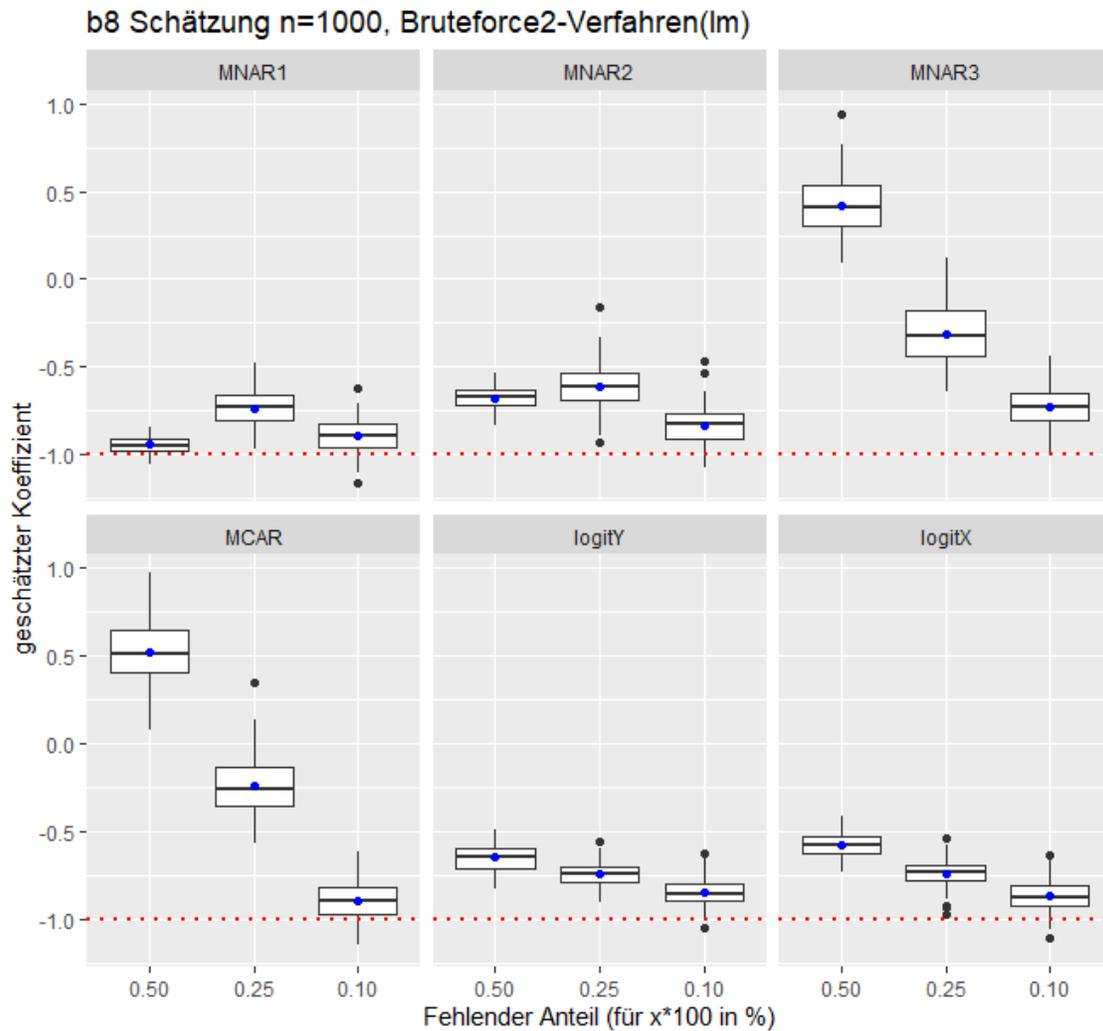


Abbildung 52: Es werden die geschätzten Werte für β_8 in Boxplotform dargestellt. Diese werden nach den Fehlendmechanismen gruppiert. Der blaue Punkt gibt den Mittelwert der geschätzten Werte an. Die rote horizontale Linie repräsentiert den echten Wert des Koeffizienten.

C Gütekriterien-Tabellen zu linearen Modellen

C.1 Normales Verfahren

	Fehlendmechanismus MCAR mit normalen Verfahren (lm)								
	n=50			n=200			n=1000		
	50%	25%	10%	50%	25%	10%	50%	25%	10%
MSE_b1	0.0726	0.0361	0.0301	0.0110	0.0065	0.0058	0.0015	0.0011	0.0009
MSE_b2	0.0534	0.0274	0.0202	0.0092	0.0063	0.0050	0.0025	0.0015	0.0013
MSE_b3	0.2476	0.1316	0.0945	0.0479	0.0269	0.0217	0.0086	0.0049	0.0041
MSE_b4	1.1477	0.7939	0.6975	0.0476	0.0386	0.0331	0.0114	0.0072	0.0069
MSE_b5	0.2399	0.1007	0.0898	0.0323	0.0214	0.0180	0.0053	0.0038	0.0033
MSE_b6	0.1388	0.0553	0.0466	0.0164	0.0102	0.0079	0.0028	0.0021	0.0021
MSE_b7	0.4949	0.2269	0.2013	0.0514	0.0310	0.0265	0.0072	0.0066	0.0057
MSE_b8	1.1255	0.6203	0.5853	0.0475	0.0381	0.0285	0.0125	0.0074	0.0076
Bias_b1	0.0176	-0.0082	-0.0012	-0.0056	-0.0011	-0.0038	0.0037	0.0014	0.0062
Bias_b2	-0.0059	0.0078	0.0030	-0.0006	0.0086	-0.0047	0.0044	0.0005	-0.0006
Bias_b3	0.0419	-0.0053	-0.0015	0.0322	0.0284	0.0260	-0.0001	-0.0026	-0.0053
Bias_b4	-0.0164	-0.0068	-0.0139	-0.0014	-0.0362	-0.0279	-0.0150	-0.0000	-0.0104
Bias_b5	-0.0330	0.0169	-0.0008	-0.0135	0.0008	-0.0007	-0.0036	0.0013	-0.0069
Bias_b6	0.0344	-0.0170	0.0056	0.0143	-0.0015	0.0055	0.0043	0.0032	0.0100
Bias_b7	-0.0592	-0.0099	-0.0193	-0.0049	0.0128	0.0034	-0.0031	-0.0036	-0.0071
Bias_b8	0.0397	0.0110	-0.0209	-0.0011	0.0114	-0.0170	0.0135	-0.0067	-0.0243
Var_b1	0.0723	0.0361	0.0301	0.0109	0.0065	0.0058	0.0015	0.0011	0.0009
Var_b2	0.0534	0.0273	0.0202	0.0092	0.0063	0.0049	0.0025	0.0015	0.0013
Var_b3	0.2458	0.1316	0.0945	0.0468	0.0261	0.0210	0.0086	0.0049	0.0041
Var_b4	1.1474	0.7938	0.6973	0.0476	0.0373	0.0324	0.0112	0.0072	0.0068
Var_b5	0.2388	0.1004	0.0898	0.0322	0.0214	0.0180	0.0053	0.0038	0.0033
Var_b6	0.1376	0.0551	0.0466	0.0162	0.0102	0.0079	0.0028	0.0020	0.0020
Var_b7	0.4914	0.2268	0.2009	0.0514	0.0309	0.0264	0.0072	0.0065	0.0056
Var_b8	1.1239	0.6202	0.5849	0.0475	0.0380	0.0282	0.0123	0.0074	0.0070

Tabelle 1: Es werden MSE, Bias und Varianz für die einzelnen Koeffizienten (b1, ..., b8) unter unterschiedlich großen Fehlend-Anteilen und unter der jeweiligen Datensatzgröße gezeigt.

Fehlendmechanismus MCAR mit normalen Verfahren (ridge, $\lambda=0$)

	n=50			n=200			n=1000		
	50%	25%	10%	50%	25%	10%	50%	25%	10%
MSE_b1	0.0813	0.0433	0.0333	0.0118	0.0068	0.0059	0.0015	0.0012	0.0009
MSE_b2	0.0613	0.0288	0.0210	0.0096	0.0066	0.0047	0.0026	0.0016	0.0014
MSE_b3	0.2588	0.1371	0.0944	0.0503	0.0267	0.0209	0.0084	0.0050	0.0042
MSE_b4	2.3485	1.3575	1.4450	0.0801	0.0569	0.0451	0.0160	0.0111	0.0095
MSE_b5	0.2947	0.1346	0.1076	0.0386	0.0281	0.0254	0.0064	0.0047	0.0044
MSE_b6	0.1558	0.0664	0.0526	0.0166	0.0106	0.0087	0.0032	0.0023	0.0023
MSE_b7	0.6518	0.3178	0.2563	0.0677	0.0448	0.0401	0.0104	0.0076	0.0068
MSE_b8	1.8403	0.6381	0.5898	0.0884	0.0455	0.0479	0.0212	0.0137	0.0126
Bias_b1	0.0232	-0.0023	-0.0008	-0.0052	-0.0004	-0.0044	0.0045	0.0001	0.0050
Bias_b2	0.0015	0.0037	0.0047	-0.0005	0.0071	-0.0038	0.0039	0.0020	0.0008
Bias_b3	0.0588	-0.0069	-0.0007	0.0319	0.0252	0.0277	-0.0008	0.0005	-0.0026
Bias_b4	0.0645	-0.0077	0.0029	0.0028	-0.0408	-0.0248	-0.0186	0.0116	0.0011
Bias_b5	-0.0385	0.0011	-0.0055	-0.0172	-0.0082	-0.0025	-0.0090	0.0044	-0.0037
Bias_b6	0.0405	-0.0081	0.0091	0.0169	0.0027	0.0071	0.0070	0.0025	0.0093
Bias_b7	-0.0726	-0.0349	-0.0266	-0.0079	-0.0002	0.0037	-0.0094	0.0030	-0.0003
Bias_b8	0.1667	0.0018	-0.0251	-0.0006	-0.0001	-0.0139	0.0076	0.0059	-0.0108
Var_b1	0.0808	0.0433	0.0333	0.0118	0.0068	0.0059	0.0015	0.0012	0.0009
Var_b2	0.0613	0.0288	0.0210	0.0096	0.0065	0.0046	0.0025	0.0016	0.0014
Var_b3	0.2553	0.1371	0.0944	0.0493	0.0260	0.0201	0.0084	0.0050	0.0042
Var_b4	2.3443	1.3574	1.4450	0.0801	0.0552	0.0445	0.0157	0.0109	0.0095
Var_b5	0.2933	0.1346	0.1076	0.0383	0.0280	0.0254	0.0063	0.0047	0.0044
Var_b6	0.1541	0.0663	0.0525	0.0163	0.0106	0.0086	0.0031	0.0023	0.0023
Var_b7	0.6465	0.3166	0.2556	0.0676	0.0448	0.0401	0.0103	0.0076	0.0068
Var_b8	1.8125	0.6381	0.5891	0.0884	0.0455	0.0477	0.0211	0.0136	0.0125

Tabelle 2: Es werden MSE, Bias und Varianz für die einzelnen Koeffizienten (b1, ..., b8) unter unterschiedlich großen Fehlend-Anteilen und unter der jeweiligen Datensatzgröße gezeigt.

Fehlendmechanismus MCAR mit normalen Verfahren (ridge)

	n=50			n=200			n=1000		
	50%	25%	10%	50%	25%	10%	50%	25%	10%
MSE_b1	0.0879	0.0433	0.0343	0.0148	0.0090	0.0069	0.0108	0.0072	0.0062
MSE_b2	0.0853	0.0379	0.0281	0.0484	0.0301	0.0252	0.0329	0.0237	0.0191
MSE_b3	0.2136	0.1174	0.0806	0.0662	0.0354	0.0268	0.0136	0.0089	0.0070
MSE_b4	1.0570	1.0124	1.0837	0.0811	0.0516	0.0393	0.0223	0.0199	0.0149
MSE_b5	1.6078	1.0500	0.8352	1.4310	1.1171	0.9289	0.8552	0.6317	0.5268
MSE_b6	0.8735	0.5768	0.4725	0.6808	0.5113	0.4337	0.3942	0.2937	0.2463
MSE_b7	2.4537	1.7183	1.3721	1.5883	1.2665	1.0836	0.8366	0.6196	0.5145
MSE_b8	1.6266	1.2569	1.1221	0.6285	0.4936	0.4634	0.2922	0.2308	0.2094
Bias_b1	0.1274	0.0717	0.0735	0.0653	0.0541	0.0429	0.0972	0.0781	0.0738
Bias_b2	-0.1845	-0.1223	-0.1024	-0.2004	-0.1574	-0.1468	-0.1755	-0.1494	-0.1338
Bias_b3	0.0305	0.0020	0.0081	-0.1341	-0.1078	-0.0951	-0.0747	-0.0639	-0.0576
Bias_b4	-0.0046	-0.2078	-0.2096	0.0261	-0.0399	-0.0236	0.0837	0.0997	0.0791
Bias_b5	-1.2382	-1.0014	-0.8914	-1.1907	-1.0521	-0.9584	-0.9231	-0.7931	-0.7241
Bias_b6	0.9137	0.7439	0.6729	0.8211	0.7119	0.6558	0.6264	0.5407	0.4948
Bias_b7	-1.4896	-1.2530	-1.1141	-1.2472	-1.1150	-1.0311	-0.9113	-0.7842	-0.7141
Bias_b8	-1.0404	-0.9390	-0.8840	-0.7506	-0.6785	-0.6558	-0.5273	-0.4693	-0.4474
Var_b1	0.0716	0.0381	0.0289	0.0106	0.0061	0.0051	0.0014	0.0011	0.0008
Var_b2	0.0512	0.0230	0.0176	0.0082	0.0053	0.0037	0.0021	0.0014	0.0012
Var_b3	0.2127	0.1174	0.0805	0.0482	0.0237	0.0177	0.0080	0.0048	0.0037
Var_b4	1.0569	0.9692	1.0398	0.0805	0.0501	0.0388	0.0153	0.0099	0.0086
Var_b5	0.0747	0.0471	0.0406	0.0132	0.0101	0.0104	0.0031	0.0027	0.0025
Var_b6	0.0387	0.0234	0.0197	0.0065	0.0044	0.0037	0.0018	0.0014	0.0015
Var_b7	0.2349	0.1483	0.1309	0.0329	0.0231	0.0205	0.0061	0.0047	0.0045
Var_b8	0.5441	0.3752	0.3407	0.0652	0.0331	0.0332	0.0142	0.0106	0.0092

Tabelle 3: Es werden MSE, Bias und Varianz für die einzelnen Koeffizienten (b1, ..., b8) unter unterschiedlich großen Fehlend-Anteilen und unter der jeweiligen Datensatzgröße gezeigt.

Fehlendmechanismus MNAR1 mit normalen Verfahren (lm)

	n=50			n=200			n=1000		
	50%	25%	10%	50%	25%	10%	50%	25%	10%
MSE_b1	0.0590	0.0328	0.0299	0.0264	0.0075	0.0060	0.0217	0.0014	0.0010
MSE_b2	0.1015	0.0350	0.0196	0.0926	0.0064	0.0055	0.0763	0.0019	0.0013
MSE_b3	0.2262	0.1256	0.0992	0.0655	0.0263	0.0240	0.0354	0.0055	0.0036
MSE_b4	2.8073	0.7974	0.6224	0.0917	0.0454	0.0389	0.0124	0.0077	0.0071
MSE_b5	0.2191	0.1131	0.0801	0.1269	0.0227	0.0195	0.1037	0.0049	0.0037
MSE_b6	0.1350	0.0633	0.0391	0.0808	0.0110	0.0093	0.0829	0.0030	0.0021
MSE_b7	0.3444	0.2326	0.1899	0.1334	0.0338	0.0298	0.0659	0.0073	0.0056
MSE_b8	1.6798	0.8808	0.5292	0.1211	0.0509	0.0328	0.0452	0.0108	0.0071
Bias_b1	0.0955	0.0120	-0.0050	0.1264	-0.0084	-0.0072	0.1414	0.0172	0.0027
Bias_b2	-0.1449	0.0290	0.0100	-0.2779	-0.0101	-0.0023	-0.2698	-0.0135	-0.0018
Bias_b3	-0.0502	0.0434	0.0085	-0.1772	0.0087	0.0255	-0.1692	-0.0214	-0.0078
Bias_b4	-0.7052	-0.1128	0.0022	0.0935	-0.0449	-0.0278	-0.0179	-0.0156	-0.0129
Bias_b5	-0.2453	-0.0263	0.0123	-0.2917	-0.0236	-0.0011	-0.3129	-0.0212	-0.0043
Bias_b6	0.2265	0.0418	-0.0068	0.2421	0.0194	0.0019	0.2817	0.0225	0.0074
Bias_b7	-0.1137	-0.0089	-0.0014	-0.2686	-0.0153	0.0038	-0.2345	-0.0174	-0.0060
Bias_b8	0.4057	-0.0832	-0.0323	-0.1885	-0.0387	-0.0100	-0.1777	-0.0564	-0.0197
Var_b1	0.0499	0.0327	0.0299	0.0104	0.0074	0.0059	0.0017	0.0011	0.0010
Var_b2	0.0805	0.0341	0.0195	0.0154	0.0063	0.0055	0.0035	0.0017	0.0013
Var_b3	0.2237	0.1237	0.0991	0.0341	0.0262	0.0233	0.0068	0.0050	0.0036
Var_b4	2.3099	0.7847	0.6224	0.0830	0.0434	0.0381	0.0121	0.0074	0.0069
Var_b5	0.1589	0.1124	0.0800	0.0418	0.0221	0.0195	0.0058	0.0044	0.0037
Var_b6	0.0836	0.0616	0.0391	0.0222	0.0106	0.0093	0.0035	0.0025	0.0021
Var_b7	0.3315	0.2326	0.1899	0.0613	0.0336	0.0298	0.0109	0.0070	0.0056
Var_b8	1.5152	0.8739	0.5282	0.0856	0.0494	0.0327	0.0136	0.0076	0.0067

Tabelle 4: Es werden MSE, Bias und Varianz für die einzelnen Koeffizienten (b1, ..., b8) unter unterschiedlich großen Fehlend-Anteilen und unter der jeweiligen Datensatzgröße gezeigt.

Fehlendmechanismus MNAR1 mit normalen Verfahren (ridge, $\lambda=0$)									
	n=50			n=200			n=1000		
	50%	25%	10%	50%	25%	10%	50%	25%	10%
MSE_b1	0.0575	0.0369	0.0347	0.0249	0.0073	0.0062	0.0200	0.0013	0.0011
MSE_b2	0.1045	0.0368	0.0210	0.0832	0.0068	0.0056	0.0825	0.0019	0.0014
MSE_b3	0.2298	0.1280	0.1014	0.0437	0.0252	0.0226	0.0303	0.0056	0.0036
MSE_b4	5.4958	2.6707	1.4585	0.1653	0.0561	0.0515	0.0368	0.0108	0.0104
MSE_b5	0.2226	0.1393	0.1008	0.0867	0.0316	0.0263	0.0821	0.0056	0.0046
MSE_b6	0.1363	0.0709	0.0455	0.0713	0.0121	0.0101	0.0809	0.0032	0.0024
MSE_b7	0.3302	0.3049	0.2559	0.0816	0.0555	0.0452	0.0344	0.0085	0.0069
MSE_b8	5.3441	1.9612	0.5366	0.2612	0.0705	0.0494	0.0633	0.0141	0.0113
Bias_b1	0.0924	0.0173	-0.0073	0.1212	-0.0123	-0.0083	0.1354	0.0149	0.0015
Bias_b2	-0.1578	0.0240	0.0103	-0.2601	-0.0044	-0.0006	-0.2811	-0.0111	-0.0003
Bias_b3	-0.0605	0.0456	0.0078	-0.0953	0.0224	0.0289	-0.1537	-0.0172	-0.0053
Bias_b4	0.0373	-0.2243	0.0121	0.2641	-0.0153	-0.0183	0.1415	0.0053	-0.0021
Bias_b5	-0.2468	-0.0393	0.0113	-0.2072	-0.0054	0.0016	-0.2752	-0.0140	-0.0012
Bias_b6	0.2280	0.0498	-0.0055	0.2242	0.0131	0.0020	0.2782	0.0203	0.0067
Bias_b7	-0.0941	-0.0255	0.0010	-0.1048	0.0184	0.0106	-0.1454	-0.0037	0.0006
Bias_b8	1.0297	0.0873	-0.0368	0.3062	0.0139	0.0006	0.1760	-0.0276	-0.0066
Var_b1	0.0490	0.0366	0.0346	0.0102	0.0072	0.0061	0.0017	0.0011	0.0011
Var_b2	0.0796	0.0362	0.0209	0.0155	0.0068	0.0056	0.0035	0.0018	0.0014
Var_b3	0.2262	0.1260	0.1013	0.0346	0.0247	0.0218	0.0067	0.0053	0.0036
Var_b4	5.4944	2.6204	1.4584	0.0955	0.0559	0.0511	0.0167	0.0107	0.0104
Var_b5	0.1617	0.1378	0.1006	0.0438	0.0315	0.0263	0.0064	0.0054	0.0046
Var_b6	0.0844	0.0684	0.0454	0.0210	0.0120	0.0101	0.0035	0.0027	0.0023
Var_b7	0.3214	0.3043	0.2559	0.0706	0.0552	0.0451	0.0133	0.0084	0.0069
Var_b8	4.2838	1.9536	0.5353	0.1674	0.0703	0.0494	0.0323	0.0133	0.0113

Tabelle 5: Es werden MSE, Bias und Varianz für die einzelnen Koeffizienten (b1, ..., b8) unter unterschiedlich großen Fehlend-Anteilen und unter der jeweiligen Datensatzgröße gezeigt.

Fehlendmechanismus MNAR1 mit normalen Verfahren (ridge)

	n=50			n=200			n=1000		
	50%	25%	10%	50%	25%	10%	50%	25%	10%
MSE_b1	0.0887	0.0424	0.0334	0.0805	0.0092	0.0071	0.0653	0.0103	0.0058
MSE_b2	0.2511	0.0452	0.0272	0.2799	0.0351	0.0252	0.2847	0.0312	0.0194
MSE_b3	0.2652	0.1170	0.0828	0.1107	0.0354	0.0279	0.0671	0.0111	0.0067
MSE_b4	1.1221	1.1953	1.1472	0.2950	0.0528	0.0440	0.0938	0.0185	0.0156
MSE_b5	1.6123	1.0139	0.8127	1.6863	1.1314	0.9338	1.1335	0.6512	0.5246
MSE_b6	1.0243	0.5988	0.4537	1.0119	0.5418	0.4312	0.7174	0.3124	0.2438
MSE_b7	1.5627	1.4802	1.3170	1.6345	1.2757	1.0811	0.7432	0.6097	0.5107
MSE_b8	2.3167	1.6438	1.0930	0.2110	0.5294	0.4468	0.0475	0.2658	0.2052
Bias_b1	0.2147	0.1032	0.0661	0.2685	0.0536	0.0404	0.2528	0.0965	0.0696
Bias_b2	-0.4270	-0.1173	-0.0998	-0.5187	-0.1724	-0.1429	-0.5315	-0.1728	-0.1348
Bias_b3	-0.2489	-0.0034	0.0147	-0.2896	-0.1238	-0.0925	-0.2483	-0.0791	-0.0589
Bias_b4	0.6145	-0.1704	-0.2148	0.4684	0.0143	-0.0149	0.2833	0.0947	0.0772
Bias_b5	-1.2512	-0.9819	-0.8795	-1.2946	-1.0586	-0.9610	-1.0634	-0.8053	-0.7226
Bias_b6	0.9994	0.7585	0.6609	1.0026	0.7329	0.6534	0.8459	0.5574	0.4924
Bias_b7	-1.1855	-1.1578	-1.0884	-1.2672	-1.1183	-1.0288	-0.8574	-0.7773	-0.7117
Bias_b8	-0.4639	-1.0818	-0.8731	-0.3075	-0.6949	-0.6421	-0.1566	-0.5058	-0.4440
Var_b1	0.0426	0.0317	0.0290	0.0084	0.0064	0.0055	0.0014	0.0010	0.0009
Var_b2	0.0688	0.0315	0.0172	0.0108	0.0054	0.0048	0.0022	0.0014	0.0012
Var_b3	0.2033	0.1170	0.0826	0.0269	0.0201	0.0193	0.0054	0.0049	0.0033
Var_b4	0.7444	1.1662	1.1010	0.0756	0.0526	0.0437	0.0135	0.0095	0.0096
Var_b5	0.0467	0.0498	0.0392	0.0103	0.0108	0.0103	0.0026	0.0028	0.0025
Var_b6	0.0254	0.0234	0.0169	0.0068	0.0047	0.0042	0.0017	0.0017	0.0014
Var_b7	0.1572	0.1398	0.1324	0.0286	0.0251	0.0227	0.0080	0.0056	0.0042
Var_b8	2.1015	0.4735	0.3308	0.1164	0.0465	0.0345	0.0230	0.0099	0.0080

Tabelle 6: Es werden MSE, Bias und Varianz für die einzelnen Koeffizienten (b1, ..., b8) unter unterschiedlich großen Fehlend-Anteilen und unter der jeweiligen Datensatzgröße gezeigt.

Fehlendmechanismus MNAR2 mit normalen Verfahren (lm)

	n=50			n=200			n=1000		
	50%	25%	10%	50%	25%	10%	50%	25%	10%
MSE_b1	0.0605	0.0408	0.0323	0.0430	0.0107	0.0079	0.0268	0.0075	0.0056
MSE_b2	0.1240	0.0449	0.0308	0.1275	0.0334	0.0248	0.1248	0.0291	0.0188
MSE_b3	0.1993	0.1111	0.0872	0.1077	0.0382	0.0301	0.0203	0.0092	0.0066
MSE_b4	1.1338	1.1259	0.9267	0.0956	0.0494	0.0385	0.0537	0.0191	0.0156
MSE_b5	1.2906	0.9715	0.8275	1.5756	1.1170	0.9275	0.9904	0.6516	0.5204
MSE_b6	0.7849	0.5445	0.4626	0.8135	0.5174	0.4317	0.5338	0.3052	0.2403
MSE_b7	2.1102	1.6779	1.3812	1.8367	1.3009	1.0913	0.7615	0.6098	0.5003
MSE_b8	1.6142	1.3204	0.9993	0.5588	0.5416	0.4439	0.1339	0.2369	0.2011
Bias_b1	0.1274	0.0781	0.0583	0.1820	0.0640	0.0491	0.1601	0.0798	0.0689
Bias_b2	-0.2588	-0.1325	-0.0977	-0.3431	-0.1719	-0.1448	-0.3511	-0.1664	-0.1331
Bias_b3	-0.1205	0.0031	0.0339	-0.2625	-0.1271	-0.0940	-0.1198	-0.0664	-0.0560
Bias_b4	0.3614	-0.0653	-0.1887	0.1510	0.0091	-0.0216	0.1984	0.0981	0.0766
Bias_b5	-1.1172	-0.9603	-0.8852	-1.2506	-1.0517	-0.9578	-0.9935	-0.8053	-0.7197
Bias_b6	0.8723	0.7235	0.6660	0.8974	0.7161	0.6541	0.7293	0.5508	0.4890
Bias_b7	-1.3783	-1.2216	-1.1044	-1.3427	-1.1285	-1.0334	-0.8675	-0.7769	-0.7042
Bias_b8	-0.7397	-1.0090	-0.8508	-0.6679	-0.7059	-0.6397	-0.3347	-0.4764	-0.4391
Var_b1	0.0442	0.0347	0.0289	0.0099	0.0066	0.0055	0.0012	0.0011	0.0008
Var_b2	0.0570	0.0273	0.0213	0.0097	0.0039	0.0039	0.0015	0.0014	0.0011
Var_b3	0.1848	0.1111	0.0860	0.0388	0.0220	0.0213	0.0060	0.0048	0.0034
Var_b4	1.0032	1.1216	0.8911	0.0728	0.0493	0.0380	0.0143	0.0095	0.0097
Var_b5	0.0424	0.0493	0.0438	0.0116	0.0110	0.0101	0.0034	0.0031	0.0025
Var_b6	0.0240	0.0210	0.0191	0.0081	0.0046	0.0039	0.0019	0.0018	0.0012
Var_b7	0.2105	0.1856	0.1615	0.0340	0.0273	0.0234	0.0090	0.0062	0.0044
Var_b8	1.0671	0.3023	0.2755	0.1127	0.0432	0.0347	0.0218	0.0100	0.0083

Tabelle 7: Es werden MSE, Bias und Varianz für die einzelnen Koeffizienten (b1, ..., b8) unter unterschiedlich großen Fehlend-Anteilen und unter der jeweiligen Datensatzgröße gezeigt.

Fehlendmechanismus MNAR2 mit normalen Verfahren (ridge, $\lambda=0$)									
	n=50			n=200			n=1000		
	50%	25%	10%	50%	25%	10%	50%	25%	10%
MSE_b1	0.0543	0.0397	0.0325	0.0149	0.0073	0.0063	0.0042	0.0011	0.0009
MSE_b2	0.0785	0.0343	0.0267	0.0258	0.0048	0.0046	0.0177	0.0018	0.0013
MSE_b3	0.2375	0.1422	0.0984	0.0492	0.0249	0.0253	0.0077	0.0057	0.0039
MSE_b4	5.3736	1.6723	1.1809	0.0839	0.0578	0.0472	0.0225	0.0114	0.0109
MSE_b5	0.1430	0.1374	0.1076	0.0633	0.0295	0.0280	0.0251	0.0061	0.0045
MSE_b6	0.0791	0.0608	0.0486	0.0360	0.0105	0.0099	0.0193	0.0034	0.0021
MSE_b7	0.4187	0.3658	0.2956	0.1043	0.0516	0.0465	0.0185	0.0099	0.0070
MSE_b8	4.3774	0.7167	0.4659	0.1905	0.0624	0.0511	0.0433	0.0150	0.0108
Bias_b1	0.0340	-0.0008	-0.0143	0.0646	0.0013	0.0021	0.0527	0.0026	0.0007
Bias_b2	-0.0702	-0.0006	0.0099	-0.1131	-0.0079	-0.0011	-0.1245	-0.0086	0.0006
Bias_b3	-0.0386	0.0302	0.0109	-0.0595	0.0259	0.0257	-0.0312	-0.0040	-0.0028
Bias_b4	-0.1976	0.1163	0.0266	0.0274	-0.0265	-0.0201	0.0765	0.0125	0.0002
Bias_b5	-0.0692	-0.0068	0.0136	-0.1374	-0.0145	0.0018	-0.1294	-0.0123	0.0029
Bias_b6	0.0664	0.0025	-0.0040	0.1164	0.0098	0.0042	0.1245	0.0145	0.0031
Bias_b7	-0.1002	-0.0514	-0.0030	-0.1408	-0.0093	0.0012	-0.0484	0.0022	0.0133
Bias_b8	0.8617	-0.0524	-0.0063	0.0521	-0.0167	0.0009	0.1145	-0.0014	-0.0023
Var_b1	0.0532	0.0397	0.0323	0.0107	0.0073	0.0063	0.0014	0.0011	0.0009
Var_b2	0.0735	0.0343	0.0266	0.0130	0.0047	0.0046	0.0022	0.0017	0.0013
Var_b3	0.2360	0.1413	0.0983	0.0457	0.0242	0.0246	0.0067	0.0056	0.0039
Var_b4	5.3345	1.6588	1.1801	0.0832	0.0571	0.0468	0.0166	0.0112	0.0109
Var_b5	0.1382	0.1374	0.1075	0.0444	0.0293	0.0280	0.0083	0.0060	0.0045
Var_b6	0.0747	0.0608	0.0486	0.0225	0.0104	0.0099	0.0038	0.0032	0.0021
Var_b7	0.4087	0.3632	0.2956	0.0844	0.0515	0.0465	0.0162	0.0099	0.0068
Var_b8	3.6349	0.7140	0.4658	0.1878	0.0621	0.0511	0.0302	0.0150	0.0108

Tabelle 8: Es werden MSE, Bias und Varianz für die einzelnen Koeffizienten (b1, ..., b8) unter unterschiedlich großen Fehlend-Anteilen und unter der jeweiligen Datensatzgröße gezeigt.

Fehlendmechanismus MNAR2 mit normalen Verfahren (ridge)

	n=50			n=200			n=1000		
	50%	25%	10%	50%	25%	10%	50%	25%	10%
MSE_b1	0.0605	0.0408	0.0323	0.0430	0.0107	0.0079	0.0268	0.0075	0.0056
MSE_b2	0.1240	0.0449	0.0308	0.1275	0.0334	0.0248	0.1248	0.0291	0.0188
MSE_b3	0.1993	0.1111	0.0872	0.1077	0.0382	0.0301	0.0203	0.0092	0.0066
MSE_b4	1.1338	1.1259	0.9267	0.0956	0.0494	0.0385	0.0537	0.0191	0.0156
MSE_b5	1.2906	0.9715	0.8275	1.5756	1.1170	0.9275	0.9904	0.6516	0.5204
MSE_b6	0.7849	0.5445	0.4626	0.8135	0.5174	0.4317	0.5338	0.3052	0.2403
MSE_b7	2.1102	1.6779	1.3812	1.8367	1.3009	1.0913	0.7615	0.6098	0.5003
MSE_b8	1.6142	1.3204	0.9993	0.5588	0.5416	0.4439	0.1339	0.2369	0.2011
Bias_b1	0.1274	0.0781	0.0583	0.1820	0.0640	0.0491	0.1601	0.0798	0.0689
Bias_b2	-0.2588	-0.1325	-0.0977	-0.3431	-0.1719	-0.1448	-0.3511	-0.1664	-0.1331
Bias_b3	-0.1205	0.0031	0.0339	-0.2625	-0.1271	-0.0940	-0.1198	-0.0664	-0.0560
Bias_b4	0.3614	-0.0653	-0.1887	0.1510	0.0091	-0.0216	0.1984	0.0981	0.0766
Bias_b5	-1.1172	-0.9603	-0.8852	-1.2506	-1.0517	-0.9578	-0.9935	-0.8053	-0.7197
Bias_b6	0.8723	0.7235	0.6660	0.8974	0.7161	0.6541	0.7293	0.5508	0.4890
Bias_b7	-1.3783	-1.2216	-1.1044	-1.3427	-1.1285	-1.0334	-0.8675	-0.7769	-0.7042
Bias_b8	-0.7397	-1.0090	-0.8508	-0.6679	-0.7059	-0.6397	-0.3347	-0.4764	-0.4391
Var_b1	0.0442	0.0347	0.0289	0.0099	0.0066	0.0055	0.0012	0.0011	0.0008
Var_b2	0.0570	0.0273	0.0213	0.0097	0.0039	0.0039	0.0015	0.0014	0.0011
Var_b3	0.1848	0.1111	0.0860	0.0388	0.0220	0.0213	0.0060	0.0048	0.0034
Var_b4	1.0032	1.1216	0.8911	0.0728	0.0493	0.0380	0.0143	0.0095	0.0097
Var_b5	0.0424	0.0493	0.0438	0.0116	0.0110	0.0101	0.0034	0.0031	0.0025
Var_b6	0.0240	0.0210	0.0191	0.0081	0.0046	0.0039	0.0019	0.0018	0.0012
Var_b7	0.2105	0.1856	0.1615	0.0340	0.0273	0.0234	0.0090	0.0062	0.0044
Var_b8	1.0671	0.3023	0.2755	0.1127	0.0432	0.0347	0.0218	0.0100	0.0083

Tabelle 9: Es werden MSE, Bias und Varianz für die einzelnen Koeffizienten (b1, ..., b8) unter unterschiedlich großen Fehlend-Anteilen und unter der jeweiligen Datensatzgröße gezeigt.

