

8 Auswertungsverfahren

Hinweis: Die Übungsaufgaben beziehen sich sowohl auf die im Buch behandelte Theorie als auch auf die in den Online-Materialien dargelegten Analysen. Für den Großteil der folgenden Aufgaben benötigen Sie Erfahrung im Umgang mit der Statistiksoftware R. Sollten Sie dies noch nicht haben, können Sie sich mit Hilfe von Lehrbüchern (z.B. Obszela & Baierl, 2020) oder Materialien im Netz mit dem Programm vertraut machen.

- (1) Ein Klinikleiter möchte die Wirksamkeit der in der Klinik angebotenen Selbsthilfegruppen evaluieren. Dazu möchte er die Symptomschwere von Patient:innen die während ihres Klinikaufenthaltes an den angebotenen Selbsthilfegruppen teilgenommen haben, im Vergleich zu jenen, die nicht teilgenommen haben, zum Zeitpunkt der Entlassung erfassen.
- (a) Nennen und erläutern Sie die Bedingungen für die Unverfälschtheit des geschätzten Effekts anhand des vorliegenden Beispiels.

Musterantwort

- (1) Positivität: Für Personen mit jeglicher Kombination an Kovariatenausprägungen muss es theoretisch möglich sein, sowohl in die Interventions- als auch die Kontrollgruppe zu gelangen. In diesem Beispiel müssen also unabhängig von ihren spezifischen Kovariatenausprägungen alle Patient:innen eine Chance haben, sowohl an den zusätzlichen Selbsthilfegruppen teilzunehmen als auch, nicht an den Gruppen teilzunehmen ($0 < P(X = 1|Z = z) < 1$). Diese Wahrscheinlichkeit muss nicht groß sein, sie muss nur größer als null sein.
- (2) Bedingte Unabhängigkeit: Die möglichen Outcomes (Y^0, Y^1), also hier die hypothetische Symptomschwere der Patient:innen bei Teilnahme und bei Nicht-Teilnahme an den Selbsthilfegruppen, muss unabhängig von der Zuweisung zu den Selbsthilfegruppen X sein. Dabei genügt es, wenn sie gegeben einem Kovariatenvektor Z unabhängig sind, also gegeben einem Set von für die Studie relevanten Kovariaten (z.B. Art der Erkrankung und Motivation): $(Y^0, Y^1) \perp X|Z$.
- (3) Stable-Unit-Treatment-Value-Annahme (SUTVA): 1.: Die möglichen Outcomes einer Person Y^1 und Y^0 , also hier die hypothetische Symptomschwere der Patient:innen bei Teilnahme und bei Nicht-Teilnahme an den Selbsthilfegruppen, müssen unabhängig von der Teilnahme oder Nicht-Teilnahme anderer Patient:innen an den Selbsthilfegruppen sein. Das heißt, es sollte für die möglichen Outcomes keinen Unterschied machen, ob andere Personen ebenfalls der Selbsthilfe- oder der Kontrollgruppe zugeordnet sind. (keine Interferenz). 2.: Das Konzept und die Umsetzung der Selbsthilfegruppen darf sich nicht zwischen den Patient:innen unterscheiden. Ebenso darf es keine Unterschiede in der Umsetzung der Kontrollbedingung zwischen den Patient:innen geben, also in der Behandlung der Personen, die nicht an den zusätzlichen Selbsthilfegruppen teilnehmen (keine versteckte Variation der Intervention).
- (b) Wodurch könnten diese hier gefährdet sein? Und was kann der Klinikleiter tun, um diese Bedingungen so weit wie möglich zu erfüllen?

Musterantwort

- (1) Positivität: Diese Bedingung könnte beispielsweise verletzt sein, wenn Patient:innen mit bestimmten Störungsbildern von der Teilnahme an der Selbsthilfegruppe ausgeschlossen würden. Auch könnte sie zum Beispiel dann verletzt sein, wenn Patient:innen mit besonders hoher Symptomschwere automatisch an den Selbsthilfegruppen teilnehmen müssten.
- Zunächst könnte der Klinikleiter dafür sorgen, dass keine Patient:innen systematisch von den Selbsthilfegruppen ausgeschlossen oder diesen immer zugewiesen werden. Zudem könnte er in der statistischen Analyse prüfen, ob es Personen mit bestimmten Kovariatenausprägungen (oder mit bestimmten Propensity-Scores) gibt, die ausschließlich an den Selbsthilfegruppen teilnehmen oder

ausschließlich nicht teilnehmen. Diese Personen könnte er dann von der Analyse ausschließen (Overlap, Abschnitt 8.2.3).

- (2) Bedingte Unabhängigkeit: Verletzt sein könnte die Bedingung zum Beispiel dann, wenn Personen, bei denen die Selbsthilfegruppe besonders gut wirken würde (d.h. große individuelle kausale Effekte $\delta_i = Y^1 - Y^0$) eher die Selbsthilfegruppe wählen als Personen mit gleicher Kovariatenausprägung $Z = z$ (z.B. gleicher hoher Motivation und Erkrankungsart), bei denen die Selbsthilfegruppe aber nicht so gut wirken würde (d.h. kleine individuelle kausale Effekte $\delta_i = Y^1 - Y^0$). In diesem Fall wäre die Wahrscheinlichkeit der Teilnahme an den Selbsthilfegruppen (gegeben der Motivation) abhängig von der Symptomschwere nach Teilnahme (Y^0) und ohne Teilnahme an einer Selbsthilfegruppe (Y^0).

Um die Annahme der bedingten Unabhängigkeit zu erfüllen, müsste der Klinikleiter die Patient:innen randomisiert der Teilnahme oder Nicht-Teilnahme an den Selbsthilfegruppen zuweisen. Sollte eine Randomisierung nicht möglich sein, könnte der Klinikleiter darüber nachdenken und recherchieren, welche Kovariaten sowohl die Wahl der Teilnahme an einer Selbsthilfegruppe als auch die Symptomschwere beeinflussen könnten. Idealerweise sollte er alle relevanten Kovariaten vor Beginn der Studie erheben und sie im Anschluss an die Studie in die statistische Analyse einbeziehen.

