

Methods 4

Bayesian Computational Modelling

Portfolio Assignment 2

Group Assignment – Study Group 3

Emma Olsen(202006507@post.au.dk)

Rikke Uldbæk (202007501@post.au.dk)

Jørgen Højlund Wibe (201807750@post.au.dk)

Otto Sejrsild Santesson (202004347@post.au.dk)

Students of Cognitive Science, University of Aarhus,

Jens Chr. Skous Vej 2, 8000, Aarhus, Denmark

Methods 4 – Portfolio Assignment 2

Study Group 3

- *Type:* Group assignment
- *Due:* 3 April 2022, 23:59

Hello CogSci's :)

In this portfolio, you are asked to do four tasks:

- Make a DAG for something
- Simulate data that fits the DAG
- Use linear models to confirm that the DAG fits the data
- Mess it up.

Each of the four tasks have some sub-steps.

Report briefly what you find, for example in a markdown document, for example called report.md so that the poor TA can easily get an overview before looking in your code :)

Then you can also make a (brief!) explanation of the phenomenon you are DAGGIN, simulating and modelling.

Looking forward!

Task 1: The DAG

- **Come up with an** incredibly interesting and scientifically important made-up **example** for a phenomenon to investigate. Decide on two variables (an outcome and a predictor) that you would like to investigate the relation between. If in doubt, you **can be inspired by Peter's amazing example** on the next page.

In our example, we wish to investigate the effect of the continuous predictor corona fear (CF) on the outcome variable energy level (also continuous).

Predictor variable: *corona fear_ (continuous variable). This variable depends on OCD, as we expect OCD to affect your corona fear in such way that a higher level of OCD increases your level of corona fear.*

Outcome variable: *energy level_ (continuous variable), depends on germ level as we expect the level of germs to determine your energy level. With a low level of germs, the germs will not take up much energy and nutrients from your body, but with a higher number of germs in your system, the germs will take more nutrients and energy.*

OCD: *someone's level of OCD describes as a continuous variable for the sake of this assignment. This variable is independent from all other variables in our DAG but acts as a common cause of corona fear and sanitizing.*

Sanitizing: *a continuous variable dependent on both corona fear and OCD. It is dependent on corona fear as we expect that you will sanitize more as your fear of corona increases. It is dependent on OCD as a typical symptom of OCD is an abnormal fear of germs, which we assume will result in highly frequent use of sanitizer. Thus, we assume that the amount of sanitizing will increase with the level of OCD.*

Social exposure: a measure of how much social mingling you engage in (a combined measure of social circle size, social activities etc.). This variable is also dependent on corona fear, as we assume that a person who is really scared of Corona would isolate more and thus have a much lower social exposure, whereas a person who didn't fear corona at all wouldn't worry about meeting people and thus have a correspondingly larger social exposure.

Germ level: a continuous variable describing the level of vira germs in your body. For the sake of simplicity, we assume that the level of germs present in your body can be described as a normally distributed continuous variable. This variable is dependent on the social exposure variable as well as sanitizing. It depends on social exposure, as a lot of social exposure implies more exposure to germs. It depends on sanitizing as sanitizing works by killing most germs, preventing them from entering your body.

- **Make a DAG** for the phenomenon. Make it medium complicated: that means, make sure there are some different kinds of relations (see next step). Change it if you don't get anything interesting for the next steps.

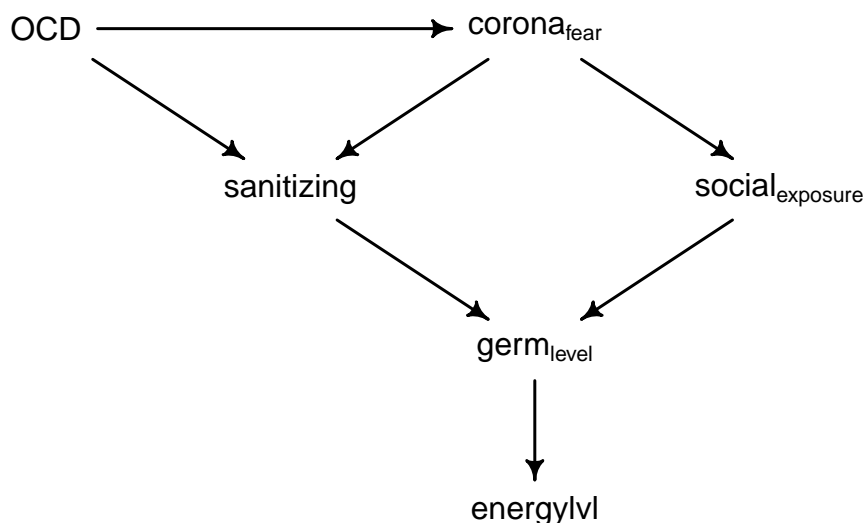
Draw it somehow (on paper, in R, laser engraved in diamond).

Code it in dagitty (this is a nice tool: <http://dagitty.net/dags.html>)

```
dag_corona <- dagitty("dag {
  corona_fear -> sanitizing
  sanitizing-> germ_level
  OCD -> sanitizing
  OCD -> corona_fear
  social_exposure -> germ_level
  corona_fear -> social_exposure
  germ_level -> energylvl
}")

# Plotting the DAG
coordinates(dag_corona) <- list(x = c(corona_fear = -1, OCD = -3,
  social_exposure = 0, sanitizing = -2, germ_level = -1, energylvl = -1),
  y = c(corona_fear = -1, OCD = -1, social_exposure = 0, sanitizing = 0,
  germ_level = 1, energylvl = 2))

drawdag(dag_corona)
```



- Find **elemental forms of variable relations** in the DAG (i.e., forks, pipes, colliders, and their descendants).

Pipe - (Corona Fear -> Social Exposure -> Germ Level)

- (Corona Fear -> Sanitizing -> Germ Level)
- (OCD -> Corona Fear -> Social Exposure)
- (OCD -> Corona Fear -> Sanitizing)
- (OCD -> Sanitizing -> Germ Level)
- (Sanitizing -> Germ Level -> EnergyLevel)
- (Social Exposure -> Germ Level -> Energy Level)

Fork - (Sanitizing <- Corona Fear -> Social Exposure)

- (Corona Fear <- OCD -> Sanitizing)

Colider - (Sanitizing -> Germ Level <- Social Exposure)

- (Corona Fear-> Sanitizing <- OCD)

Descendant - (OCD -> Corona Fear -> Social Exposure, Corona Fear -> Sanitizing)

- Find out **what variables to include (and not include)** in a multiple linear regression to avoid ‘back door’ (AKA non-causal) paths. Do this first with your eyes and your mind. Then you can use dagitty’s function `adjustmentSets()`.

```
# Shutting the backdoor - analyzing the graph to block the
# backdoor
adjustmentSets(dag_corona, exposure = "corona_fear", outcome = "energylvl")
```

```
## { OCD }
```

OCD is a backdoor that we should control for (i.e., include in a our model) in order to find the full causal effect of corona fear on energy level.