Fehlendmechanismus MNAR3 mit normalen Verfahren (lm)

	n=50			n=200			n=1000		
	50%	25%	10%	50%	25%	10%	50%	25%	10%
MSE_b1	0.0460	0.0335	0.0292	0.0242	0.0097	0.0063	0.0023	0.0013	0.0009
MSE_b2	0.0480	0.0255	0.0215	0.0125	0.0066	0.0048	0.0133	0.0038	0.0015
MSE_b3	0.1936	0.1265	0.0837	0.0448	0.0301	0.0247	0.0063	0.0045	0.0039
MSE_b4	0.6993	0.6090	0.4556	0.0779	0.0412	0.0318	0.0172	0.0084	0.0070
MSE_b5	0.2029	0.1145	0.0769	0.0336	0.0228	0.0191	0.0157	0.0076	0.0038
MSE_b6	0.1069	0.0585	0.0392	0.0172	0.0104	0.0087	0.0094	0.0037	0.0020
MSE_b7	0.4022	0.2314	0.1689	0.0620	0.0322	0.0284	0.0111	0.0077	0.0055
MSE_b8	0.9032	0.5287	0.3878	0.0717	0.0422	0.0327	0.0161	0.0092	0.0065
Bias_b1	-0.0331	-0.0125	-0.0110	-0.1010	-0.0364	-0.0107	-0.0326	-0.0157	-0.0053
Bias_b2	0.0802	0.0443	0.0168	0.0710	0.0389	0.0179	0.1071	0.0475	0.0167
Bias_b3	-0.0806	-0.0293	-0.0034	0.0312	0.0385	0.0352	0.0084	0.0073	-0.0009
Bias_b4	-0.0968	-0.0090	0.0114	-0.1304	-0.0603	-0.0229	-0.0677	-0.0258	-0.0113
Bias_b5	0.1158	0.0504	0.0158	0.0411	0.0255	0.0121	0.0990	0.0492	0.0163
Bias_b6	-0.1167	-0.0521	-0.0131	-0.0441	-0.0264	-0.0080	-0.0810	-0.0378	-0.0109
Bias_b7	0.0511	-0.0004	-0.0086	0.0771	0.0170	0.0088	0.0327	0.0184	0.0085
Bias_b8	0.1062	0.0077	-0.0314	-0.0925	-0.0120	-0.0102	-0.0592	-0.0354	-0.0171
Var_b1	0.0449	0.0333	0.0291	0.0140	0.0084	0.0061	0.0012	0.0010	0.0008
Var_b2	0.0416	0.0235	0.0212	0.0075	0.0051	0.0045	0.0019	0.0015	0.0012
Var_b3	0.1871	0.1257	0.0837	0.0438	0.0286	0.0234	0.0062	0.0045	0.0039
Var_b4	0.6899	0.6089	0.4555	0.0609	0.0375	0.0312	0.0126	0.0077	0.0069
Var_b5	0.1895	0.1120	0.0767	0.0319	0.0222	0.0190	0.0059	0.0051	0.0036
Var_b6	0.0933	0.0558	0.0391	0.0153	0.0097	0.0086	0.0028	0.0023	0.0019
Var_b7	0.3996	0.2314	0.1688	0.0560	0.0319	0.0284	0.0100	0.0074	0.0055
Var_b8	0.8919	0.5286	0.3868	0.0631	0.0420	0.0326	0.0126	0.0080	0.0062

Tabelle 10: Es werden MSE, Bias und Varianz für die einzelnen Koeffizienten (b1, ..., b8) unter unterschiedlich großen Fehlend-Anteilen und unter der jeweiligen Datensatzgröße gezeigt.

Fehlendmechanismus MNAR3 mit normalen Verfahren (ridge, $\lambda=0$)									
	n=50			n=200			n=1000		
	50%	25%	10%	50%	25%	10%	50%	25%	10%
MSE_b1	0.0453	0.0359	0.0343	0.0235	0.0101	0.0062	0.0022	0.0013	0.0009
MSE_b2	0.0527	0.0270	0.0218	0.0109	0.0065	0.0048	0.0134	0.0040	0.0017
MSE_b3	0.1948	0.1239	0.0840	0.0450	0.0302	0.0242	0.0067	0.0047	0.0040
MSE_b4	1.9742	1.4492	1.0839	0.1315	0.0640	0.0431	0.0220	0.0117	0.0098
MSE_b5	0.2447	0.1442	0.0998	0.0452	0.0312	0.0259	0.0164	0.0091	0.0049
MSE_b6	0.1193	0.0691	0.0470	0.0170	0.0109	0.0097	0.0092	0.0040	0.0022
MSE_b7	0.4785	0.3129	0.2409	0.0858	0.0473	0.0431	0.0124	0.0093	0.0062
MSE_b8	0.9244	0.5262	0.3962	0.1313	0.0682	0.0493	0.0222	0.0143	0.0100
Bias_b1	-0.0510	-0.0199	-0.0116	-0.0968	-0.0357	-0.0102	-0.0324	-0.0162	-0.0061
Bias_b2	0.0991	0.0498	0.0180	0.0576	0.0367	0.0164	0.1066	0.0482	0.0177
Bias_b3	-0.0853	-0.0288	-0.0014	0.0081	0.0342	0.0322	0.0072	0.0085	0.0009
Bias_b4	0.1684	0.0811	0.0258	-0.1979	-0.0720	-0.0299	-0.0707	-0.0208	-0.0040
Bias_b5	0.1207	0.0551	0.0106	-0.0137	0.0126	0.0025	0.0931	0.0487	0.0178
Bias_b6	-0.1089	-0.0518	-0.0093	-0.0203	-0.0203	-0.0033	-0.0782	-0.0372	-0.0110
Bias_b7	0.0624	0.0098	-0.0147	-0.0142	-0.0012	-0.0043	0.0261	0.0198	0.0124
Bias_b8	0.0924	0.0089	-0.0351	-0.1903	-0.0320	-0.0236	-0.0649	-0.0309	-0.0093
Var_b1	0.0427	0.0355	0.0342	0.0142	0.0088	0.0060	0.0011	0.0010	0.0009
Var_b2	0.0429	0.0246	0.0215	0.0076	0.0052	0.0045	0.0021	0.0017	0.0013
Var_b3	0.1875	0.1231	0.0840	0.0449	0.0290	0.0231	0.0067	0.0046	0.0040
Var_b4	1.9458	1.4426	1.0833	0.0923	0.0588	0.0422	0.0170	0.0113	0.0098
Var_b5	0.2301	0.1411	0.0997	0.0450	0.0310	0.0259	0.0077	0.0067	0.0046
Var_b6	0.1074	0.0664	0.0469	0.0166	0.0105	0.0097	0.0031	0.0026	0.0021
Var_b7	0.4746	0.3128	0.2407	0.0856	0.0473	0.0431	0.0118	0.0090	0.0060
Var_b8	0.9159	0.5262	0.3950	0.0951	0.0672	0.0487	0.0180	0.0134	0.0099

Tabelle 11: Es werden MSE, Bias und Varianz für die einzelnen Koeffizienten (b1, ..., b8) unter unterschiedlich großen Fehlend-Anteilen und unter der jeweiligen Datensatzgröße gezeigt.

Fehlendmechanismus MNAR3 mit normalen Verfahren (ridge)

	n=50			n=200			n=1000		
	50%	25%	10%	50%	25%	10%	50%	25%	10%
MSE_b1	0.0432	0.0340	0.0330	0.0184	0.0074	0.0062	0.0013	0.0028	0.0039
MSE_b2	0.0351	0.0201	0.0225	0.0103	0.0122	0.0170	0.0018	0.0066	0.0117
MSE_b3	0.1586	0.1077	0.0764	0.0461	0.0314	0.0270	0.0076	0.0060	0.0060
MSE_b4	1.4203	1.1020	0.8825	0.1492	0.0616	0.0386	0.0159	0.0117	0.0129
MSE_b5	1.4918	1.0276	0.8386	1.3566	1.0471	0.9056	0.7042	0.5624	0.4980
MSE_b6	0.7788	0.5456	0.4597	0.5504	0.4515	0.4076	0.2674	0.2362	0.2214
MSE_b7	2.3318	1.7003	1.3824	1.8393	1.3329	1.1011	0.8932	0.6245	0.5092
MSE_b8	1.8793	1.2487	1.0222	1.1450	0.6182	0.4857	0.4678	0.2948	0.2154
Bias_b1	-0.0568	0.0187	0.0473	-0.0833	-0.0033	0.0257	0.0169	0.0422	0.0554
Bias_b2	0.0797	-0.0251	-0.0718	-0.0704	-0.0912	-0.1156	-0.0154	-0.0729	-0.1028
Bias_b3	0.0376	0.0229	0.0358	-0.0897	-0.0767	-0.0784	-0.0347	-0.0408	-0.0491
Bias_b4	-0.3266	-0.2701	-0.2314	-0.2719	-0.1046	-0.0491	-0.0217	0.0434	0.0645
Bias_b5	-1.1990	-0.9900	-0.8940	-1.1589	-1.0177	-0.9460	-0.8372	-0.7478	-0.7039
Bias_b6	0.8701	0.7251	0.6654	0.7384	0.6689	0.6352	0.5157	0.4846	0.4691
Bias_b7	-1.4572	-1.2449	-1.1191	-1.3426	-1.1439	-1.0389	-0.9417	-0.7871	-0.7110
Bias_b8	-1.1928	-0.9800	-0.8825	-1.0436	-0.7572	-0.6720	-0.6750	-0.5339	-0.4561
Var_b1	0.0400	0.0336	0.0308	0.0114	0.0074	0.0055	0.0010	0.0010	0.0008
Var_b2	0.0288	0.0194	0.0173	0.0053	0.0039	0.0037	0.0016	0.0013	0.0011
Var_b3	0.1571	0.1072	0.0751	0.0381	0.0255	0.0209	0.0064	0.0043	0.0036
Var_b4	1.3137	1.0290	0.8289	0.0753	0.0506	0.0362	0.0155	0.0098	0.0087
Var_b5	0.0542	0.0475	0.0394	0.0136	0.0114	0.0106	0.0034	0.0033	0.0025
Var_b6	0.0217	0.0199	0.0169	0.0053	0.0040	0.0041	0.0015	0.0014	0.0013
Var_b7	0.2085	0.1504	0.1300	0.0367	0.0243	0.0217	0.0064	0.0050	0.0037
Var_b8	0.4565	0.2883	0.2434	0.0560	0.0448	0.0341	0.0122	0.0098	0.0074

Tabelle 12: Es werden MSE, Bias und Varianz für die einzelnen Koeffizienten (b1, ..., b8) unter unterschiedlich großen Fehlend-Anteilen und unter der jeweiligen Datensatzgröße gezeigt.

Fehlendmechanismus LogitY mit normalen Verfahren (lm)

	n=50			n=200			n=1000		
	50%	25%	10%	50%	25%	10%	50%	25%	10%
MSE_b1	0.0583	0.0335	0.0281	0.0179	0.0087	0.0060	0.0089	0.0037	0.0015
MSE_b2	0.1311	0.0531	0.0358	0.0394	0.0152	0.0072	0.0300	0.0143	0.0046
MSE_b3	0.2597	0.1263	0.0980	0.0462	0.0296	0.0252	0.0156	0.0090	0.0050
MSE_b4	1.8942	1.0380	0.8900	0.0723	0.0458	0.0306	0.0155	0.0085	0.0066
MSE_b5	0.1996	0.1176	0.0950	0.0740	0.0337	0.0206	0.0463	0.0211	0.0076
MSE_b6	0.1051	0.0625	0.0455	0.0417	0.0201	0.0114	0.0342	0.0157	0.0053
MSE_b7	0.4301	0.2778	0.2098	0.1006	0.0498	0.0321	0.0360	0.0168	0.0073
MSE_b8	0.9651	0.9035	0.7220	0.0942	0.0371	0.0325	0.0295	0.0093	0.0067
Bias_b1	0.0896	0.0364	0.0099	0.0825	0.0551	0.0217	0.0860	0.0522	0.0266
Bias_b2	-0.1702	-0.1092	-0.0626	-0.1578	-0.0907	-0.0445	-0.1645	-0.1123	-0.0571
Bias_b3	-0.0280	-0.0509	-0.0547	-0.0869	-0.0394	-0.0067	-0.0973	-0.0593	-0.0339
Bias_b4	-0.8702	-0.1730	0.0249	0.0296	0.0157	-0.0107	-0.0391	-0.0070	-0.0022
Bias_b5	-0.2053	-0.0646	-0.0182	-0.1873	-0.1006	-0.0408	-0.1972	-0.1297	-0.0623
Bias_b6	0.1843	0.0718	0.0148	0.1474	0.0859	0.0425	0.1732	0.1147	0.0574
Bias_b7	-0.1906	-0.0578	-0.0241	-0.1801	-0.1121	-0.0528	-0.1435	-0.0951	-0.0433
Bias_b8	0.7162	0.1505	0.0215	-0.0912	0.0045	0.0261	-0.1209	-0.0289	0.0072
Var_b1	0.0503	0.0322	0.0280	0.0111	0.0056	0.0055	0.0015	0.0010	0.0008
Var_b2	0.1021	0.0412	0.0318	0.0145	0.0070	0.0052	0.0030	0.0017	0.0014
Var_b3	0.2590	0.1237	0.0950	0.0386	0.0281	0.0251	0.0061	0.0054	0.0039
Var_b4	1.1369	1.0081	0.8894	0.0714	0.0456	0.0305	0.0140	0.0085	0.0066
Var_b5	0.1574	0.1134	0.0946	0.0389	0.0236	0.0189	0.0074	0.0043	0.0037
Var_b6	0.0712	0.0574	0.0453	0.0199	0.0127	0.0096	0.0042	0.0025	0.0021
Var_b7	0.3938	0.2745	0.2092	0.0682	0.0372	0.0294	0.0154	0.0078	0.0054
Var_b8	0.4522	0.8808	0.7215	0.0859	0.0371	0.0318	0.0148	0.0084	0.0067

Tabelle 13: Es werden MSE, Bias und Varianz für die einzelnen Koeffizienten (b1, ..., b8) unter unterschiedlich großen Fehlend-Anteilen und unter der jeweiligen Datensatzgröße gezeigt.

Fehlendmechanismus LogitY mit normalen Verfahren (ridge, $\lambda=0$)									
	n=50			n=200			n=1000		
	50%	25%	10%	50%	25%	10%	50%	25%	10%
MSE_b1	0.0557	0.0363	0.0312	0.0168	0.0082	0.0060	0.0078	0.0031	0.0014
MSE_b2	0.1295	0.0539	0.0362	0.0347	0.0140	0.0068	0.0298	0.0131	0.0042
MSE_b3	0.2593	0.1278	0.0998	0.0395	0.0279	0.0237	0.0128	0.0075	0.0046
MSE_b4	4.7552	3.2074	2.0994	0.1037	0.0551	0.0405	0.0238	0.0163	0.0104
MSE_b5	0.2008	0.1235	0.1053	0.0647	0.0361	0.0270	0.0345	0.0159	0.0073
MSE_b6	0.1053	0.0630	0.0485	0.0371	0.0192	0.0121	0.0313	0.0138	0.0051
MSE_b7	0.4404	0.3131	0.2892	0.0849	0.0558	0.0432	0.0233	0.0117	0.0075
MSE_b8	4.5638	2.3110	0.8886	0.1711	0.0752	0.0507	0.0445	0.0248	0.0149
Bias_b1	0.0826	0.0335	0.0099	0.0755	0.0496	0.0197	0.0791	0.0457	0.0233
Bias_b2	-0.1680	-0.1092	-0.0628	-0.1443	-0.0845	-0.0419	-0.1640	-0.1066	-0.0535
Bias_b3	-0.0281	-0.0511	-0.0530	-0.0447	-0.0202	-0.0004	-0.0813	-0.0469	-0.0273
Bias_b4	-0.1311	-0.1931	0.1097	0.1173	0.0540	0.0039	0.0795	0.0622	0.0287
Bias_b5	-0.1984	-0.0640	-0.0218	-0.1318	-0.0742	-0.0332	-0.1609	-0.1014	-0.0494
Bias_b6	0.1806	0.0720	0.0176	0.1294	0.0769	0.0406	0.1643	0.1052	0.0529
Bias_b7	-0.1655	-0.0510	-0.0267	-0.0813	-0.0661	-0.0376	-0.0723	-0.0460	-0.0207
Bias_b8	1.4212	0.4235	0.0512	0.1070	0.0784	0.0484	0.1172	0.0811	0.0493
Var_b1	0.0489	0.0352	0.0311	0.0111	0.0058	0.0056	0.0016	0.0010	0.0009
Var_b2	0.1013	0.0420	0.0323	0.0139	0.0068	0.0051	0.0029	0.0017	0.0014
Var_b3	0.2585	0.1252	0.0970	0.0375	0.0275	0.0237	0.0062	0.0053	0.0039
Var_b4	4.7380	3.1701	2.0873	0.0899	0.0522	0.0405	0.0175	0.0125	0.0096
Var_b5	0.1614	0.1194	0.1049	0.0473	0.0306	0.0259	0.0086	0.0056	0.0048
Var_b6	0.0727	0.0578	0.0482	0.0203	0.0133	0.0104	0.0043	0.0028	0.0023
Var_b7	0.4130	0.3105	0.2885	0.0783	0.0514	0.0418	0.0181	0.0096	0.0070
Var_b8	2.5439	2.1317	0.8860	0.1597	0.0690	0.0483	0.0308	0.0182	0.0125

Tabelle 14: Es werden MSE, Bias und Varianz für die einzelnen Koeffizienten (b1, ..., b8) unter unterschiedlich großen Fehlend-Anteilen und unter der jeweiligen Datensatzgröße gezeigt.

Fehlendmechanismus LogitY mit normalen Verfahren (ridge)

	n=50			n=200			n=1000		
	50%	25%	10%	50%	25%	10%	50%	25%	10%
MSE_b1	0.0810	0.0485	0.0358	0.0494	0.0247	0.0117	0.0373	0.0213	0.0112
MSE_b2	0.2463	0.1109	0.0670	0.1566	0.0896	0.0493	0.1639	0.0955	0.0472
MSE_b3	0.2347	0.1188	0.0856	0.0931	0.0664	0.0454	0.0354	0.0209	0.0124
MSE_b4	0.8027	1.2388	1.3521	0.1502	0.0703	0.0394	0.0603	0.0392	0.0217
MSE_b5	1.4925	0.9492	0.8343	1.5457	1.2144	1.0054	1.0124	0.7824	0.6025
MSE_b6	0.9384	0.5863	0.4785	0.8286	0.6272	0.4935	0.5805	0.4267	0.3050
MSE_b7	1.9150	1.4001	1.2843	1.6494	1.3414	1.1360	0.7487	0.6230	0.5233
MSE_b8	0.7059	0.9843	0.9409	0.4112	0.3496	0.3529	0.1089	0.1167	0.1285
Bias_b1	0.1987	0.1304	0.0951	0.1986	0.1409	0.0832	0.1896	0.1428	0.1022
Bias_b2	-0.4117	-0.2802	-0.2055	-0.3832	-0.2900	-0.2127	-0.4022	-0.3070	-0.2148
Bias_b3	-0.1866	-0.1110	-0.0767	-0.2464	-0.2080	-0.1558	-0.1723	-0.1275	-0.0950
Bias_b4	0.6101	0.0510	0.0674	0.2648	0.1572	0.0752	0.2140	0.1693	0.1138
Bias_b5	-1.2013	-0.9525	-0.8912	-1.2379	-1.0975	-0.9975	-1.0043	-0.8831	-0.7745
Bias_b6	0.9556	0.7525	0.6791	0.9065	0.7888	0.6991	0.7604	0.6521	0.5510
Bias_b7	-1.3035	-1.1174	-1.0639	-1.2712	-1.1481	-1.0559	-0.8592	-0.7857	-0.7204
Bias_b8	-0.0554	-0.7056	-0.7341	-0.5597	-0.5510	-0.5676	-0.2928	-0.3218	-0.3445
Var_b1	0.0415	0.0315	0.0268	0.0100	0.0048	0.0048	0.0014	0.0009	0.0008
Var_b2	0.0768	0.0324	0.0247	0.0098	0.0055	0.0041	0.0021	0.0013	0.0011
Var_b3	0.1999	0.1065	0.0797	0.0324	0.0232	0.0211	0.0057	0.0047	0.0034
Var_b4	0.4305	1.2362	1.3475	0.0801	0.0456	0.0338	0.0145	0.0105	0.0087
Var_b5	0.0493	0.0421	0.0401	0.0134	0.0099	0.0104	0.0038	0.0025	0.0027
Var_b6	0.0253	0.0201	0.0173	0.0068	0.0049	0.0047	0.0023	0.0015	0.0014
Var_b7	0.2159	0.1514	0.1524	0.0335	0.0233	0.0210	0.0106	0.0057	0.0043
Var_b8	0.7029	0.4864	0.4020	0.0980	0.0460	0.0308	0.0232	0.0131	0.0099

Tabelle 15: Es werden MSE, Bias und Varianz für die einzelnen Koeffizienten (b1, ..., b8) unter unterschiedlich großen Fehlend-Anteilen und unter der jeweiligen Datensatzgröße gezeigt.

Fehlendmechanismus LogitX mit normalen Verfahren (lm)

	n=50			n=200			n=1000		
	50%	25%	10%	50%	25%	10%	50%	25%	10%
MSE_b1	0.0551	0.0303	0.0282	0.0142	0.0061	0.0060	0.0019	0.0011	0.0009
MSE_b2	0.0996	0.0410	0.0269	0.0150	0.0074	0.0050	0.0044	0.0020	0.0017
MSE_b3	0.2122	0.1277	0.0938	0.0389	0.0303	0.0266	0.0081	0.0054	0.0045
MSE_b4	1.3171	0.9918	0.8424	0.0885	0.0383	0.0293	0.0172	0.0093	0.0074
MSE_b5	0.1593	0.1166	0.0871	0.0311	0.0236	0.0200	0.0075	0.0045	0.0031
MSE_b6	0.0802	0.0574	0.0414	0.0156	0.0114	0.0097	0.0044	0.0025	0.0018
MSE_b7	0.3661	0.2496	0.1813	0.0654	0.0378	0.0297	0.0152	0.0065	0.0053
MSE_b8	1.0635	0.8635	0.6912	0.0823	0.0326	0.0295	0.0167	0.0091	0.0077
Bias_b1	-0.0036	0.0109	-0.0097	-0.0074	-0.0087	-0.0058	-0.0020	0.0006	0.0001
Bias_b2	0.0219	-0.0052	0.0088	-0.0068	0.0032	0.0026	-0.0013	-0.0013	-0.0001
Bias_b3	0.0454	0.0176	0.0154	0.0199	0.0312	0.0326	0.0061	-0.0010	-0.0036
Bias_b4	-0.7057	-0.1598	0.0122	-0.0068	-0.0170	-0.0211	-0.0063	-0.0003	-0.0022
Bias_b5	-0.0080	-0.0002	0.0127	0.0169	0.0148	0.0121	0.0001	-0.0016	0.0017
Bias_b6	0.0014	0.0024	-0.0066	-0.0149	-0.0073	-0.0033	0.0052	0.0044	0.0029
Bias_b7	-0.0484	-0.0423	-0.0075	0.0006	0.0059	0.0077	0.0140	0.0023	0.0038
Bias_b8	0.6986	0.1744	-0.0419	0.0278	0.0081	0.0054	-0.0208	-0.0089	-0.0076
Var_b1	0.0551	0.0302	0.0281	0.0142	0.0060	0.0060	0.0019	0.0011	0.0009
Var_b2	0.0991	0.0409	0.0268	0.0149	0.0074	0.0050	0.0044	0.0020	0.0017
Var_b3	0.2101	0.1274	0.0936	0.0385	0.0293	0.0255	0.0081	0.0054	0.0044
Var_b4	0.8191	0.9663	0.8423	0.0885	0.0380	0.0289	0.0172	0.0093	0.0074
Var_b5	0.1593	0.1166	0.0870	0.0308	0.0234	0.0199	0.0075	0.0045	0.0031
Var_b6	0.0802	0.0574	0.0414	0.0154	0.0114	0.0097	0.0044	0.0025	0.0018
Var_b7	0.3638	0.2478	0.1813	0.0654	0.0378	0.0296	0.0150	0.0065	0.0053
Var_b8	0.5755	0.8330	0.6895	0.0815	0.0325	0.0295	0.0162	0.0091	0.0076

Tabelle 16: Es werden MSE, Bias und Varianz für die einzelnen Koeffizienten (b1, ..., b8) unter unterschiedlich großen Fehlend-Anteilen und unter der jeweiligen Datensatzgröße gezeigt.

Fehlendmechanismus LogitX mit normalen Verfahren (ridge, $\lambda=0$)									
	n=50			n=200			n=1000		
	50%	25%	10%	50%	25%	10%	50%	25%	10%
MSE_b1	0.0546	0.0353	0.0323	0.0140	0.0061	0.0061	0.0020	0.0011	0.0009
MSE_b2	0.1000	0.0414	0.0267	0.0151	0.0071	0.0050	0.0044	0.0019	0.0017
MSE_b3	0.2143	0.1289	0.0938	0.0414	0.0294	0.0247	0.0082	0.0055	0.0044
MSE_b4	4.6512	4.1211	2.1522	0.1051	0.0542	0.0409	0.0221	0.0125	0.0101
MSE_b5	0.1637	0.1220	0.0986	0.0400	0.0321	0.0256	0.0083	0.0057	0.0041
MSE_b6	0.0815	0.0579	0.0454	0.0167	0.0124	0.0102	0.0045	0.0028	0.0021
MSE_b7	0.4113	0.3047	0.2482	0.0861	0.0568	0.0400	0.0169	0.0089	0.0063
MSE_b8	4.4791	3.0250	0.6846	0.1825	0.0577	0.0490	0.0354	0.0186	0.0139
Bias_b1	0.0003	0.0083	-0.0074	-0.0070	-0.0078	-0.0058	-0.0019	0.0002	-0.0003
Bias_b2	0.0224	-0.0058	0.0100	-0.0070	0.0025	0.0023	-0.0018	-0.0009	0.0003
Bias_b3	0.0472	0.0171	0.0176	0.0229	0.0286	0.0323	0.0060	0.0001	-0.0026
Bias_b4	-0.0088	-0.2907	-0.0116	-0.0028	-0.0202	-0.0209	-0.0056	0.0058	0.0026
Bias_b5	-0.0183	-0.0001	0.0048	0.0149	0.0053	0.0073	-0.0029	-0.0022	0.0014
Bias_b6	0.0078	0.0036	-0.0017	-0.0127	-0.0024	-0.0004	0.0073	0.0054	0.0036
Bias_b7	-0.0689	-0.0358	-0.0208	0.0012	-0.0061	0.0027	0.0117	0.0034	0.0050
Bias_b8	1.2930	0.5952	-0.0436	0.0455	-0.0035	0.0029	-0.0186	-0.0017	-0.0027
Var_b1	0.0546	0.0352	0.0323	0.0139	0.0061	0.0061	0.0020	0.0011	0.0009
Var_b2	0.0995	0.0413	0.0266	0.0151	0.0071	0.0049	0.0044	0.0019	0.0017
Var_b3	0.2120	0.1286	0.0935	0.0409	0.0286	0.0237	0.0081	0.0055	0.0044
Var_b4	4.6512	4.0366	2.1521	0.1051	0.0538	0.0405	0.0221	0.0124	0.0101
Var_b5	0.1634	0.1220	0.0986	0.0397	0.0321	0.0256	0.0083	0.0057	0.0041
Var_b6	0.0814	0.0578	0.0454	0.0166	0.0124	0.0102	0.0045	0.0028	0.0021
Var_b7	0.4065	0.3034	0.2478	0.0861	0.0568	0.0400	0.0167	0.0089	0.0063
Var_b8	2.8073	2.6708	0.6827	0.1805	0.0577	0.0490	0.0350	0.0186	0.0138

Tabelle 17: Es werden MSE, Bias und Varianz für die einzelnen Koeffizienten (b1, ..., b8) unter unterschiedlich großen Fehlend-Anteilen und unter der jeweiligen Datensatzgröße gezeigt.

Fehlendmechanismus LogitX mit normalen Verfahren (ridge)

	n=50			n=200			n=1000		
	50%	25%	10%	50%	25%	10%	50%	25%	10%
MSE_b1	0.0688	0.0412	0.0343	0.0301	0.0140	0.0090	0.0173	0.0118	0.0073
MSE_b2	0.1310	0.0676	0.0405	0.0818	0.0516	0.0334	0.0754	0.0484	0.0287
MSE_b3	0.1988	0.1059	0.0805	0.0743	0.0525	0.0379	0.0163	0.0123	0.0090
MSE_b4	0.7743	1.2069	1.4940	0.1171	0.0554	0.0380	0.0395	0.0260	0.0178
MSE_b5	1.3190	0.8942	0.7992	1.4140	1.1317	0.9546	0.8503	0.6824	0.5482
MSE_b6	0.7782	0.5195	0.4620	0.6984	0.5534	0.4526	0.4345	0.3400	0.2635
MSE_b7	1.9249	1.4532	1.2684	1.6829	1.2952	1.0901	0.7091	0.5990	0.5072
MSE_b8	0.8971	1.0118	1.0310	0.5621	0.4413	0.4125	0.2225	0.1832	0.1696
Bias_b1	0.1255	0.1040	0.0807	0.1341	0.0925	0.0587	0.1253	0.1037	0.0805
Bias_b2	-0.2419	-0.1918	-0.1391	-0.2647	-0.2150	-0.1710	-0.2691	-0.2167	-0.1655
Bias_b3	-0.1319	-0.0639	-0.0099	-0.1979	-0.1690	-0.1295	-0.0974	-0.0870	-0.0724
Bias_b4	0.5951	0.1359	-0.0759	0.1614	0.0803	0.0542	0.1430	0.1239	0.0923
Bias_b5	-1.1240	-0.9196	-0.8720	-1.1846	-1.0586	-0.9721	-0.9202	-0.8243	-0.7389
Bias_b6	0.8651	0.7052	0.6673	0.8321	0.7405	0.6698	0.6574	0.5818	0.5121
Bias_b7	-1.3157	-1.1409	-1.0688	-1.2838	-1.1259	-1.0346	-0.8364	-0.7706	-0.7092
Bias_b8	-0.0891	-0.7165	-0.7947	-0.6588	-0.6340	-0.6135	-0.4427	-0.4100	-0.3990
Var_b1	0.0531	0.0303	0.0278	0.0121	0.0055	0.0056	0.0016	0.0011	0.0008
Var_b2	0.0725	0.0308	0.0212	0.0117	0.0053	0.0042	0.0030	0.0015	0.0013
Var_b3	0.1814	0.1019	0.0804	0.0351	0.0240	0.0211	0.0068	0.0047	0.0038
Var_b4	0.4202	1.1885	1.4882	0.0910	0.0489	0.0351	0.0191	0.0107	0.0093
Var_b5	0.0557	0.0485	0.0388	0.0108	0.0110	0.0095	0.0036	0.0029	0.0023
Var_b6	0.0297	0.0221	0.0167	0.0060	0.0051	0.0040	0.0023	0.0016	0.0013
Var_b7	0.1940	0.1516	0.1262	0.0347	0.0275	0.0198	0.0095	0.0052	0.0042
Var_b8	0.8892	0.4984	0.3994	0.1281	0.0394	0.0361	0.0266	0.0151	0.0104

Tabelle 18: Es werden MSE, Bias und Varianz für die einzelnen Koeffizienten (b1, ..., b8) unter unterschiedlich großen Fehlend-Anteilen und unter der jeweiligen Datensatzgröße gezeigt.