- (3) Stable-Unit-Treatment-Value-Annahme (SUTVA): Die SUTVA könnte zum Beispiel dann verletzt sein, wenn einige Patient:innen, die an der Selbsthilfegruppe teilnehmen, ihre Erkenntnisse und Methoden aus der Selbsthilfegruppe an die Patient:innen weitergeben, die nicht an einer Selbsthilfegruppe teilnehmen (Interferenz). Damit würden Patient:innen in der Kontrollbedingung indirekt ebenfalls von den Selbsthilfegruppen profitieren. Ein anderes Beispiel für Interferenz könnte auftreten, wenn Patient:innen besonders von der Selbsthilfegruppe profitieren, wenn andere Personen der Selbsthilfegruppe besonders reflektiert sind.

Auch könnte sie beispielsweise dann verletzt sein, wenn die Selbsthilfegruppen sich dadurch unterscheiden würden, welchen Fokus, welches Programm oder welchen zeitlichen Rahmen die Gruppen haben und dies in der statistischen Analyse nicht berücksichtigt würde (versteckte Variation der Intervention).

Spillover-Effekte und andere Arten der Beeinflussung der individuellen kausalen Effekte durch (Nicht-)Aufnahme anderer Patient:innen in die Selbsthilfegruppen (Interferenz) können in der geplanten Studie wahrscheinlich kaum verhindert werden. Da einige Patient:innen, die an Selbsthilfegruppen teilnehmen, in der Klinik sicherlich Kontakt zu anderen Patient:innen haben, die nicht teilnehmen, wird es wahrscheinlich zu einem Austausch zwischen den beiden Bedingungen kommen. Auch wird es sich bei Gruppeninterventionen nicht verhindern lassen, dass die Gruppenzusammensetzung Einfluss auf das Ergebnis hat. Diese Einschränkung sollte der Klinikleiter bei der Interpretation der Studienergebnisse berücksichtigen.

Dass sich die Durchführung der Selbsthilfegruppen untereinander nicht unterscheidet (keine versteckte Variation der Intervention), könnte der Klinikleiter bereits vor Beginn der Studie planen, indem er zum Beispiel sicherstellt, dass die Selbsthilfegruppen sich möglichst wenig in Inhalt, Dauer, Tageszeit, Programm usw. unterscheiden. Zusätzlich könnte er die Ähnlichkeit der Selbsthilfegruppen durch Implementationskontrollen prüfen.

- (2) Sie möchten den Effekt einer Online-Therapie im Vergleich zu einer Therapie in Präsenz (*therapie*) auf depressive Symptome (*depression*) untersuchen. Aus ethischen und praktischen Gründen können Sie keine Randomisierung vornehmen, sondern die Patient:innen entscheiden selbst, an welcher Therapieform sie teilnehmen wollen. Als Kovariate haben Sie das Alter der Personen (*alter*) erfasst. Die Daten finden Sie unter dem Namen *therapie.rda* in den Online-Materialien.

- (a) Wenden Sie mithilfe von EffectLiteR eine generalisierte ANCOVA auf die Daten an und schätzen Sie sowohl den durchschnittlichen kausalen Effekt (ATE), den durchschnittlichen kausalen Effekt der

Behandelten (ATT) als auch den durchschnittlichen kausalen Effekt der Unbehandelten (ATU) und interpretieren Sie diese.

Musterantwort

Die depressive Symptomatik nach der Therapie (depression) soll auf die Therapieform (therapie) und das Alter der Teilnehmenden (alter) zurückgeführt werden. Dabei dient die Gruppe, die Therapie in Präsenz erhält, als Kontrollgruppe (therapie=0) und die Gruppe, die Online-Therapie erhält, als Interventionsgruppe (therapie=1). Da es sich um eine generalisierte ANCOVA handelt, lassen wir Interaktionen zwischen der Interventionsvariable (therapie) und dem Alter (alter) zu:

```
# Laden des Paketes
library(EffectLiteR)

# Einlesen der Daten
load("therapie.rda")

# Effektschätzung
genANC <- effectLite(y = "depression",
                     x = "therapie",
                     control = "0",
                     z = "alter",
                     data = therapie_dat,
                     interactions = "all")

print(genANC)

...
## ----- Average Effects -----
##
##           Estimate      SE  Est./SE  p-value  Effect Size
## E[g1(Z)]      -1.13   0.543   -2.08   0.0378   -0.253
##
## ----- Effects given a Treatment Condition -----
##
##           Estimate      SE  Est./SE  p-value  Effect Size
## E[g1(Z)|X=0]   -0.506   0.590   -0.857   0.3914   -0.114
## E[g1(Z)|X=1]  -1.463   0.571   -2.562   0.0104   -0.329
```

Aus dem Output können wir ablesen:

- ATE (durchschnittlicher kausaler Effekt):

$$E[g_1(Z)] = \widehat{ATE} = -1,13 \text{ (SE} = 0,543; p = ,038; \Delta = -0,253)$$

Im Durchschnitt führt die Teilnahme an der Online-Therapie im Vergleich zur Teilnahme an der Therapie in Präsenz zu einer um 1,13 Punkte geringeren depressiven Symptomatik. Der Effekt ist signifikant und entspricht einer standardisierten Effektstärke (Glass' Δ) von -0,253 (klein).

- ATU (durchschnittlicher kausaler Effekt der Unbehandelten):

$$E[g_1(Z)|X = 0] = \widehat{ATU} = -0,506 \text{ (SE} = 0,509; p = ,391; \Delta = -0,114)$$

Für Personen welche die Therapie in Präsenz erhalten, würde die Online-Therapie im Vergleich zur Präsenztherapie im Mittel zu einer 0,506 Punkten geringeren Depressivität führen. Der Effekt ist nicht signifikant und entspricht einer standardisierten Effektstärke (Glass' Δ) von -0,114 (klein).

- ATT (durchschnittlicher kausaler Effekt der Behandelten):

$$E[g_1(Z)|X = 1] = \widehat{ATT} = -1,463 \text{ (SE} = 0,571; p = ,010; \Delta = -0,329)$$

Für Personen, welche die Online-Therapie erhalten, führt die Online-Therapie im Mittel zu 1,463 Punkten geringerer Depressivität im Vgl. dazu, wenn diese Personen Präsenztherapie erhalten. Der Effekt ist signifikant und entspricht einer standardisierten Effektstärke (Glass' Δ) von -0,329 (klein).

- (b) Zeichnen Sie die Regressionsgeraden per Hand in ein Diagramm und fügen Sie die in (a) geschätzten kausalen Effekte ein.