- Find out which **conditional independencies** the DAG implies. First with the mind, then with dagitty’s function `impliedConditionalIndependencies()`.

```
# Deriving our DAG’s conditional independencies
impliedConditionalIndependencies(dag_corona)
```

```
## OCD _||_ enrg | grm_
## OCD _||_ enrg | sntz, scl_
## OCD _||_ enrg | crn_, sntz
## OCD _||_ grm_ | sntz, scl_
## OCD _||_ grm_ | crn_, sntz
## OCD _||_ scl_ | crn_
## crn_ _||_ enrg | grm_
## crn_ _||_ enrg | sntz, scl_
## crn_ _||_ grm_ | sntz, scl_
## enrg _||_ sntz | grm_
## enrg _||_ scl_ | grm_
## sntz _||_ scl_ | crn_
```

The above output shows the testable implications of our DAG, which we will go through in task 3 to test if any of them can be falsified - i.e. to know if the DAG is compatible with the data. If any of the conditional independencies are not fulfilled, we can falsify our DAG.

- Find the full list of **Markov equivalent** DAGs. Use dagitty's function `equivalentGraphs()`.

The `equivalentGraphs()` function outputs a set of DAGs with the same conditional independencies is known as a Markov equivalence set.

```
# Finding equivalent DAGs
```

```
equivalentDAGs(dag_corona)
```

```
## [[1]]
## dag {
##   OCD [pos="-3.000,-1.000"]
##   corona_fear [pos="-1.000,-1.000"]
##   energylvl [pos="-1.000,2.000"]
##   germ_level [pos="-1.000,1.000"]
##   sanitizing [pos="-2.000,0.000"]
##   social_exposure [pos="0.000,0.000"]
##   OCD -> corona_fear
##   OCD -> sanitizing
##   corona_fear -> sanitizing
##   corona_fear -> social_exposure
##   germ_level -> energylvl
##   sanitizing -> germ_level
##   social_exposure -> germ_level
## }
##
## [[2]]
## dag {
##   OCD [pos="-3.000,-1.000"]
##   corona_fear [pos="-1.000,-1.000"]
##   energylvl [pos="-1.000,2.000"]
##   germ_level [pos="-1.000,1.000"]
##   sanitizing [pos="-2.000,0.000"]
##   social_exposure [pos="0.000,0.000"]
##   OCD -> corona_fear
##   OCD -> sanitizing
##   corona_fear -> social_exposure
##   germ_level -> energylvl
##   sanitizing -> corona_fear
##   sanitizing -> germ_level
##   social_exposure -> germ_level
## }
##
## [[3]]
## dag {
##   OCD [pos="-3.000,-1.000"]
##   corona_fear [pos="-1.000,-1.000"]
##   energylvl [pos="-1.000,2.000"]
##   germ_level [pos="-1.000,1.000"]
##   sanitizing [pos="-2.000,0.000"]
##   social_exposure [pos="0.000,0.000"]
##   OCD -> corona_fear
```

```

## corona_fear -> social_exposure
## germ_level -> energylvl
## sanitizing -> OCD
## sanitizing -> corona_fear
## sanitizing -> germ_level
## social_exposure -> germ_level
## }
##
## [[4]]
## dag {
## OCD [pos="-3.000,-1.000"]
## corona_fear [pos="-1.000,-1.000"]
## energylvl [pos="-1.000,2.000"]
## germ_level [pos="-1.000,1.000"]
## sanitizing [pos="-2.000,0.000"]
## social_exposure [pos="0.000,0.000"]
## OCD -> sanitizing
## corona_fear -> OCD
## corona_fear -> sanitizing
## corona_fear -> social_exposure
## germ_level -> energylvl
## sanitizing -> germ_level
## social_exposure -> germ_level
## }
##
## [[5]]
## dag {
## OCD [pos="-3.000,-1.000"]
## corona_fear [pos="-1.000,-1.000"]
## energylvl [pos="-1.000,2.000"]
## germ_level [pos="-1.000,1.000"]
## sanitizing [pos="-2.000,0.000"]
## social_exposure [pos="0.000,0.000"]
## OCD -> sanitizing
## corona_fear -> OCD
## corona_fear -> sanitizing
## germ_level -> energylvl
## sanitizing -> germ_level
## social_exposure -> corona_fear
## social_exposure -> germ_level
## }
##
## [[6]]
## dag {
## OCD [pos="-3.000,-1.000"]
## corona_fear [pos="-1.000,-1.000"]
## energylvl [pos="-1.000,2.000"]
## germ_level [pos="-1.000,1.000"]
## sanitizing [pos="-2.000,0.000"]
## social_exposure [pos="0.000,0.000"]
## corona_fear -> OCD
## corona_fear -> sanitizing
## corona_fear -> social_exposure
## germ_level -> energylvl

```

```

## sanitizing -> OCD
## sanitizing -> germ_level
## social_exposure -> germ_level
## }
##
## [[7]]
## dag {
## OCD [pos="-3.000,-1.000"]
## corona_fear [pos="-1.000,-1.000"]
## energylvl [pos="-1.000,2.000"]
## germ_level [pos="-1.000,1.000"]
## sanitizing [pos="-2.000,0.000"]
## social_exposure [pos="0.000,0.000"]
## corona_fear -> OCD
## corona_fear -> sanitizing
## germ_level -> energylvl
## sanitizing -> OCD
## sanitizing -> germ_level
## social_exposure -> corona_fear
## social_exposure -> germ_level
## }
##
## [[8]]
## dag {
## OCD [pos="-3.000,-1.000"]
## corona_fear [pos="-1.000,-1.000"]
## energylvl [pos="-1.000,2.000"]
## germ_level [pos="-1.000,1.000"]
## sanitizing [pos="-2.000,0.000"]
## social_exposure [pos="0.000,0.000"]
## corona_fear -> OCD
## corona_fear -> social_exposure
## germ_level -> energylvl
## sanitizing -> OCD
## sanitizing -> corona_fear
## sanitizing -> germ_level
## social_exposure -> germ_level
## }

```

The list outputted above shows us 8 DAGs that would have the exact same conditional independencies as our DAG, dag_corona, has. However, many of the causal relations between variables in the DAGs listed above would be incompatible with our scientific knowledge and we would have no reason to build statistical models based on these DAGs. For instance, we would have no reason to assume that sanitizing will have a causal effect on OCD, as suggested by e.g. DAG 3, 6, 7, and 8 above.