C.2 Bruteforce1

Fehlendmechanismus MCAR mit Bruteforce1-Verfahren (lm)									
	n=50			n=200			n=1000		
	50%	25%	10%	50%	25%	10%	50%	25%	10%
MSE_b1	0.3559	0.1067	0.0507	0.2837	0.0843	0.0226	0.2676	0.0707	0.0156
MSE_b2	1.1214	0.2965	0.0658	1.0863	0.2497	0.0613	1.0026	0.2624	0.0620
MSE_b3	0.6126	0.3424	0.1498	0.3434	0.1305	0.0557	0.2557	0.0776	0.0291
MSE_b4	3.8360	2.1812	1.3184	0.3504	0.1442	0.0709	0.2250	0.0827	0.0114
MSE_b5	1.2072	0.3831	0.1213	1.2259	0.3384	0.1005	1.0102	0.2655	0.0858
MSE_b6	1.1706	0.3674	0.0717	1.1561	0.2960	0.0681	1.0014	0.2643	0.0647
MSE_b7	0.7708	0.4534	0.2749	0.4378	0.1494	0.1148	0.2985	0.0862	0.0567
MSE_b8	1.9914	1.9249	1.3347	0.3298	0.1115	0.0492	0.2479	0.0645	0.0159
Bias_b1	0.4586	0.1793	0.0747	0.5172	0.2625	0.1273	0.5144	0.2609	0.1175
Bias_b2	-1.0206	-0.4960	-0.1626	-1.0325	-0.4846	-0.2306	-0.9995	-0.5086	-0.2427
Bias_b3	-0.5502	-0.2825	-0.1618	-0.5164	-0.2562	-0.1327	-0.4889	-0.2549	-0.1428
Bias_b4	1.0547	0.4039	-0.2602	0.4811	0.2401	0.0414	0.4453	0.2558	-0.0058
Bias_b5	-0.9600	-0.4104	-0.0376	-1.0752	-0.5418	-0.2481	-0.9990	-0.5045	-0.2815
Bias_b6	0.9781	0.4458	0.0463	1.0579	0.5200	0.2235	0.9979	0.5094	0.2464
Bias_b7	-0.3121	-0.1489	-0.0223	-0.5641	-0.2650	-0.2323	-0.5269	-0.2599	-0.2123
Bias_b8	0.5361	0.3146	0.4723	0.4456	0.1940	0.0069	0.4613	0.2159	0.0242
Var_b1	0.1456	0.0746	0.0451	0.0162	0.0154	0.0064	0.0030	0.0026	0.0018
Var_b2	0.0798	0.0505	0.0394	0.0203	0.0149	0.0081	0.0036	0.0037	0.0031
Var_b3	0.3098	0.2626	0.1236	0.0767	0.0649	0.0381	0.0167	0.0127	0.0087
Var_b4	2.7236	2.0180	1.2507	0.1189	0.0865	0.0692	0.0267	0.0173	0.0113
Var_b5	0.2857	0.2147	0.1199	0.0699	0.0449	0.0389	0.0122	0.0110	0.0065
Var_b6	0.2139	0.1687	0.0696	0.0370	0.0256	0.0181	0.0055	0.0048	0.0040
Var_b7	0.6734	0.4312	0.2744	0.1196	0.0792	0.0608	0.0208	0.0187	0.0116
Var_b8	1.7040	1.8260	1.1116	0.1312	0.0738	0.0492	0.0351	0.0179	0.0153

Tabelle 19: Es werden MSE, Bias und Varianz für die einzelnen Koeffizienten (b1, ..., b8) unter unterschiedlich großen Fehlend-Anteilen und unter der jeweiligen Datensatzgröße gezeigt.

Fehlendmechanismus MCAR mit Bruteforce1-Verfahren (ridge, $\lambda=0$)

	n=50			n=200			n=1000		
	50%	25%	10%	50%	25%	10%	50%	25%	10%
MSE_b1	0.4252	0.1111	0.0422	0.2827	0.0850	0.0202	0.2629	0.0692	0.0122
MSE_b2	1.1432	0.2959	0.0560	1.0760	0.2476	0.0542	0.9885	0.2607	0.0539
MSE_b3	0.5835	0.3305	0.1493	0.3290	0.1399	0.0529	0.2459	0.0759	0.0218
MSE_b4	5.5221	5.2473	3.0347	0.4302	0.1849	0.0882	0.2538	0.0954	0.0310
MSE_b5	1.4122	0.4210	0.1122	1.1706	0.3524	0.0875	0.9668	0.2592	0.0571
MSE_b6	1.2542	0.3975	0.0745	1.1284	0.2994	0.0600	0.9821	0.2609	0.0538
MSE_b7	0.9319	0.5210	0.3034	0.3859	0.1644	0.1079	0.2609	0.0781	0.0261
MSE_b8	1.9590	1.9444	1.4380	0.4810	0.1938	0.1305	0.3201	0.0864	0.0576
Bias_b1	0.5134	0.1732	0.0468	0.5149	0.2602	0.1180	0.5097	0.2576	0.1021
Bias_b2	-1.0380	-0.5013	-0.1461	-1.0277	-0.4830	-0.2157	-0.9927	-0.5069	-0.2253
Bias_b3	-0.5290	-0.2814	-0.1502	-0.5012	-0.2559	-0.1025	-0.4796	-0.2507	-0.1162
Bias_b4	0.5330	0.3186	-0.0891	0.5308	0.2610	0.1204	0.4818	0.2772	0.1127
Bias_b5	-1.0672	-0.4297	-0.0036	-1.0424	-0.5364	-0.1939	-0.9760	-0.4955	-0.2209
Bias_b6	1.0229	0.4546	0.0335	1.0435	0.5204	0.2012	0.9880	0.5057	0.2221
Bias_b7	-0.5118	-0.1983	0.0573	-0.5114	-0.2525	-0.1469	-0.4854	-0.2417	-0.1136
Bias_b8	0.5249	0.3397	0.5072	0.5135	0.2141	0.1147	0.5256	0.2394	0.1740
Var_b1	0.1617	0.0811	0.0401	0.0176	0.0173	0.0063	0.0031	0.0028	0.0018
Var_b2	0.0659	0.0446	0.0346	0.0199	0.0143	0.0076	0.0031	0.0037	0.0031
Var_b3	0.3037	0.2514	0.1268	0.0778	0.0744	0.0424	0.0159	0.0131	0.0083
Var_b4	5.2380	5.1458	3.0268	0.1484	0.1168	0.0737	0.0217	0.0185	0.0183
Var_b5	0.2733	0.2363	0.1122	0.0841	0.0647	0.0499	0.0142	0.0136	0.0083
Var_b6	0.2079	0.1908	0.0734	0.0394	0.0286	0.0196	0.0060	0.0052	0.0045
Var_b7	0.6699	0.4817	0.3001	0.1244	0.1007	0.0863	0.0253	0.0197	0.0132
Var_b8	1.6835	1.8290	1.1807	0.2173	0.1480	0.1174	0.0439	0.0291	0.0273

Tabelle 20: Es werden MSE, Bias und Varianz für die einzelnen Koeffizienten (b1, ..., b8) unter unterschiedlich großen Fehlend-Anteilen und unter der jeweiligen Datensatzgröße gezeigt.

Fehlendmechanismus MCAR mit Bruteforce1-Verfahren (ridge)

	n=50			n=200			n=1000		
	50%	25%	10%	50%	25%	10%	50%	25%	10%
MSE_b1	0.5811	0.1754	0.0554	0.3684	0.0975	0.0275	0.2779	0.0854	0.0238
MSE_b2	2.3553	0.7437	0.1312	1.4754	0.3302	0.1058	1.0466	0.3215	0.1006
MSE_b3	0.6715	0.3278	0.1338	0.4238	0.1690	0.0717	0.2576	0.0865	0.0307
MSE_b4	1.1092	1.9760	1.8193	0.4256	0.1627	0.0747	0.2724	0.1156	0.0431
MSE_b5	3.4256	2.0561	0.9651	2.5648	1.2002	0.8771	1.3620	0.7306	0.5474
MSE_b6	2.9432	1.5740	0.6106	2.1162	0.8032	0.4897	1.2407	0.5530	0.3259
MSE_b7	1.5892	1.7631	1.3212	1.1123	0.7587	0.9126	0.4747	0.3729	0.4062
MSE_b8	0.3125	0.7735	0.8300	0.1787	0.1399	0.2260	0.2050	0.0260	0.0445
Bias_b1	0.7312	0.3468	0.1332	0.5959	0.2856	0.1470	0.5246	0.2878	0.1487
Bias_b2	-1.5239	-0.8242	-0.2924	-1.2072	-0.5629	-0.3155	-1.0215	-0.5641	-0.3130
Bias_b3	-0.7761	-0.4033	-0.1508	-0.6119	-0.3225	-0.1846	-0.4919	-0.2709	-0.1521
Bias_b4	0.5287	0.1027	-0.2760	0.5761	0.2469	0.1033	0.5005	0.3128	0.1618
Bias_b5	-1.8442	-1.4035	-0.9459	-1.5890	-1.0793	-0.9253	-1.1617	-0.8497	-0.7367
Bias_b6	1.7075	1.2193	0.7443	1.4465	0.8867	0.6934	1.1114	0.7412	0.5683
Bias_b7	-1.2313	-1.2699	-1.0741	-1.0287	-0.8373	-0.9314	-0.6733	-0.5991	-0.6304
Bias_b8	0.2886	-0.3732	-0.3981	0.2820	-0.1467	-0.3798	0.4021	0.0171	-0.1486
Var_b1	0.0465	0.0551	0.0377	0.0133	0.0159	0.0059	0.0027	0.0026	0.0016
Var_b2	0.0330	0.0643	0.0457	0.0181	0.0133	0.0063	0.0031	0.0033	0.0027
Var_b3	0.0692	0.1652	0.1111	0.0494	0.0650	0.0377	0.0156	0.0131	0.0076
Var_b4	0.8297	1.9655	1.7431	0.0937	0.1018	0.0641	0.0219	0.0177	0.0170
Var_b5	0.0246	0.0864	0.0705	0.0398	0.0353	0.0210	0.0124	0.0087	0.0047
Var_b6	0.0275	0.0872	0.0566	0.0238	0.0170	0.0089	0.0055	0.0037	0.0029
Var_b7	0.0730	0.1506	0.1676	0.0540	0.0577	0.0451	0.0213	0.0141	0.0088
Var_b8	0.2292	0.6342	0.6716	0.0992	0.1184	0.0818	0.0433	0.0257	0.0225

Tabelle 21: Es werden MSE, Bias und Varianz für die einzelnen Koeffizienten (b1, ..., b8) unter unterschiedlich großen Fehlend-Anteilen und unter der jeweiligen Datensatzgröße gezeigt.

Fehlendmechanismus MNAR1 mit Bruteforce1-Verfahren (lm)

	n=50			n=200			n=1000		
	50%	25%	10%	50%	25%	10%	50%	25%	10%
MSE_b1	0.7181	0.1392	0.0618	1.6961	0.1254	0.0245	1.5101	0.0971	0.0148
MSE_b2	2.4481	0.3704	0.0696	4.9656	0.3605	0.0614	6.3475	0.3975	0.0610
MSE_b3	1.0741	0.3282	0.1247	2.3417	0.2676	0.0626	1.9990	0.1342	0.0257
MSE_b4	2.5756	2.8806	2.0761	0.2039	0.1206	0.0811	0.0291	0.0108	0.0110
MSE_b5	1.8744	0.2398	0.1262	7.8839	0.6112	0.1132	8.8141	0.5388	0.0878
MSE_b6	1.4688	0.1864	0.0712	5.3207	0.4211	0.0720	6.3088	0.3931	0.0638
MSE_b7	1.0048	0.3303	0.2811	7.4264	0.5687	0.1343	4.9604	0.2861	0.0493
MSE_b8	4.1440	2.7678	1.6613	0.1026	0.1176	0.0616	0.0908	0.0216	0.0142
Bias_b1	0.7977	0.2655	0.0899	1.2983	0.3377	0.1265	1.2278	0.3051	0.1141
Bias_b2	-1.5203	-0.5426	-0.1865	-2.2258	-0.5876	-0.2229	-2.5179	-0.6249	-0.2418
Bias_b3	-0.9555	-0.4170	-0.1518	-1.5223	-0.4672	-0.1472	-1.4107	-0.3434	-0.1346
Bias_b4	-1.1712	-0.4566	-0.2877	0.4050	0.1074	0.0494	-0.1377	-0.0198	-0.0148
Bias_b5	-1.2521	-0.2428	-0.0433	-2.8031	-0.7335	-0.2670	-2.9658	-0.7243	-0.2842
Bias_b6	1.1152	0.2468	0.0757	2.3031	0.6163	0.2313	2.5099	0.6202	0.2457
Bias_b7	-0.8327	-0.0969	-0.0021	-2.7154	-0.6833	-0.2466	-2.2221	-0.5165	-0.2005
Bias_b8	1.8473	0.9926	0.4569	-0.2712	0.0752	0.0098	-0.2849	0.0649	0.0220
Var_b1	0.0818	0.0687	0.0537	0.0104	0.0113	0.0085	0.0026	0.0040	0.0018
Var_b2	0.1368	0.0760	0.0348	0.0113	0.0152	0.0117	0.0078	0.0070	0.0025
Var_b3	0.1610	0.1543	0.1016	0.0242	0.0494	0.0409	0.0089	0.0163	0.0076
Var_b4	1.2039	2.6721	1.9934	0.0398	0.1091	0.0786	0.0102	0.0104	0.0107
Var_b5	0.3067	0.1809	0.1243	0.0265	0.0732	0.0419	0.0182	0.0143	0.0070
Var_b6	0.2250	0.1255	0.0655	0.0166	0.0413	0.0185	0.0093	0.0085	0.0034
Var_b7	0.3114	0.3209	0.2811	0.0531	0.1018	0.0735	0.0226	0.0193	0.0091
Var_b8	0.7317	1.7824	1.4526	0.0290	0.1120	0.0615	0.0097	0.0173	0.0137

Tabelle 22: Es werden MSE, Bias und Varianz für die einzelnen Koeffizienten (b1, ..., b8) unter unterschiedlich großen Fehlend-Anteilen und unter der jeweiligen Datensatzgröße gezeigt.

Fehlendmechanismus MNAR1 mit Bruteforce1-Verfahren (ridge, $\lambda=0$)

	n=50			n=200			n=1000		
	50%	25%	10%	50%	25%	10%	50%	25%	10%
MSE_b1	0.3921	0.0853	0.0508	1.2949	0.1022	0.0223	1.1301	0.0753	0.0116
MSE_b2	2.1934	0.3139	0.0603	4.0392	0.3116	0.0542	5.3583	0.3406	0.0530
MSE_b3	1.0657	0.3490	0.1269	1.1007	0.1901	0.0559	1.1393	0.0854	0.0172
MSE_b4	3.1301	7.7146	4.6083	2.3864	0.2244	0.1205	1.5134	0.0978	0.0302
MSE_b5	1.0185	0.1325	0.1126	3.7072	0.3913	0.0878	4.9083	0.3277	0.0549
MSE_b6	1.0896	0.1439	0.0691	3.7820	0.3354	0.0584	4.8914	0.3163	0.0520
MSE_b7	0.2790	0.2910	0.2938	1.6372	0.2756	0.1115	1.0842	0.0831	0.0170
MSE_b8	4.6989	2.8980	1.7101	2.1896	0.4145	0.1335	2.2953	0.2430	0.0582
Bias_b1	0.5750	0.1861	0.0631	1.1341	0.3014	0.1157	1.0621	0.2681	0.0984
Bias_b2	-1.4375	-0.5039	-0.1757	-2.0076	-0.5449	-0.2068	-2.3134	-0.5780	-0.2244
Bias_b3	-0.9573	-0.4346	-0.1548	-1.0384	-0.3694	-0.1118	-1.0630	-0.2648	-0.1009
Bias_b4	0.8354	0.2918	-0.0661	1.5260	0.3328	0.1382	1.2241	0.2833	0.1084
Bias_b5	-0.8936	-0.1111	-0.0079	-1.9190	-0.5570	-0.1965	-2.2125	-0.5602	-0.2187
Bias_b6	0.9485	0.1871	0.0586	1.9411	0.5437	0.2014	2.2100	0.5553	0.2200
Bias_b7	-0.1539	0.1324	0.0640	-1.2591	-0.3893	-0.1329	-1.0324	-0.2504	-0.0937
Bias_b8	1.9810	1.1001	0.4804	1.4606	0.4328	0.1501	1.5109	0.4592	0.1825
Var_b1	0.0615	0.0507	0.0468	0.0087	0.0114	0.0089	0.0021	0.0034	0.0019
Var_b2	0.1269	0.0600	0.0294	0.0089	0.0147	0.0114	0.0067	0.0066	0.0026
Var_b3	0.1494	0.1602	0.1029	0.0225	0.0537	0.0434	0.0094	0.0153	0.0070
Var_b4	2.4322	7.6295	4.6040	0.0578	0.1137	0.1014	0.0149	0.0176	0.0184
Var_b5	0.2199	0.1202	0.1125	0.0248	0.0812	0.0492	0.0132	0.0139	0.0071
Var_b6	0.1899	0.1089	0.0657	0.0142	0.0397	0.0179	0.0074	0.0079	0.0036
Var_b7	0.2553	0.2735	0.2897	0.0519	0.1241	0.0939	0.0183	0.0203	0.0082
Var_b8	0.7746	1.6878	1.4793	0.0562	0.2272	0.1110	0.0125	0.0321	0.0249

Tabelle 23: Es werden MSE, Bias und Varianz für die einzelnen Koeffizienten (b1, ..., b8) unter unterschiedlich großen Fehlend-Anteilen und unter der jeweiligen Datensatzgröße gezeigt.

Fehlendmechanismus MNAR1 mit Bruteforce1-Verfahren (ridge)

	n=50			n=200			n=1000		
	50%	25%	10%	50%	25%	10%	50%	25%	10%
MSE_b1	0.6754	0.1789	0.0688	1.0734	0.1164	0.0284	1.0960	0.0911	0.0225
MSE_b2	3.0258	0.7581	0.1410	4.0106	0.4114	0.1041	4.9749	0.4060	0.0990
MSE_b3	0.8824	0.3191	0.1156	1.0206	0.2324	0.0760	1.0975	0.0954	0.0261
MSE_b4	1.2230	3.3147	2.5745	1.2761	0.2102	0.1073	1.3580	0.1168	0.0426
MSE_b5	3.4285	1.7886	0.9922	3.9640	1.2579	0.8859	4.3397	0.8051	0.5472
MSE_b6	3.1633	1.3571	0.6668	3.9812	0.8621	0.4924	4.4742	0.6095	0.3239
MSE_b7	1.2360	1.5365	1.3063	1.1756	1.0095	0.9112	0.9028	0.3655	0.3799
MSE_b8	1.0190	0.7091	0.9860	1.2547	0.1888	0.2063	2.0892	0.0889	0.0403
Bias_b1	0.8029	0.3620	0.1584	1.0354	0.3258	0.1425	1.0463	0.2963	0.1440
Bias_b2	-1.7231	-0.8176	-0.3234	-2.0025	-0.6305	-0.3081	-2.2287	-0.6328	-0.3112
Bias_b3	-0.9222	-0.4654	-0.1543	-1.0094	-0.4288	-0.1957	-1.0451	-0.2859	-0.1391
Bias_b4	0.8231	-0.0114	-0.2903	1.1254	0.3294	0.1246	1.1613	0.3183	0.1588
Bias_b5	-1.8437	-1.3033	-0.9603	-1.9909	-1.0994	-0.9303	-2.0820	-0.8927	-0.7370
Bias_b6	1.7638	1.1196	0.7820	1.9952	0.9138	0.6960	2.1144	0.7774	0.5673
Bias_b7	-1.0959	-1.1943	-1.0686	-1.0828	-0.9679	-0.9298	-0.9469	-0.5915	-0.6118
Bias_b8	0.9339	0.0353	-0.4238	1.1164	0.0728	-0.3489	1.4401	0.2513	-0.1434
Var_b1	0.0307	0.0478	0.0437	0.0013	0.0103	0.0081	0.0014	0.0033	0.0018
Var_b2	0.0567	0.0896	0.0364	0.0006	0.0139	0.0092	0.0077	0.0056	0.0021
Var_b3	0.0319	0.1025	0.0918	0.0017	0.0485	0.0378	0.0052	0.0137	0.0068
Var_b4	0.5455	3.3146	2.4903	0.0096	0.1017	0.0918	0.0093	0.0155	0.0174
Var_b5	0.0293	0.0898	0.0701	0.0003	0.0493	0.0204	0.0047	0.0083	0.0041
Var_b6	0.0523	0.1035	0.0553	0.0002	0.0271	0.0080	0.0037	0.0052	0.0021
Var_b7	0.0350	0.1101	0.1643	0.0031	0.0726	0.0466	0.0063	0.0157	0.0056
Var_b8	0.1468	0.7078	0.8064	0.0084	0.1835	0.0845	0.0152	0.0257	0.0197

Tabelle 24: Es werden MSE, Bias und Varianz für die einzelnen Koeffizienten (b1, ..., b8) unter unterschiedlich großen Fehlend-Anteilen und unter der jeweiligen Datensatzgröße gezeigt.

Fehlendmechanismus MNAR2 mit Bruteforce1-Verfahren (lm)

	n=50			n=200			n=1000		
	50%	25%	10%	50%	25%	10%	50%	25%	10%
MSE_b1	0.5859	0.1663	0.0588	1.0048	0.1179	0.0257	0.9403	0.0812	0.0139
MSE_b2	2.3821	0.4103	0.0752	3.2486	0.3602	0.0600	3.8451	0.3716	0.0584
MSE_b3	1.1359	0.4115	0.1632	1.8471	0.2215	0.0567	1.2569	0.1289	0.0226
MSE_b4	2.8691	2.5719	1.6130	0.1320	0.1451	0.0605	0.0484	0.0135	0.0115
MSE_b5	1.2765	0.1924	0.1334	4.7147	0.5043	0.0885	5.5908	0.4911	0.0681
MSE_b6	1.0426	0.1519	0.0606	3.2728	0.3640	0.0603	3.9676	0.3620	0.0529
MSE_b7	0.8897	0.3419	0.2996	4.5060	0.4999	0.1038	3.4043	0.2598	0.0428
MSE_b8	5.7090	2.9935	0.9422	0.0834	0.1016	0.0555	0.0370	0.0344	0.0181
Bias_b1	0.7085	0.2552	0.0740	0.9949	0.3261	0.1293	0.9680	0.2782	0.1101
Bias_b2	-1.4988	-0.5783	-0.1818	-1.7989	-0.5860	-0.2224	-1.9596	-0.6052	-0.2367
Bias_b3	-0.9745	-0.4740	-0.1374	-1.3430	-0.4205	-0.1495	-1.1155	-0.3437	-0.1227
Bias_b4	-1.1439	-0.4061	0.0148	0.2839	0.1906	0.0592	-0.1695	0.0118	-0.0015
Bias_b5	-1.0113	-0.2165	-0.0656	-2.1610	-0.6725	-0.2463	-2.3614	-0.6899	-0.2504
Bias_b6	0.9401	0.2414	0.0628	1.8019	0.5787	0.2174	1.9895	0.5951	0.2239
Bias_b7	-0.7151	-0.0704	-0.0351	-2.1026	-0.6440	-0.2256	-1.8387	-0.4887	-0.1783
Bias_b8	2.1266	1.0848	0.2485	-0.1250	0.0982	0.0483	-0.1346	0.1302	0.0551
Var_b1	0.0839	0.1012	0.0533	0.0150	0.0116	0.0090	0.0033	0.0038	0.0018
Var_b2	0.1358	0.0758	0.0422	0.0124	0.0168	0.0105	0.0051	0.0054	0.0024
Var_b3	0.1861	0.1868	0.1443	0.0435	0.0447	0.0343	0.0125	0.0108	0.0076
Var_b4	1.5605	2.4069	1.6128	0.0514	0.1088	0.0570	0.0197	0.0133	0.0115
Var_b5	0.2538	0.1455	0.1291	0.0449	0.0520	0.0279	0.0146	0.0152	0.0054
Var_b6	0.1589	0.0937	0.0567	0.0260	0.0292	0.0130	0.0095	0.0079	0.0027
Var_b7	0.3784	0.3370	0.2983	0.0850	0.0852	0.0529	0.0235	0.0210	0.0110
Var_b8	1.1868	1.8167	0.8805	0.0677	0.0919	0.0531	0.0189	0.0174	0.0150

Tabelle 25: Es werden MSE, Bias und Varianz für die einzelnen Koeffizienten (b1, ..., b8) unter unterschiedlich großen Fehlend-Anteilen und unter der jeweiligen Datensatzgröße gezeigt.

Fehlendmechanismus MNAR2 mit Bruteforce1-Verfahren (ridge, $\lambda=0$)

	n=50			n=200			n=1000		
	50%	25%	10%	50%	25%	10%	50%	25%	10%
MSE_b1	0.3019	0.1040	0.0497	0.8023	0.1010	0.0236	0.7049	0.0648	0.0115
MSE_b2	2.0843	0.3507	0.0619	2.7553	0.3120	0.0546	3.2440	0.3248	0.0513
MSE_b3	1.1635	0.3929	0.1646	1.1007	0.1675	0.0534	0.7185	0.0839	0.0170
MSE_b4	3.7663	6.5359	4.4192	1.1243	0.2859	0.0807	0.8453	0.0964	0.0264
MSE_b5	0.6285	0.1185	0.1027	2.5613	0.3269	0.0799	3.1450	0.3115	0.0454
MSE_b6	0.7280	0.1235	0.0532	2.4778	0.2954	0.0549	3.0858	0.2983	0.0445
MSE_b7	0.4062	0.3954	0.2754	1.4090	0.2562	0.0995	0.8366	0.0895	0.0217
MSE_b8	6.4098	3.1805	0.9458	1.1896	0.3614	0.1365	1.6790	0.2617	0.0639
Bias_b1	0.4838	0.1639	0.0367	0.8869	0.2962	0.1205	0.8378	0.2480	0.0975
Bias_b2	-1.3998	-0.5388	-0.1642	-1.6562	-0.5442	-0.2107	-1.7999	-0.5656	-0.2211
Bias_b3	-0.9806	-0.4555	-0.1322	-1.0267	-0.3339	-0.1270	-0.8401	-0.2732	-0.0965
Bias_b4	0.8986	0.3935	0.3102	1.0259	0.3990	0.1146	0.9051	0.2775	0.1047
Bias_b5	-0.6401	-0.0775	-0.0098	-1.5812	-0.5152	-0.2121	-1.7703	-0.5454	-0.1971
Bias_b6	0.7620	0.1766	0.0389	1.5650	0.5147	0.2038	1.7544	0.5398	0.2036
Bias_b7	0.0031	0.2142	0.0738	-1.1449	-0.3829	-0.1599	-0.9025	-0.2629	-0.0903
Bias_b8	2.2768	1.1474	0.2619	1.0320	0.4258	0.1327	1.2845	0.4708	0.1888
Var_b1	0.0678	0.0771	0.0483	0.0158	0.0132	0.0090	0.0031	0.0033	0.0019
Var_b2	0.1248	0.0604	0.0349	0.0122	0.0159	0.0102	0.0044	0.0049	0.0024
Var_b3	0.2020	0.1854	0.1471	0.0465	0.0559	0.0373	0.0127	0.0093	0.0077
Var_b4	2.9589	6.3811	4.3229	0.0718	0.1267	0.0676	0.0261	0.0194	0.0155
Var_b5	0.2188	0.1125	0.1026	0.0610	0.0615	0.0349	0.0110	0.0141	0.0065
Var_b6	0.1474	0.0923	0.0517	0.0285	0.0305	0.0134	0.0077	0.0069	0.0030
Var_b7	0.4062	0.3495	0.2700	0.0982	0.1096	0.0739	0.0220	0.0204	0.0135
Var_b8	1.2261	1.8640	0.8772	0.1246	0.1801	0.1189	0.0291	0.0400	0.0283

Tabelle 26: Es werden MSE, Bias und Varianz für die einzelnen Koeffizienten (b1, ..., b8) unter unterschiedlich großen Fehlend-Anteilen und unter der jeweiligen Datensatzgröße gezeigt.

Fehlendmechanismus MNAR2 mit Bruteforce1-Verfahren (ridge)

	n=50			n=200			n=1000		
	50%	25%	10%	50%	25%	10%	50%	25%	10%
MSE_b1	0.6131	0.1789	0.0572	0.8896	0.1117	0.0299	0.8273	0.0789	0.0228
MSE_b2	2.9344	0.8145	0.1321	3.4038	0.4077	0.1068	3.5651	0.3867	0.0975
MSE_b3	0.9097	0.3728	0.1356	1.0321	0.2000	0.0794	0.8302	0.0978	0.0246
MSE_b4	1.1670	2.3970	2.3942	1.0137	0.2684	0.0689	0.9050	0.1143	0.0381
MSE_b5	3.2381	1.7652	0.9652	3.7657	1.1568	0.9022	3.8098	0.7835	0.5230
MSE_b6	3.0314	1.3682	0.6018	3.5527	0.7985	0.4918	3.6454	0.5856	0.3095
MSE_b7	1.2713	1.4969	1.2680	1.4473	0.9994	0.9506	1.0609	0.3730	0.3859
MSE_b8	1.1169	0.8275	0.8319	0.8539	0.1404	0.2224	1.1752	0.0962	0.0405
Bias_b1	0.7641	0.3470	0.1264	0.9416	0.3170	0.1484	0.9081	0.2752	0.1448
Bias_b2	-1.6994	-0.8523	-0.3033	-1.8435	-0.6273	-0.3137	-1.8872	-0.6183	-0.3090
Bias_b3	-0.9347	-0.4948	-0.1187	-1.0122	-0.3902	-0.2134	-0.9082	-0.2986	-0.1328
Bias_b4	0.7653	0.0819	0.0091	1.0009	0.3836	0.1021	0.9470	0.3084	0.1559
Bias_b5	-1.7908	-1.2969	-0.9485	-1.9396	-1.0577	-0.9411	-1.9510	-0.8800	-0.7206
Bias_b6	1.7306	1.1308	0.7474	1.8839	0.8818	0.6963	1.9081	0.7621	0.5547
Bias_b7	-1.1028	-1.1618	-1.0498	-1.1979	-0.9624	-0.9513	-1.0276	-0.5980	-0.6138
Bias_b8	0.9617	0.0256	-0.5802	0.9135	0.0493	-0.3690	1.0796	0.2490	-0.1381
Var_b1	0.0292	0.0585	0.0412	0.0031	0.0112	0.0079	0.0026	0.0032	0.0018
Var_b2	0.0465	0.0881	0.0401	0.0054	0.0141	0.0084	0.0036	0.0044	0.0020
Var_b3	0.0360	0.1280	0.1216	0.0076	0.0478	0.0338	0.0054	0.0087	0.0070
Var_b4	0.5812	2.3903	2.3941	0.0118	0.1212	0.0585	0.0082	0.0192	0.0138
Var_b5	0.0310	0.0833	0.0655	0.0038	0.0381	0.0166	0.0036	0.0090	0.0036
Var_b6	0.0364	0.0895	0.0432	0.0038	0.0209	0.0069	0.0044	0.0047	0.0019
Var_b7	0.0551	0.1471	0.1660	0.0124	0.0732	0.0457	0.0049	0.0154	0.0091
Var_b8	0.1921	0.8269	0.4953	0.0195	0.1380	0.0863	0.0097	0.0342	0.0215

Tabelle 27: Es werden MSE, Bias und Varianz für die einzelnen Koeffizienten (b1, ..., b8) unter unterschiedlich großen Fehlend-Anteilen und unter der jeweiligen Datensatzgröße gezeigt.

Fehlendmechanismus MNAR3 mit Bruteforce1-Verfahren (lm)

	n=50			n=200			n=1000		
	50%	25%	10%	50%	25%	10%	50%	25%	10%
MSE_b1	0.0704	0.0592	0.0386	0.0180	0.0097	0.0067	0.0080	0.0028	0.0012
MSE_b2	0.0756	0.0382	0.0201	0.0223	0.0085	0.0056	0.0297	0.0087	0.0026
MSE_b3	0.1736	0.1537	0.0989	0.0293	0.0271	0.0261	0.0150	0.0076	0.0048
MSE_b4	0.5913	0.6220	0.5229	0.0507	0.0430	0.0356	0.0155	0.0105	0.0082
MSE_b5	0.1530	0.1268	0.1011	0.0509	0.0269	0.0218	0.0351	0.0120	0.0048
MSE_b6	0.0620	0.0579	0.0527	0.0287	0.0145	0.0100	0.0274	0.0088	0.0031
MSE_b7	0.4411	0.2857	0.2193	0.0583	0.0339	0.0316	0.0196	0.0105	0.0062
MSE_b8	0.7814	0.4881	0.4151	0.0546	0.0411	0.0293	0.0241	0.0122	0.0072
Bias_b1	-0.0605	-0.0341	-0.0141	0.1000	0.0527	0.0209	0.0803	0.0384	0.0177
Bias_b2	-0.2082	-0.1059	-0.0347	-0.1291	-0.0598	-0.0231	-0.1682	-0.0858	-0.0353
Bias_b3	-0.0127	-0.0462	-0.0241	0.0159	0.0066	0.0237	-0.0877	-0.0471	-0.0166
Bias_b4	0.0710	0.0684	0.0485	0.0753	0.0388	0.0102	0.0424	0.0119	0.0077
Bias_b5	0.1628	0.0748	0.0489	-0.1467	-0.0697	-0.0351	-0.1706	-0.0860	-0.0360
Bias_b6	-0.0767	-0.0213	-0.0277	0.1331	0.0673	0.0351	0.1580	0.0808	0.0358
Bias_b7	0.3220	0.1544	0.0724	-0.1532	-0.0686	-0.0398	-0.0883	-0.0430	-0.0178
Bias_b8	0.6350	0.2576	0.0648	0.0186	0.0071	0.0043	-0.0414	-0.0154	-0.0134
Var_b1	0.0668	0.0580	0.0384	0.0080	0.0070	0.0063	0.0016	0.0014	0.0009
Var_b2	0.0323	0.0270	0.0189	0.0056	0.0049	0.0051	0.0014	0.0014	0.0013
Var_b3	0.1734	0.1515	0.0983	0.0290	0.0270	0.0255	0.0073	0.0054	0.0046
Var_b4	0.5863	0.6173	0.5205	0.0450	0.0415	0.0355	0.0137	0.0104	0.0081
Var_b5	0.1265	0.1212	0.0988	0.0293	0.0221	0.0205	0.0060	0.0046	0.0036
Var_b6	0.0561	0.0574	0.0520	0.0110	0.0100	0.0088	0.0024	0.0022	0.0018
Var_b7	0.3375	0.2619	0.2141	0.0348	0.0292	0.0300	0.0118	0.0086	0.0059
Var_b8	0.3781	0.4218	0.4109	0.0543	0.0411	0.0293	0.0224	0.0120	0.0070

Tabelle 28: Es werden MSE, Bias und Varianz für die einzelnen Koeffizienten (b1, ..., b8) unter unterschiedlich großen Fehlend-Anteilen und unter der jeweiligen Datensatzgröße gezeigt.