Musterantwort

```
## ----- Regression Model -----
##
## E(Y|X,Z) = g0(Z) + g1(Z)*I_X=1
## g0(Z) = g000 + g001 * Z1
## g1(Z) = g100 + g101 * Z1
##
## Intercept Function g0(Z) [Reference group: 0]
##
## Coefficient Estimate SE Est./SE p-value
## g000 24.435 2.092 11.680 0
## g001 -0.252 0.051 -4.982 0
##
## Effect Function g1(Z) [therapie: 1 vs. 0]
##
## Coefficient Estimate SE Est./SE p-value
## g100 -15.179 2.438 -6.227 0
## g101 0.362 0.060 6.037 0
##
```

Aus dem Output lässt sich folgende Regressionsgleichung ablesen:

$$E(Y|X,Z) = g_0(Z) + g_1(Z) * X$$

$$E(Y|X,Z) = 24,435 - 0,252 * Z + (-15,179 + 0,362 * Z) * X$$

$$E(Y|X,Z) = 24,435 - 0,252 * Z - 15,179 * X + 0,362 * Z * X$$

Für die Kontrollgruppe (Therapie in Präsenz) ergibt sich durch das Einsetzen von $X = 0$:

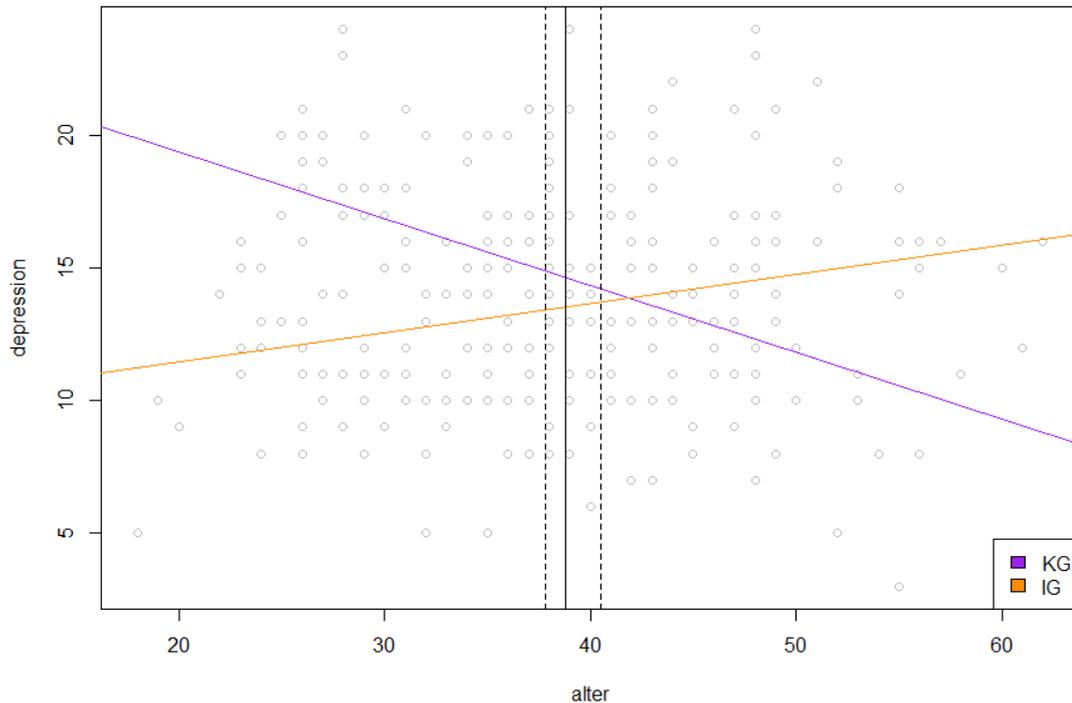
$$E(Y|X,Z) = 24,435 - 0,252 * Z$$

Für die Interventionsgruppe (Online-Therapie) ergibt sich durch das Einsetzen von $X = 1$:

$$E(Y|X,Z) = 24,435 - 0,252 * Z - 15,179 + 0,362 * Z$$

$$E(Y|X,Z) = 9,256 + 0,110 * Z$$

Zeichnet man die beiden Regressionsgeraden in ein Diagramm, sehen sie in etwa wie folgt aus, wobei hier die lilafarbene Gerade die geschätzte depressive Symptomatik in Abhängigkeit des Alters in der Präsenztherapie und die orangefarbene geschätzte depressive Symptomatik in Abhängigkeit des Alters in der Online-Therapie kennzeichnet:



Aus der Abbildung wird deutlich, dass die depressive Symptomatik bei jüngeren Personen bei Teilnahme an der Online-Therapie geringer ist als bei Teilnahme an der Therapie in Präsenz. Bei älteren Personen verhält es sich genau andersherum: Ihre depressive Symptomatik nach der Therapie ist bei Teilnahme an der Therapie in Präsenz geringer als bei Teilnahme an der Online-Therapie. Jüngere Personen scheinen also eher von der Online-Therapie zu profitieren, während ältere Personen eher von der Therapie in Präsenz profitieren.

Zum Einzeichnen der geschätzten kausalen Effekte benötigen wir den Mittelwert des Alters, zum einen für die Gesamtstichprobe und zum anderen für die einzelnen Therapiebedingungen:

```
mean(therapie_dat$alter)

## [1] 38.76981

tapply(therapie_dat$alter, therapie_dat$therapie, mean)

##      0      1
## 40.48387 37.84302
```

Personen in der Präsenztherapie sind im Mittel 40,48 Jahre alt, während Personen in der Online-Therapie im Mittel 37,84 Jahre alt sind. Im Mittel sind die Personen der gesamten Stichprobe 38,77 Jahre alt.

Die geschätzten kausalen Effekte entsprechen den Abständen der beiden Regressionsgeraden an verschiedenen Stellen:

- Der geschätzte ATE ist der Unterschied der orangefarbenen und der lilafarbenen Geraden an der Stelle des Mittelwerts des Alters der Gesamtstichprobe (38,77 Jahre, hier als durchgezogene Linie dargestellt).
- Der geschätzte ATU ist der Unterschied der orangefarbenen und der lilafarbenen Geraden an der Stelle des Mittelwerts des Alters der Gruppe, die Therapie in Präsenz erhalten hat (40,48 Jahre, die rechte gestrichelte Linie).