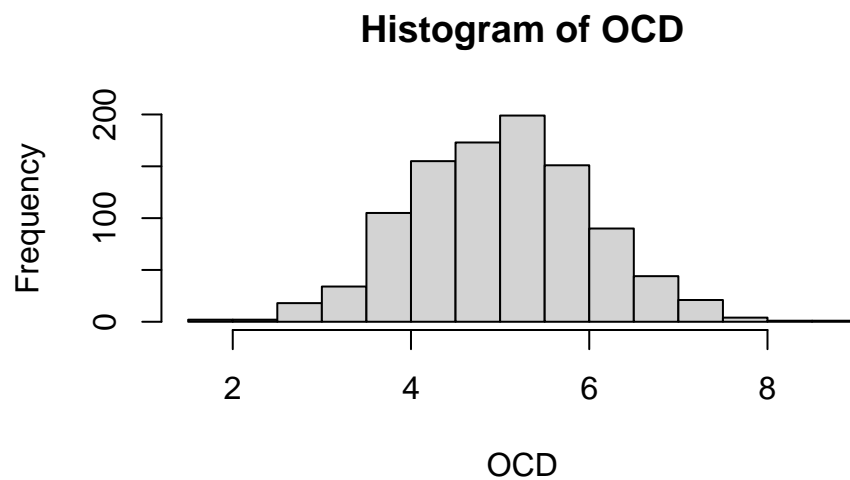
Task 2: The data

- **Simulate some data that fits the DAG.** There are many ways to do this. A simple way is just to sample one variable from a normal distribution which has another variable as mean. McElreath does this in the book a few times, and you can use this as inspiration.

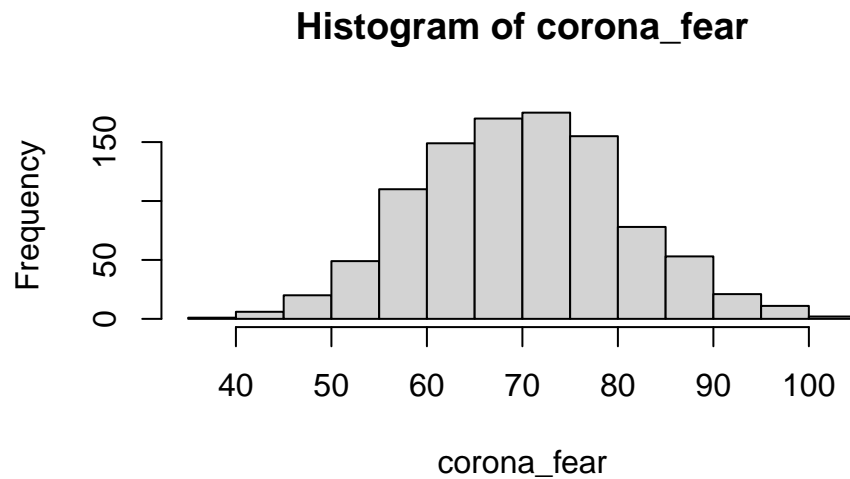
```
set.seed(3)
N = 1000
```

Simulating variables

```
# OCD
OCD <- rnorm(N, mean = 5, sd = 1)
hist(OCD)
```

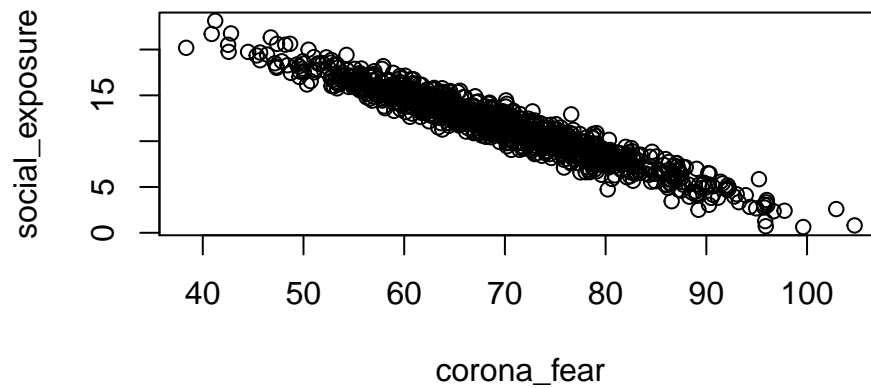


```
# Corona fear
corona_fear <- rnorm(N, mean = 20 + (10 * OCD), sd = 4)
hist(corona_fear)
```



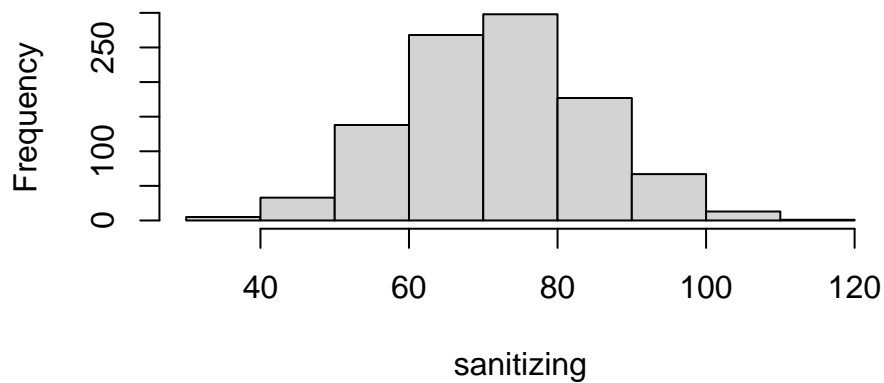

```
# Social exposure
social_exposure <- rnorm(N, mean = (35 - (corona_fear/3)), sd = 1)
```

Plot: corona fear and social exposure

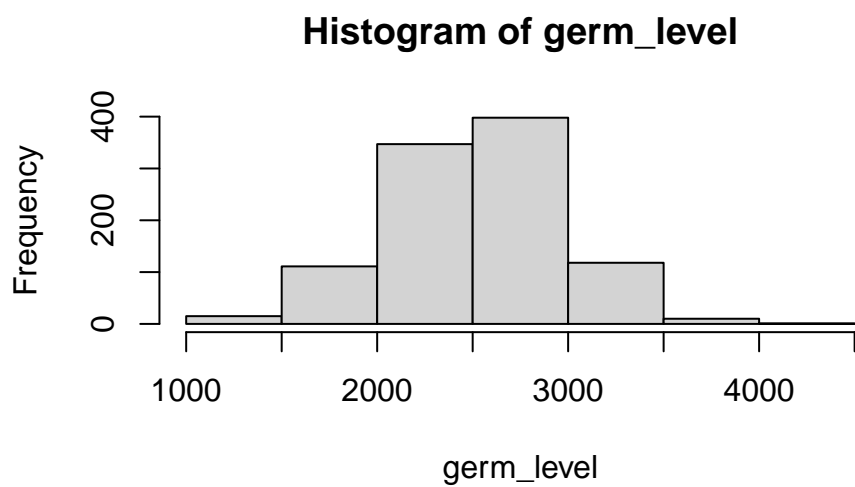


```
# Sanitizing
sanitizing <- rnorm(N, mean = (corona_fear - 8) + (OCD * 2),
  sd = 2)
hist(sanitizing)
```