Fehlendmechanismus MNAR3 mit Bruteforce1-Verfahren (ridge, $\lambda=0$)

	n=50			n=200			n=1000		
	50%	25%	10%	50%	25%	10%	50%	25%	10%
MSE_b1	0.0884	0.0698	0.0482	0.0156	0.0091	0.0067	0.0061	0.0026	0.0012
MSE_b2	0.0721	0.0387	0.0205	0.0155	0.0068	0.0056	0.0274	0.0078	0.0024
MSE_b3	0.1638	0.1613	0.1015	0.0377	0.0297	0.0243	0.0106	0.0069	0.0048
MSE_b4	0.8844	1.0436	1.0052	0.0973	0.0613	0.0469	0.0285	0.0144	0.0110
MSE_b5	0.2173	0.1663	0.1232	0.0398	0.0296	0.0261	0.0249	0.0111	0.0054
MSE_b6	0.0816	0.0684	0.0571	0.0216	0.0141	0.0108	0.0236	0.0079	0.0032
MSE_b7	0.6293	0.4381	0.3036	0.0662	0.0480	0.0424	0.0110	0.0092	0.0064
MSE_b8	0.7700	0.4737	0.4350	0.1133	0.0698	0.0483	0.0212	0.0148	0.0106
Bias_b1	-0.0585	-0.0328	-0.0128	0.0819	0.0429	0.0131	0.0697	0.0354	0.0168
Bias_b2	-0.2078	-0.1120	-0.0345	-0.1008	-0.0467	-0.0203	-0.1606	-0.0787	-0.0320
Bias_b3	-0.0350	-0.0555	-0.0284	0.0624	0.0314	0.0352	-0.0667	-0.0369	-0.0128
Bias_b4	0.0374	0.0745	0.0408	0.2107	0.0993	0.0337	0.1105	0.0479	0.0208
Bias_b5	0.1460	0.0838	0.0530	-0.0477	-0.0251	-0.0190	-0.1373	-0.0718	-0.0287
Bias_b6	-0.0642	-0.0288	-0.0305	0.0964	0.0502	0.0289	0.1453	0.0748	0.0343
Bias_b7	0.2918	0.1509	0.0766	0.0088	0.0072	-0.0100	-0.0341	-0.0167	-0.0049
Bias_b8	0.6150	0.2678	0.0696	0.2148	0.1109	0.0343	0.0433	0.0280	0.0069
Var_b1	0.0849	0.0687	0.0481	0.0089	0.0073	0.0066	0.0013	0.0013	0.0009
Var_b2	0.0289	0.0261	0.0193	0.0053	0.0046	0.0051	0.0016	0.0016	0.0014
Var_b3	0.1626	0.1582	0.1007	0.0339	0.0287	0.0230	0.0062	0.0055	0.0046
Var_b4	0.8830	1.0380	1.0035	0.0529	0.0514	0.0458	0.0163	0.0121	0.0106
Var_b5	0.1960	0.1593	0.1204	0.0376	0.0290	0.0257	0.0060	0.0059	0.0045
Var_b6	0.0775	0.0675	0.0562	0.0123	0.0116	0.0099	0.0025	0.0023	0.0020
Var_b7	0.5441	0.4154	0.2977	0.0661	0.0480	0.0423	0.0098	0.0089	0.0063
Var_b8	0.3918	0.4020	0.4301	0.0671	0.0575	0.0471	0.0194	0.0140	0.0106

Tabelle 29: Es werden MSE, Bias und Varianz für die einzelnen Koeffizienten (b1, ..., b8) unter unterschiedlich großen Fehlend-Anteilen und unter der jeweiligen Datensatzgröße gezeigt.

Fehlendmechanismus MNAR3 mit Bruteforce1-Verfahren (ridge)

	n=50			n=200			n=1000		
	50%	25%	10%	50%	25%	10%	50%	25%	10%
MSE_b1	0.0755	0.0654	0.0470	0.0208	0.0116	0.0082	0.0129	0.0074	0.0056
MSE_b2	0.2529	0.0885	0.0320	0.0422	0.0273	0.0232	0.0528	0.0289	0.0198
MSE_b3	0.1347	0.1298	0.0838	0.0299	0.0279	0.0244	0.0138	0.0093	0.0071
MSE_b4	0.5354	0.7480	0.8280	0.0767	0.0522	0.0423	0.0344	0.0213	0.0165
MSE_b5	1.1327	0.8583	0.7149	0.5569	0.6383	0.7370	0.3097	0.3486	0.3955
MSE_b6	0.7407	0.5153	0.4053	0.3144	0.3231	0.3544	0.1815	0.1789	0.1899
MSE_b7	1.4077	1.2496	1.1188	0.5765	0.7165	0.8515	0.2144	0.2916	0.3701
MSE_b8	0.3686	0.6050	0.8101	0.1251	0.2200	0.3071	0.0648	0.0996	0.1371
Bias_b1	0.1077	0.0707	0.0576	0.1118	0.0705	0.0499	0.1085	0.0790	0.0692
Bias_b2	-0.4800	-0.2627	-0.1246	-0.1931	-0.1524	-0.1371	-0.2262	-0.1655	-0.1361
Bias_b3	-0.1456	-0.0553	0.0086	-0.0136	-0.0590	-0.0636	-0.0912	-0.0681	-0.0553
Bias_b4	-0.0626	-0.1274	-0.1583	0.1848	0.0791	0.0188	0.1465	0.1004	0.0858
Bias_b5	-1.0229	-0.8708	-0.8110	-0.7328	-0.7894	-0.8516	-0.5527	-0.5875	-0.6265
Bias_b6	0.8405	0.6892	0.6158	0.5547	0.5635	0.5911	0.4239	0.4211	0.4343
Bias_b7	-1.0706	-0.9869	-0.9719	-0.7338	-0.8285	-0.9097	-0.4540	-0.5339	-0.6045
Bias_b8	-0.3842	-0.5880	-0.7365	-0.2603	-0.4238	-0.5228	-0.2207	-0.2966	-0.3587
Var_b1	0.0638	0.0604	0.0437	0.0083	0.0067	0.0057	0.0012	0.0012	0.0008
Var_b2	0.0225	0.0195	0.0164	0.0049	0.0041	0.0044	0.0016	0.0015	0.0012
Var_b3	0.1135	0.1268	0.0837	0.0297	0.0245	0.0203	0.0054	0.0046	0.0041
Var_b4	0.5315	0.7317	0.8030	0.0425	0.0459	0.0420	0.0129	0.0112	0.0092
Var_b5	0.0864	0.1000	0.0572	0.0199	0.0152	0.0119	0.0042	0.0035	0.0030
Var_b6	0.0343	0.0403	0.0261	0.0067	0.0056	0.0050	0.0018	0.0015	0.0013
Var_b7	0.2616	0.2757	0.1741	0.0380	0.0301	0.0240	0.0082	0.0065	0.0047
Var_b8	0.2210	0.2593	0.2678	0.0573	0.0404	0.0338	0.0161	0.0117	0.0084

Tabelle 30: Es werden MSE, Bias und Varianz für die einzelnen Koeffizienten (b1, ..., b8) unter unterschiedlich großen Fehlend-Anteilen und unter der jeweiligen Datensatzgröße gezeigt.

Fehlendmechanismus LogitY mit Bruteforce1-Verfahren (lm)

	n=50			n=200			n=1000		
	50%	25%	10%	50%	25%	10%	50%	25%	10%
MSE_b1	1.0342	0.6190	0.4237	0.9409	0.4397	0.0853	0.7770	0.3895	0.1033
MSE_b2	3.7502	2.7459	1.3282	2.8165	1.3867	0.3575	3.3592	1.7994	0.5120
MSE_b3	1.5376	1.7421	1.2326	1.5492	0.8933	0.2542	1.0544	0.5287	0.1699
MSE_b4	2.8501	2.7647	2.6349	0.2031	0.2631	0.1837	0.0223	0.0187	0.0182
MSE_b5	2.9936	0.8151	0.4351	4.2142	2.0853	0.7211	4.6657	2.3898	0.6709
MSE_b6	2.3192	0.7253	0.3310	2.9495	1.5197	0.4813	3.3412	1.7801	0.4961
MSE_b7	1.8984	0.4647	0.3592	4.0071	1.9947	0.6854	2.6658	1.2929	0.3603
MSE_b8	5.3796	8.6015	8.6492	0.0634	0.1074	0.0714	0.0109	0.0949	0.1345
Bias_b1	0.9933	0.7624	0.6262	0.9596	0.6509	0.2735	0.8783	0.6209	0.3168
Bias_b2	-1.9202	-1.6356	-1.1106	-1.6729	-1.1653	-0.5849	-1.8288	-1.3363	-0.7118
Bias_b3	-1.1775	-1.2493	-1.0006	-1.2257	-0.9199	-0.4609	-1.0216	-0.7187	-0.4012
Bias_b4	-1.4796	-1.3434	-1.0745	0.3791	0.4380	0.3539	-0.1035	0.0083	0.0019
Bias_b5	-1.6796	-0.8239	-0.5591	-2.0407	-1.4239	-0.8177	-2.1535	-1.5388	-0.8098
Bias_b6	1.4865	0.8081	0.5071	1.7086	1.2207	0.6684	1.8235	1.3287	0.6991
Bias_b7	-1.2545	-0.3828	-0.2513	-1.9801	-1.3869	-0.7738	-1.6247	-1.1251	-0.5871
Bias_b8	2.1559	2.7711	2.6436	-0.0540	0.1717	0.0770	-0.0355	0.2765	0.3341
Var_b1	0.0476	0.0378	0.0315	0.0201	0.0160	0.0105	0.0056	0.0040	0.0029
Var_b2	0.0631	0.0707	0.0948	0.0179	0.0288	0.0154	0.0146	0.0135	0.0053
Var_b3	0.1511	0.1815	0.2314	0.0469	0.0471	0.0418	0.0108	0.0122	0.0090
Var_b4	0.6609	0.9599	1.4803	0.0594	0.0713	0.0584	0.0116	0.0186	0.0182
Var_b5	0.1726	0.1362	0.1225	0.0498	0.0578	0.0525	0.0283	0.0219	0.0152
Var_b6	0.1094	0.0724	0.0739	0.0303	0.0296	0.0344	0.0161	0.0147	0.0074
Var_b7	0.3245	0.3181	0.2961	0.0861	0.0712	0.0867	0.0261	0.0271	0.0156
Var_b8	0.7316	0.9228	1.6606	0.0604	0.0779	0.0655	0.0096	0.0184	0.0229

Tabelle 31: Es werden MSE, Bias und Varianz für die einzelnen Koeffizienten (b1, ..., b8) unter unterschiedlich großen Fehlend-Anteilen und unter der jeweiligen Datensatzgröße gezeigt.

Fehlendmechanismus LogitY mit Bruteforce1-Verfahren (ridge, $\lambda=0$)

	n=50			n=200			n=1000		
	50%	25%	10%	50%	25%	10%	50%	25%	10%
MSE_b1	0.5906	0.3049	0.1521	0.7536	0.3650	0.0707	0.5842	0.3068	0.0832
MSE_b2	3.3810	2.4169	1.1119	2.3600	1.2076	0.3155	2.8516	1.5731	0.4571
MSE_b3	1.5702	1.7725	1.3013	0.8887	0.6058	0.1901	0.6019	0.3389	0.1187
MSE_b4	1.9900	2.2679	3.9477	1.2618	0.7882	0.3551	0.7587	0.3771	0.1044
MSE_b5	1.7602	0.3216	0.1516	2.2594	1.2835	0.5168	2.6724	1.4991	0.4531
MSE_b6	1.7551	0.4471	0.1617	2.2320	1.2193	0.4067	2.6269	1.4550	0.4176
MSE_b7	0.5188	0.5688	0.7875	1.2238	0.8580	0.3963	0.6448	0.4004	0.1419
MSE_b8	6.0658	9.5922	9.8102	1.1888	0.8045	0.2789	1.5269	1.1225	0.5227
Bias_b1	0.7344	0.5074	0.3421	0.8579	0.5923	0.2440	0.7613	0.5505	0.2831
Bias_b2	-1.8199	-1.5302	-1.0075	-1.5306	-1.0866	-0.5478	-1.6850	-1.2491	-0.6724
Bias_b3	-1.1898	-1.2592	-1.0313	-0.9177	-0.7480	-0.3817	-0.7701	-0.5722	-0.3315
Bias_b4	0.8982	0.9814	1.4953	1.0922	0.8409	0.5487	0.8630	0.5959	0.2760
Bias_b5	-1.2600	-0.4083	-0.1035	-1.4858	-1.1040	-0.6713	-1.6290	-1.2165	-0.6621
Bias_b6	1.2855	0.6087	0.2906	1.4844	1.0904	0.6086	1.6168	1.2008	0.6406
Bias_b7	-0.4446	0.4104	0.6324	-1.0591	-0.8643	-0.5319	-0.7913	-0.6147	-0.3527
Bias_b8	2.3067	2.9499	2.8522	1.0481	0.7888	0.3683	1.2261	1.0431	0.6898
Var_b1	0.0513	0.0475	0.0351	0.0177	0.0142	0.0112	0.0047	0.0038	0.0031
Var_b2	0.0689	0.0754	0.0969	0.0172	0.0270	0.0154	0.0124	0.0127	0.0049
Var_b3	0.1545	0.1869	0.2377	0.0465	0.0464	0.0444	0.0089	0.0114	0.0089
Var_b4	1.1831	1.3047	1.7119	0.0689	0.0811	0.0541	0.0139	0.0221	0.0282
Var_b5	0.1725	0.1550	0.1409	0.0517	0.0646	0.0662	0.0187	0.0192	0.0148
Var_b6	0.1027	0.0766	0.0772	0.0284	0.0302	0.0363	0.0129	0.0132	0.0072
Var_b7	0.3211	0.4004	0.3876	0.1021	0.1110	0.1134	0.0186	0.0226	0.0175
Var_b8	0.7452	0.8901	1.6750	0.0903	0.1823	0.1433	0.0235	0.0346	0.0468

Tabelle 32: Es werden MSE, Bias und Varianz für die einzelnen Koeffizienten (b1, ..., b8) unter unterschiedlich großen Fehlend-Anteilen und unter der jeweiligen Datensatzgröße gezeigt.

Fehlendmechanismus LogitY mit Bruteforce1-Verfahren (ridge)

	n=50			n=200			n=1000		
	50%	25%	10%	50%	25%	10%	50%	25%	10%
MSE_b1	0.8502	0.6895	0.4536	0.8274	0.4194	0.0841	0.6732	0.3202	0.1017
MSE_b2	3.7113	3.2338	2.1352	3.0595	1.5191	0.4132	3.0991	1.6301	0.5337
MSE_b3	1.0630	1.0884	0.9019	0.9217	0.6708	0.2336	0.6940	0.3488	0.1327
MSE_b4	0.9609	0.8965	1.2447	1.0761	0.7397	0.3233	0.8199	0.3892	0.1196
MSE_b5	3.7865	3.3043	2.7859	3.6146	2.5150	1.4924	3.3656	1.8838	1.0221
MSE_b6	3.6420	3.1627	2.4824	3.3548	2.0788	0.9691	3.1630	1.7081	0.7645
MSE_b7	1.0878	1.0241	1.1721	1.5222	1.6744	1.2687	0.9406	0.6002	0.4999
MSE_b8	1.3735	1.3871	1.4407	0.7816	0.4058	0.1100	1.1449	0.9173	0.2541
Bias_b1	0.9197	0.8236	0.6565	0.9058	0.6381	0.2718	0.8164	0.5625	0.3147
Bias_b2	-1.9248	-1.7916	-1.4364	-1.7452	-1.2198	-0.6310	-1.7557	-1.2720	-0.7276
Bias_b3	-1.0255	-1.0325	-0.9175	-0.9546	-0.7996	-0.4406	-0.8285	-0.5817	-0.3523
Bias_b4	0.9448	0.8519	0.8786	1.0286	0.8266	0.5260	0.9003	0.6063	0.3074
Bias_b5	-1.9443	-1.8107	-1.6517	-1.8991	-1.5705	-1.2085	-1.8288	-1.3668	-1.0068
Bias_b6	1.9066	1.7724	1.5608	1.8291	1.4318	0.9743	1.7735	1.3025	0.8719
Bias_b7	-1.0380	-0.9914	-1.0334	-1.2264	-1.2711	-1.0969	-0.9631	-0.7631	-0.6990
Bias_b8	1.1496	1.1279	1.0363	0.8707	0.5366	0.0087	1.0644	0.9403	0.4588
Var_b1	0.0043	0.0112	0.0226	0.0069	0.0122	0.0102	0.0067	0.0037	0.0027
Var_b2	0.0065	0.0240	0.0718	0.0139	0.0312	0.0150	0.0164	0.0121	0.0043
Var_b3	0.0114	0.0223	0.0601	0.0105	0.0314	0.0395	0.0076	0.0104	0.0086
Var_b4	0.0683	0.1709	0.4728	0.0180	0.0564	0.0466	0.0093	0.0217	0.0251
Var_b5	0.0063	0.0258	0.0577	0.0079	0.0485	0.0319	0.0211	0.0158	0.0085
Var_b6	0.0070	0.0215	0.0461	0.0093	0.0288	0.0199	0.0176	0.0117	0.0043
Var_b7	0.0104	0.0412	0.1043	0.0183	0.0587	0.0656	0.0131	0.0179	0.0113
Var_b8	0.0519	0.1150	0.3669	0.0235	0.1179	0.1099	0.0120	0.0331	0.0436

Tabelle 33: Es werden MSE, Bias und Varianz für die einzelnen Koeffizienten (b1, ..., b8) unter unterschiedlich großen Fehlend-Anteilen und unter der jeweiligen Datensatzgröße gezeigt.

Fehlendmechanismus LogitX mit Bruteforce1-Verfahren (lm)

	n=50			n=200			n=1000		
	50%	25%	10%	50%	25%	10%	50%	25%	10%
MSE_b1	1.0250	0.5515	0.4164	0.9269	0.3539	0.0690	0.7184	0.3474	0.0848
MSE_b2	3.7500	2.4601	1.1637	2.6642	1.2042	0.2868	3.1579	1.6161	0.4270
MSE_b3	1.4746	1.8480	1.0377	1.4334	0.7818	0.2367	0.9686	0.4967	0.1407
MSE_b4	2.5287	2.8129	2.2746	0.2141	0.2021	0.1579	0.0122	0.0207	0.0137
MSE_b5	2.6068	0.7008	0.3571	3.9366	1.9170	0.6221	4.2717	2.1121	0.5215
MSE_b6	2.0248	0.5626	0.2927	2.7399	1.4147	0.4121	3.0948	1.5729	0.3907
MSE_b7	1.5021	0.4511	0.2587	3.7650	1.8349	0.5563	2.3302	1.1274	0.2807
MSE_b8	5.5512	9.7883	8.3212	0.0641	0.1086	0.0632	0.0139	0.1352	0.1346
Bias_b1	0.9804	0.7203	0.6217	0.9522	0.5830	0.2356	0.8442	0.5861	0.2868
Bias_b2	-1.9246	-1.5450	-1.0389	-1.6238	-1.0862	-0.5166	-1.7719	-1.2663	-0.6490
Bias_b3	-1.1674	-1.2933	-0.9238	-1.1758	-0.8594	-0.4329	-0.9770	-0.6964	-0.3622
Bias_b4	-1.3915	-1.2975	-1.1709	0.3965	0.3872	0.3259	-0.0543	0.0342	0.0111
Bias_b5	-1.5700	-0.7444	-0.5067	-1.9693	-1.3590	-0.7642	-2.0590	-1.4463	-0.7163
Bias_b6	1.3935	0.6944	0.4828	1.6458	1.1716	0.6255	1.7531	1.2484	0.6207
Bias_b7	-1.1054	-0.3820	-0.1995	-1.9205	-1.3196	-0.7035	-1.5176	-1.0508	-0.5175
Bias_b8	2.2155	2.9571	2.6702	0.0774	0.2041	0.0639	0.0399	0.3396	0.3405
Var_b1	0.0639	0.0328	0.0299	0.0203	0.0141	0.0135	0.0057	0.0039	0.0026
Var_b2	0.0461	0.0730	0.0844	0.0276	0.0243	0.0200	0.0183	0.0126	0.0058
Var_b3	0.1117	0.1754	0.1843	0.0508	0.0432	0.0492	0.0141	0.0117	0.0095
Var_b4	0.5925	1.1294	0.9036	0.0570	0.0522	0.0517	0.0092	0.0196	0.0136
Var_b5	0.1418	0.1468	0.1004	0.0584	0.0701	0.0380	0.0322	0.0204	0.0085
Var_b6	0.0829	0.0803	0.0597	0.0312	0.0421	0.0209	0.0216	0.0145	0.0054
Var_b7	0.2802	0.3052	0.2189	0.0768	0.0935	0.0614	0.0269	0.0234	0.0129
Var_b8	0.6428	1.0440	1.1914	0.0581	0.0670	0.0591	0.0123	0.0199	0.0186

Tabelle 34: Es werden MSE, Bias und Varianz für die einzelnen Koeffizienten (b1, ..., b8) unter unterschiedlich großen Fehlend-Anteilen und unter der jeweiligen Datensatzgröße gezeigt.

Fehlendmechanismus LogitX mit Bruteforce1-Verfahren (ridge, $\lambda=0$)

	n=50			n=200			n=1000		
	50%	25%	10%	50%	25%	10%	50%	25%	10%
MSE_b1	0.5953	0.2528	0.1470	0.7493	0.2961	0.0565	0.5488	0.2710	0.0686
MSE_b2	3.3938	2.1682	0.9516	2.2485	1.0564	0.2486	2.6934	1.4151	0.3816
MSE_b3	1.5307	1.8782	1.0813	0.8401	0.5435	0.1710	0.5652	0.3198	0.1006
MSE_b4	1.9932	2.5247	3.1194	1.2517	0.6237	0.3513	0.7173	0.3680	0.0837
MSE_b5	1.5092	0.2956	0.1306	2.1056	1.2457	0.4157	2.4804	1.3243	0.3546
MSE_b6	1.5232	0.3397	0.1329	2.0631	1.1574	0.3393	2.4470	1.2862	0.3302
MSE_b7	0.4194	0.6349	0.7607	1.1807	0.8515	0.2779	0.5664	0.3413	0.1107
MSE_b8	6.3667	11.0754	9.5945	1.3804	0.6991	0.2889	1.5078	1.1681	0.4692
Bias_b1	0.7258	0.4571	0.3359	0.8553	0.5314	0.2069	0.7376	0.5169	0.2572
Bias_b2	-1.8293	-1.4454	-0.9281	-1.4911	-1.0170	-0.4799	-1.6364	-1.1847	-0.6135
Bias_b3	-1.1890	-1.3057	-0.9456	-0.8875	-0.7098	-0.3436	-0.7443	-0.5567	-0.3019
Bias_b4	0.9053	1.0337	1.3514	1.0773	0.7368	0.5300	0.8364	0.5866	0.2494
Bias_b5	-1.1635	-0.3236	-0.0515	-1.4315	-1.0854	-0.6095	-1.5684	-1.1443	-0.5885
Bias_b6	1.1994	0.4917	0.2591	1.4264	1.0581	0.5646	1.5589	1.1288	0.5702
Bias_b7	-0.3314	0.4394	0.6719	-1.0440	-0.8648	-0.4467	-0.7421	-0.5699	-0.3114
Bias_b8	2.3925	3.1638	2.8919	1.1277	0.7481	0.3825	1.2188	1.0648	0.6595
Var_b1	0.0685	0.0438	0.0341	0.0177	0.0137	0.0138	0.0048	0.0037	0.0024
Var_b2	0.0476	0.0791	0.0903	0.0251	0.0222	0.0183	0.0156	0.0115	0.0053
Var_b3	0.1170	0.1734	0.1871	0.0524	0.0397	0.0529	0.0112	0.0099	0.0095
Var_b4	1.1736	1.4561	1.2930	0.0910	0.0808	0.0704	0.0177	0.0240	0.0215
Var_b5	0.1556	0.1909	0.1279	0.0563	0.0676	0.0442	0.0206	0.0150	0.0083
Var_b6	0.0845	0.0980	0.0658	0.0283	0.0379	0.0205	0.0168	0.0120	0.0052
Var_b7	0.3095	0.4418	0.3092	0.0908	0.1036	0.0784	0.0157	0.0165	0.0138
Var_b8	0.6429	1.0659	1.2314	0.1087	0.1394	0.1426	0.0222	0.0344	0.0343

Tabelle 35: Es werden MSE, Bias und Varianz für die einzelnen Koeffizienten (b1, ..., b8) unter unterschiedlich großen Fehlend-Anteilen und unter der jeweiligen Datensatzgröße gezeigt.

Fehlendmechanismus LogitX mit Bruteforce1-Verfahren (ridge)

	n=50			n=200			n=1000		
	50%	25%	10%	50%	25%	10%	50%	25%	10%
MSE_b1	0.8540	0.6159	0.4352	0.8225	0.3502	0.0670	0.6279	0.2889	0.0862
MSE_b2	3.7347	3.0145	2.0039	2.9307	1.3561	0.3345	2.9231	1.4665	0.4565
MSE_b3	1.0492	1.1031	0.8377	0.8892	0.6171	0.2149	0.6353	0.3300	0.1144
MSE_b4	0.9671	0.8674	0.9205	1.0676	0.6080	0.3152	0.7611	0.3826	0.1044
MSE_b5	3.7726	3.1171	2.7699	3.5310	2.4689	1.3889	3.1630	1.7062	0.9261
MSE_b6	3.6370	2.9187	2.4673	3.2344	1.9964	0.8993	2.9578	1.5341	0.6741
MSE_b7	1.0918	1.0446	1.1794	1.5503	1.6972	1.1357	0.8655	0.5460	0.4785
MSE_b8	1.3265	1.8543	1.3049	0.8310	0.3230	0.1136	1.1180	0.9400	0.2003
Bias_b1	0.9211	0.7772	0.6466	0.9030	0.5808	0.2343	0.7881	0.5343	0.2898
Bias_b2	-1.9313	-1.7276	-1.3897	-1.7053	-1.1521	-0.5640	-1.7039	-1.2065	-0.6723
Bias_b3	-1.0212	-1.0401	-0.8920	-0.9349	-0.7657	-0.4084	-0.7905	-0.5657	-0.3255
Bias_b4	0.9520	0.8039	0.8228	1.0203	0.7423	0.5042	0.8646	0.5996	0.2878
Bias_b5	-1.9407	-1.7544	-1.6531	-1.8759	-1.5568	-1.1687	-1.7717	-1.3020	-0.9601
Bias_b6	1.9056	1.6992	1.5615	1.7948	1.4008	0.9424	1.7133	1.2346	0.8193
Bias_b7	-1.0389	-0.9937	-1.0580	-1.2382	-1.2831	-1.0432	-0.9231	-0.7307	-0.6846
Bias_b8	1.1321	1.2591	1.0280	0.8934	0.4833	0.0025	1.0507	0.9515	0.4099
Var_b1	0.0055	0.0119	0.0171	0.0070	0.0129	0.0121	0.0067	0.0035	0.0022
Var_b2	0.0050	0.0300	0.0728	0.0225	0.0288	0.0163	0.0198	0.0110	0.0045
Var_b3	0.0063	0.0212	0.0420	0.0151	0.0307	0.0481	0.0105	0.0100	0.0085
Var_b4	0.0608	0.2212	0.2434	0.0265	0.0569	0.0611	0.0136	0.0231	0.0216
Var_b5	0.0065	0.0391	0.0371	0.0122	0.0454	0.0229	0.0240	0.0110	0.0043
Var_b6	0.0058	0.0314	0.0290	0.0130	0.0341	0.0111	0.0225	0.0098	0.0029
Var_b7	0.0125	0.0571	0.0600	0.0170	0.0509	0.0475	0.0134	0.0120	0.0098
Var_b8	0.0448	0.2689	0.2480	0.0328	0.0895	0.1136	0.0139	0.0346	0.0323

Tabelle 36: Es werden MSE, Bias und Varianz für die einzelnen Koeffizienten (b1, ..., b8) unter unterschiedlich großen Fehlend-Anteilen und unter der jeweiligen Datensatzgröße gezeigt.

C.3 Bruteforce2

Fehlendmechanismus MCAR mit Bruteforce2-Verfahren (lm)									
	n=50			n=200			n=1000		
	50%	25%	10%	50%	25%	10%	50%	25%	10%
MSE_b1	0.4333	0.1617	0.0475	0.2240	0.0560	0.0088	0.1785	0.0478	0.0022
MSE_b2	1.0056	0.2861	0.0371	0.8949	0.2090	0.0075	0.7740	0.2205	0.0031
MSE_b3	1.0610	0.6867	0.1657	0.2174	0.1439	0.0326	0.1207	0.0543	0.0077
MSE_b4	6.9897	2.4192	0.8612	1.2784	0.3969	0.0437	1.6421	0.4652	0.0126
MSE_b5	1.2308	0.5956	0.1719	0.5364	0.1983	0.0279	0.3168	0.0937	0.0083
MSE_b6	1.2373	0.5201	0.0838	0.8345	0.2462	0.0128	0.6771	0.1865	0.0040
MSE_b7	1.2182	0.9387	0.3652	0.2484	0.1746	0.0398	0.0841	0.0602	0.0131
MSE_b8	2.6780	1.8984	0.9536	1.9276	0.5358	0.0561	2.3586	0.6059	0.0243
Bias_b1	0.3514	0.0583	-0.0280	0.4270	0.1816	-0.0183	0.4153	0.2005	-0.0258
Bias_b2	-0.9330	-0.4537	-0.0493	-0.9286	-0.4276	0.0062	-0.8751	-0.4628	0.0306
Bias_b3	-0.5632	-0.2903	-0.1136	-0.2834	-0.1163	0.0037	-0.3072	-0.1601	0.0278
Bias_b4	2.1636	0.9061	-0.1491	1.0765	0.5367	0.0277	1.2642	0.6560	0.0210
Bias_b5	-0.7997	-0.2782	0.1780	-0.6517	-0.3360	0.0165	-0.5374	-0.2662	0.0506
Bias_b6	0.8745	0.3896	-0.1231	0.8774	0.4378	-0.0031	0.8159	0.4185	-0.0346
Bias_b7	0.0738	0.1318	0.2040	0.2007	0.0764	0.0281	0.1948	0.1393	0.0565
Bias_b8	0.7128	0.3001	0.3948	1.3293	0.6639	0.0920	1.5228	0.7594	0.1071
Var_b1	0.3098	0.1583	0.0467	0.0418	0.0230	0.0084	0.0060	0.0076	0.0015
Var_b2	0.1352	0.0803	0.0346	0.0326	0.0262	0.0075	0.0083	0.0063	0.0021
Var_b3	0.7439	0.6024	0.1528	0.1371	0.1304	0.0326	0.0263	0.0287	0.0069
Var_b4	2.3087	1.5982	0.8390	0.1196	0.1088	0.0430	0.0437	0.0349	0.0122
Var_b5	0.5913	0.5182	0.1403	0.1116	0.0854	0.0276	0.0280	0.0229	0.0058
Var_b6	0.4725	0.3683	0.0687	0.0648	0.0546	0.0128	0.0114	0.0113	0.0028
Var_b7	1.2128	0.9213	0.3235	0.2081	0.1687	0.0390	0.0462	0.0408	0.0099
Var_b8	2.1699	1.8083	0.7978	0.1605	0.0950	0.0476	0.0398	0.0292	0.0128

Tabelle 37: Es werden MSE, Bias und Varianz für die einzelnen Koeffizienten (b1, ..., b8) unter unterschiedlich großen Fehlend-Anteilen und unter der jeweiligen Datensatzgröße gezeigt.