- Der geschätzte ATT ist der Unterschied der orangefarbenen und der lilafarbenen Geraden an der Stelle des Mittelwerts des Alters der Gruppe, die Online-Therapie erhalten hat (37,84 Jahre, die linke gestrichelte Linie).

Im Mittel führt die Online-Therapie zu geringeren Depressivitätswerten als die Therapie in Präsenz (ATE). Personen, die besonders stark von der Online-Therapie profitieren würden, wählen diese auch eher (ATT ist größer als ATU).

- (3) Der Datensatz *therapie.rda* aus der vorherigen Aufgabe enthält zudem noch die Kovariaten Symptombelastung vorher (*belastung*) und Motivation (*motivation*). Führen Sie ein Propensity-Score-Matching durch und schätzen Sie den durchschnittlichen kausalen Effekt. Spezifizieren Sie dazu zunächst das Modell zur Schätzung des Propensity-Score (PS) und achten Sie darauf Overlap und Balance für alle Kovariaten herzustellen.

Musterantwort

Schätzung der Propensity-Scores

```
# Laden der Pakete
library(MatchIt)
library(cobalt)

# Laden der Daten
load("df.rda")

# Matching vornehmen
match_ps <- matchit(therapie ~ alter + belastung + motivation,
                    data = therapie_dat,
                    method = "nearest",
                    distance = "glm",
                    link = "logit",
                    ratio = 1,
                    replace = FALSE,
                    caliper = 0.1,
                    discard = "both")

# Ergebnisse anzeigen lassen
summary(match_ps)
```

Overlap

Durch das Argument „discard = "both"“ werden Daten, die nicht Teil der Region of Common Support sind, ausgeschlossen, womit Overlap gegeben ist.

Balance

```
...
##      Summary      of      Balance      for      Matched      Data:
##      Means Treated Means Control Std. Mean Diff. Var. Ratio eCDF Mean
## distance      0.5619      0.5487      0.0617      1.0654      0.0202
## alter      40.2174      39.8261      0.0442      1.0871      0.0222
## belastung      17.2391      17.2174      0.0050      0.9728      0.0299
## motivation      4.1739      4.1087      0.0478      1.0978      0.0059
...

```

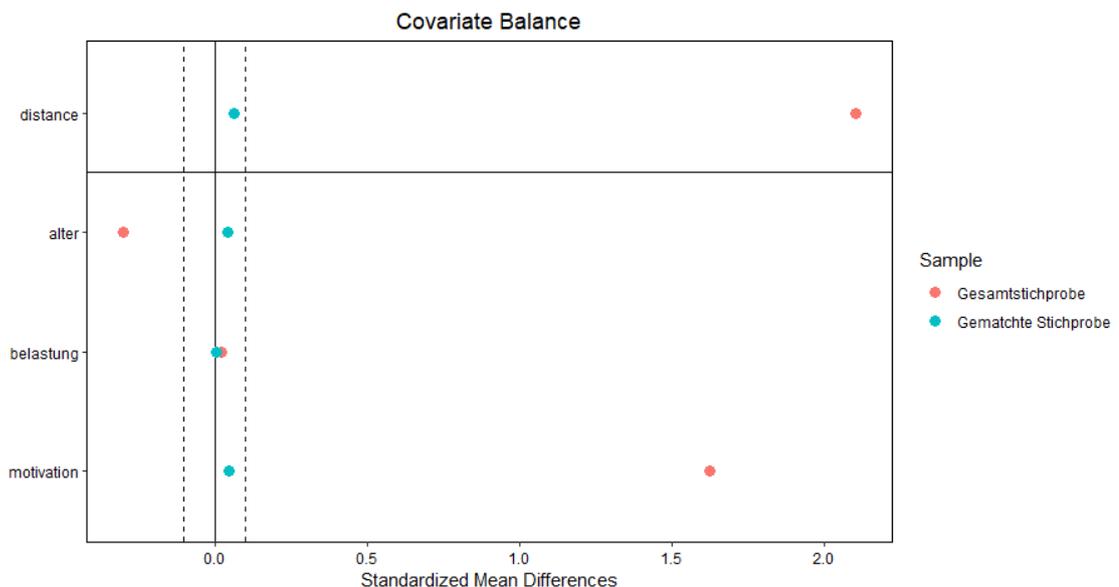
Im Mittel unterscheiden sich nach dem Matching die Propensity-Scores (*distance*) zwischen den beiden Interventionsgruppen kaum (Cohen's $d = 0,062$). Der Unterschied liegt innerhalb der von Steiner et al. (2010) vorgeschlagenen Grenzen (maximaler Unterschied von 0,1 Standardabweichung). Das Varianzverhältnis liegt bei 1,065 und damit ebenfalls innerhalb der von Steiner et al. (2010) definierten Grenzen von 0,8 bis 1,25.

Der Unterschied zwischen Präsenz- und Onlinetherapiegruppe bezüglich der Kovariaten Alter, Belastung und Motivation verschwindet durch das Matching ebenfalls fast vollkommen ($d_{\text{alter}} = 0,044$; $d_{\text{belastung}} = 0,005$; $d_{\text{motivation}} = 0,048$). Die entsprechenden Varianzverhältnisse betragen nach dem Matching 1,087; 0,973 und 1,098.

Diese Mittelwertsdifferenzen und Varianzverhältnisse können auch grafisch dargestellt werden:

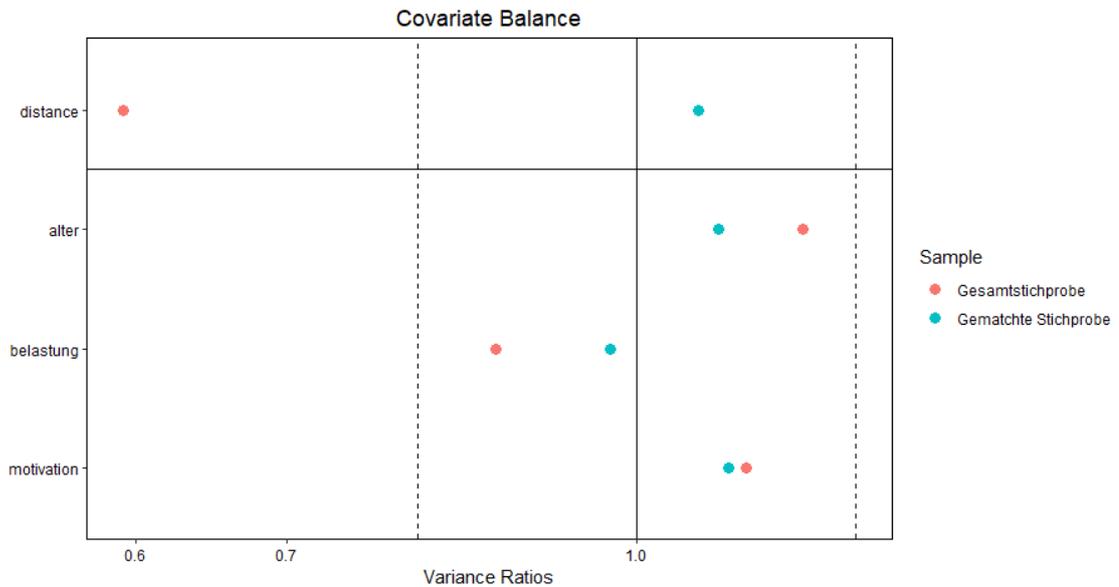
```
# Balance für Mittelwerte
love.plot(match_ps,
  stats = "mean.diffs",
  thresholds = c(-0.1, 0.1),
  sample.names = c("Gesamtstichprobe", "Gematchte Stichprobe"))

```



Nach dem Matchen (türkisfarbene Punkte) befinden sich die Mittelwertsdifferenzen der Propensity-Scores und aller Kovariaten innerhalb der Grenzen nach Steiner et al. (2010); vor dem Matchen (rote Punkte) lagen die Differenzen der Propensity-Scores und der Kovariaten Alter und Motivation noch außerhalb der Grenzen.

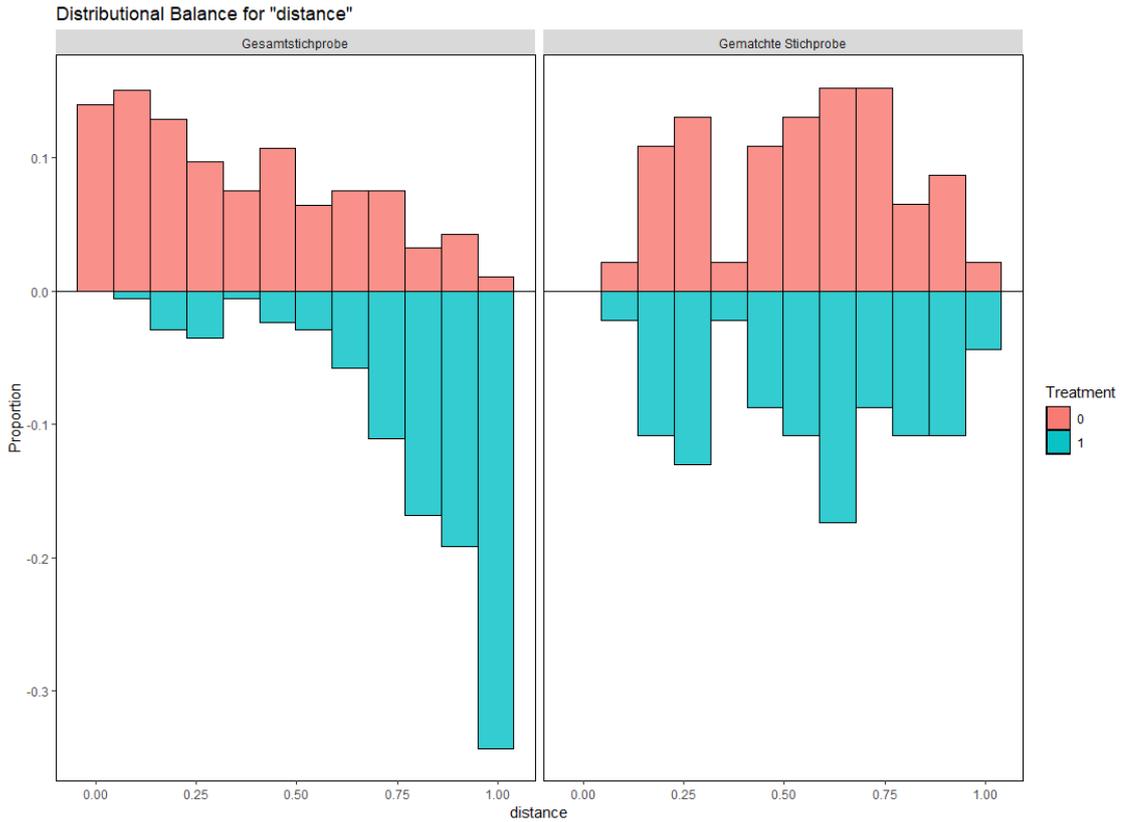
```
# Balance für Varianzen
love.plot(match_ps,
  stats = "variance.ratios",
  thresholds = c(0.8, 1.25),
  sample.names = c("Gesamtstichprobe", "Gematchte Stichprobe"))
```



Nach dem Matchen (türkisfarbene Punkte) befinden sich die Varianzverhältnisse der Propensity-Scores und aller Kovariaten innerhalb der Grenzen nach Steiner et al. (2010), vor dem Matchen (rote Punkte) lagen die Varianzverhältnisse der Propensity-Scores noch außerhalb der Grenzen.