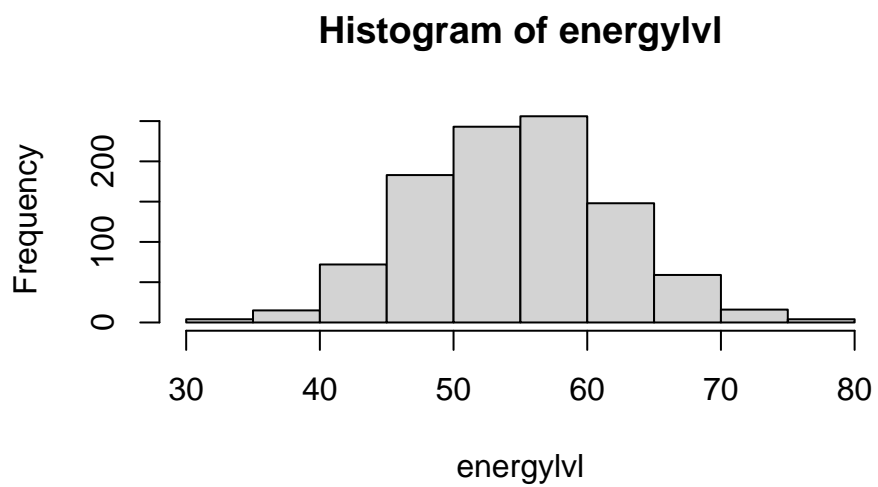
Histogram of sanitizing



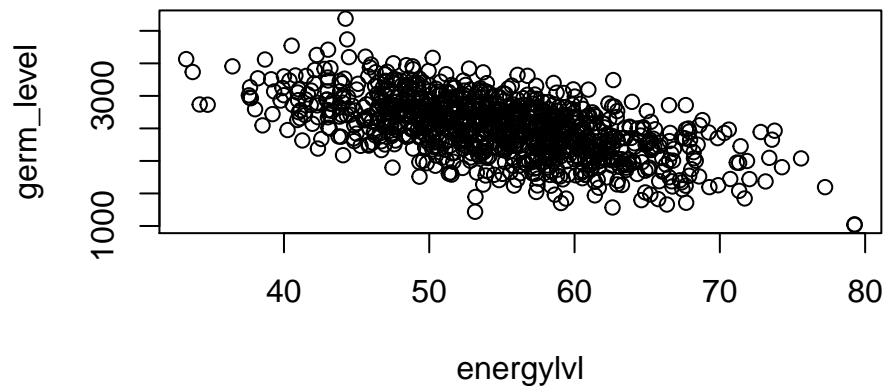
```
# Germ level
germ_level <- rnorm(N, mean = 8000 - ((sanitizing * 60) + (100 *
  social_exposure)), sd = 100)
hist(germ_level)
```



```
# Energy level  
energylvl <- rnorm(N, mean = 80 - ((germ_level/100)), sd = 6)  
hist(energylvl)
```



Energy level and germ level



Adding the simulated variables to a tibble.

```
df <- tibble(germ_level, sanitizing, energylvl, social_exposure,
             corona_fear, OCD)
```

Standardizing variables.

```
df <- df %>%
  mutate(germ_std = scale(germ_level), sanitizing_std = scale(sanitizing),
         energylvl_std = scale(energylvl), social_exposure_std = scale(social_exposure),
         corona_fear_std = scale(corona_fear), OCD_std = scale(OCD))
```

Task 3: Statistics

- Run **multiple linear regressions** to **test the conditional independencies implied by your DAG**. Make sure to avoid backdoor paths. See that the linear model shows the conditional independencies implied by your DAG, implying that the data and the DAG are compatible (if the linear model doesn't show the conditional independencies implied by the DAG, the data and the DAG doesn't fit).

```
impliedConditionalIndependencies(dag_corona)
```

```
## OCD _||_ enrg | grm_
## OCD _||_ enrg | sntz, scl_
## OCD _||_ enrg | crn_, sntz
## OCD _||_ grm_ | sntz, scl_
## OCD _||_ grm_ | crn_, sntz
## OCD _||_ scl_ | crn_
## crn_ _||_ enrg | grm_
## crn_ _||_ enrg | sntz, scl_
## crn_ _||_ grm_ | sntz, scl_
## enrg _||_ sntz | grm_
## enrg _||_ scl_ | grm_
## sntz _||_ scl_ | crn_
```

Testing the conditional independencies

From the output above, we will now test the 12 conditional independencies one by one.

Model 1: OCD // enrg | germ

Testing the implication that the causal coefficient of energy_level becomes significantly close to zero when stratifying by germ level in a model that has OCD as outcome variable.

```
adjustmentSets(dag_corona, exposure = "energylvl", outcome = "OCD") # Adding { germ_level }
```

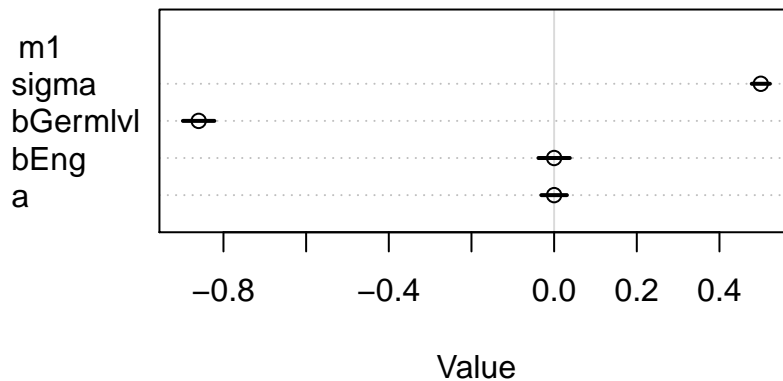
```
## { corona_fear, sanitizing }  
## { sanitizing, social_exposure }  
## { germ_level }
```

```
m1 <- quap(alist(OCD_std ~ dnorm(mu, sigma), mu <- a + bEng *  
  energylvl_std + bGermlvl * germ_std, a ~ dnorm(0, 0.2), bEng ~  
  dnorm(0, 1), bGermlvl ~ dnorm(0, 1), sigma ~ dexp(1)), data = df)
```

```
precis(m1)
```

##		mean	sd	5.5%	94.5%
## a		1.388363e-07	0.01579213	-0.02523874	0.02523901
## bEng		2.101325e-03	0.01942447	-0.02894274	0.03314539
## bGermlvl		-8.638228e-01	0.01942448	-0.89486692	-0.83277878
## sigma		5.009551e-01	0.01119728	0.48305970	0.51885055

```
plot(coeftab(m1), by.model = TRUE)
```



The plot above shows the posterior mean for our model's parameters sigma, bGermlvl, bEng and our intercept, alpha (a). The posterior means are shown by the points and the 89% compatibility intervals are shown by the solid horizontal lines. We can interpret the posterior mean of the parameter bEng as: once we know a person's germ level, there is no additional predictive power in also knowing the person's energy level.