Fehlendmechanismus MCAR mit Bruteforce2-Verfahren (ridge, $\lambda=0$)

	n=50			n=200			n=1000		
	50%	25%	10%	50%	25%	10%	50%	25%	10%
MSE_b1	0.6691	0.2132	0.0519	0.3028	0.0720	0.0084	0.2678	0.0704	0.0020
MSE_b2	1.1188	0.3085	0.0357	1.1102	0.2542	0.0076	0.9868	0.2814	0.0028
MSE_b3	1.0645	0.6883	0.1651	0.4177	0.1931	0.0347	0.2825	0.0991	0.0076
MSE_b4	3.2350	2.9302	1.9270	0.4530	0.2517	0.0617	0.3170	0.1079	0.0178
MSE_b5	1.8530	0.8675	0.2063	1.3489	0.4282	0.0366	0.9687	0.2660	0.0078
MSE_b6	1.4734	0.6174	0.0977	1.1930	0.3367	0.0137	0.9855	0.2692	0.0037
MSE_b7	1.8680	1.5076	0.4511	0.5982	0.3184	0.0595	0.2961	0.0961	0.0118
MSE_b8	2.4286	1.9332	0.9594	0.4140	0.2046	0.0740	0.2996	0.0935	0.0227
Bias_b1	0.5210	0.1193	-0.0442	0.5057	0.2189	-0.0090	0.5111	0.2502	-0.0204
Bias_b2	-0.9977	-0.4779	-0.0416	-1.0370	-0.4765	-0.0054	-0.9891	-0.5239	0.0238
Bias_b3	-0.5443	-0.2761	-0.1072	-0.5137	-0.2333	-0.0226	-0.5031	-0.2620	0.0185
Bias_b4	0.6370	0.3390	0.0029	0.5201	0.2689	-0.0277	0.4986	0.2543	-0.0220
Bias_b5	-1.0668	-0.3887	0.2022	-1.0932	-0.5461	-0.0303	-0.9637	-0.4895	0.0250
Bias_b6	1.0007	0.4453	-0.1347	1.0572	0.5241	0.0168	0.9855	0.5069	-0.0239
Bias_b7	-0.4473	-0.0687	0.2511	-0.5270	-0.2671	-0.0463	-0.4766	-0.2128	0.0190
Bias_b8	0.6226	0.2552	0.4202	0.4687	0.2511	0.0031	0.5068	0.2309	0.0488
Var_b1	0.3977	0.1989	0.0499	0.0470	0.0240	0.0083	0.0066	0.0078	0.0016
Var_b2	0.1234	0.0800	0.0340	0.0348	0.0271	0.0076	0.0086	0.0069	0.0022
Var_b3	0.7683	0.6121	0.1536	0.1539	0.1387	0.0342	0.0293	0.0304	0.0073
Var_b4	2.8293	2.8153	1.9270	0.1825	0.1794	0.0609	0.0685	0.0432	0.0173
Var_b5	0.7150	0.7165	0.1655	0.1538	0.1300	0.0357	0.0399	0.0264	0.0072
Var_b6	0.4720	0.4192	0.0795	0.0753	0.0620	0.0134	0.0143	0.0122	0.0032
Var_b7	1.6679	1.5029	0.3881	0.3205	0.2471	0.0573	0.0689	0.0508	0.0114
Var_b8	2.0410	1.8681	0.7828	0.1943	0.1415	0.0740	0.0428	0.0402	0.0203

Tabelle 38: Es werden MSE, Bias und Varianz für die einzelnen Koeffizienten (b1, ..., b8) unter unterschiedlich großen Fehlend-Anteilen und unter der jeweiligen Datensatzgröße gezeigt.

Fehlendmechanismus MCAR mit Bruteforce2-Verfahren (ridge)

	n=50			n=200			n=1000		
	50%	25%	10%	50%	25%	10%	50%	25%	10%
MSE_b1	0.6344	0.2181	0.0460	0.3450	0.0806	0.0080	0.2820	0.0864	0.0030
MSE_b2	2.1720	0.6774	0.0530	1.3511	0.3312	0.0234	1.0455	0.3463	0.0101
MSE_b3	0.7979	0.5243	0.1365	0.4602	0.2134	0.0464	0.2933	0.1105	0.0072
MSE_b4	0.8858	1.3742	1.4943	0.4526	0.2252	0.0542	0.3323	0.1229	0.0176
MSE_b5	3.4670	2.0278	0.6601	2.3916	1.2561	0.8093	1.3500	0.7514	0.3990
MSE_b6	2.8382	1.5202	0.3850	1.9038	0.8264	0.3649	1.2379	0.5732	0.1734
MSE_b7	1.7497	1.9613	1.0919	1.0929	0.8943	0.9847	0.4992	0.3777	0.4095
MSE_b8	0.4139	0.7398	0.8240	0.1768	0.1123	0.3835	0.1916	0.0326	0.1406
Bias_b1	0.7100	0.2757	0.0346	0.5507	0.2423	0.0283	0.5252	0.2814	0.0396
Bias_b2	-1.4416	-0.7590	-0.1416	-1.1461	-0.5549	-0.1308	-1.0187	-0.5835	-0.0906
Bias_b3	-0.7646	-0.3899	-0.0670	-0.5857	-0.2989	-0.1259	-0.5146	-0.2861	-0.0297
Bias_b4	0.5406	0.1132	-0.2034	0.5590	0.2610	-0.0382	0.5172	0.2885	0.0481
Bias_b5	-1.8433	-1.3104	-0.7601	-1.5128	-1.0896	-0.8912	-1.1494	-0.8569	-0.6285
Bias_b6	1.6652	1.1476	0.5847	1.3580	0.8891	0.5989	1.1080	0.7516	0.4141
Bias_b7	-1.2350	-1.1408	-0.9284	-0.9455	-0.8581	-0.9775	-0.6658	-0.5829	-0.6338
Bias_b8	0.3252	-0.3394	-0.5065	0.2576	-0.1174	-0.5739	0.3962	0.0036	-0.3529
Var_b1	0.1303	0.1421	0.0448	0.0417	0.0219	0.0072	0.0062	0.0072	0.0014
Var_b2	0.0938	0.1013	0.0330	0.0375	0.0233	0.0063	0.0078	0.0058	0.0019
Var_b3	0.2132	0.3722	0.1320	0.1172	0.1241	0.0306	0.0285	0.0286	0.0063
Var_b4	0.5935	1.3614	1.4530	0.1401	0.1571	0.0527	0.0648	0.0396	0.0153
Var_b5	0.0694	0.3108	0.0824	0.1029	0.0690	0.0150	0.0289	0.0172	0.0040
Var_b6	0.0653	0.2033	0.0432	0.0596	0.0360	0.0062	0.0102	0.0082	0.0019
Var_b7	0.2244	0.6600	0.2301	0.1988	0.1580	0.0292	0.0559	0.0379	0.0077
Var_b8	0.3081	0.6247	0.5674	0.1104	0.0985	0.0542	0.0347	0.0326	0.0161

Tabelle 39: Es werden MSE, Bias und Varianz für die einzelnen Koeffizienten (b1, ..., b8) unter unterschiedlich großen Fehlend-Anteilen und unter der jeweiligen Datensatzgröße gezeigt.

Fehlendmechanismus MNAR1 mit Bruteforce2-Verfahren (lm)

	n=50			n=200			n=1000		
	50%	25%	10%	50%	25%	10%	50%	25%	10%
MSE_b1	0.3016	0.0845	0.0563	0.6377	0.0095	0.0096	0.4768	0.0036	0.0017
MSE_b2	1.0918	0.0952	0.0370	1.9242	0.0189	0.0076	2.1025	0.0059	0.0034
MSE_b3	0.7825	0.3569	0.1963	0.9961	0.0532	0.0353	0.6258	0.0094	0.0084
MSE_b4	1.3425	1.3900	1.3573	0.1267	0.0751	0.0451	0.0033	0.0166	0.0093
MSE_b5	0.4677	0.3047	0.1664	2.9393	0.0522	0.0353	2.8362	0.0137	0.0090
MSE_b6	0.3935	0.1827	0.0930	2.0370	0.0322	0.0130	2.0589	0.0065	0.0046
MSE_b7	0.3317	0.4741	0.3920	2.7673	0.0791	0.0448	1.5423	0.0197	0.0125
MSE_b8	4.4329	2.4660	1.2223	0.0067	0.0996	0.0535	0.0053	0.0807	0.0206
Bias_b1	0.4772	0.0392	-0.0092	0.7973	0.0194	-0.0028	0.6903	-0.0356	-0.0201
Bias_b2	-0.9808	-0.2029	-0.0338	-1.3864	-0.0579	0.0050	-1.4498	0.0225	0.0353
Bias_b3	-0.8227	-0.3537	-0.1125	-0.9944	-0.1048	0.0041	-0.7903	0.0260	0.0305
Bias_b4	-0.8483	-0.2797	-0.1730	0.3411	0.1343	0.0246	-0.0378	0.0545	0.0252
Bias_b5	-0.4065	0.3053	0.1132	-1.7127	-0.0423	0.0339	-1.6837	0.0537	0.0502
Bias_b6	0.4189	-0.2113	-0.0569	1.4260	0.0623	-0.0125	1.4346	-0.0328	-0.0361
Bias_b7	-0.1916	0.2421	0.1052	-1.6606	-0.0499	0.0182	-1.2406	0.0695	0.0537
Bias_b8	2.0184	1.2134	0.3808	-0.0002	0.1895	0.1194	0.0576	0.2657	0.1075
Var_b1	0.0739	0.0830	0.0563	0.0021	0.0092	0.0096	0.0003	0.0024	0.0013
Var_b2	0.1299	0.0540	0.0359	0.0022	0.0156	0.0076	0.0005	0.0054	0.0022
Var_b3	0.1057	0.2318	0.1837	0.0073	0.0422	0.0353	0.0013	0.0087	0.0074
Var_b4	0.6228	1.3118	1.3273	0.0104	0.0571	0.0444	0.0019	0.0136	0.0087
Var_b5	0.3025	0.2114	0.1536	0.0058	0.0504	0.0341	0.0013	0.0108	0.0065
Var_b6	0.2181	0.1381	0.0897	0.0034	0.0284	0.0128	0.0007	0.0054	0.0033
Var_b7	0.2950	0.4155	0.3809	0.0098	0.0766	0.0445	0.0032	0.0148	0.0096
Var_b8	0.3590	0.9938	1.0773	0.0067	0.0637	0.0392	0.0020	0.0101	0.0090

Tabelle 40: Es werden MSE, Bias und Varianz für die einzelnen Koeffizienten (b1, ..., b8) unter unterschiedlich großen Fehlend-Anteilen und unter der jeweiligen Datensatzgröße gezeigt.

Fehlendmechanismus MNAR1 mit Bruteforce2-Verfahren (ridge, $\lambda=0$)

	n=50			n=200			n=1000		
	50%	25%	10%	50%	25%	10%	50%	25%	10%
MSE_b1	0.1497	0.0796	0.0588	0.5051	0.0095	0.0094	0.3620	0.0027	0.0016
MSE_b2	0.9487	0.0788	0.0340	1.6151	0.0205	0.0074	1.8000	0.0051	0.0031
MSE_b3	0.8100	0.3545	0.1935	0.5459	0.0602	0.0361	0.3679	0.0088	0.0080
MSE_b4	1.8588	4.2540	3.0097	0.8992	0.0829	0.0639	0.4757	0.0183	0.0146
MSE_b5	0.2505	0.3722	0.1844	1.5532	0.0679	0.0419	1.6544	0.0115	0.0081
MSE_b6	0.2656	0.2007	0.0951	1.5288	0.0363	0.0139	1.6312	0.0055	0.0040
MSE_b7	0.3967	0.6599	0.4589	0.7968	0.1216	0.0703	0.3766	0.0174	0.0104
MSE_b8	4.9636	2.5820	1.2226	0.8713	0.0906	0.0688	1.0247	0.0435	0.0174
Bias_b1	0.2978	-0.0290	-0.0200	0.7091	0.0320	0.0062	0.6014	-0.0248	-0.0145
Bias_b2	-0.9144	-0.1806	-0.0294	-1.2700	-0.0772	-0.0070	-1.3415	0.0112	0.0296
Bias_b3	-0.8428	-0.3564	-0.1093	-0.7340	-0.1442	-0.0200	-0.6053	0.0064	0.0211
Bias_b4	0.8090	0.3196	-0.0837	0.9407	0.0445	-0.0336	0.6879	-0.0290	-0.0161
Bias_b5	-0.1190	0.4143	0.1205	-1.2431	-0.1270	-0.0203	-1.2856	0.0062	0.0238
Bias_b6	0.2795	-0.2667	-0.0581	1.2351	0.0981	0.0102	1.2769	-0.0135	-0.0252
Bias_b7	0.3575	0.4440	0.1275	-0.8844	-0.1866	-0.0677	-0.6105	-0.0040	0.0139
Bias_b8	2.1447	1.2843	0.3833	0.9291	0.0289	0.0224	1.0112	0.1581	0.0504
Var_b1	0.0610	0.0788	0.0584	0.0022	0.0085	0.0094	0.0003	0.0021	0.0014
Var_b2	0.1126	0.0462	0.0332	0.0022	0.0145	0.0073	0.0005	0.0050	0.0022
Var_b3	0.0997	0.2275	0.1816	0.0071	0.0394	0.0357	0.0015	0.0088	0.0076
Var_b4	1.2044	4.1518	3.0027	0.0143	0.0810	0.0627	0.0025	0.0175	0.0143
Var_b5	0.2363	0.2005	0.1698	0.0079	0.0518	0.0415	0.0016	0.0114	0.0075
Var_b6	0.1874	0.1295	0.0917	0.0034	0.0266	0.0138	0.0007	0.0053	0.0034
Var_b7	0.2689	0.4628	0.4426	0.0146	0.0868	0.0657	0.0039	0.0174	0.0102
Var_b8	0.3636	0.9326	1.0756	0.0080	0.0898	0.0683	0.0020	0.0185	0.0149

Tabelle 41: Es werden MSE, Bias und Varianz für die einzelnen Koeffizienten (b1, ..., b8) unter unterschiedlich großen Fehlend-Anteilen und unter der jeweiligen Datensatzgröße gezeigt.

Fehlendmechanismus MNAR1 mit Bruteforce2-Verfahren (ridge)

	n=50			n=200			n=1000		
	50%	25%	10%	50%	25%	10%	50%	25%	10%
MSE_b1	0.2535	0.0773	0.0542	0.5147	0.0119	0.0102	0.3827	0.0029	0.0033
MSE_b2	1.3664	0.1738	0.0495	1.7059	0.0486	0.0234	1.8736	0.0137	0.0092
MSE_b3	0.7109	0.2754	0.1557	0.5891	0.0895	0.0455	0.3847	0.0095	0.0075
MSE_b4	0.9663	2.8402	2.3861	0.8626	0.0730	0.0572	0.4993	0.0171	0.0158
MSE_b5	1.7949	0.7751	0.7355	2.3130	0.8538	0.7954	2.1607	0.3780	0.4016
MSE_b6	1.5488	0.5295	0.4534	2.0216	0.4102	0.3594	1.9603	0.1632	0.1733
MSE_b7	1.0738	1.2747	1.2755	1.3861	1.1036	1.0216	0.6370	0.3930	0.4130
MSE_b8	1.1328	0.6326	0.9755	0.5339	0.3073	0.3597	0.7788	0.0703	0.1372
Bias_b1	0.4604	0.0880	0.0585	0.7160	0.0656	0.0428	0.6184	0.0313	0.0453
Bias_b2	-1.1297	-0.3423	-0.1360	-1.3054	-0.1899	-0.1303	-1.3686	-0.0960	-0.0856
Bias_b3	-0.8013	-0.3210	-0.0810	-0.7633	-0.2352	-0.1222	-0.6191	-0.0391	-0.0290
Bias_b4	0.5712	0.0726	-0.2819	0.9218	0.0303	-0.0498	0.7050	0.0361	0.0517
Bias_b5	-1.3006	-0.7801	-0.8082	-1.5196	-0.9121	-0.8821	-1.4696	-0.6092	-0.6302
Bias_b6	1.2057	0.6385	0.6360	1.4212	0.6302	0.5947	1.3999	0.3998	0.4137
Bias_b7	-0.9691	-0.9718	-1.0091	-1.1737	-1.0283	-0.9938	-0.7962	-0.6174	-0.6373
Bias_b8	0.9366	0.1086	-0.5047	0.7270	-0.4952	-0.5576	0.8815	-0.2343	-0.3540
Var_b1	0.0414	0.0695	0.0508	0.0020	0.0076	0.0084	0.0003	0.0020	0.0013
Var_b2	0.0903	0.0567	0.0310	0.0020	0.0125	0.0064	0.0005	0.0044	0.0019
Var_b3	0.0689	0.1723	0.1491	0.0064	0.0341	0.0306	0.0014	0.0080	0.0067
Var_b4	0.6400	2.8349	2.3066	0.0129	0.0721	0.0547	0.0023	0.0158	0.0131
Var_b5	0.1033	0.1664	0.0824	0.0037	0.0218	0.0174	0.0011	0.0068	0.0045
Var_b6	0.0951	0.1218	0.0488	0.0018	0.0131	0.0057	0.0005	0.0034	0.0022
Var_b7	0.1345	0.3302	0.2572	0.0086	0.0461	0.0339	0.0030	0.0117	0.0069
Var_b8	0.2557	0.6208	0.7208	0.0054	0.0620	0.0488	0.0017	0.0154	0.0119

Tabelle 42: Es werden MSE, Bias und Varianz für die einzelnen Koeffizienten (b1, ..., b8) unter unterschiedlich großen Fehlend-Anteilen und unter der jeweiligen Datensatzgröße gezeigt.

Fehlendmechanismus MNAR2 mit Bruteforce2-Verfahren (lm)

	n=50			n=200			n=1000		
	50%	25%	10%	50%	25%	10%	50%	25%	10%
MSE_b1	0.2654	0.1114	0.0707	0.3553	0.0205	0.0107	0.2280	0.0025	0.0019
MSE_b2	1.0637	0.1517	0.0407	1.0696	0.0307	0.0078	1.1273	0.0113	0.0025
MSE_b3	0.8943	0.4190	0.2535	0.6038	0.0517	0.0342	0.2934	0.0135	0.0081
MSE_b4	1.4953	1.5132	1.1358	0.1741	0.1217	0.0509	0.0100	0.0383	0.0159
MSE_b5	0.3514	0.2740	0.1608	1.3777	0.0608	0.0320	1.4253	0.0119	0.0065
MSE_b6	0.2733	0.1483	0.0769	1.0502	0.0409	0.0161	1.0747	0.0070	0.0031
MSE_b7	0.3541	0.5798	0.3178	1.3334	0.0933	0.0505	0.6931	0.0332	0.0156
MSE_b8	5.8344	2.5953	0.9056	0.0754	0.1517	0.0701	0.1078	0.1657	0.0402
Bias_b1	0.4078	0.0587	0.0150	0.5925	0.0726	0.0268	0.4769	-0.0001	-0.0137
Bias_b2	-0.9824	-0.2601	-0.0675	-1.0327	-0.1324	-0.0215	-1.0613	-0.0836	-0.0154
Bias_b3	-0.7847	-0.3583	-0.1556	-0.7655	-0.0487	-0.0046	-0.5394	0.0231	0.0256
Bias_b4	-0.7457	-0.1806	0.1204	0.3969	0.2551	0.0726	0.0759	0.1558	0.0561
Bias_b5	-0.2875	0.2517	0.1149	-1.1687	-0.0908	0.0081	-1.1930	0.0034	0.0205
Bias_b6	0.3398	-0.1493	-0.0802	1.0216	0.1154	0.0287	1.0361	0.0383	-0.0029
Bias_b7	-0.0470	0.3189	0.1530	-1.1468	-0.1030	-0.0107	-0.8299	0.0912	0.0581
Bias_b8	2.2618	1.2174	0.2679	0.2446	0.3001	0.1435	0.3227	0.3844	0.1690
Var_b1	0.0991	0.1079	0.0705	0.0043	0.0153	0.0100	0.0006	0.0025	0.0017
Var_b2	0.0986	0.0841	0.0362	0.0031	0.0132	0.0073	0.0008	0.0043	0.0023
Var_b3	0.2786	0.2907	0.2292	0.0178	0.0493	0.0341	0.0024	0.0129	0.0074
Var_b4	0.9393	1.4806	1.1213	0.0166	0.0566	0.0457	0.0042	0.0140	0.0127
Var_b5	0.2688	0.2107	0.1476	0.0119	0.0526	0.0319	0.0021	0.0119	0.0061
Var_b6	0.1579	0.1260	0.0705	0.0065	0.0276	0.0153	0.0011	0.0056	0.0031
Var_b7	0.3519	0.4780	0.2944	0.0182	0.0827	0.0504	0.0045	0.0249	0.0122
Var_b8	0.7184	1.1131	0.8339	0.0156	0.0616	0.0495	0.0037	0.0179	0.0116

Tabelle 43: Es werden MSE, Bias und Varianz für die einzelnen Koeffizienten (b1, ..., b8) unter unterschiedlich großen Fehlend-Anteilen und unter der jeweiligen Datensatzgröße gezeigt.

Fehlendmechanismus MNAR2 mit Bruteforce2-Verfahren (ridge, $\lambda=0$)

	n=50			n=200			n=1000		
	50%	25%	10%	50%	25%	10%	50%	25%	10%
MSE_b1	0.1401	0.1009	0.0776	0.3038	0.0222	0.0117	0.1814	0.0026	0.0018
MSE_b2	0.9246	0.1226	0.0382	0.9495	0.0351	0.0085	0.9960	0.0138	0.0028
MSE_b3	0.9307	0.4200	0.2547	0.4169	0.0569	0.0330	0.1881	0.0131	0.0076
MSE_b4	2.4269	4.2420	2.7779	0.5245	0.1012	0.0663	0.2574	0.0225	0.0179
MSE_b5	0.2381	0.4013	0.1899	0.8804	0.1001	0.0428	0.9250	0.0162	0.0083
MSE_b6	0.1842	0.1817	0.0845	0.8607	0.0531	0.0184	0.8925	0.0098	0.0036
MSE_b7	0.6902	0.9601	0.4562	0.5973	0.1835	0.0884	0.2206	0.0256	0.0155
MSE_b8	6.4255	2.7273	0.8970	0.5465	0.1144	0.0706	0.7833	0.0819	0.0306
Bias_b1	0.2298	-0.0171	-0.0037	0.5472	0.0849	0.0354	0.4252	0.0135	-0.0071
Bias_b2	-0.9133	-0.2285	-0.0575	-0.9729	-0.1522	-0.0313	-0.9976	-0.0995	-0.0229
Bias_b3	-0.8042	-0.3636	-0.1555	-0.6332	-0.0897	-0.0285	-0.4308	-0.0062	0.0110
Bias_b4	0.8211	0.5038	0.2742	0.7062	0.1561	0.0240	0.5018	0.0437	0.0053
Bias_b5	-0.0012	0.3742	0.1376	-0.9289	-0.1800	-0.0353	-0.9604	-0.0630	-0.0107
Bias_b6	0.2036	-0.2083	-0.0892	0.9238	0.1536	0.0468	0.9441	0.0651	0.0098
Bias_b7	0.5072	0.5597	0.2029	-0.7514	-0.2435	-0.0826	-0.4628	-0.0122	0.0111
Bias_b8	2.3844	1.2703	0.2795	0.7224	0.1326	0.0630	0.8821	0.2323	0.1022
Var_b1	0.0873	0.1007	0.0776	0.0045	0.0150	0.0104	0.0006	0.0024	0.0017
Var_b2	0.0905	0.0704	0.0349	0.0031	0.0119	0.0075	0.0008	0.0039	0.0023
Var_b3	0.2839	0.2878	0.2305	0.0160	0.0489	0.0321	0.0025	0.0131	0.0075
Var_b4	1.7528	3.9882	2.7027	0.0258	0.0769	0.0657	0.0057	0.0206	0.0179
Var_b5	0.2381	0.2613	0.1710	0.0175	0.0677	0.0415	0.0027	0.0123	0.0082
Var_b6	0.1427	0.1383	0.0766	0.0073	0.0295	0.0162	0.0011	0.0055	0.0035
Var_b7	0.4330	0.6469	0.4150	0.0327	0.1242	0.0816	0.0064	0.0255	0.0154
Var_b8	0.7403	1.1136	0.8189	0.0247	0.0968	0.0666	0.0053	0.0279	0.0201

Tabelle 44: Es werden MSE, Bias und Varianz für die einzelnen Koeffizienten (b1, ..., b8) unter unterschiedlich großen Fehlend-Anteilen und unter der jeweiligen Datensatzgröße gezeigt.

Fehlendmechanismus MNAR2 mit Bruteforce2-Verfahren (ridge)

	n=50			n=200			n=1000		
	50%	25%	10%	50%	25%	10%	50%	25%	10%
MSE_b1	0.2336	0.0955	0.0724	0.3150	0.0270	0.0143	0.2039	0.0066	0.0041
MSE_b2	1.3687	0.2069	0.0567	1.0562	0.0774	0.0304	1.0873	0.0415	0.0194
MSE_b3	0.7664	0.3306	0.2023	0.4681	0.0735	0.0452	0.2067	0.0141	0.0080
MSE_b4	1.2750	2.7756	2.0501	0.4918	0.0870	0.0578	0.2854	0.0288	0.0216
MSE_b5	1.6899	0.7720	0.7044	1.7704	0.8892	0.7994	1.5516	0.4090	0.4171
MSE_b6	1.4706	0.5309	0.3993	1.4273	0.4474	0.3925	1.2878	0.2041	0.1903
MSE_b7	1.0399	1.0937	1.1179	1.3914	1.1646	1.0355	0.5628	0.3592	0.3940
MSE_b8	1.5458	0.7318	0.9022	0.2098	0.2030	0.3098	0.4771	0.0391	0.1002
Bias_b1	0.4125	0.0866	0.0694	0.5576	0.1155	0.0705	0.4510	0.0656	0.0509
Bias_b2	-1.1385	-0.3789	-0.1566	-1.0264	-0.2587	-0.1548	-1.0424	-0.1950	-0.1321
Bias_b3	-0.7815	-0.3349	-0.1097	-0.6734	-0.1791	-0.1305	-0.4522	-0.0483	-0.0345
Bias_b4	0.6212	0.2676	0.0391	0.6844	0.1406	0.0072	0.5293	0.0988	0.0725
Bias_b5	-1.2439	-0.7625	-0.7887	-1.3275	-0.9270	-0.8843	-1.2449	-0.6336	-0.6423
Bias_b6	1.1707	0.6504	0.6021	1.1931	0.6580	0.6204	1.1344	0.4475	0.4338
Bias_b7	-0.8947	-0.8053	-0.9319	-1.1713	-1.0454	-0.9948	-0.7471	-0.5839	-0.6196
Bias_b8	1.0705	0.1581	-0.5848	0.4402	-0.3678	-0.5122	0.6876	-0.1282	-0.2913
Var_b1	0.0634	0.0880	0.0675	0.0041	0.0137	0.0094	0.0006	0.0023	0.0015
Var_b2	0.0725	0.0633	0.0322	0.0027	0.0104	0.0064	0.0007	0.0034	0.0020
Var_b3	0.1557	0.2185	0.1902	0.0146	0.0414	0.0281	0.0022	0.0117	0.0068
Var_b4	0.8891	2.7040	2.0486	0.0234	0.0673	0.0577	0.0053	0.0191	0.0163
Var_b5	0.1427	0.1906	0.0824	0.0081	0.0299	0.0175	0.0017	0.0075	0.0046
Var_b6	0.1000	0.1078	0.0367	0.0038	0.0144	0.0076	0.0009	0.0039	0.0022
Var_b7	0.2395	0.4452	0.2495	0.0195	0.0718	0.0459	0.0046	0.0183	0.0101
Var_b8	0.4000	0.7068	0.5602	0.0160	0.0677	0.0474	0.0044	0.0227	0.0153

Tabelle 45: Es werden MSE, Bias und Varianz für die einzelnen Koeffizienten (b1, ..., b8) unter unterschiedlich großen Fehlend-Anteilen und unter der jeweiligen Datensatzgröße gezeigt.

Fehlendmechanismus MNAR3 mit Bruteforce2-Verfahren (lm)

	n=50			n=200			n=1000		
	50%	25%	10%	50%	25%	10%	50%	25%	10%
MSE_b1	0.2684	0.2239	0.0979	0.0410	0.0262	0.0174	0.0158	0.0071	0.0020
MSE_b2	0.3365	0.1500	0.0443	0.0103	0.0125	0.0093	0.0162	0.0066	0.0028
MSE_b3	0.9128	0.7232	0.2611	0.0671	0.0969	0.0500	0.0816	0.0322	0.0128
MSE_b4	1.5205	1.9117	0.8773	0.7847	0.2424	0.0936	0.5319	0.1744	0.0402
MSE_b5	1.7004	0.7484	0.2874	0.3349	0.1246	0.0485	0.0533	0.0224	0.0080
MSE_b6	0.7635	0.3909	0.1753	0.0690	0.0425	0.0221	0.0073	0.0073	0.0031
MSE_b7	2.7658	1.3812	0.5636	0.2417	0.1626	0.0687	0.3690	0.1099	0.0269
MSE_b8	7.8091	3.0066	1.0407	1.9197	0.5614	0.1414	2.0439	0.4968	0.0902
Bias_b1	-0.3310	-0.1591	-0.0599	0.1548	0.0741	0.0354	-0.1121	-0.0558	-0.0140
Bias_b2	-0.5264	-0.2741	-0.0829	0.0053	0.0034	0.0168	-0.1132	-0.0585	-0.0250
Bias_b3	-0.5800	-0.3377	-0.1351	0.0867	0.0802	0.0098	0.2637	0.1273	0.0552
Bias_b4	0.3270	0.2286	0.0811	0.8510	0.3884	0.1532	0.7134	0.3940	0.1554
Bias_b5	1.2084	0.6207	0.2329	0.5332	0.2601	0.1011	0.1991	0.0883	0.0404
Bias_b6	-0.8016	-0.4027	-0.1598	-0.2198	-0.1039	-0.0373	-0.0457	-0.0175	-0.0059
Bias_b7	1.4169	0.7724	0.2832	0.4028	0.2084	0.0806	0.5836	0.2832	0.1117
Bias_b8	2.6689	1.2336	0.4092	1.3618	0.6870	0.2835	1.4213	0.6847	0.2750
Var_b1	0.1588	0.1986	0.0943	0.0170	0.0208	0.0161	0.0032	0.0039	0.0018
Var_b2	0.0595	0.0749	0.0374	0.0103	0.0125	0.0090	0.0034	0.0032	0.0021
Var_b3	0.5763	0.6092	0.2428	0.0595	0.0905	0.0499	0.0121	0.0160	0.0098
Var_b4	1.4136	1.8595	0.8708	0.0605	0.0916	0.0702	0.0230	0.0191	0.0161
Var_b5	0.2401	0.3631	0.2332	0.0505	0.0569	0.0383	0.0136	0.0146	0.0064
Var_b6	0.1209	0.2287	0.1498	0.0207	0.0317	0.0207	0.0052	0.0070	0.0031
Var_b7	0.7583	0.7847	0.4834	0.0795	0.1192	0.0622	0.0285	0.0298	0.0145
Var_b8	0.6863	1.4848	0.8733	0.0653	0.0894	0.0610	0.0239	0.0281	0.0146

Tabelle 46: Es werden MSE, Bias und Varianz für die einzelnen Koeffizienten (b1, ..., b8) unter unterschiedlich großen Fehlend-Anteilen und unter der jeweiligen Datensatzgröße gezeigt.

Fehlendmechanismus MNAR3 mit Bruteforce2-Verfahren (ridge, $\lambda=0$)

	n=50			n=200			n=1000		
	50%	25%	10%	50%	25%	10%	50%	25%	10%
MSE_b1	0.3078	0.2532	0.1062	0.0782	0.0364	0.0192	0.0052	0.0049	0.0019
MSE_b2	0.3490	0.1483	0.0455	0.0223	0.0156	0.0092	0.0417	0.0129	0.0039
MSE_b3	0.8846	0.7230	0.2652	0.0987	0.0959	0.0545	0.0268	0.0200	0.0110
MSE_b4	1.2242	2.5196	1.8010	0.1394	0.1580	0.0906	0.0577	0.0461	0.0250
MSE_b5	1.6514	0.8060	0.3312	0.0688	0.0800	0.0555	0.0265	0.0212	0.0090
MSE_b6	0.7439	0.4035	0.1836	0.0222	0.0323	0.0230	0.0109	0.0092	0.0038
MSE_b7	2.9482	1.6627	0.7134	0.2814	0.1987	0.0989	0.0495	0.0430	0.0191
MSE_b8	7.6070	2.9849	1.0406	0.2331	0.1868	0.1226	0.5176	0.1523	0.0421
Bias_b1	-0.2952	-0.1472	-0.0584	0.2476	0.1229	0.0524	-0.0449	-0.0242	-0.0011
Bias_b2	-0.5405	-0.2764	-0.0851	-0.1135	-0.0574	-0.0067	-0.1952	-0.0978	-0.0407
Bias_b3	-0.5714	-0.3341	-0.1348	-0.1807	-0.0582	-0.0433	0.1243	0.0609	0.0281
Bias_b4	0.0289	0.0645	0.0436	0.2326	0.0680	0.0356	0.1700	0.1279	0.0501
Bias_b5	1.1399	0.5885	0.2207	0.0457	-0.0015	0.0011	-0.0998	-0.0600	-0.0201
Bias_b6	-0.7666	-0.3854	-0.1509	-0.0218	0.0035	0.0043	0.0743	0.0422	0.0189
Bias_b7	1.2961	0.7249	0.2654	-0.3883	-0.2179	-0.0784	0.1132	0.0524	0.0166
Bias_b8	2.6154	1.2137	0.4008	0.4036	0.1900	0.0936	0.7020	0.3381	0.1333
Var_b1	0.2207	0.2315	0.1028	0.0169	0.0213	0.0164	0.0032	0.0043	0.0019
Var_b2	0.0569	0.0719	0.0382	0.0094	0.0123	0.0091	0.0036	0.0033	0.0023
Var_b3	0.5582	0.6114	0.2470	0.0661	0.0925	0.0527	0.0114	0.0163	0.0102
Var_b4	1.2234	2.5154	1.7991	0.0853	0.1534	0.0893	0.0288	0.0297	0.0224
Var_b5	0.3519	0.4596	0.2825	0.0667	0.0800	0.0555	0.0165	0.0176	0.0086
Var_b6	0.1562	0.2550	0.1608	0.0217	0.0323	0.0229	0.0054	0.0074	0.0035
Var_b7	1.2682	1.1372	0.6429	0.1306	0.1512	0.0927	0.0367	0.0403	0.0188
Var_b8	0.7668	1.5119	0.8799	0.0702	0.1507	0.1138	0.0249	0.0380	0.0243

Tabelle 47: Es werden MSE, Bias und Varianz für die einzelnen Koeffizienten (b1, ..., b8) unter unterschiedlich großen Fehlend-Anteilen und unter der jeweiligen Datensatzgröße gezeigt.