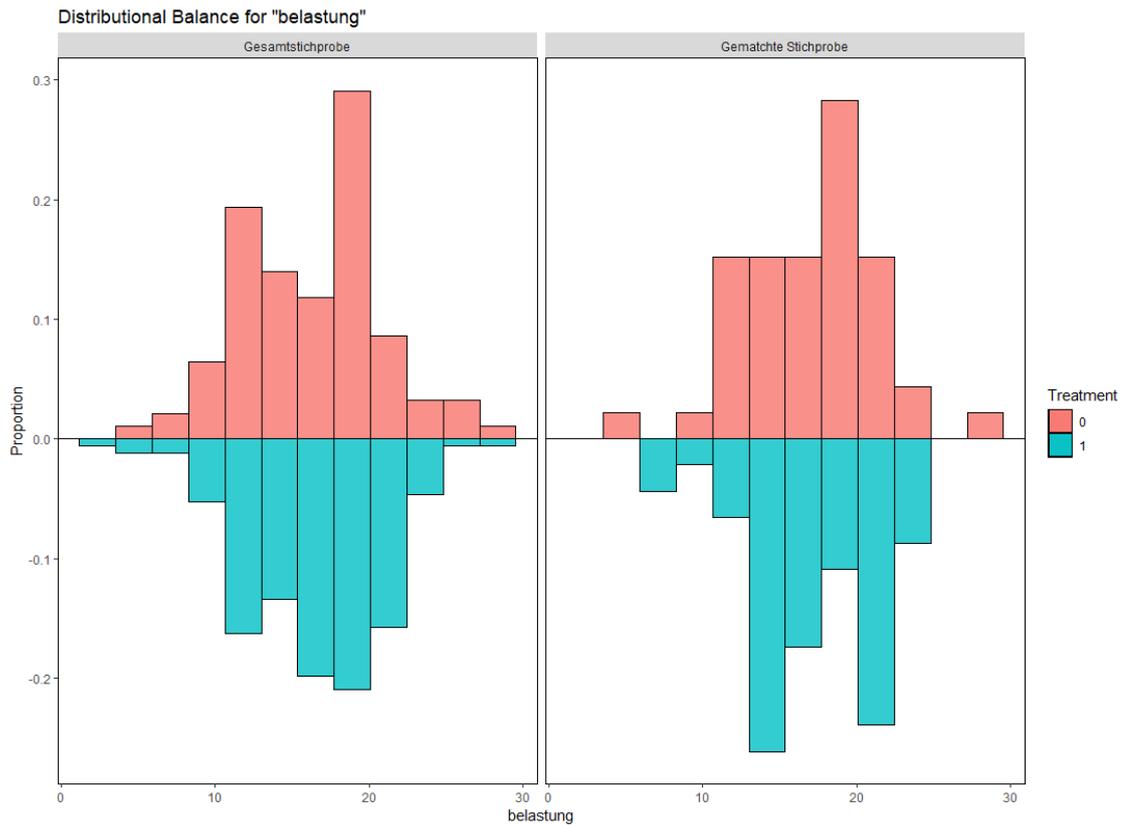
Mittelwerte und Varianzen beschreiben nur einen Teil der Verteilung metrischer Kovariaten. Für eine genauere Überprüfung der Gleichheit der Verteilungen der Kovariaten nach dem Matchen, können die Verteilungen selbst verglichen werden.

```
# Balance für die Verteilung der Propensity Scores
bal.plot(match_ps,
  which = "both",
  type = "histogram",
  mirror = TRUE,
  sample.names = c("Gesamtstichprobe", "Gematchte Stichprobe"))
```



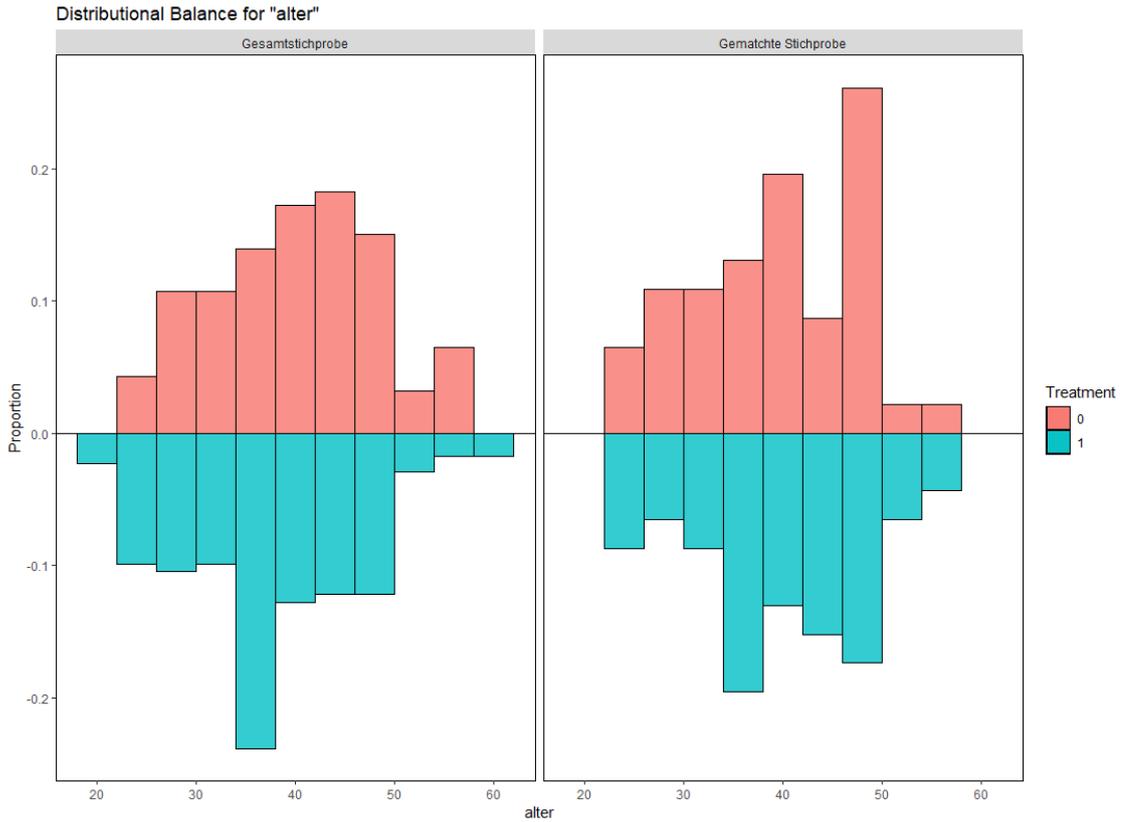
Vor dem Matchen (linke Seite der Abbildung) sieht man eine deutliche Imbalance der Verteilung der Propensity-Scores zwischen Präsenztherapiegruppe (rote Balken) und Onlinetherapiegruppe (türkisfarbene Balken). Nach dem Matchen (rechte Seite der Abbildung) ist die Verteilung der Propensity-Scores balanciert.

```
# Balance für die Verteilung von Belastung
bal.plot(match_ps,
  var.name = "belastung",
  which = "both",
  type = "histogram",
  mirror = TRUE,
  sample.names = c("Gesamtstichprobe", "Gematchte Stichprobe"))
```