The posterior mean for energy level (bEng) is 0 with very little probability of both sides of zero. This does not violate the implication that energy level becomes independent from OCD when stratifying by germ level.

Model 2: OCD // enrg | sntz, scl_

Testing the implication that the causal coefficient of energy_level becomes significantly close to zero when stratifying by sanitizing and social exposure in a model that has OCD as outcome variable.

```
adjustmentSets(dag_corona, exposure = "energylvl", outcome = "OCD") # Adding { sanitizing, social_expo

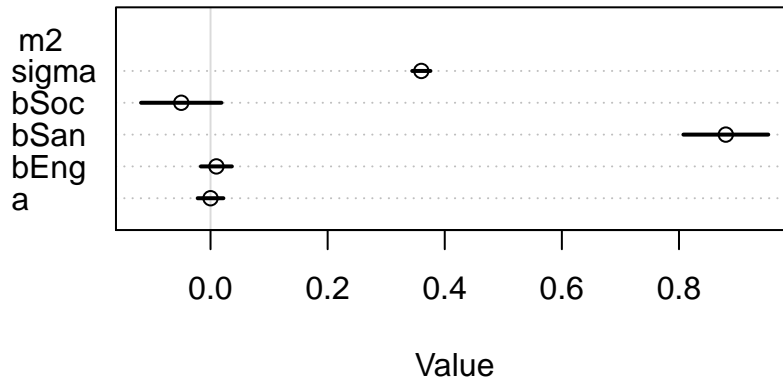
## { corona_fear, sanitizing }
## { sanitizing, social_exposure }
## { germ_level }

m2 <- quap(alist(OCD_std ~ dnorm(mu, sigma), mu <- a + bEng *
  energylvl_std + bSan * sanitizing_std + bSoc * social_exposure_std,
  a ~ dnorm(0, 0.2), bEng ~ dnorm(0, 1), bSan ~ dnorm(0, 1),
  bSoc ~ dnorm(0, 1), sigma ~ dexp(1) # OBS : why exp? It doesn't change anything changing it to dno
),
  data = df)

precis(m2)
```

```
##           mean          sd      5.5%      94.5%
## a      -2.964072e-09 0.011210641 -0.017916773 0.017916767
## bEng    1.442049e-02 0.013519283 -0.007185931 0.036026920
## bSan    8.776202e-01 0.036986486  0.818508687 0.936731784
## bSoc   -5.207536e-02 0.035014376 -0.108035094 0.003884377
## sigma   3.550698e-01 0.007937291  0.342384522 0.367755169
```

```
plot(coeftab(m2), by.model = TRUE)
```



The plot above shows the posterior mean for our model's parameters σ , $bSoc$, $bSan$, $bEng$ and our intercept, α (a). The posterior means are shown by the points and the 89% compatibility intervals are shown by the solid horizontal lines. We can interpret the posterior mean of the parameter $bEng$ as: once we know sanitizing and social exposure, there is no additional predictive power in also knowing the person's energy level.

The posterior mean for energy level (b_{Eng}) is very close to 0 (.01) with very little probability of both sides of zero ($SD = .01$). This does not violate the implication that energy level becomes independent from OCD when stratifying by germ level by energy level, social exposure and sanitizing.

Model 3: OCD // enrg | crn_, sntz

Testing the implication that the causal coefficient of energy_level becomes significantly close to zero when stratifying by sanitizing and corona fear in a model that has OCD as outcome variable.

```
adjustmentSets(dag_corona, exposure = "energylvl", outcome = "OCD") # No additional to add
```

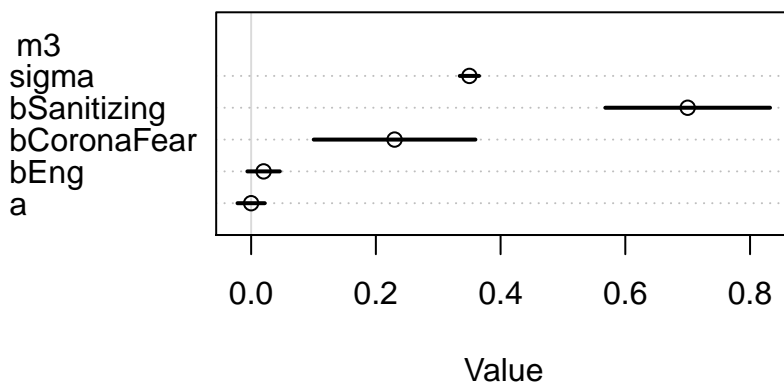
```
## { corona_fear, sanitizing }
## { sanitizing, social_exposure }
## { germ_level }
```

```
m3 <- quap(alist(OCD_std ~ dnorm(mu, sigma), mu <- a + bEng *
  energylvl_std + bCoronaFear * corona_fear_std + bSanitizing *
  sanitizing_std, a ~ dnorm(0, 0.2), bEng ~ dnorm(0, 1), bCoronaFear ~
  dnorm(0, 1), bSanitizing ~ dnorm(0, 1), sigma ~ dexp(1)),
  data = df)
```

```
precis(m3)
```

##		mean	sd	5.5%	94.5%
## a		4.788040e-07	0.011158935	-0.017833654	0.01783461
## bEng		1.620590e-02	0.013285733	-0.005027271	0.03743906
## bCoronaFear		2.270883e-01	0.066159223	0.121353062	0.33282350
## bSanitizing		7.021110e-01	0.067397412	0.594396927	0.80982509
## sigma		3.534271e-01	0.007900576	0.340800407	0.36605370

```
plot(coeftab(m3), by.model = TRUE)
```



As seen in the above plot, energy level becomes independent from OCF (i.e., the model estimates the coefficient b_{Eng} to be very close to zero (-.03)) when stratifying by energy level, corona fear and sanitizing.