Fehlendmechanismus MNAR3 mit Bruteforce2-Verfahren (ridge)

	n=50			n=200			n=1000		
	50%	25%	10%	50%	25%	10%	50%	25%	10%
MSE_b1	0.2121	0.1907	0.0874	0.0844	0.0407	0.0217	0.0029	0.0047	0.0047
MSE_b2	0.4962	0.2139	0.0631	0.0527	0.0377	0.0239	0.0763	0.0381	0.0229
MSE_b3	0.6955	0.5759	0.2150	0.1216	0.1029	0.0670	0.0178	0.0152	0.0098
MSE_b4	0.7947	1.8545	1.3723	0.1152	0.1441	0.0791	0.0732	0.0593	0.0319
MSE_b5	0.4338	0.6156	0.6372	0.5351	0.6608	0.7443	0.3615	0.3797	0.4023
MSE_b6	0.2637	0.3936	0.3732	0.2478	0.3033	0.3446	0.1653	0.1720	0.1854
MSE_b7	1.1460	1.2441	1.1197	1.5000	1.2087	1.0383	0.1771	0.2779	0.3659
MSE_b8	1.8186	0.9737	0.8129	0.0618	0.2316	0.3085	0.1559	0.0317	0.0791
Bias_b1	-0.1445	-0.0440	0.0188	0.2624	0.1467	0.0846	0.0003	0.0256	0.0542
Bias_b2	-0.6705	-0.3997	-0.1771	-0.2107	-0.1659	-0.1280	-0.2704	-0.1879	-0.1451
Bias_b3	-0.5009	-0.2850	-0.0908	-0.2525	-0.1466	-0.1406	0.0851	0.0198	-0.0174
Bias_b4	-0.1177	-0.1375	-0.1585	0.2013	0.0465	0.0201	0.2168	0.1792	0.1092
Bias_b5	-0.2805	-0.5404	-0.7034	-0.7088	-0.7902	-0.8486	-0.5920	-0.6073	-0.6304
Bias_b6	0.3145	0.4649	0.5395	0.4861	0.5351	0.5775	0.4017	0.4084	0.4279
Bias_b7	-0.3688	-0.6404	-0.8629	-1.1915	-1.0611	-0.9944	-0.3868	-0.4991	-0.5937
Bias_b8	1.1166	0.1152	-0.4885	-0.1254	-0.3566	-0.4816	0.3678	-0.0174	-0.2455
Var_b1	0.1912	0.1887	0.0871	0.0155	0.0191	0.0146	0.0029	0.0040	0.0017
Var_b2	0.0467	0.0541	0.0317	0.0083	0.0102	0.0075	0.0031	0.0028	0.0019
Var_b3	0.4446	0.4947	0.2068	0.0578	0.0814	0.0472	0.0105	0.0148	0.0095
Var_b4	0.7809	1.8356	1.3471	0.0747	0.1419	0.0787	0.0262	0.0272	0.0199
Var_b5	0.3551	0.3236	0.1425	0.0327	0.0363	0.0242	0.0111	0.0109	0.0049
Var_b6	0.1648	0.1775	0.0821	0.0115	0.0169	0.0111	0.0040	0.0052	0.0023
Var_b7	1.0100	0.8339	0.3751	0.0802	0.0827	0.0495	0.0276	0.0289	0.0135
Var_b8	0.5718	0.9605	0.5742	0.0460	0.1045	0.0766	0.0207	0.0314	0.0188

Tabelle 48: Es werden MSE, Bias und Varianz für die einzelnen Koeffizienten (b1, ..., b8) unter unterschiedlich großen Fehlend-Anteilen und unter der jeweiligen Datensatzgröße gezeigt.

Fehlendmechanismus LogitY mit Bruteforce2-Verfahren (lm)

	n=50			n=200			n=1000		
	50%	25%	10%	50%	25%	10%	50%	25%	10%
MSE_b1	0.5109	0.2850	0.2007	0.2493	0.0712	0.0136	0.1473	0.0453	0.0118
MSE_b2	1.9597	1.0086	0.4567	0.7847	0.2633	0.0600	0.7342	0.2309	0.0558
MSE_b3	1.0232	0.8158	0.4639	0.4535	0.1990	0.0578	0.1877	0.0694	0.0193
MSE_b4	1.3423	1.2203	1.0350	0.1405	0.0917	0.0601	0.0128	0.0057	0.0065
MSE_b5	0.9113	0.2642	0.1863	1.0917	0.4519	0.1599	0.8882	0.2792	0.0719
MSE_b6	0.7482	0.2354	0.1376	0.8112	0.3276	0.0979	0.6821	0.2154	0.0540
MSE_b7	0.4796	0.1812	0.1859	1.0359	0.4482	0.1477	0.4169	0.1450	0.0409
MSE_b8	5.0862	4.8858	3.7123	0.0667	0.0379	0.0257	0.1290	0.0700	0.0282
Bias_b1	0.6937	0.5095	0.4150	0.4880	0.2532	0.0929	0.3792	0.2072	0.1029
Bias_b2	-1.3875	-0.9825	-0.6387	-0.8773	-0.5022	-0.2336	-0.8494	-0.4728	-0.2306
Bias_b3	-0.9722	-0.8563	-0.5903	-0.6563	-0.4168	-0.1851	-0.4240	-0.2501	-0.1240
Bias_b4	-1.0644	-0.9616	-0.7418	0.3445	0.2569	0.1919	0.0876	0.0281	0.0024
Bias_b5	-0.8998	-0.4337	-0.3172	-1.0328	-0.6577	-0.3635	-0.9323	-0.5174	-0.2579
Bias_b6	0.8290	0.4355	0.3070	0.8918	0.5613	0.2868	0.8181	0.4554	0.2257
Bias_b7	-0.5746	-0.1692	-0.1359	-1.0021	-0.6434	-0.3358	-0.6343	-0.3671	-0.1852
Bias_b8	2.2272	2.1519	1.7412	0.2204	0.1307	0.0194	0.3502	0.2554	0.1482
Var_b1	0.0297	0.0255	0.0285	0.0111	0.0071	0.0050	0.0035	0.0024	0.0012
Var_b2	0.0346	0.0432	0.0488	0.0151	0.0111	0.0055	0.0127	0.0074	0.0026
Var_b3	0.0780	0.0825	0.1155	0.0228	0.0252	0.0235	0.0079	0.0069	0.0039
Var_b4	0.2093	0.2956	0.4847	0.0218	0.0257	0.0233	0.0052	0.0049	0.0065
Var_b5	0.1016	0.0761	0.0856	0.0249	0.0194	0.0277	0.0191	0.0116	0.0054
Var_b6	0.0609	0.0457	0.0433	0.0159	0.0125	0.0157	0.0129	0.0080	0.0031
Var_b7	0.1494	0.1526	0.1674	0.0318	0.0343	0.0349	0.0145	0.0103	0.0066
Var_b8	0.1260	0.2553	0.6805	0.0181	0.0208	0.0253	0.0063	0.0048	0.0063

Tabelle 49: Es werden MSE, Bias und Varianz für die einzelnen Koeffizienten (b1, ..., b8) unter unterschiedlich großen Fehlend-Anteilen und unter der jeweiligen Datensatzgröße gezeigt.

Fehlendmechanismus LogitY mit Bruteforce2-Verfahren (ridge, $\lambda=0$)

	n=50			n=200			n=1000		
	50%	25%	10%	50%	25%	10%	50%	25%	10%
MSE_b1	0.2714	0.1260	0.0801	0.2156	0.0609	0.0118	0.1191	0.0371	0.0099
MSE_b2	1.7413	0.8658	0.3627	0.7005	0.2328	0.0535	0.6545	0.2080	0.0502
MSE_b3	1.0590	0.8483	0.4847	0.3226	0.1481	0.0462	0.1248	0.0495	0.0147
MSE_b4	0.8369	1.1853	1.7678	0.3902	0.2022	0.1016	0.1666	0.0443	0.0164
MSE_b5	0.4243	0.1052	0.0970	0.7257	0.3184	0.1297	0.5916	0.1953	0.0530
MSE_b6	0.5061	0.1260	0.0691	0.6746	0.2782	0.0864	0.5731	0.1849	0.0472
MSE_b7	0.2004	0.4163	0.4562	0.4932	0.2472	0.1099	0.1418	0.0623	0.0220
MSE_b8	5.7494	5.5724	4.3083	0.3968	0.1680	0.0612	0.5864	0.2305	0.0752
Bias_b1	0.4852	0.3095	0.2181	0.4537	0.2326	0.0815	0.3412	0.1871	0.0934
Bias_b2	-1.3067	-0.9083	-0.5624	-0.8293	-0.4720	-0.2191	-0.8026	-0.4488	-0.2189
Bias_b3	-0.9918	-0.8750	-0.6078	-0.5510	-0.3552	-0.1545	-0.3447	-0.2085	-0.1047
Bias_b4	0.7960	0.8321	1.0185	0.5936	0.4084	0.2676	0.3979	0.1902	0.0802
Bias_b5	-0.5609	-0.1115	-0.0006	-0.8416	-0.5447	-0.3087	-0.7619	-0.4312	-0.2176
Bias_b6	0.6672	0.2794	0.1557	0.8144	0.5162	0.2650	0.7505	0.4218	0.2101
Bias_b7	0.0737	0.4554	0.4718	-0.6826	-0.4503	-0.2427	-0.3654	-0.2297	-0.1205
Bias_b8	2.3716	2.3050	1.8966	0.6041	0.3596	0.1296	0.7591	0.4671	0.2495
Var_b1	0.0360	0.0302	0.0325	0.0097	0.0068	0.0052	0.0027	0.0021	0.0012
Var_b2	0.0338	0.0407	0.0464	0.0127	0.0101	0.0054	0.0104	0.0066	0.0023
Var_b3	0.0752	0.0827	0.1153	0.0190	0.0219	0.0223	0.0061	0.0061	0.0038
Var_b4	0.2034	0.4930	0.7305	0.0379	0.0354	0.0300	0.0083	0.0081	0.0100
Var_b5	0.1097	0.0928	0.0970	0.0174	0.0218	0.0344	0.0112	0.0094	0.0056
Var_b6	0.0609	0.0479	0.0448	0.0113	0.0117	0.0162	0.0099	0.0070	0.0031
Var_b7	0.1950	0.2089	0.2336	0.0272	0.0445	0.0510	0.0083	0.0096	0.0075
Var_b8	0.1250	0.2596	0.7112	0.0318	0.0387	0.0444	0.0103	0.0124	0.0130

Tabelle 50: Es werden MSE, Bias und Varianz für die einzelnen Koeffizienten (b1, ..., b8) unter unterschiedlich großen Fehlend-Anteilen und unter der jeweiligen Datensatzgröße gezeigt.

Fehlendmechanismus LogitY mit Bruteforce2-Verfahren (ridge)

	n=50			n=200			n=1000		
	50%	25%	10%	50%	25%	10%	50%	25%	10%
MSE_b1	0.4139	0.2207	0.1432	0.2298	0.0734	0.0188	0.1408	0.0557	0.0228
MSE_b2	2.2003	1.2236	0.6044	0.8053	0.3253	0.1139	0.7441	0.2884	0.1050
MSE_b3	0.9254	0.7397	0.4185	0.3758	0.1996	0.0797	0.1424	0.0648	0.0249
MSE_b4	0.6418	0.6372	0.9669	0.3653	0.1819	0.0881	0.1930	0.0656	0.0293
MSE_b5	2.3849	1.6706	1.4174	1.6627	1.3553	1.1342	1.2201	0.8399	0.6415
MSE_b6	2.1187	1.4227	1.1263	1.2517	0.8763	0.6053	0.9652	0.5606	0.3629
MSE_b7	1.0245	0.9095	0.9797	1.3526	1.2523	1.1186	0.5032	0.5094	0.4925
MSE_b8	1.6045	1.2925	0.9086	0.1130	0.0345	0.1715	0.2923	0.0379	0.0236
Bias_b1	0.6272	0.4461	0.3420	0.4699	0.2601	0.1183	0.3719	0.2319	0.1473
Bias_b2	-1.4747	-1.0897	-0.7482	-0.8913	-0.5626	-0.3303	-0.8570	-0.5317	-0.3210
Bias_b3	-0.9431	-0.8305	-0.5779	-0.5994	-0.4241	-0.2448	-0.3697	-0.2435	-0.1469
Bias_b4	0.7525	0.6414	0.7277	0.5765	0.3866	0.2488	0.4307	0.2414	0.1429
Bias_b5	-1.5231	-1.2587	-1.1533	-1.2865	-1.1606	-1.0587	-1.1016	-0.9137	-0.7990
Bias_b6	1.4415	1.1717	1.0385	1.1159	0.9331	0.7733	0.9789	0.7457	0.6008
Bias_b7	-0.9646	-0.8658	-0.8933	-1.1557	-1.1079	-1.0460	-0.7056	-0.7099	-0.6984
Bias_b8	1.2338	1.0567	0.7117	0.2924	-0.0640	-0.3712	0.5317	0.1596	-0.1124
Var_b1	0.0206	0.0216	0.0263	0.0090	0.0058	0.0048	0.0024	0.0019	0.0011
Var_b2	0.0257	0.0360	0.0447	0.0109	0.0087	0.0048	0.0097	0.0058	0.0020
Var_b3	0.0360	0.0500	0.0845	0.0164	0.0197	0.0198	0.0057	0.0055	0.0033
Var_b4	0.0755	0.2258	0.4373	0.0330	0.0324	0.0262	0.0076	0.0074	0.0089
Var_b5	0.0651	0.0864	0.0873	0.0075	0.0084	0.0134	0.0065	0.0051	0.0031
Var_b6	0.0407	0.0499	0.0478	0.0065	0.0056	0.0073	0.0069	0.0045	0.0019
Var_b7	0.0939	0.1599	0.1818	0.0169	0.0249	0.0245	0.0053	0.0055	0.0047
Var_b8	0.0824	0.1760	0.4021	0.0275	0.0304	0.0337	0.0097	0.0125	0.0109

Tabelle 51: Es werden MSE, Bias und Varianz für die einzelnen Koeffizienten (b1, ..., b8) unter unterschiedlich großen Fehlend-Anteilen und unter der jeweiligen Datensatzgröße gezeigt.

Fehlendmechanismus LogitX mit Bruteforce2-Verfahren (lm)

	n=50			n=200			n=1000		
	50%	25%	10%	50%	25%	10%	50%	25%	10%
MSE_b1	0.4514	0.2542	0.1935	0.2127	0.0470	0.0115	0.1026	0.0279	0.0070
MSE_b2	1.8412	0.8347	0.3727	0.6373	0.1880	0.0396	0.5552	0.1489	0.0328
MSE_b3	1.1060	0.7558	0.3572	0.3945	0.1486	0.0434	0.1280	0.0454	0.0147
MSE_b4	1.2534	1.1228	1.0725	0.1589	0.0767	0.0538	0.0165	0.0070	0.0062
MSE_b5	0.6956	0.2222	0.1481	0.8823	0.3288	0.1149	0.6522	0.1797	0.0418
MSE_b6	0.6074	0.1810	0.1163	0.6609	0.2404	0.0680	0.5080	0.1407	0.0313
MSE_b7	0.3556	0.1722	0.1532	0.8750	0.3224	0.1069	0.2764	0.0947	0.0255
MSE_b8	5.4331	4.8392	3.6280	0.1114	0.0381	0.0310	0.1828	0.0739	0.0252
Bias_b1	0.6466	0.4790	0.4097	0.4492	0.1950	0.0656	0.3138	0.1607	0.0767
Bias_b2	-1.3454	-0.8927	-0.5732	-0.7862	-0.4163	-0.1777	-0.7366	-0.3746	-0.1732
Bias_b3	-1.0203	-0.8219	-0.5138	-0.6128	-0.3473	-0.1400	-0.3455	-0.2001	-0.0992
Bias_b4	-1.0284	-0.8863	-0.8161	0.3697	0.2326	0.1702	0.1029	0.0277	0.0031
Bias_b5	-0.7728	-0.3670	-0.2837	-0.9201	-0.5441	-0.3064	-0.7953	-0.4108	-0.1932
Bias_b6	0.7392	0.3626	0.2869	0.7990	0.4692	0.2396	0.7033	0.3634	0.1691
Bias_b7	-0.4250	-0.1382	-0.1187	-0.9073	-0.5332	-0.2764	-0.5095	-0.2904	-0.1392
Bias_b8	2.2873	2.1361	1.7584	0.3054	0.1317	-0.0001	0.4214	0.2616	0.1345
Var_b1	0.0334	0.0248	0.0257	0.0109	0.0090	0.0072	0.0041	0.0021	0.0012
Var_b2	0.0310	0.0378	0.0442	0.0191	0.0147	0.0080	0.0127	0.0086	0.0028
Var_b3	0.0651	0.0803	0.0932	0.0190	0.0279	0.0238	0.0086	0.0054	0.0048
Var_b4	0.1958	0.3374	0.4065	0.0222	0.0226	0.0248	0.0059	0.0062	0.0062
Var_b5	0.0984	0.0875	0.0676	0.0357	0.0328	0.0210	0.0196	0.0109	0.0045
Var_b6	0.0610	0.0495	0.0339	0.0226	0.0203	0.0106	0.0134	0.0086	0.0028
Var_b7	0.1750	0.1531	0.1392	0.0517	0.0381	0.0305	0.0167	0.0104	0.0062
Var_b8	0.2013	0.2761	0.5359	0.0181	0.0208	0.0310	0.0052	0.0054	0.0072

Tabelle 52: Es werden MSE, Bias und Varianz für die einzelnen Koeffizienten (b1, ..., b8) unter unterschiedlich großen Fehlend-Anteilen und unter der jeweiligen Datensatzgröße gezeigt.

Fehlendmechanismus LogitX mit Bruteforce2-Verfahren (ridge, $\lambda=0$)

	n=50			n=200			n=1000		
	50%	25%	10%	50%	25%	10%	50%	25%	10%
MSE_b1	0.2334	0.1123	0.0750	0.1876	0.0400	0.0102	0.0854	0.0234	0.0060
MSE_b2	1.6359	0.6963	0.2898	0.5825	0.1676	0.0339	0.5033	0.1354	0.0299
MSE_b3	1.1470	0.7818	0.3801	0.3052	0.1128	0.0351	0.0884	0.0333	0.0115
MSE_b4	0.9724	1.3866	1.6425	0.3453	0.1620	0.0964	0.1225	0.0303	0.0132
MSE_b5	0.3014	0.1145	0.0913	0.6389	0.2376	0.0885	0.4620	0.1305	0.0321
MSE_b6	0.3937	0.0989	0.0568	0.5680	0.2059	0.0577	0.4389	0.1221	0.0277
MSE_b7	0.2619	0.4720	0.4499	0.5017	0.1904	0.0750	0.1068	0.0446	0.0151
MSE_b8	6.1360	5.4629	4.1450	0.3702	0.1288	0.0684	0.5355	0.1847	0.0569
Bias_b1	0.4432	0.2777	0.2111	0.4226	0.1779	0.0548	0.2868	0.1473	0.0702
Bias_b2	-1.2678	-0.8127	-0.4957	-0.7523	-0.3930	-0.1633	-0.7018	-0.3574	-0.1655
Bias_b3	-1.0392	-0.8368	-0.5329	-0.5378	-0.2954	-0.1102	-0.2862	-0.1688	-0.0845
Bias_b4	0.8329	0.9440	0.9712	0.5521	0.3522	0.2444	0.3344	0.1445	0.0594
Bias_b5	-0.4393	-0.0456	0.0340	-0.7831	-0.4537	-0.2519	-0.6705	-0.3491	-0.1655
Bias_b6	0.5759	0.2083	0.1343	0.7421	0.4327	0.2175	0.6545	0.3384	0.1582
Bias_b7	0.2154	0.4791	0.4911	-0.6773	-0.3803	-0.1865	-0.3124	-0.1906	-0.0932
Bias_b8	2.4367	2.2790	1.8966	0.5815	0.3073	0.1109	0.7249	0.4163	0.2070
Var_b1	0.0370	0.0352	0.0304	0.0090	0.0084	0.0071	0.0031	0.0017	0.0011
Var_b2	0.0285	0.0359	0.0441	0.0166	0.0131	0.0073	0.0107	0.0076	0.0025
Var_b3	0.0670	0.0815	0.0961	0.0160	0.0255	0.0230	0.0064	0.0048	0.0044
Var_b4	0.2787	0.4955	0.6992	0.0404	0.0380	0.0367	0.0107	0.0095	0.0097
Var_b5	0.1084	0.1124	0.0901	0.0257	0.0318	0.0250	0.0124	0.0087	0.0047
Var_b6	0.0621	0.0555	0.0388	0.0173	0.0187	0.0104	0.0106	0.0076	0.0027
Var_b7	0.2156	0.2424	0.2087	0.0430	0.0458	0.0403	0.0092	0.0083	0.0064
Var_b8	0.1987	0.2690	0.5480	0.0320	0.0343	0.0561	0.0100	0.0114	0.0140

Tabelle 53: Es werden MSE, Bias und Varianz für die einzelnen Koeffizienten (b1, ..., b8) unter unterschiedlich großen Fehlend-Anteilen und unter der jeweiligen Datensatzgröße gezeigt.

Fehlendmechanismus LogitX mit Bruteforce2-Verfahren (ridge)

	n=50			n=200			n=1000		
	50%	25%	10%	50%	25%	10%	50%	25%	10%
MSE_b1	0.3744	0.1957	0.1357	0.2011	0.0503	0.0148	0.1055	0.0398	0.0168
MSE_b2	2.1092	1.0350	0.5180	0.6840	0.2523	0.0846	0.5913	0.2048	0.0754
MSE_b3	0.9807	0.6730	0.3307	0.3537	0.1603	0.0615	0.1039	0.0467	0.0206
MSE_b4	0.6158	0.7484	0.8484	0.3212	0.1440	0.0826	0.1459	0.0467	0.0234
MSE_b5	2.1914	1.5712	1.3635	1.5993	1.2456	1.0695	1.0863	0.7448	0.5869
MSE_b6	1.9737	1.2993	1.0917	1.1433	0.7793	0.5597	0.8207	0.4703	0.3166
MSE_b7	0.9294	0.9281	0.9548	1.4147	1.2031	1.0608	0.4743	0.4973	0.4799
MSE_b8	1.7107	1.2473	0.7900	0.0904	0.0504	0.2164	0.2361	0.0193	0.0394
Bias_b1	0.5940	0.4124	0.3338	0.4392	0.2067	0.0923	0.3202	0.1957	0.1255
Bias_b2	-1.4449	-1.0010	-0.6858	-0.8183	-0.4905	-0.2800	-0.7628	-0.4451	-0.2707
Bias_b3	-0.9742	-0.7898	-0.5093	-0.5825	-0.3715	-0.2043	-0.3127	-0.2056	-0.1288
Bias_b4	0.7283	0.7129	0.6834	0.5343	0.3328	0.2254	0.3689	0.1949	0.1225
Bias_b5	-1.4585	-1.2146	-1.1331	-1.2605	-1.1102	-1.0292	-1.0388	-0.8605	-0.7644
Bias_b6	1.3911	1.1155	1.0253	1.0654	0.8775	0.7449	0.9019	0.6823	0.5611
Bias_b7	-0.9030	-0.8728	-0.8969	-1.1803	-1.0859	-1.0195	-0.6845	-0.7020	-0.6897
Bias_b8	1.2611	1.0229	0.6901	0.2508	-0.1412	-0.4113	0.4741	0.0893	-0.1660
Var_b1	0.0215	0.0255	0.0243	0.0082	0.0075	0.0063	0.0030	0.0015	0.0010
Var_b2	0.0215	0.0330	0.0477	0.0145	0.0117	0.0062	0.0094	0.0067	0.0022
Var_b3	0.0316	0.0492	0.0713	0.0143	0.0223	0.0198	0.0061	0.0045	0.0040
Var_b4	0.0854	0.2402	0.3813	0.0357	0.0332	0.0318	0.0097	0.0087	0.0084
Var_b5	0.0642	0.0959	0.0796	0.0103	0.0130	0.0103	0.0072	0.0043	0.0027
Var_b6	0.0386	0.0549	0.0404	0.0083	0.0093	0.0049	0.0072	0.0048	0.0017
Var_b7	0.1141	0.1663	0.1504	0.0215	0.0238	0.0214	0.0057	0.0045	0.0042
Var_b8	0.1203	0.2010	0.3137	0.0275	0.0305	0.0473	0.0114	0.0113	0.0118

Tabelle 54: Es werden MSE, Bias und Varianz für die einzelnen Koeffizienten (b1, ..., b8) unter unterschiedlich großen Fehlend-Anteilen und unter der jeweiligen Datensatzgröße gezeigt.

D Koeffizientenboxplots zur logistischen Regression

D.1 Bruteforce2 mit logistischer Regression

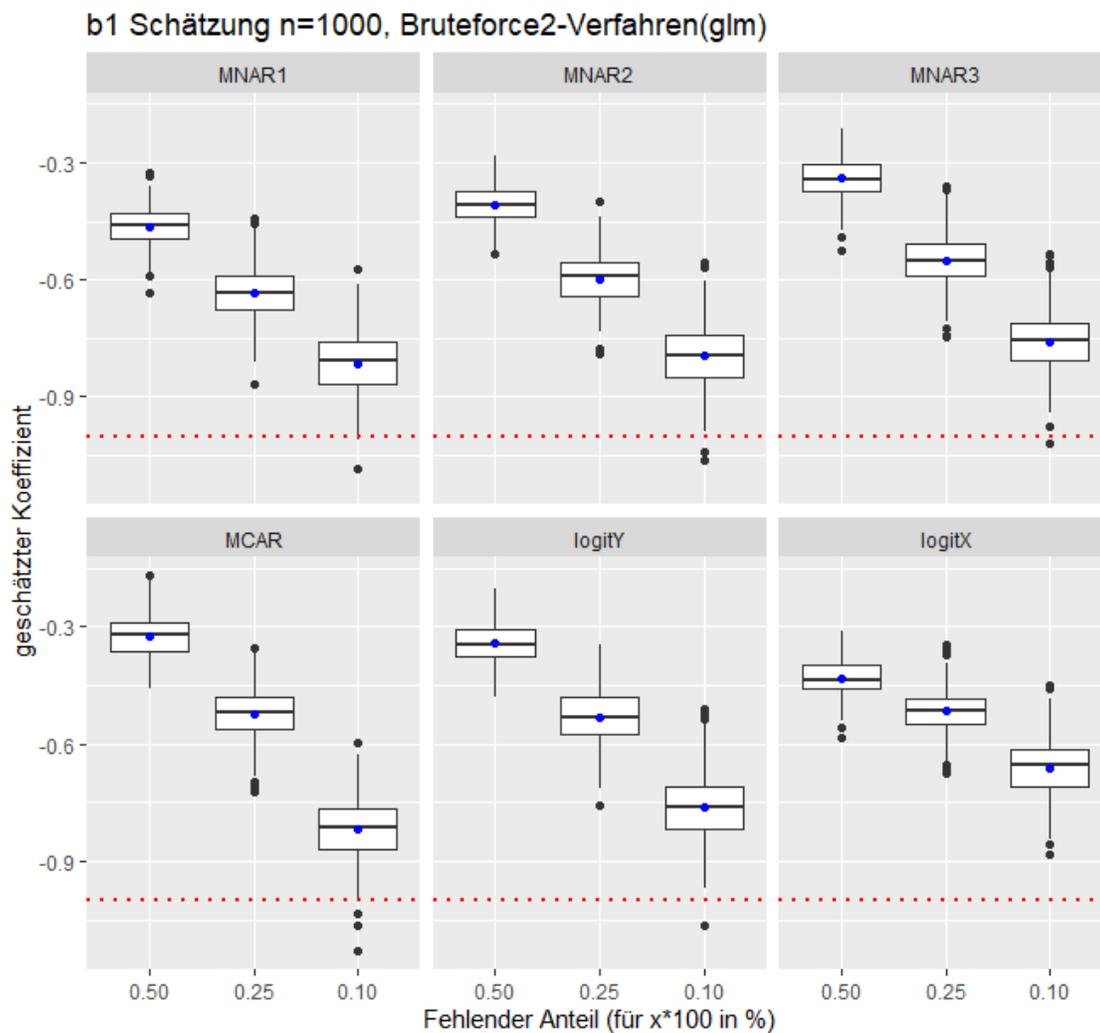


Abbildung 53: Es werden die geschätzten Werte für β_1 in Boxplotform dargestellt. Diese werden nach den Fehlendmechanismen gruppiert. Der blaue Punkt gibt den Mittelwert der geschätzten Werte an. Die rote horizontale Linie repräsentiert den echten Wert des Koeffizienten.

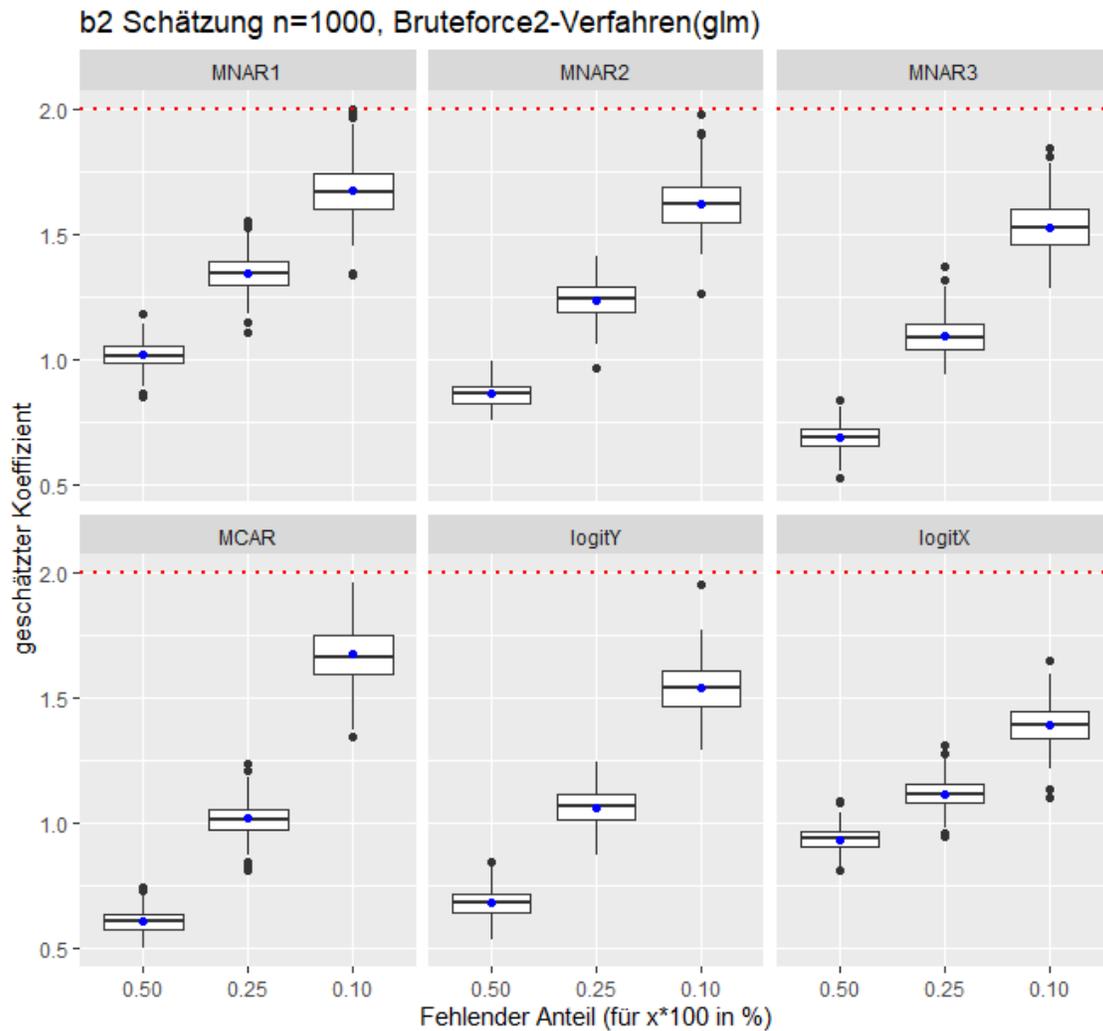


Abbildung 54: Es werden die geschätzten Werte für β_2 in Boxplotform dargestellt. Diese werden nach den Fehlendmechanismen gruppiert. Der blaue Punkt gibt den Mittelwert der geschätzten Werte an. Die rote horizontale Linie repräsentiert den echten Wert des Koeffizienten.

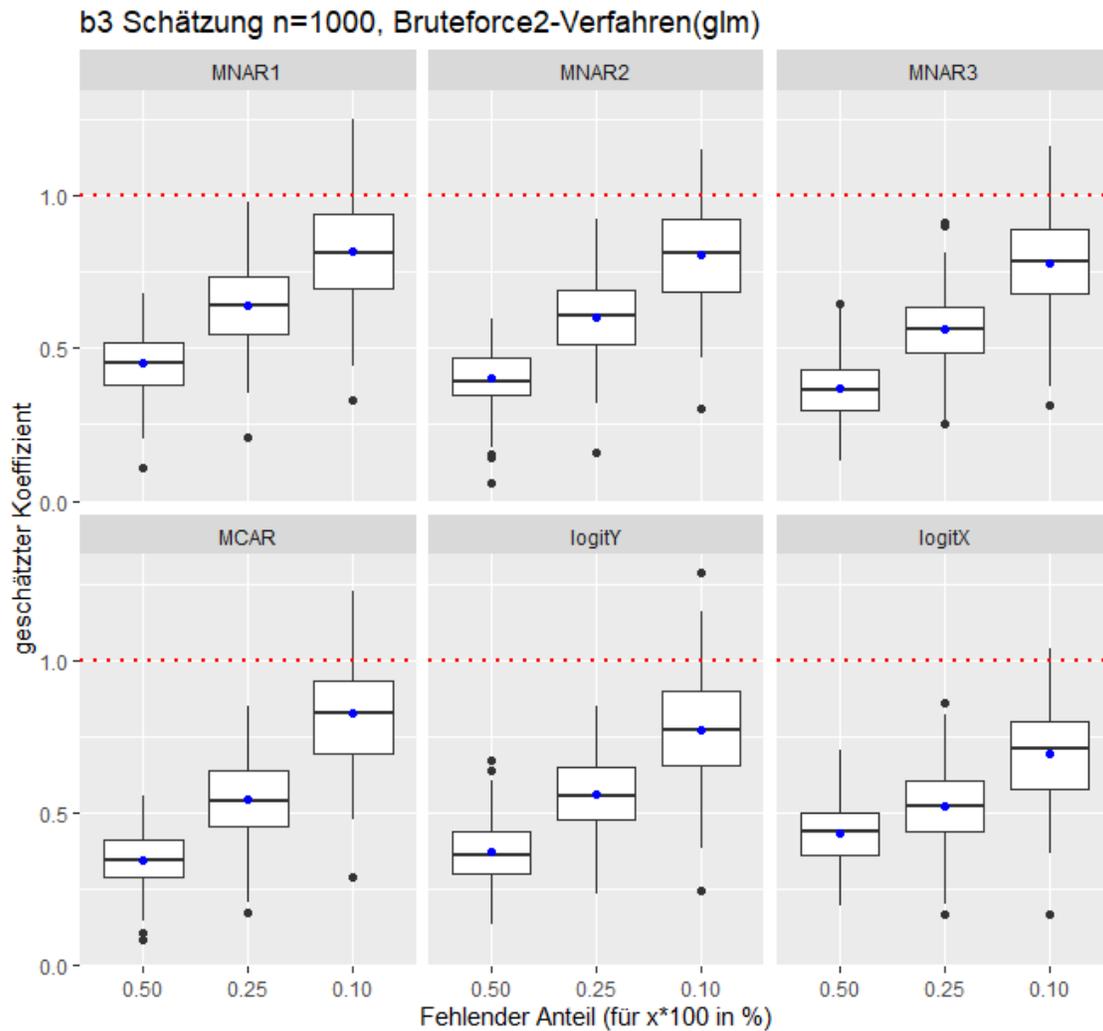


Abbildung 55: Es werden die geschätzten Werte für β_3 in Boxplotform dargestellt. Diese werden nach den Fehlendmechanismen gruppiert. Der blaue Punkt gibt den Mittelwert der geschätzten Werte an. Die rote horizontale Linie repräsentiert den echten Wert des Koeffizienten.