Die Verteilung der Kovariate Belastung unterscheidet sich sowohl vor dem Matchen (linke Seite der Abbildung) als auch nach dem Matchen (rechte Seite der Abbildung) kaum zwischen Präsenztherapiegruppe (rote Balken) und Onlinetherapiegruppe (türkisfarbene Balken). Sie ist balanciert.

```
# Balance für die Verteilung von Alter
bal.plot(match_ps,
  var.name = "alter",
  which = "both",
  type = "histogram",
  mirror = TRUE,
  sample.names = c("Gesamtstichprobe", "Gematchte Stichprobe"))
```



Vor dem Matchen (linke Seite der Abbildung) gibt es leichte Unterschiede in der Verteilung der Kovariate Alter zwischen Präsenztherapiegruppe (rote Balken) und Onlinetherapiegruppe (türkisfarbene Balken). Nach dem Matchen (rechte Seite der Abbildung) ist die Verteilung von Alter balanciert.

```
# Balance für die Verteilung von Motivation
bal.plot(match_ps,
  var.name = "motivation",
  which = "both",
  type = "histogram",
  mirror = TRUE,
  sample.names = c("Gesamtstichprobe", "Gematchte Stichprobe"))
```


Der geschätzte durchschnittliche kausale Effekt der gematchten Stichprobe beträgt $-0,8043$. Wir schätzen also, dass die Online-Therapie für Personen in der gematchten Stichprobe im Durchschnitt zu einer um $0,8043$ Punkte niedrigeren depressiven Symptomatik nach der Therapie führt als die Therapie in Präsenz.

- (4) In einer Schule wird ein spezielles Förderprogramm für naturwissenschaftlich hoch interessierte Schüler:innen angeboten. Dazu wird das Interesse der Schüler:innen (*interesse*, Wertebereich von 0 bis 10) gemessen und alle Schüler:innen über einem Interesse von 7 Punkten dem Förderprogramm zugewiesen (*programm*=1 Förderung, *programm*=0 Kontrollbedingung). Nach einem Schuljahr wird das naturwissenschaftliche Wissen (*wissen*) aller Schüler:innen gemessen. Sie finden die Daten unter dem Dateinamen *nawi.rda* in den Online-Materialien.

(a) Welches Untersuchungsdesign wurde hier umgesetzt?

Musterantwort

Es handelt sich um ein Regressions-Diskontinuitäts-Design.

- (b) Spezifizieren Sie ein lineares Modell und schätzen Sie den lokalen durchschnittlichen kausalen Effekt. Interpretieren Sie Ihr Ergebnis.

Musterantwort

```
# Daten einlesen
load("nawi.rda")

# Schätzung des Effektes
m1 <- lm(wissen ~ interesse + programm, data = nawi)

summary(m1)

##
## Call:
## lm(formula = wissen ~ interesse + programm, data = nawi)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -46.513 -11.106  -0.977  12.742  45.276
##
## Coefficients:
##              Estimate Std. Error t value      Pr(>|t|)
## (Intercept)  52.5780     4.2783  12.289 <0.000000000000002 ***
## interesse   -0.9699     0.8923  -1.087     0.2783
## programm     9.6336     4.6538   2.070     0.0397 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 17.99 on 213 degrees of freedom
## Multiple R-squared:  0.02075,    Adjusted R-squared:  0.01155
## F-statistic: 2.257 on 2 and 213 DF,  p-value: 0.1072
```

Der geschätzte lokale durchschnittliche Effekt (LATE) des Förderprogramms für Personen mit einem Interesse von 7 beträgt $9,63$ ($SE = 4,654$; $p = ,040$). Das bedeutet, dass Personen mit einem Interesse von 7 Punkten im Mittel um $9,63$ Punkte höheres Wissen haben, wenn sie an dem Förderprogramm teilnehmen, als wenn sie nicht teilnehmen.

- (c) Erläutern Sie eine Möglichkeit, wie man mit einer möglichen Fehlspezifikation der funktionalen Form umgehen kann.

Musterantwort

Eine Möglichkeit, mit einer möglichen Fehlspezifikation der funktionalen Form umzugehen, ist die Schätzung des lokalen durchschnittlichen Treatmenteffekts (LATE) auf Basis eines Teildatensatzes (subset), welcher nur Personen enthält, deren Werte auf der Variable interesse nahe dem Cutoff von sieben Punkten liegen. Wie weit individuelle Werte auf der Variablen 'Interesse' maximal vom Cutoff abweichen dürfen, um im Teildatensatz beibehalten zu werden, können wir entweder selbst festlegen, oder über eine automatische Schätzung der optimalen Bandbreite bestimmen. Eine weitere Möglichkeit besteht in der Spezifikation einer anderen funktionalen Form (z.B. eine quadratische anstatt einer linearen Regression). Die Umsetzung dieser Möglichkeiten in R wird in den Online-Materialien demonstriert.

- (5) Sie wollen den Effekt des Tragens einer Gesichtsmaske (*maske*) auf die Konzentrationsfähigkeit (*konzentration*) untersuchen und führen dazu eine experimentelle Studie durch. Eine Gruppe junger, gesunder Menschen soll sich einen Tag lang in ihrem gewohnten Umfeld bewegen ohne eine FFP2-Maske zu tragen, während eine andere Gruppe konsequent eine FFP-2 Maske tragen soll. Die Zuweisung zu den Gruppen (*zuweisung*) erfolgt zufällig. In Ihrer Studie haben sich nicht alle Teilnehmenden an die Aufforderung gehalten. Nachdem Sie die Personen befragt haben, stellen Sie fest, dass einige der Personen, die eine FFP2-Maske hätten tragen sollen, dies nicht getan haben, während einige Personen, die keine Maske tragen sollten, dennoch eine verwendet haben. Den Datensatz finden Sie unter dem Namen *ffp2.rda* in den Online-Materialien.
- (a) Berechnen und interpretieren Sie die entsprechenden kausalen Effekte mithilfe des

- i. Per-Protocol-,

Musterantwort

```
# Daten einlesen
load("ffp2.rda")

# Datensatz mit Personen erstellen, die sich an die Zuweisung gehalten haben
maske_pp <- ffp2[ffp2$zuweisung == ffp2$maske,]

# Effekt schätzen
mod_perProtocol <- lm(konz ~ maske, data = maske_pp)

summary(mod_perProtocol)$coef
```

##	Estimate	Std. Error	t value	Pr(> t)
## (Intercept)	4.5833333	0.1119628	40.936208	5.374656e-141
## maske	0.2730496	0.1591793	1.715359	8.709899e-02

Personen, die eine Maske getragen haben, haben im Mittel eine um 0,27 Punkte höhere Konzentrationsfähigkeit als Personen, die keine Maske getragen haben. Dieser Effekt ist nicht signifikant ($SE = 1,98$; $t(1) = 1,72$; $p = 0,087$).