The plot above shows the posterior mean for our model's parameters σ , $b_{\text{Sanitizing}}$, $b_{\text{CoronaFear}}$, b_{Eng} and our intercept, α (a). The posterior means are shown by the points and the 89% compatibility intervals are shown by the solid horizontal lines. We can interpret the posterior mean of the parameter b_{Eng} as: once we know sanitizing and corona fear, there is very little additional predictive power in also knowing the person's energy level.

The posterior mean for energy level (b_{Eng}) is very close to 0 (-.03). This does not violate the implication that energy level becomes independent from OCD when stratifying by germ level by energy level, social exposure and sanitizing.

Model 4: OCD // grm_ | sntz, scl_

Testing the implication that the causal coefficient of germ level becomes significantly close to zero when stratifying by sanitizing and social exposure in a model that has OCD as outcome variable.

```
adjustmentSets(dag_corona, exposure = "germ_level", outcome = "OCD") # No additional to add
```

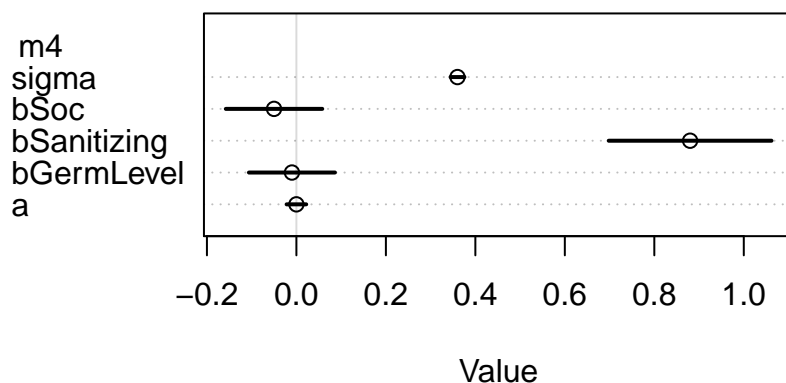
```
## { corona_fear, sanitizing }
## { sanitizing, social_exposure }
```

```
m4 <- quap(alist(OCD_std ~ dnorm(mu, sigma), mu <- a + bGermLevel *
  germ_std + bSanitizing * sanitizing_std + bSoc * social_exposure_std,
  a ~ dnorm(0, 0.2), bGermLevel ~ dnorm(0, 1), bSanitizing ~
    dnorm(0, 1), bSoc ~ dnorm(0, 1), sigma ~ dexp(1)), data = df)
```

```
precis(m4)
```

##		mean	sd	5.5%	94.5%
##	a	-6.627457e-08	0.011216862	-0.01792678	0.01792665
##	bGermLevel	-5.735589e-03	0.049062759	-0.08414735	0.07267618
##	bSanitizing	8.823438e-01	0.092858261	0.73393838	1.03074925
##	bSoc	-4.950360e-02	0.055035147	-0.13746040	0.03845319
##	sigma	3.552675e-01	0.007941738	0.34257507	0.36795993

```
plot(coeftab(m4), by.model = TRUE)
```



As seen in the above plot, germ level becomes independent from OCD (i.e., the model estimates the coefficient $b_{\text{GermLevel}}$ to be very close to zero (.05)) when stratifying by sanitizing, and social exposure.

The plot above shows the posterior mean for our model's parameters σ , b_{Soc} , $b_{\text{Sanitizing}}$, $b_{\text{GermLevel}}$ and our intercept, α (a). The posterior means are shown by the points and the 89% compatibility intervals are shown by the solid horizontal lines. We can interpret the posterior mean of the parameter $b_{\text{GermLevel}}$ as: once we know sanitizing and social exposure, there is very little additional predictive power in also knowing the person's germ level.

The posterior mean for germ level ($b_{\text{GermLevel}}$) is very close to 0 (.05 with very little probability of both sides of zero. This does not violate the implication that germ level becomes independent from OCD when stratifying by sanitizing and social exposure.

Model 5: OCD || grm__ | crn__, sntz

Testing the implication that causal coefficient of germ level becomes significantly close to zero when stratifying by corona fear and sanitizing in a model that has OCD as outcome variable.

```
adjustmentSets(dag_corona, exposure = "germ_level", outcome = "OCD") # No additional to add
```

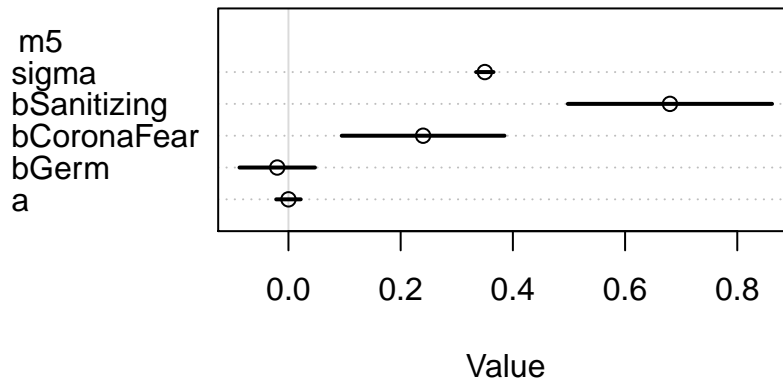
```
## { corona_fear, sanitizing }
## { sanitizing, social_exposure }
```

```
m5 <- quap(alist(OCD_std ~ dnorm(mu, sigma), mu <- a + bGerm *
  germ_std + bCoronaFear * corona_fear_std + bSanitizing *
  sanitizing_std, a ~ dnorm(0, 0.2), bGerm ~ dnorm(0, 1), bCoronaFear ~
  dnorm(0, 1), bSanitizing ~ dnorm(0, 1), sigma ~ dexp(1)),
  data = df)
```

```
precis(m5)
```

##		mean	sd	5.5%	94.5%
## a		-8.124094e-07	0.01116498	-0.01784460	0.01784297
## bGerm		-2.248399e-02	0.03432662	-0.07734456	0.03237657
## bCoronaFear		2.391195e-01	0.07403326	0.12080008	0.35743896
## bSanitizing		6.779187e-01	0.09284279	0.52953798	0.82629941
## sigma		3.536190e-01	0.00790488	0.34098544	0.36625249

```
plot(coeftab(m5), by.model = TRUE)
```

As seen in the above plot, germ level becomes independent from OCD (i.e., the model estimates the coefficient b_{Germ} to be very close to zero (-.02) when stratifying by sanitizing, and corona fear.

The plot above shows the posterior mean for our model's parameters σ , $b_{\text{CoronaFear}}$, $b_{\text{Sanitizing}}$, b_{Germ} and our intercept, α (a). The posterior means are shown by the points and the 89% compatibility intervals are shown by the solid horizontal lines. We can interpret the posterior mean of the parameter b_{Germ} as: once we know sanitizing and corona fear, there is very little additional predictive power in also knowing the person's germ level.