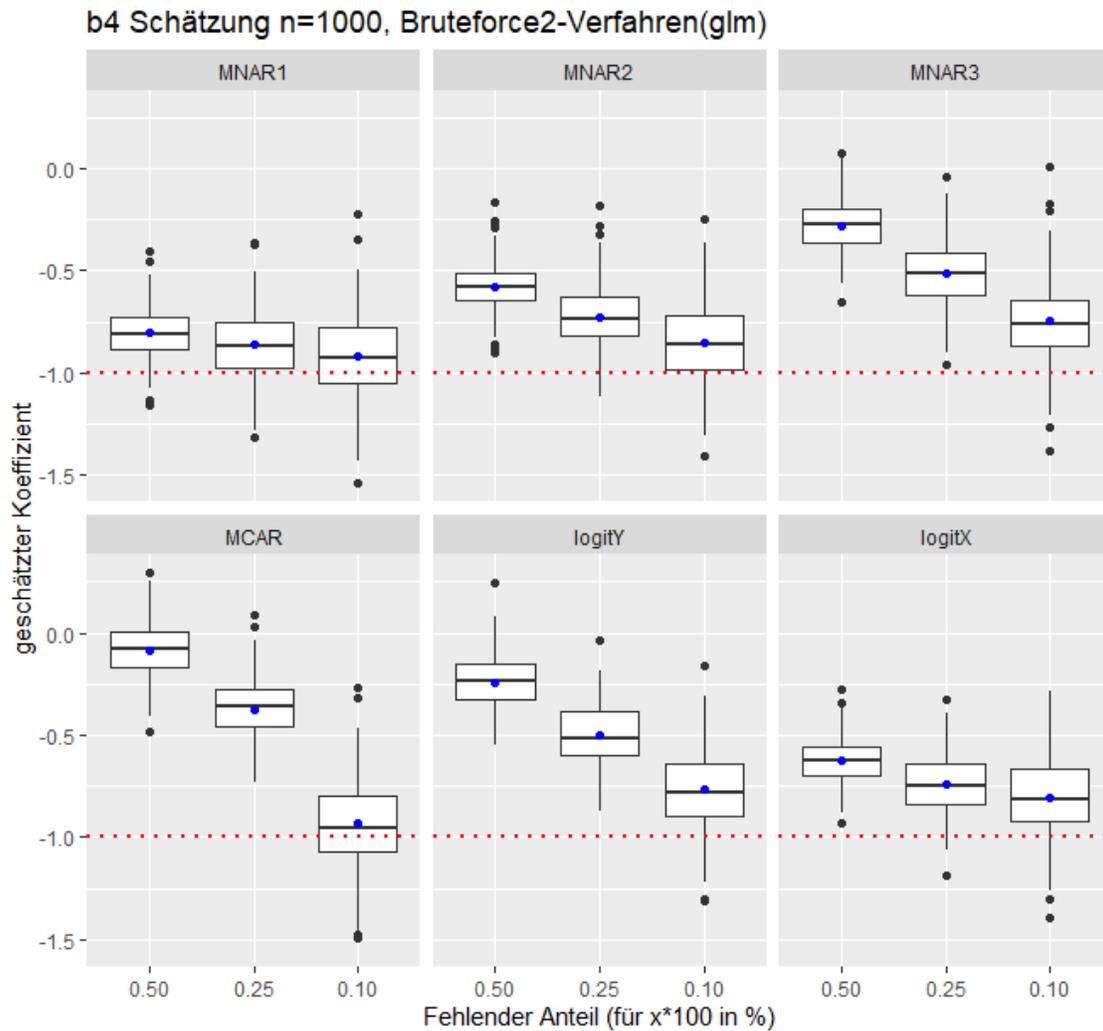


Abbildung 56: Es werden die geschätzten Werte für β_4 in Boxplotform dargestellt. Diese werden nach den Fehlendmechanismen gruppiert. Der blaue Punkt gibt den Mittelwert der geschätzten Werte an. Die rote horizontale Linie repräsentiert den echten Wert des Koeffizienten.

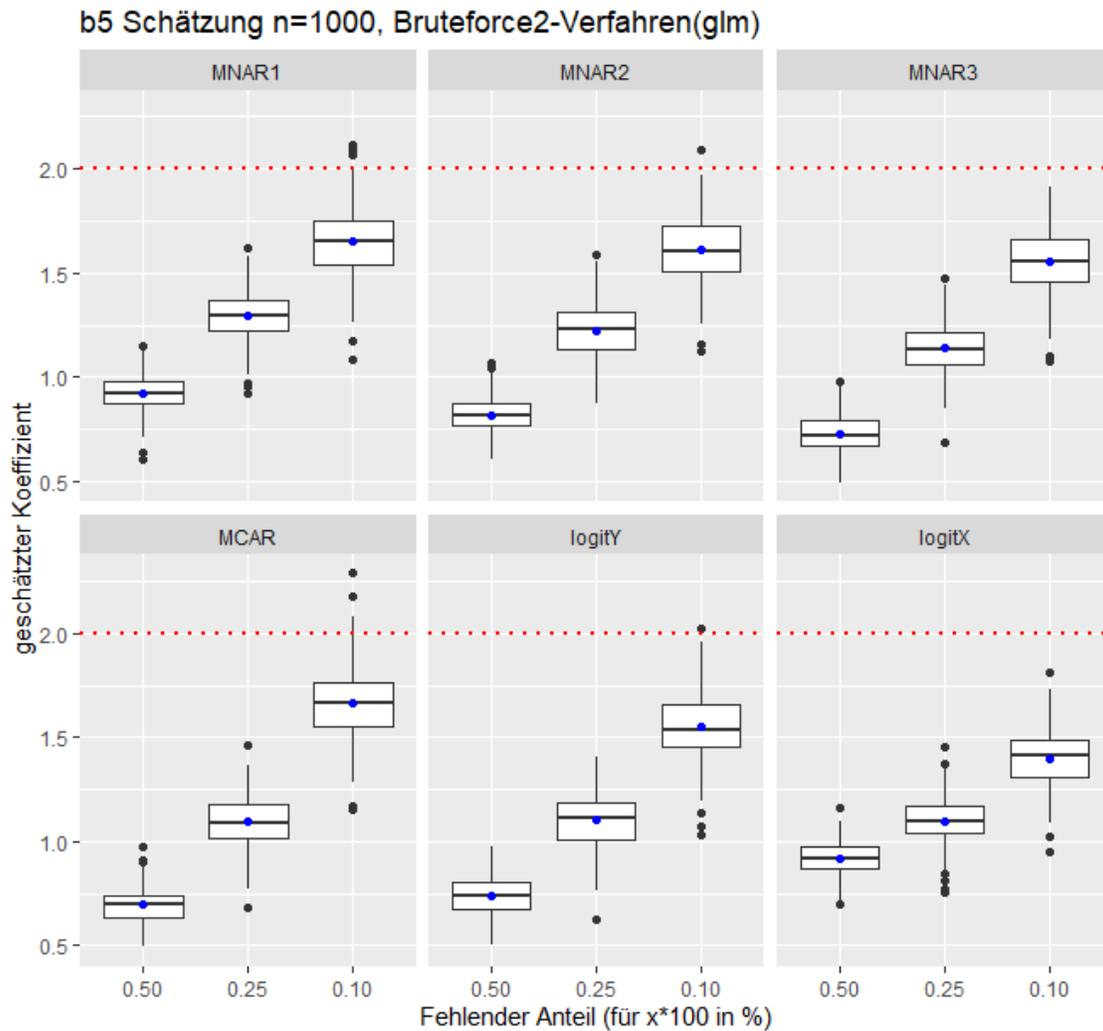


Abbildung 57: Es werden die geschätzten Werte für β_5 in Boxplotform dargestellt. Diese werden nach den Fehlendmechanismen gruppiert. Der blaue Punkt gibt den Mittelwert der geschätzten Werte an. Die rote horizontale Linie repräsentiert den echten Wert des Koeffizienten.

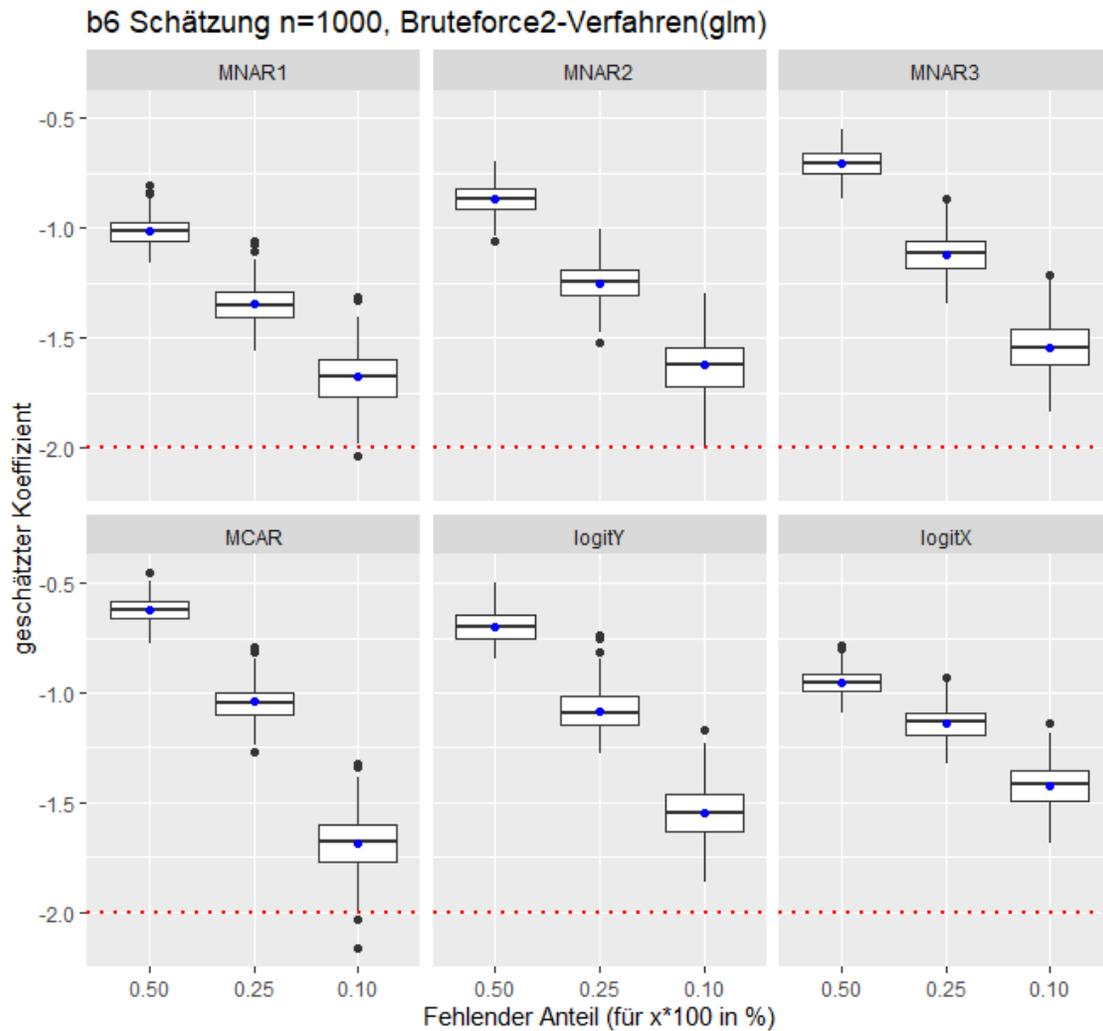


Abbildung 58: Es werden die geschätzten Werte für β_6 in Boxplotform dargestellt. Diese werden nach den Fehlendmechanismen gruppiert. Der blaue Punkt gibt den Mittelwert der geschätzten Werte an. Die rote horizontale Linie repräsentiert den echten Wert des Koeffizienten.

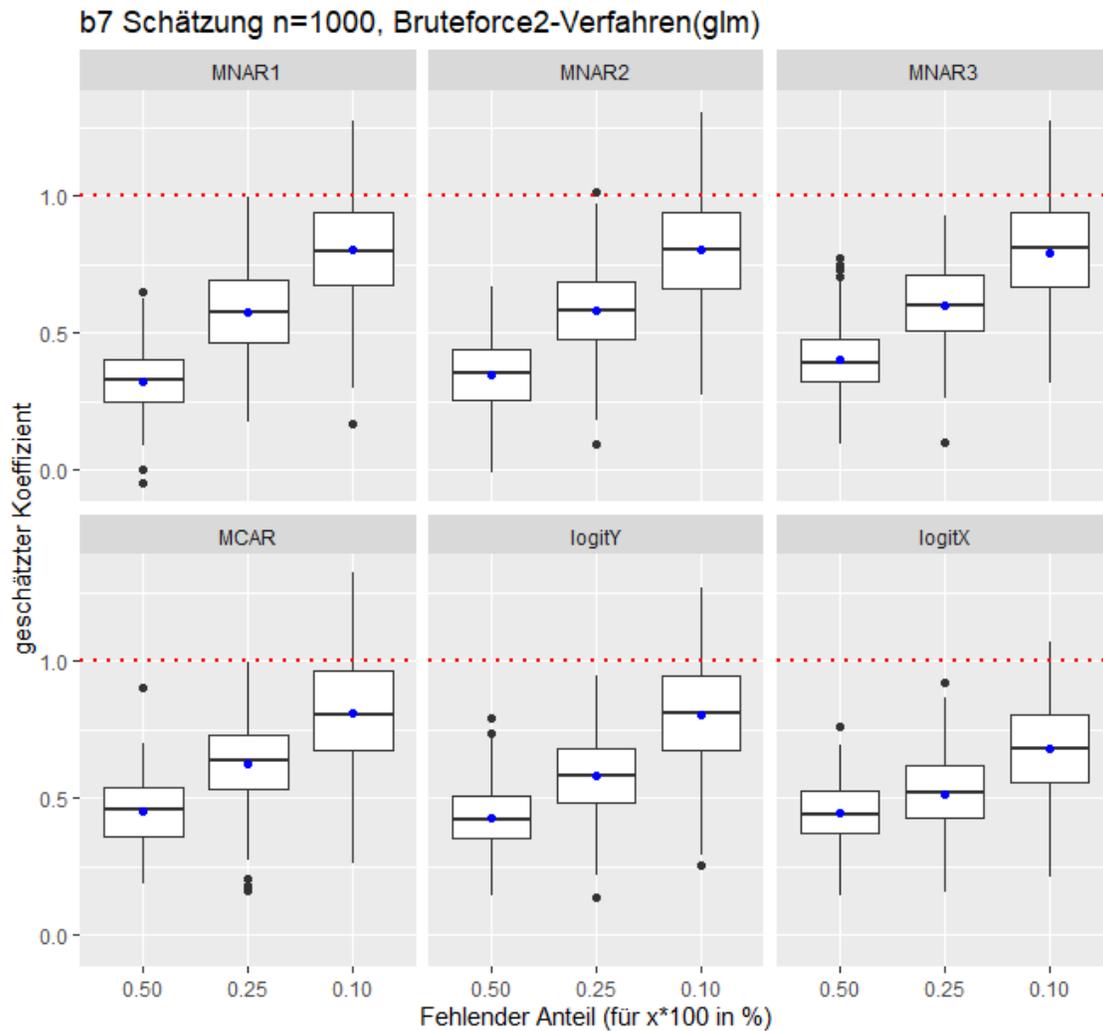


Abbildung 59: Es werden die geschätzten Werte für β_7 in Boxplotform dargestellt. Diese werden nach den Fehlendmechanismen gruppiert. Der blaue Punkt gibt den Mittelwert der geschätzten Werte an. Die rote horizontale Linie repräsentiert den echten Wert des Koeffizienten.

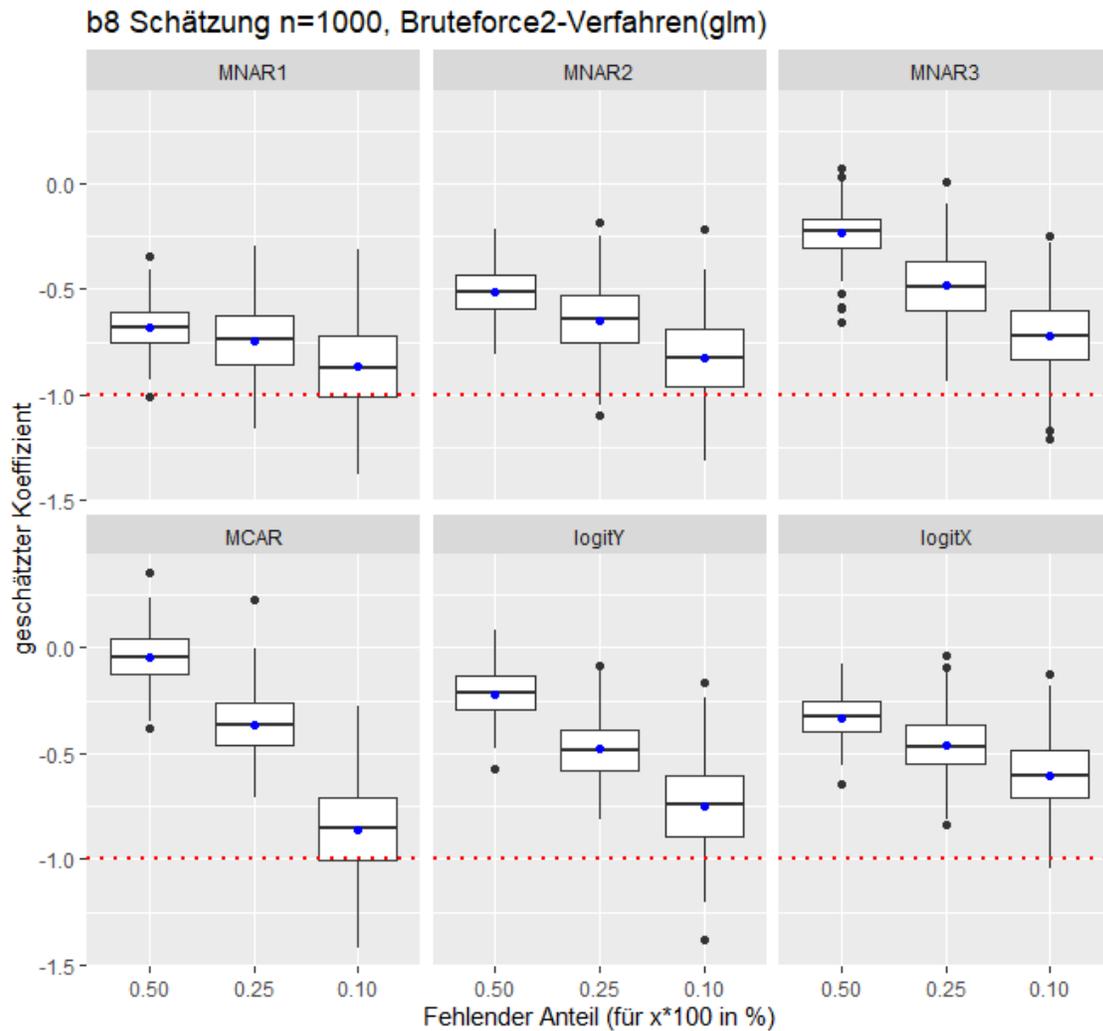


Abbildung 60: Es werden die geschätzten Werte für β_8 in Boxplotform dargestellt. Diese werden nach den Fehlendmechanismen gruppiert. Der blaue Punkt gibt den Mittelwert der geschätzten Werte an. Die rote horizontale Linie repräsentiert den echten Wert des Koeffizienten.

E Gütekriterien-Tabellen zur logistischen Regression

E.1 Normales Verfahren

Fehlendmechanismus MCAR mit normalen Verfahren (glm)						
	n=200			n=1000		
	50%	25%	10%	50%	25%	10%
MSE_b1	0.6267	0.2065	0.1303	0.0313	0.0205	0.0162
MSE_b2	1.6825	0.4105	0.2778	0.0575	0.0392	0.0288
MSE_b3	2.3758	0.4919	0.3197	0.0998	0.0625	0.0469
MSE_b4	1.9509	0.9625	0.5537	0.1890	0.1000	0.0834
MSE_b5	1.7380	0.7800	0.3927	0.1142	0.0859	0.0570
MSE_b6	1.5107	0.5661	0.3177	0.0697	0.0565	0.0368
MSE_b7	1.6415	0.9187	0.4968	0.1642	0.0986	0.0705
MSE_b8	2.5719	0.9146	0.6155	0.1938	0.1080	0.0896
Bias_b1	-0.3611	-0.1761	-0.1391	-0.0388	-0.0194	-0.0043
Bias_b2	0.6652	0.3123	0.2048	0.0487	0.0451	0.0346
Bias_b3	0.4026	0.2064	0.0456	0.0165	0.0435	0.0189
Bias_b4	-0.2388	-0.2106	-0.2432	-0.0462	-0.0391	-0.0896
Bias_b5	0.5942	0.3447	0.1807	0.0908	0.0700	0.0183
Bias_b6	-0.6364	-0.3317	-0.2191	-0.0729	-0.0637	-0.0376
Bias_b7	0.2120	0.1815	0.0665	0.0780	0.0735	-0.0156
Bias_b8	-0.3369	-0.1234	-0.1610	-0.0255	-0.0201	-0.1081
Var_b1	0.4964	0.1755	0.1109	0.0298	0.0201	0.0162
Var_b2	1.2400	0.3130	0.2358	0.0552	0.0371	0.0276
Var_b3	2.2137	0.4492	0.3177	0.0995	0.0606	0.0466
Var_b4	1.8938	0.9182	0.4945	0.1869	0.0985	0.0754
Var_b5	1.3849	0.6611	0.3600	0.1060	0.0810	0.0567
Var_b6	1.1057	0.4561	0.2697	0.0644	0.0524	0.0354
Var_b7	1.5966	0.8858	0.4924	0.1581	0.0932	0.0703
Var_b8	2.4584	0.8994	0.5896	0.1931	0.1076	0.0779

Tabelle 55: Es werden MSE, Bias und Varianz für die einzelnen Koeffizienten (b1, ..., b8) unter unterschiedlich großen Fehlend-Anteilen und unter der jeweiligen Datensatzgröße gezeigt.

Fehlendmechanismus MNAR1 mit normalen Verfahren (glm)						
	n=200			n=1000		
	50%	25%	10%	50%	25%	10%
MSE_b1	4092.9068	0.1839	0.1183	0.0930	0.0210	0.0158
MSE_b2	6950.8009	0.4633	0.2472	0.1741	0.0318	0.0282
MSE_b3	6404.1689	0.5282	0.2875	0.3459	0.0644	0.0459
MSE_b4	11896.4026	0.9485	0.5235	1.0929	0.1306	0.0852
MSE_b5	12813.9063	0.5742	0.3523	0.5372	0.0751	0.0575
MSE_b6	9530.0196	0.4889	0.2730	0.1968	0.0441	0.0352
MSE_b7	15844.0174	0.7877	0.4771	1.2676	0.1129	0.0737
MSE_b8	24838.1650	1.1606	0.6169	2.3862	0.2073	0.0910
Bias_b1	-15.1783	-0.1903	-0.1263	0.0591	0.0261	0.0002
Bias_b2	27.8868	0.2927	0.2046	0.0086	0.0091	0.0283
Bias_b3	0.1232	0.0749	0.0550	-0.1953	-0.0275	0.0094
Bias_b4	-52.9142	-0.4196	-0.2054	-0.8502	-0.1799	-0.0770
Bias_b5	-13.0433	0.1727	0.1657	-0.5404	-0.0667	0.0183
Bias_b6	-21.2954	-0.2867	-0.2048	0.1421	0.0028	-0.0305
Bias_b7	-34.8232	-0.0210	0.0370	-0.9377	-0.1377	0.0009
Bias_b8	-75.3782	-0.5481	-0.1845	-1.3594	-0.3221	-0.1225
Var_b1	3862.5248	0.1476	0.1023	0.0896	0.0203	0.0158
Var_b2	6173.1250	0.3776	0.2053	0.1740	0.0317	0.0274
Var_b3	6404.1537	0.5226	0.2845	0.3077	0.0637	0.0458
Var_b4	9096.4934	0.7724	0.4813	0.3700	0.0982	0.0793
Var_b5	12643.7775	0.5443	0.3248	0.2452	0.0706	0.0572
Var_b6	9076.5263	0.4067	0.2311	0.1766	0.0441	0.0343
Var_b7	14631.3655	0.7873	0.4757	0.3884	0.0940	0.0737
Var_b8	19156.2931	0.8602	0.5829	0.5383	0.1035	0.0760

Tabelle 56: Es werden MSE, Bias und Varianz für die einzelnen Koeffizienten (b1, ..., b8) unter unterschiedlich großen Fehlend-Anteilen und unter der jeweiligen Datensatzgröße gezeigt.

Fehlendmechanismus MNAR2 mit normalen Verfahren (glm)						
	n=200			n=1000		
	50%	25%	10%	50%	25%	10%
MSE_b1	349828.5990	0.1947	0.1209	0.0419	0.0187	0.0145
MSE_b2	5437370.0978	0.4342	0.2885	0.0719	0.0370	0.0290
MSE_b3	1322975.3858	0.5072	0.3166	0.1621	0.0684	0.0507
MSE_b4	12935536.0150	0.9530	0.6013	0.4764	0.1082	0.0769
MSE_b5	6010632.0828	0.6449	0.4118	0.1992	0.0776	0.0566
MSE_b6	5536667.2296	0.5337	0.3223	0.0834	0.0450	0.0377
MSE_b7	2222961.2361	0.9062	0.4835	0.4521	0.1000	0.0677
MSE_b8	440898.0018	0.9813	0.5717	0.9912	0.1703	0.0869
Bias_b1	-79.4931	-0.1762	-0.1172	0.0417	0.0115	0.0013
Bias_b2	231.8916	0.3245	0.2126	0.0117	0.0234	0.0259
Bias_b3	108.6147	0.0969	0.0819	-0.1200	-0.0293	0.0130
Bias_b4	-320.4960	-0.4184	-0.1935	-0.5034	-0.1557	-0.0638
Bias_b5	232.3705	0.2110	0.1859	-0.2695	-0.0485	0.0258
Bias_b6	-233.7252	-0.3074	-0.2173	0.0559	-0.0024	-0.0356
Bias_b7	107.4632	0.0687	0.0399	-0.5098	-0.1013	0.0170
Bias_b8	-29.2127	-0.4422	-0.1544	-0.8625	-0.2495	-0.0929
Var_b1	343509.4489	0.1636	0.1072	0.0402	0.0185	0.0145
Var_b2	5383596.3884	0.3289	0.2433	0.0718	0.0365	0.0283
Var_b3	1311178.2406	0.4978	0.3099	0.1477	0.0676	0.0506
Var_b4	12832818.3008	0.7780	0.5638	0.2230	0.0839	0.0729
Var_b5	5956636.0156	0.6004	0.3772	0.1266	0.0753	0.0559
Var_b6	5482039.7456	0.4392	0.2751	0.0802	0.0450	0.0364
Var_b7	2211412.8875	0.9015	0.4819	0.1921	0.0897	0.0675
Var_b8	440044.6214	0.7857	0.5479	0.2473	0.1080	0.0782

Tabelle 57: Es werden MSE, Bias und Varianz für die einzelnen Koeffizienten (b1, ..., b8) unter unterschiedlich großen Fehlend-Anteilen und unter der jeweiligen Datensatzgröße gezeigt.

Fehlendmechanismus MNAR3 mit normalen Verfahren (glm)

	n=200			n=1000		
	50%	25%	10%	50%	25%	10%
MSE_b1	1380.0907	0.1708	0.1169	0.0336	0.0232	0.0150
MSE_b2	1536.5682	0.3811	0.2762	0.0785	0.0449	0.0293
MSE_b3	37.7452	0.4899	0.3771	0.1126	0.0627	0.0470
MSE_b4	669.2478	0.7153	0.5074	0.2007	0.1119	0.0794
MSE_b5	2359.7846	0.5530	0.3687	0.1422	0.0723	0.0631
MSE_b6	1867.6256	0.4086	0.2882	0.0913	0.0501	0.0360
MSE_b7	244.0041	0.8267	0.4426	0.1677	0.0848	0.0740
MSE_b8	2441.8559	0.7663	0.5247	0.3564	0.1362	0.0873
Bias_b1	-3.4834	-0.1590	-0.1304	-0.0186	-0.0235	-0.0077
Bias_b2	4.0377	0.2726	0.2236	0.0628	0.0513	0.0340
Bias_b3	0.3263	0.0978	0.1364	0.0346	0.0122	0.0241
Bias_b4	-3.0483	-0.2400	-0.1579	-0.2180	-0.0872	-0.0382
Bias_b5	4.8208	0.2216	0.2019	-0.0089	0.0343	0.0450
Bias_b6	-4.4685	-0.2872	-0.2332	-0.0658	-0.0577	-0.0451
Bias_b7	1.3484	0.0396	0.0845	-0.1196	-0.0098	0.0361
Bias_b8	-5.0683	-0.2262	-0.1404	-0.3987	-0.1289	-0.0533
Var_b1	1367.9568	0.1455	0.0999	0.0333	0.0226	0.0149
Var_b2	1520.2654	0.3068	0.2262	0.0745	0.0423	0.0282
Var_b3	37.6388	0.4804	0.3585	0.1114	0.0626	0.0464
Var_b4	659.9556	0.6577	0.4824	0.1532	0.1043	0.0779
Var_b5	2336.5449	0.5039	0.3279	0.1421	0.0712	0.0610
Var_b6	1847.6580	0.3261	0.2339	0.0870	0.0468	0.0340
Var_b7	242.1858	0.8252	0.4355	0.1534	0.0847	0.0727
Var_b8	2416.1685	0.7151	0.5050	0.1974	0.1196	0.0844

Tabelle 58: Es werden MSE, Bias und Varianz für die einzelnen Koeffizienten (b1, ..., b8) unter unterschiedlich großen Fehlend-Anteilen und unter der jeweiligen Datensatzgröße gezeigt.

Fehlendmechanismus LogitY mit normalen Verfahren (glm)

	n=200			n=1000		
	50%	25%	10%	50%	25%	10%
MSE_b1	0.9966	0.2186	0.1269	0.0324	0.0212	0.0162
MSE_b2	3.6089	0.4375	0.2878	0.0604	0.0357	0.0263
MSE_b3	1.7540	0.5449	0.3683	0.0821	0.0684	0.0486
MSE_b4	3.1453	0.8839	0.5554	0.1933	0.1234	0.0726
MSE_b5	2.7223	0.5811	0.3637	0.1143	0.0716	0.0613
MSE_b6	2.8675	0.4962	0.2777	0.0749	0.0444	0.0347
MSE_b7	2.7706	0.7641	0.4415	0.1633	0.0985	0.0731
MSE_b8	3.2049	0.9107	0.6398	0.2242	0.1382	0.0845
Bias_b1	-0.4207	-0.1968	-0.1385	-0.0015	-0.0057	-0.0060
Bias_b2	0.8019	0.3144	0.2288	0.0551	0.0336	0.0308
Bias_b3	0.3230	0.1474	0.1160	-0.0220	0.0105	0.0197
Bias_b4	-0.7080	-0.3061	-0.1709	-0.1766	-0.1077	-0.0501
Bias_b5	0.6164	0.2874	0.1936	0.0138	0.0246	0.0359
Bias_b6	-0.7472	-0.3309	-0.2125	-0.0539	-0.0368	-0.0365
Bias_b7	0.2467	0.0867	0.0586	-0.0737	0.0033	0.0297
Bias_b8	-0.6891	-0.2610	-0.1754	-0.2631	-0.1579	-0.0661
Var_b1	0.8197	0.1798	0.1077	0.0324	0.0212	0.0162
Var_b2	2.9658	0.3386	0.2354	0.0573	0.0345	0.0254
Var_b3	1.6497	0.5232	0.3548	0.0816	0.0683	0.0483
Var_b4	2.6440	0.7902	0.5262	0.1621	0.1118	0.0701
Var_b5	2.3423	0.4985	0.3262	0.1141	0.0710	0.0600
Var_b6	2.3092	0.3867	0.2325	0.0720	0.0430	0.0334
Var_b7	2.7097	0.7565	0.4381	0.1579	0.0985	0.0722
Var_b8	2.7301	0.8426	0.6090	0.1550	0.1133	0.0801

Tabelle 59: Es werden MSE, Bias und Varianz für die einzelnen Koeffizienten (b1, ..., b8) unter unterschiedlich großen Fehlend-Anteilen und unter der jeweiligen Datensatzgröße gezeigt.