Da nur noch die Personen einbezogen wurden, die sich entsprechend der Anweisung verhalten haben, treten Selektionseffekte auf. Da diese in dieser Analyse nicht kontrolliert wurden (z.B. durch ANCOVA oder PS-Verfahren), ist eine kausale Interpretierbarkeit des geschätzten Effektes nicht unbedingt gegeben.

ii. As-Treated-

Musterantwort

```
# Effekt schätzen
mod_asTreat <- lm(konz ~ maske, data = ffp2)

summary(mod_asTreat)$coef

##              Estimate Std. Error  t value    Pr(>|t|)
## (Intercept) 4.5598291  0.1017583 44.810380 7.454528e-171
## maske       0.5860937  0.1440623  4.068334 5.562227e-05
```

Personen, die eine Maske tragen, haben im Mittel eine um 0,58 Punkte höhere Konzentrationsfähigkeit als Personen, die keine Maske tragen. Dieser Effekt ist signifikant ($SE = 0,14$; $t(1) = 4,07$; $p < ,001$), sollte allerdings ebenfalls nicht kausal interpretiert werden, da das tatsächliche Tragen einer Gesichtsmaske nicht randomisiert wurde und somit Selektionseffekte auftreten können.

iii. Intention-To-Treat- sowie

Musterantwort

```
# Effekt schätzen
mod_itt <- lm(konz ~ zuweisung, data = ffp2)

summary(mod_itt)$coef

##              Estimate Std. Error  t value    Pr(>|t|)
## (Intercept) 4.9198312  0.1027990 47.858724 8.824737e-182
## zuweisung  -0.1372225  0.1464818 -0.936789 3.493534e-01
```

Die Anweisung, eine Gesichtsmaske zu tragen, führt im Vergleich zur Anweisung, keine Maske zu tragen, zu einer um 0,14 Punkte niedrigeren Konzentrationsfähigkeit. Dieser Effekt (der kausale Effekt der Zuweisung) ist nicht signifikant ($SE = 0,15$; $t(1) = -0,94$; $p = 0,35$). Er ist kausal interpretierbar, da die Anweisung zum Tragen einer Maske randomisiert erfolgte.

iv. Instrumenten-Variablen-Ansatzes.

Musterantwort

```
# Effekt schätzen
sem_mod <- tsls(konz ~ maske,
                instruments = ~ zuweisung,
                data = ffp2)

summary(sem_mod)$coef

##              Estimate Std. Error  t value    Pr(>|t|)
## (Intercept) 4.9613518  0.1397856 35.4925933 0.0000000
## maske       -0.2186751  0.2371793 -0.9219823 0.3570159
```

Der geschätzte kausale Effekt des Tragens einer Gesichtsmaske auf die Konzentrationsfähigkeit für solche Personen, die sich an jegliche Anweisung halten würden (d.h. eine Maske tragen würden, wenn sie dazu angewiesen werden und keine Maske tragen, wenn sie nicht angewiesen

werden), beträgt -0,22 Punkte (Complier-Average-Treatment-Effekt). Das bedeutet, dass bei Complyern das Tragen einer Gesichtsmaske im Mittel zu einer um 0,22 Punkte geringeren Konzentrationsfähigkeit führt. Dieser Effekt ist nicht signifikant ($SE = 0,24$; $t(1) = -0,92$; $p = 0,36$). Wenn alle notwendigen Annahmen des Instrumenten-Variablen-Ansatzes erfüllt sind, ist der Effekt kausal interpretierbar.

- (b) Geben Sie für den Instrumenten-Variablen-Ansatz an, welche Annahmen gemacht werden und schätzen Sie ab, wie plausibel diese in Ihrer Untersuchung sind.

Die Annahmen, die gelten müssen, damit der anhand des Instrumenten-Variablen-Ansatzes geschätzte Effekt kausal interpretierbar ist, sind Relevanz, Monotonie und Exogenität.

- (1) Relevanz beschreibt die Annahme, dass es zwischen der Zuweisung zum Tragen der Maske und dem tatsächlichen Tragen der Maske einen Zusammenhang gibt, es also zumindest einige Personen gibt, die sich an die Zuweisung halten.

Die Annahme lässt sich leicht prüfen, indem wir die Korrelation zwischen der Zuweisung und dem Tragen der Gesichtsmaske betrachten:

```
cor(ffp2$zuweisung, ffp2$maske)
```

```
## [1] 0.6274488
```

Die Korrelation beträgt $\phi = 0,63$, somit ist die Annahme der Relevanz erfüllt.

- (2) Monotonie beschreibt die Annahme, dass es im Datensatz keine Defiers geben darf, also keine Personen, die sich immer entgegen der Anweisung zum Tragen oder Nicht-Tragen der Gesichtsmaske verhalten würden.

Wir können nicht mit Sicherheit sagen, dass es solche Personen nicht gibt, allerdings erscheint es im vorliegenden Beispiel eher unplausibel, dass sich bestimmte Personen immer gegenteilig zur Anweisung verhalten würden. Wahrscheinlich gibt es einige Personen, die unabhängig von der Anweisung zum Tragen oder Nicht-Tragen nie eine Maske tragen würden (Never-Taker) und einige Personen, die unabhängig von der Anweisung immer eine Maske tragen würden (Always-Taker). Dass eine Person bei der Anweisung zum Nicht-Tragen der Maske doch eine Maske tragen, aber bei der Anweisung zum Tragen der Maske keine Maske tragen würde, erscheint weniger plausibel, ist jedoch nicht völlig auszuschließen.

- (3) Exogenität beschreibt die Annahme, dass a) das Instrument, also die Anweisung zum Tragen oder Nicht-Tragen der Maske, nur indirekt über die tatsächliche Interventionsvariable, also das tatsächliche Tragen oder Nicht-Tragen der Maske mit dem Outcome, also der Konzentrationsfähigkeit zusammenhängt und b) das Instrument unabhängig von unbeobachteten konfundierenden Variablen ist.

Da die Zuweisung der Personen zur Anweisung zum (Nicht-)Tragen der Gesichtsmasken randomisiert erfolgt ist, sind sowohl a) als auch b) erfüllt, womit Exogenität gegeben ist.