The posterior mean for germ level (b_{Germ}) is very close to 0 (-.02) with very little probability of both sides of zero. This does not violate the implication that germ level becomes independent from OCD when stratifying by sanitizing and corona fear.

Model 6: OCD || scl_ | crn_

Testing the implication that the causal coefficient of social exposure becomes significantly close to zero when stratifying by corona fear in a model that has OCD as outcome variable.

```
adjustmentSets(dag_corona, exposure = "social_exposure", outcome = "OCD") # No additional to add
```

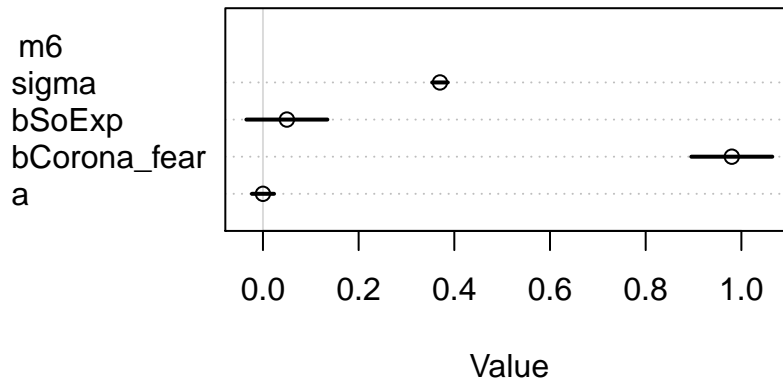
```
## { corona_fear }
```

```
m6 <- quap(alist(OCD_std ~ dnorm(mu, sigma), mu <- a + bCorona_fear *
  corona_fear_std + bSoExp * social_exposure_std, a ~ dnorm(0,
  0.2), bCorona_fear ~ dnorm(0, 1), bSoExp ~ dnorm(0, 1), sigma ~
  dexp(1)), data = df)
```

```
precis(m6)
```

	mean	sd	5.5%	94.5%
## a	4.882448e-07	0.011807503	-0.01887018	0.01887116
## bCorona_fear	9.782469e-01	0.043205535	0.90919614	1.04729772
## bSoExp	5.322389e-02	0.043205529	-0.01582689	0.12227467
## sigma	3.740384e-01	0.008361211	0.36067561	0.38740127

```
plot(coeftab(m6), by.model = TRUE)
```



The plot above shows the posterior mean for our model's parameters σ , $bSoExp$, $bCorona_fear$ and our intercept, α (a). The posterior means are shown by the points and the 89% compatibility intervals are shown by the solid horizontal lines. We can interpret the posterior mean of the parameter $bSoExp$ as: once we know a person's corona fear, there is very little additional predictive power in also knowing the person's social exposure.

The posterior mean for social exposure ($bSoExp$) is very close to 0 (.05) with very little probability of both sides of zero. This does not violate the implication that social exposure becomes independent from OCD when stratifying by corona fear.

Model 7: crn__ || enrg | grm__

Testing the implication that the causal coefficient energy level becomes significantly close to zero when stratifying by germ level in a model that has corona fear as outcome variable.

```
adjustmentSets(dag_corona, exposure = "energylvl", outcome = "corona_fear") # No additional to add
```

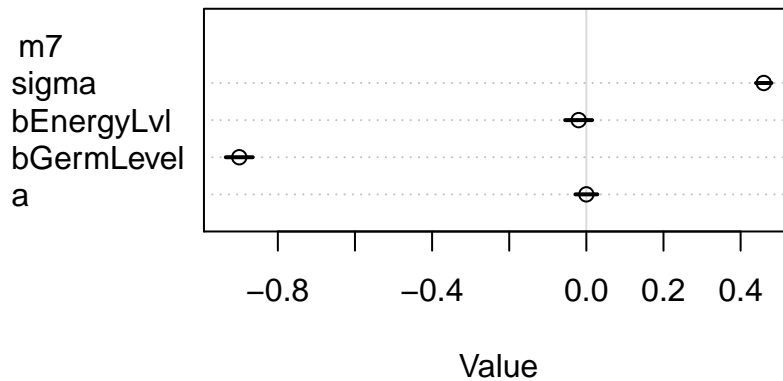
```
## { sanitizing, social_exposure }
## { germ_level }
```

```
m7 <- quap(alist(corona_fear_std ~ dnorm(mu, sigma), mu <- a +
  bGermLevel * germ_std + bEnergyLvl * energylvl_std, a ~ dnorm(0,
  0.2), bGermLevel ~ dnorm(0, 1), bEnergyLvl ~ dnorm(0, 1),
  sigma ~ dexp(1)), data = df)
```

```
precis(m7)
```

```
##               mean          sd      5.5%      94.5%
## a      -2.219322e-05 0.01451490 -0.02321980 0.02317541
## bGermLevel -8.965607e-01 0.01784548 -0.92508118 -0.86804013
## bEnergyLvl -1.602263e-02 0.01784548 -0.04454315 0.01249789
## sigma      4.602149e-01 0.01028696 0.44377435 0.47665545
```

```
plot(coeftab(m7), by.model = TRUE)
```



The plot above shows the posterior mean for our model's parameters *sigma*, *bGermLvl*, *bEnergyLvl* and our intercept, *alpha* (*a*). The posterior means are shown by the points and the 89% compatibility intervals are shown by the solid horizontal lines. We can interpret the posterior mean of the parameter *bEnergyLvl* as: once we know a person's germ level, there is no additional predictive power in also knowing the person's energy level.

The posterior mean for energy level (*bEnergyLvl*) is -0.2 with very little probability of both sides of zero. This does not violate the implication that energy level becomes independent from corona fear when stratifying by germ level.