Fehlendmechanismus LogitX mit normalen Verfahren (glm)

	n=200			n=1000		
	50%	25%	10%	50%	25%	10%
MSE_b1	0.5237	0.1477	0.1121	0.0332	0.0167	0.0141
MSE_b2	1.0802	0.3315	0.2380	0.0536	0.0375	0.0274
MSE_b3	0.8635	0.3635	0.2663	0.1051	0.0554	0.0420
MSE_b4	1.8443	0.5326	0.4622	0.1465	0.0829	0.0662
MSE_b5	1.6465	0.4803	0.3479	0.1226	0.0709	0.0546
MSE_b6	1.3237	0.3563	0.2604	0.0745	0.0456	0.0347
MSE_b7	1.4237	0.5392	0.4292	0.1189	0.0819	0.0640
MSE_b8	1.6931	0.6248	0.5027	0.1423	0.0892	0.0714
Bias_b1	-0.2450	-0.1422	-0.1192	-0.0267	-0.0148	-0.0105
Bias_b2	0.4847	0.2521	0.1871	0.0554	0.0455	0.0324
Bias_b3	0.1969	0.1497	0.0862	0.0378	0.0148	0.0233
Bias_b4	-0.2268	-0.1476	-0.1261	-0.0543	-0.0406	-0.0195
Bias_b5	0.5337	0.2634	0.1899	0.0779	0.0790	0.0501
Bias_b6	-0.5315	-0.2539	-0.1899	-0.0651	-0.0642	-0.0439
Bias_b7	0.2463	0.1225	0.0845	0.0741	0.0758	0.0523
Bias_b8	-0.1938	-0.1355	-0.0781	-0.0262	-0.0026	-0.0134
Var_b1	0.4637	0.1275	0.0979	0.0325	0.0165	0.0140
Var_b2	0.8453	0.2680	0.2030	0.0505	0.0354	0.0263
Var_b3	0.8247	0.3411	0.2589	0.1037	0.0551	0.0414
Var_b4	1.7928	0.5108	0.4463	0.1435	0.0813	0.0658
Var_b5	1.3617	0.4109	0.3118	0.1165	0.0646	0.0521
Var_b6	1.0412	0.2919	0.2244	0.0702	0.0415	0.0328
Var_b7	1.3631	0.5242	0.4221	0.1134	0.0762	0.0613
Var_b8	1.6556	0.6065	0.4966	0.1417	0.0892	0.0712

Tabelle 60: Es werden MSE, Bias und Varianz für die einzelnen Koeffizienten (b1, ..., b8) unter unterschiedlich großen Fehlend-Anteilen und unter der jeweiligen Datensatzgröße gezeigt.

E.2 Bruteforce1

Fehlendmechanismus MCAR mit Bruteforce1-Verfahren (glm)

	n=200			n=1000		
	50%	25%	10%	50%	25%	10%
MSE_b1	0.5773	0.2893	0.1076	0.5508	0.2798	0.0999
MSE_b2	2.1470	1.0711	0.3661	2.1865	1.0904	0.3210
MSE_b3	0.6220	0.3476	0.2580	0.5495	0.2859	0.1157
MSE_b4	0.5782	0.3278	0.2102	0.5468	0.2806	0.0519
MSE_b5	2.2972	1.1435	0.4581	2.1593	1.0445	0.3815
MSE_b6	2.2746	1.1600	0.3916	2.1599	1.0557	0.3193
MSE_b7	0.7240	0.4940	0.4290	0.5500	0.2690	0.1596
MSE_b8	0.6126	0.4081	0.2617	0.5863	0.3150	0.0913
Bias_b1	0.7538	0.5212	0.2710	0.7407	0.5258	0.3053
Bias_b2	-1.4625	-1.0261	-0.5626	-1.4781	-1.0422	-0.5593
Bias_b3	-0.7577	-0.5292	-0.3599	-0.7368	-0.5218	-0.3049
Bias_b4	0.7192	0.4576	0.0255	0.7295	0.5086	0.1162
Bias_b5	-1.5053	-1.0334	-0.5762	-1.4676	-1.0160	-0.5970
Bias_b6	1.5038	1.0589	0.5614	1.4688	1.0246	0.5521
Bias_b7	-0.8191	-0.6049	-0.4674	-0.7340	-0.4981	-0.3559
Bias_b8	0.7401	0.5298	0.2389	0.7579	0.5418	0.2282
Var_b1	0.0091	0.0177	0.0341	0.0022	0.0033	0.0067
Var_b2	0.0082	0.0181	0.0496	0.0016	0.0042	0.0082
Var_b3	0.0479	0.0676	0.1285	0.0066	0.0136	0.0227
Var_b4	0.0609	0.1185	0.2096	0.0145	0.0219	0.0384
Var_b5	0.0313	0.0756	0.1260	0.0055	0.0123	0.0251
Var_b6	0.0133	0.0388	0.0764	0.0024	0.0059	0.0144
Var_b7	0.0530	0.1280	0.2105	0.0113	0.0209	0.0329
Var_b8	0.0648	0.1273	0.2046	0.0118	0.0214	0.0392

Tabelle 61: Es werden MSE, Bias und Varianz für die einzelnen Koeffizienten (b1, ..., b8) unter unterschiedlich großen Fehlend-Anteilen und unter der jeweiligen Datensatzgröße gezeigt.

Fehlendmechanismus MNAR1 mit Bruteforce1-Verfahren (glm)

	n=200			n=1000		
	50%	25%	10%	50%	25%	10%
MSE_b1	0.5436	0.2827	0.1102	0.5417	0.3003	0.1061
MSE_b2	2.0049	1.0743	0.3747	1.8137	0.9698	0.3267
MSE_b3	0.8008	0.4648	0.2394	0.5925	0.3086	0.1216
MSE_b4	0.0738	0.1543	0.2037	0.0406	0.0565	0.0580
MSE_b5	2.5424	1.2812	0.4683	2.3417	1.1740	0.3965
MSE_b6	2.1714	1.1406	0.3970	1.8663	0.9866	0.3316
MSE_b7	1.6412	0.8795	0.4230	1.0584	0.4845	0.1646
MSE_b8	0.0601	0.1819	0.2596	0.0892	0.1304	0.0844
Bias_b1	0.7335	0.5167	0.2763	0.7350	0.5446	0.3157
Bias_b2	-1.4139	-1.0272	-0.5732	-1.3462	-0.9830	-0.5633
Bias_b3	-0.8805	-0.6200	-0.3395	-0.7663	-0.5433	-0.3140
Bias_b4	0.1972	0.1219	0.0572	0.1814	0.1916	0.1384
Bias_b5	-1.5877	-1.1031	-0.6028	-1.5289	-1.0785	-0.6106
Bias_b6	1.4702	1.0540	0.5822	1.3654	0.9903	0.5651
Bias_b7	-1.2677	-0.8753	-0.5091	-1.0252	-0.6821	-0.3637
Bias_b8	0.1930	0.2793	0.2227	0.2860	0.3282	0.2186
Var_b1	0.0056	0.0158	0.0339	0.0015	0.0038	0.0064
Var_b2	0.0057	0.0192	0.0460	0.0013	0.0035	0.0094
Var_b3	0.0254	0.0804	0.1241	0.0054	0.0135	0.0230
Var_b4	0.0349	0.1395	0.2004	0.0077	0.0198	0.0389
Var_b5	0.0215	0.0644	0.1050	0.0042	0.0109	0.0236
Var_b6	0.0098	0.0297	0.0581	0.0021	0.0059	0.0122
Var_b7	0.0341	0.1134	0.1638	0.0074	0.0193	0.0323
Var_b8	0.0229	0.1039	0.2100	0.0074	0.0227	0.0366

Tabelle 62: Es werden MSE, Bias und Varianz für die einzelnen Koeffizienten (b1, ..., b8) unter unterschiedlich großen Fehlend-Anteilen und unter der jeweiligen Datensatzgröße gezeigt.

Fehlendmechanismus MNAR2 mit Bruteforce1-Verfahren (glm)

	n=200			n=1000		
	50%	25%	10%	50%	25%	10%
MSE_b1	0.5740	0.3073	0.1136	0.5629	0.2953	0.1017
MSE_b2	2.1248	1.0928	0.3772	1.9852	1.0292	0.3341
MSE_b3	0.7825	0.4579	0.2519	0.6068	0.3187	0.1152
MSE_b4	0.1526	0.1763	0.2468	0.1288	0.0928	0.0643
MSE_b5	2.5737	1.2443	0.4771	2.3749	1.1735	0.3785
MSE_b6	2.2737	1.1462	0.4036	2.0221	1.0269	0.3318
MSE_b7	1.4538	0.7578	0.4087	0.9175	0.4228	0.1445
MSE_b8	0.1332	0.2084	0.2562	0.1793	0.1692	0.0917
Bias_b1	0.7531	0.5369	0.2763	0.7492	0.5396	0.3085
Bias_b2	-1.4557	-1.0373	-0.5713	-1.4084	-1.0124	-0.5701
Bias_b3	-0.8618	-0.6011	-0.3537	-0.7752	-0.5534	-0.3047
Bias_b4	0.3241	0.2254	0.0767	0.3456	0.2682	0.1692
Bias_b5	-1.5960	-1.0885	-0.5956	-1.5393	-1.0779	-0.5989
Bias_b6	1.5037	1.0558	0.5786	1.4211	1.0104	0.5654
Bias_b7	-1.1856	-0.8068	-0.4714	-0.9525	-0.6343	-0.3358
Bias_b8	0.3065	0.3227	0.2693	0.4124	0.3821	0.2320
Var_b1	0.0068	0.0190	0.0373	0.0016	0.0042	0.0065
Var_b2	0.0058	0.0168	0.0508	0.0014	0.0043	0.0091
Var_b3	0.0397	0.0966	0.1267	0.0059	0.0125	0.0223
Var_b4	0.0476	0.1255	0.2409	0.0094	0.0209	0.0357
Var_b5	0.0265	0.0594	0.1223	0.0055	0.0117	0.0199
Var_b6	0.0125	0.0314	0.0688	0.0027	0.0061	0.0120
Var_b7	0.0480	0.1069	0.1865	0.0103	0.0205	0.0317
Var_b8	0.0393	0.1042	0.1837	0.0092	0.0232	0.0379

Tabelle 63: Es werden MSE, Bias und Varianz für die einzelnen Koeffizienten (b1, ..., b8) unter unterschiedlich großen Fehlend-Anteilen und unter der jeweiligen Datensatzgröße gezeigt.

Fehlendmechanismus MNAR3 mit Bruteforce1-Verfahren (glm)

	n=200			n=1000		
	50%	25%	10%	50%	25%	10%
MSE_b1	0.5735	0.3149	0.1048	0.5703	0.2921	0.0951
MSE_b2	2.2339	1.1172	0.3596	2.1552	1.0927	0.3393
MSE_b3	0.7065	0.4212	0.2462	0.5675	0.2992	0.1055
MSE_b4	0.3304	0.2388	0.2466	0.3460	0.1854	0.0856
MSE_b5	2.4992	1.2389	0.4410	2.2923	1.1151	0.3573
MSE_b6	2.3586	1.2025	0.3810	2.1451	1.0583	0.3280
MSE_b7	1.0314	0.6233	0.3358	0.6977	0.3352	0.1240
MSE_b8	0.3676	0.3154	0.2628	0.3875	0.2618	0.1138
Bias_b1	0.7514	0.5427	0.2691	0.7539	0.5371	0.2983
Bias_b2	-1.4920	-1.0471	-0.5593	-1.4674	-1.0432	-0.5748
Bias_b3	-0.8170	-0.5884	-0.3296	-0.7488	-0.5360	-0.2913
Bias_b4	0.5216	0.3368	0.1577	0.5794	0.4033	0.2283
Bias_b5	-1.5707	-1.0852	-0.5758	-1.5119	-1.0501	-0.5796
Bias_b6	1.5310	1.0827	0.5678	1.4635	1.0255	0.5631
Bias_b7	-0.9873	-0.7200	-0.4099	-0.8277	-0.5632	-0.3040
Bias_b8	0.5544	0.4517	0.2798	0.6136	0.4874	0.2812
Var_b1	0.0088	0.0204	0.0324	0.0019	0.0037	0.0061
Var_b2	0.0077	0.0209	0.0468	0.0020	0.0044	0.0089
Var_b3	0.0390	0.0749	0.1375	0.0067	0.0119	0.0206
Var_b4	0.0583	0.1254	0.2217	0.0103	0.0228	0.0335
Var_b5	0.0322	0.0612	0.1095	0.0065	0.0124	0.0213
Var_b6	0.0145	0.0303	0.0586	0.0032	0.0066	0.0109
Var_b7	0.0565	0.1049	0.1677	0.0126	0.0180	0.0316
Var_b8	0.0603	0.1114	0.1845	0.0109	0.0242	0.0347

Tabelle 64: Es werden MSE, Bias und Varianz für die einzelnen Koeffizienten (b1, ..., b8) unter unterschiedlich großen Fehlend-Anteilen und unter der jeweiligen Datensatzgröße gezeigt.

Fehlendmechanismus LogitY mit Bruteforce1-Verfahren (glm)

	n=200			n=1000		
	50%	25%	10%	50%	25%	10%
MSE_b1	0.5948	0.3208	0.1302	0.5584	0.3156	0.1019
MSE_b2	2.2965	1.1941	0.4400	2.1044	1.1708	0.3478
MSE_b3	0.7105	0.4223	0.2450	0.5636	0.3130	0.1165
MSE_b4	0.4376	0.2752	0.2472	0.3795	0.1930	0.0800
MSE_b5	2.5726	1.2888	0.5307	2.1895	1.2118	0.3833
MSE_b6	2.4561	1.2403	0.4794	2.1015	1.1486	0.3523
MSE_b7	0.9203	0.6202	0.3747	0.6369	0.3637	0.1294
MSE_b8	0.4778	0.3494	0.2788	0.4153	0.2477	0.1141
Bias_b1	0.7661	0.5473	0.3166	0.7461	0.5581	0.3082
Bias_b2	-1.5122	-1.0821	-0.6227	-1.4498	-1.0798	-0.5809
Bias_b3	-0.8213	-0.5781	-0.3496	-0.7463	-0.5471	-0.3034
Bias_b4	0.6160	0.3852	0.1656	0.6067	0.4165	0.2106
Bias_b5	-1.5933	-1.1087	-0.6430	-1.4772	-1.0948	-0.5973
Bias_b6	1.5624	1.0990	0.6366	1.4483	1.0684	0.5807
Bias_b7	-0.9280	-0.7188	-0.4317	-0.7907	-0.5890	-0.3101
Bias_b8	0.6480	0.4759	0.3141	0.6365	0.4784	0.2674
Var_b1	0.0080	0.0213	0.0299	0.0018	0.0042	0.0069
Var_b2	0.0097	0.0232	0.0523	0.0025	0.0049	0.0104
Var_b3	0.0360	0.0882	0.1228	0.0066	0.0136	0.0245
Var_b4	0.0582	0.1268	0.2198	0.0114	0.0195	0.0356
Var_b5	0.0339	0.0597	0.1172	0.0074	0.0133	0.0265
Var_b6	0.0149	0.0326	0.0742	0.0040	0.0072	0.0151
Var_b7	0.0592	0.1035	0.1883	0.0117	0.0168	0.0332
Var_b8	0.0579	0.1229	0.1801	0.0103	0.0188	0.0426

Tabelle 65: Es werden MSE, Bias und Varianz für die einzelnen Koeffizienten (b1, ..., b8) unter unterschiedlich großen Fehlend-Anteilen und unter der jeweiligen Datensatzgröße gezeigt.

Fehlendmechanismus LogitX mit Bruteforce1-Verfahren (glm)

	n=200			n=1000		
	50%	25%	10%	50%	25%	10%
MSE_b1	0.6052	0.4847	0.2908	0.6304	0.5110	0.2948
MSE_b2	2.3358	1.8181	0.9929	2.1913	1.7031	0.9318
MSE_b3	0.7447	0.6392	0.4470	0.6686	0.5253	0.2583
MSE_b4	0.2010	0.1478	0.1581	0.1727	0.1444	0.1180
MSE_b5	2.3477	1.7723	1.0191	2.4760	1.8512	0.9465
MSE_b6	2.2863	1.8512	1.1569	2.1838	1.6642	0.8701
MSE_b7	1.2160	1.1212	0.6981	0.8276	0.6086	0.3113
MSE_b8	0.4240	0.4960	0.4103	0.4381	0.4989	0.3906
Bias_b1	0.7719	0.6867	0.5189	0.7928	0.7129	0.5385
Bias_b2	-1.5258	-1.3434	-0.9851	-1.4798	-1.3041	-0.9629
Bias_b3	-0.8442	-0.7628	-0.6032	-0.8138	-0.7170	-0.4921
Bias_b4	0.3909	0.2528	0.1249	0.4063	0.3610	0.3035
Bias_b5	-1.5229	-1.3131	-0.9698	-1.5720	-1.3574	-0.9654
Bias_b6	1.5078	1.3528	1.0585	1.4770	1.2883	0.9287
Bias_b7	-1.0846	-1.0190	-0.7689	-0.9053	-0.7715	-0.5381
Bias_b8	0.6171	0.6556	0.5472	0.6565	0.6968	0.6049
Var_b1	0.0093	0.0132	0.0216	0.0020	0.0028	0.0048
Var_b2	0.0076	0.0134	0.0225	0.0014	0.0024	0.0047
Var_b3	0.0321	0.0573	0.0831	0.0062	0.0112	0.0162
Var_b4	0.0481	0.0838	0.1425	0.0076	0.0140	0.0259
Var_b5	0.0286	0.0480	0.0786	0.0048	0.0087	0.0146
Var_b6	0.0129	0.0212	0.0366	0.0023	0.0043	0.0076
Var_b7	0.0397	0.0829	0.1069	0.0080	0.0134	0.0217
Var_b8	0.0432	0.0662	0.1110	0.0071	0.0133	0.0247

Tabelle 66: Es werden MSE, Bias und Varianz für die einzelnen Koeffizienten (b1, ..., b8) unter unterschiedlich großen Fehlend-Anteilen und unter der jeweiligen Datensatzgröße gezeigt.

E.3 Bruteforce2

Fehlendmechanismus MCAR mit Bruteforce2-Verfahren (glm)						
	n=200			n=1000		
	50%	25%	10%	50%	25%	10%
MSE_b1	0.5171	0.2647	0.0587	0.4597	0.2329	0.0405
MSE_b2	1.8794	0.9239	0.1517	1.9383	0.9679	0.1161
MSE_b3	0.4902	0.2889	0.1891	0.4362	0.2261	0.0578
MSE_b4	0.8659	0.4928	0.2574	0.8466	0.4160	0.0485
MSE_b5	1.8831	0.9814	0.2314	1.7018	0.8337	0.1437
MSE_b6	2.0297	1.0400	0.1673	1.8933	0.9236	0.1165
MSE_b7	0.4483	0.3844	0.3047	0.3145	0.1629	0.0741
MSE_b8	1.0025	0.5493	0.2758	0.9318	0.4244	0.0657
Bias_b1	0.7106	0.4941	0.1257	0.6759	0.4781	0.1801
Bias_b2	-1.3662	-0.9493	-0.2772	-1.3916	-0.9811	-0.3237
Bias_b3	-0.6540	-0.4580	-0.1890	-0.6540	-0.4580	-0.1752
Bias_b4	0.8863	0.5870	-0.0302	0.9105	0.6263	0.0665
Bias_b5	-1.3577	-0.9427	-0.2824	-1.3018	-0.9052	-0.3369
Bias_b6	1.4186	0.9972	0.2733	1.3747	0.9573	0.3147
Bias_b7	-0.6119	-0.4852	-0.2505	-0.5467	-0.3740	-0.1886
Bias_b8	0.9599	0.6343	0.1416	0.9579	0.6325	0.1409
Var_b1	0.0122	0.0206	0.0430	0.0028	0.0043	0.0081
Var_b2	0.0130	0.0227	0.0749	0.0019	0.0053	0.0114
Var_b3	0.0625	0.0791	0.1534	0.0085	0.0164	0.0272
Var_b4	0.0804	0.1482	0.2565	0.0176	0.0237	0.0441
Var_b5	0.0397	0.0928	0.1516	0.0070	0.0144	0.0302
Var_b6	0.0172	0.0456	0.0926	0.0034	0.0072	0.0175
Var_b7	0.0739	0.1490	0.2420	0.0157	0.0231	0.0385
Var_b8	0.0811	0.1469	0.2558	0.0143	0.0243	0.0458

Tabelle 67: Es werden MSE, Bias und Varianz für die einzelnen Koeffizienten (b1, ..., b8) unter unterschiedlich großen Fehlend-Anteilen und unter der jeweiligen Datensatzgröße gezeigt.

Fehlendmechanismus MNAR1 mit Bruteforce2-Verfahren (glm)

	n=200			n=1000		
	50%	25%	10%	50%	25%	10%
MSE_b1	0.3047	0.1315	0.0595	0.2904	0.1381	0.0422
MSE_b2	1.1469	0.4740	0.1493	0.9678	0.4376	0.1185
MSE_b3	0.4427	0.2577	0.1849	0.3098	0.1472	0.0590
MSE_b4	0.1254	0.1789	0.2482	0.0525	0.0470	0.0504
MSE_b5	1.3458	0.5725	0.2312	1.1679	0.5131	0.1475
MSE_b6	1.2284	0.4971	0.1646	0.9777	0.4371	0.1193
MSE_b7	0.8147	0.4561	0.2924	0.4682	0.2029	0.0749
MSE_b8	0.1550	0.2004	0.2781	0.1149	0.0933	0.0620
Bias_b1	0.5408	0.3299	0.1271	0.5364	0.3649	0.1860
Bias_b2	-1.0645	-0.6670	-0.2823	-0.9824	-0.6575	-0.3261
Bias_b3	-0.6251	-0.4001	-0.1761	-0.5483	-0.3601	-0.1798
Bias_b4	0.2031	0.0635	-0.0060	0.1947	0.1424	0.0789
Bias_b5	-1.1399	-0.6985	-0.2969	-1.0770	-0.7055	-0.3445
Bias_b6	1.0979	0.6755	0.2831	0.9867	0.6548	0.3216
Bias_b7	-0.8669	-0.5635	-0.2737	-0.6748	-0.4236	-0.1929
Bias_b8	0.3174	0.2468	0.1277	0.3193	0.2574	0.1357
Var_b1	0.0122	0.0227	0.0433	0.0026	0.0050	0.0077
Var_b2	0.0137	0.0291	0.0696	0.0026	0.0054	0.0122
Var_b3	0.0520	0.0976	0.1539	0.0092	0.0176	0.0267
Var_b4	0.0841	0.1748	0.2482	0.0146	0.0267	0.0442
Var_b5	0.0465	0.0845	0.1431	0.0081	0.0153	0.0288
Var_b6	0.0231	0.0408	0.0844	0.0042	0.0084	0.0159
Var_b7	0.0632	0.1385	0.2175	0.0129	0.0234	0.0377
Var_b8	0.0543	0.1395	0.2618	0.0129	0.0271	0.0436

Tabelle 68: Es werden MSE, Bias und Varianz für die einzelnen Koeffizienten (b1, ..., b8) unter unterschiedlich großen Fehlend-Anteilen und unter der jeweiligen Datensatzgröße gezeigt.

Fehlendmechanismus MNAR2 mit Bruteforce2-Verfahren (glm)

	n=200			n=1000		
	50%	25%	10%	50%	25%	10%
MSE_b1	0.3720	0.1718	0.0679	0.3545	0.1641	0.0493
MSE_b2	1.3635	0.5970	0.1841	1.2980	0.5867	0.1575
MSE_b3	0.4786	0.2909	0.2017	0.3679	0.1753	0.0644
MSE_b4	0.2153	0.2035	0.2810	0.1927	0.1003	0.0600
MSE_b5	1.5270	0.6737	0.2665	1.4019	0.6212	0.1743
MSE_b6	1.4536	0.6205	0.2041	1.2923	0.5738	0.1563
MSE_b7	0.7633	0.4288	0.3003	0.4382	0.1965	0.0762
MSE_b8	0.2704	0.2411	0.2742	0.2480	0.1531	0.0735
Bias_b1	0.6001	0.3818	0.1566	0.5933	0.3990	0.2039
Bias_b2	-1.1630	-0.7565	-0.3394	-1.1382	-0.7621	-0.3815
Bias_b3	-0.6452	-0.4234	-0.2171	-0.5991	-0.4000	-0.1953
Bias_b4	0.3660	0.2148	0.0490	0.4230	0.2751	0.1436
Bias_b5	-1.2178	-0.7710	-0.3458	-1.1804	-0.7778	-0.3856
Bias_b6	1.1970	0.7613	0.3456	1.1348	0.7518	0.3754
Bias_b7	-0.8286	-0.5406	-0.2750	-0.6501	-0.4144	-0.1970
Bias_b8	0.4514	0.3347	0.1873	0.4841	0.3546	0.1763
Var_b1	0.0120	0.0260	0.0434	0.0025	0.0049	0.0077
Var_b2	0.0109	0.0248	0.0688	0.0025	0.0058	0.0120
Var_b3	0.0624	0.1116	0.1546	0.0090	0.0154	0.0262
Var_b4	0.0813	0.1574	0.2786	0.0137	0.0246	0.0394
Var_b5	0.0439	0.0793	0.1469	0.0084	0.0162	0.0257
Var_b6	0.0207	0.0409	0.0846	0.0045	0.0085	0.0153
Var_b7	0.0768	0.1365	0.2247	0.0155	0.0248	0.0374
Var_b8	0.0666	0.1291	0.2391	0.0137	0.0274	0.0425

Tabelle 69: Es werden MSE, Bias und Varianz für die einzelnen Koeffizienten (b1, ..., b8) unter unterschiedlich großen Fehlend-Anteilen und unter der jeweiligen Datensatzgröße gezeigt.

Fehlendmechanismus MNAR3 mit Bruteforce2-Verfahren (glm)

	n=200			n=1000		
	50%	25%	10%	50%	25%	10%
MSE_b1	0.4398	0.2293	0.0771	0.4362	0.2069	0.0638
MSE_b2	1.7355	0.7970	0.2404	1.7284	0.8292	0.2330
MSE_b3	0.5145	0.3066	0.2054	0.4077	0.2053	0.0714
MSE_b4	0.4977	0.3279	0.2730	0.5355	0.2602	0.1009
MSE_b5	1.7766	0.8588	0.3135	1.6254	0.7559	0.2247
MSE_b6	1.8160	0.8561	0.2605	1.6797	0.7864	0.2217
MSE_b7	0.5842	0.3857	0.2640	0.3733	0.1806	0.0780
MSE_b8	0.6105	0.4105	0.2836	0.5973	0.2996	0.1131
Bias_b1	0.6531	0.4528	0.2014	0.6583	0.4495	0.2385
Bias_b2	-1.3132	-0.8765	-0.4274	-1.3136	-0.9075	-0.4713
Bias_b3	-0.6806	-0.4708	-0.2462	-0.6300	-0.4369	-0.2201
Bias_b4	0.6398	0.3997	0.1747	0.7219	0.4839	0.2518
Bias_b5	-1.3157	-0.8800	-0.4336	-1.2714	-0.8602	-0.4473
Bias_b6	1.3398	0.9023	0.4377	1.2944	0.8818	0.4568
Bias_b7	-0.7067	-0.4978	-0.2774	-0.5974	-0.3968	-0.2041
Bias_b8	0.7306	0.5144	0.2745	0.7641	0.5219	0.2754
Var_b1	0.0133	0.0243	0.0366	0.0029	0.0048	0.0069
Var_b2	0.0110	0.0288	0.0577	0.0027	0.0057	0.0109
Var_b3	0.0513	0.0850	0.1448	0.0107	0.0145	0.0229
Var_b4	0.0883	0.1682	0.2425	0.0143	0.0260	0.0375
Var_b5	0.0456	0.0844	0.1255	0.0090	0.0160	0.0246
Var_b6	0.0210	0.0419	0.0689	0.0041	0.0089	0.0131
Var_b7	0.0848	0.1379	0.1870	0.0164	0.0231	0.0363
Var_b8	0.0767	0.1459	0.2082	0.0134	0.0272	0.0373

Tabelle 70: Es werden MSE, Bias und Varianz für die einzelnen Koeffizienten (b1, ..., b8) unter unterschiedlich großen Fehlend-Anteilen und unter der jeweiligen Datensatzgröße gezeigt.

Fehlendmechanismus LogitY mit Bruteforce2-Verfahren (glm)

	n=200			n=1000		
	50%	25%	10%	50%	25%	10%
MSE_b1	0.4935	0.2486	0.0958	0.4349	0.2258	0.0647
MSE_b2	1.8904	0.9039	0.2927	1.7387	0.8838	0.2225
MSE_b3	0.5196	0.3177	0.2127	0.4062	0.2100	0.0797
MSE_b4	0.6589	0.3858	0.2957	0.5908	0.2748	0.0927
MSE_b5	1.9813	0.9724	0.3837	1.6006	0.8247	0.2307
MSE_b6	2.0505	0.9621	0.3378	1.7009	0.8531	0.2242
MSE_b7	0.5600	0.4214	0.3003	0.3422	0.1969	0.0748
MSE_b8	0.8027	0.4296	0.3074	0.6240	0.2934	0.1078
Bias_b1	0.6937	0.4741	0.2413	0.6575	0.4699	0.2385
Bias_b2	-1.3704	-0.9373	-0.4774	-1.3173	-0.9367	-0.4597
Bias_b3	-0.6836	-0.4689	-0.2698	-0.6302	-0.4406	-0.2275
Bias_b4	0.7637	0.4870	0.1944	0.7579	0.5010	0.2339
Bias_b5	-1.3924	-0.9452	-0.4952	-1.2612	-0.8990	-0.4497
Bias_b6	1.4251	0.9598	0.5016	1.3022	0.9186	0.4559
Bias_b7	-0.6967	-0.5351	-0.2957	-0.5715	-0.4194	-0.1989
Bias_b8	0.8535	0.5389	0.3062	0.7812	0.5207	0.2508
Var_b1	0.0123	0.0238	0.0376	0.0026	0.0050	0.0078
Var_b2	0.0124	0.0254	0.0648	0.0034	0.0064	0.0112
Var_b3	0.0523	0.0978	0.1399	0.0091	0.0159	0.0279
Var_b4	0.0757	0.1487	0.2579	0.0164	0.0239	0.0380
Var_b5	0.0427	0.0789	0.1384	0.0100	0.0165	0.0284
Var_b6	0.0196	0.0408	0.0862	0.0053	0.0093	0.0163
Var_b7	0.0746	0.1350	0.2129	0.0156	0.0210	0.0353
Var_b8	0.0743	0.1392	0.2136	0.0137	0.0222	0.0448

Tabelle 71: Es werden MSE, Bias und Varianz für die einzelnen Koeffizienten (b1, ..., b8) unter unterschiedlich großen Fehlend-Anteilen und unter der jeweiligen Datensatzgröße gezeigt.

Fehlendmechanismus LogitX mit Bruteforce2-Verfahren (glm)

	n=200			n=1000		
	50%	25%	10%	50%	25%	10%
MSE_b1	0.3280	0.2345	0.1270	0.3251	0.2376	0.1220
MSE_b2	1.2182	0.8314	0.3945	1.1314	0.7812	0.3731
MSE_b3	0.3889	0.3468	0.2545	0.3339	0.2447	0.1161
MSE_b4	0.1873	0.1452	0.2100	0.1510	0.0875	0.0708
MSE_b5	1.0666	0.8181	0.4457	1.1722	0.8218	0.3788
MSE_b6	1.1659	0.8599	0.4617	1.1008	0.7535	0.3463
MSE_b7	0.5564	0.5953	0.3970	0.3203	0.2546	0.1332
MSE_b8	0.5686	0.3864	0.3123	0.4609	0.3097	0.1896
Bias_b1	0.5608	0.4641	0.3080	0.5679	0.4837	0.3402
Bias_b2	-1.0984	-0.9001	-0.5956	-1.0625	-0.8817	-0.6048
Bias_b3	-0.5832	-0.5209	-0.3779	-0.5691	-0.4794	-0.3072
Bias_b4	0.3454	0.1590	0.0432	0.3728	0.2599	0.1935
Bias_b5	-1.0141	-0.8656	-0.5792	-1.0795	-0.8997	-0.5991
Bias_b6	1.0706	0.9088	0.6374	1.0475	0.8644	0.5792
Bias_b7	-0.7050	-0.6952	-0.4858	-0.5563	-0.4862	-0.3236
Bias_b8	0.7074	0.5339	0.3768	0.6704	0.5387	0.3950
Var_b1	0.0135	0.0191	0.0321	0.0026	0.0037	0.0063
Var_b2	0.0118	0.0212	0.0397	0.0024	0.0038	0.0073
Var_b3	0.0487	0.0755	0.1116	0.0100	0.0149	0.0218
Var_b4	0.0680	0.1200	0.2082	0.0120	0.0200	0.0334
Var_b5	0.0381	0.0688	0.1101	0.0069	0.0123	0.0199
Var_b6	0.0197	0.0340	0.0554	0.0035	0.0063	0.0108
Var_b7	0.0593	0.1120	0.1609	0.0109	0.0182	0.0285
Var_b8	0.0682	0.1013	0.1704	0.0115	0.0195	0.0335

Tabelle 72: Es werden MSE, Bias und Varianz für die einzelnen Koeffizienten (b1, ..., b8) unter unterschiedlich großen Fehlend-Anteilen und unter der jeweiligen Datensatzgröße gezeigt.