Model 8: `crn__ || enrg | sntz, scl__`

Testing the implication that the causal coefficient energy level becomes significantly close to zero when stratifying by sanitizing and social exposure in a model that has corona fear as outcome variable.

```
adjustmentSets(dag_corona, exposure = "energylvl", outcome = "corona_fear") # No additional to add
```

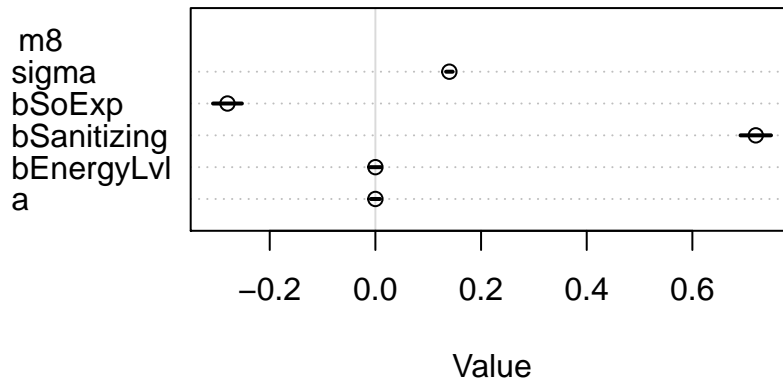
```
## { sanitizing, social_exposure }
## { germ_level }
```

```
m8 <- quap(alist(corona_fear_std ~ dnorm(mu, sigma), mu <- a +
  bEnergyLvl * energylvl_std + bSanitizing * sanitizing_std +
  bSoExp * social_exposure_std, a ~ dnorm(0, 0.2), bEnergyLvl ~
  dnorm(0, 1), bSanitizing ~ dnorm(0, 1), bSoExp ~ dnorm(0,
  1), sigma ~ dexp(1)), data = df)
```

```
precis(m8)
```

	mean	sd	5.5%	94.5%
## a	4.787490e-08	0.004468553	-0.007141563	0.007141659
## bEnergyLvl	-3.562787e-03	0.005382614	-0.012165243	0.005039669
## bSanitizing	7.183677e-01	0.014738551	0.694812599	0.741922702
## bSoExp	-2.846796e-01	0.013952484	-0.306978329	-0.262380802
## sigma	1.413433e-01	0.003159672	0.136293581	0.146393115

```
plot(coeftab(m8), by.model = TRUE)
```



The plot above shows the posterior mean for our model's parameters *sigma*, *bSoExp*, *bEnergyLvl*, *bSanitizing* and our intercept, *alpha* (*a*). The posterior means are shown by the points and the 89% compatibility intervals are shown by the solid horizontal lines. We can interpret the posterior mean of the parameter *bEnergyLvl* as: once we know a person's sanitizing and social exposure level, there is no additional predictive power in also knowing the person's energy level.

The posterior mean for energy level (*bEnergy*) is 0 with very little probability of both sides of zero. This does not violate the implication that energy level becomes independent from corona fear when stratifying by social exposure and sanitizing

Model 9: `crn_ // grm_ | sntz, scl_`

Testing the implication that the causal coefficient for germ level becomes significantly close to zero when stratifying by sanitizing and social exposure in a model that has corona fear as outcome variable.

```
adjustmentSets(dag_corona, exposure = "germ_level", outcome = "corona_fear") # No additional to add
```

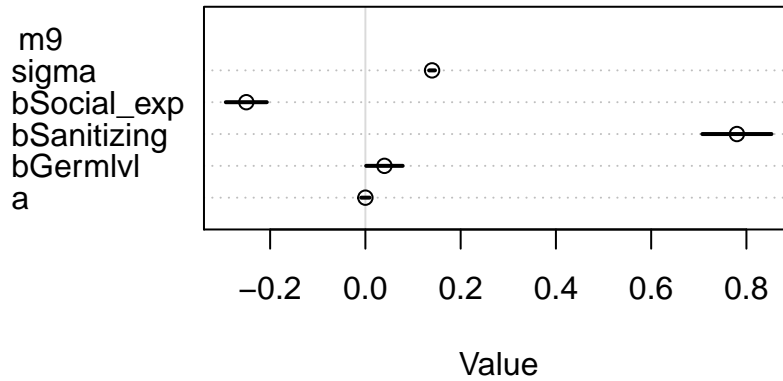
```
## { sanitizing, social_exposure }
```

```
m9 <- quap(alist(corona_fear_std ~ dnorm(mu, sigma), mu <- a +
  bGermlvl * germ_std + bSanitizing * sanitizing_std + bSocial_exp *
  social_exposure_std, a ~ dnorm(0, 0.2), bGermlvl ~ dnorm(0,
  1), bSanitizing ~ dnorm(0, 1), bSocial_exp ~ dnorm(0, 1),
  sigma ~ dexp(1)), data = df)
```

```
precis(m9)
```

	mean	sd	5.5%	94.5%
## a	5.906975e-06	0.004460488	-0.007122814	0.007134628
## bGermlvl	3.934401e-02	0.019581202	0.008049467	0.070638551
## bSanitizing	7.839733e-01	0.037086701	0.724701596	0.843245017
## bSocial_exp	-2.519762e-01	0.021969286	-0.287087376	-0.216865051
## sigma	1.410881e-01	0.003154109	0.136047230	0.146128981

```
plot(coeftab(m9), by.model = TRUE)
```



The plot above shows the posterior mean for our model's parameters σ , $b_{\text{Social_exp}}$, $b_{\text{Sanitizing}}$, b_{GermLvl} , and our intercept, α (a). The posterior means are shown by the points and the 89% compatibility intervals are shown by the solid horizontal lines. We can interpret the posterior mean of the parameter b_{GermLvl} as: once we know a person's sanitizing and social exposure level, there is no additional predictive power in also knowing the person's germ level.

The posterior mean for germ level (b_{GermLvl}) is 0.04 with very little probability of both sides of zero. This does not violate the implication that germ level becomes independent from corona fear when stratifying by sanitizing and social exposure

Model 10: `enrg // sntz | grm__`

Testing the implication that the causal coefficient for sanitizing becomes significantly close to zero when stratifying by germ level in a model that has energy level as outcome variable.

```
adjustmentSets(dag_corona, exposure = "sanitizing", outcome = "energylvl") # Adding { Corona Fear }
```

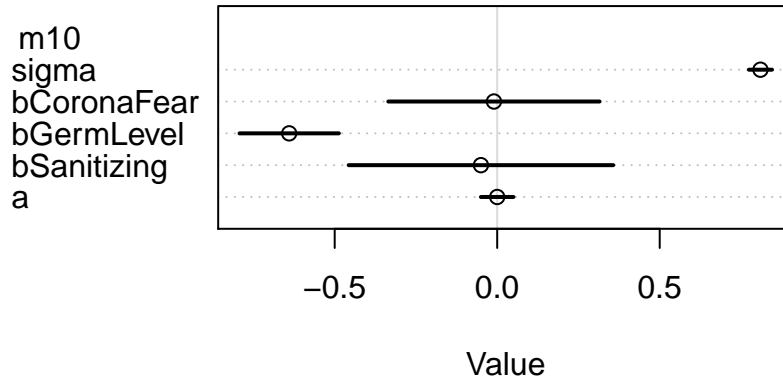
```
## { social_exposure }
## { corona_fear }
```

```
m10 <- quap(alist(energylvl_std ~ dnorm(mu, sigma), mu <- a +
  bSanitizing * sanitizing_std + bGermLevel * germ_std + bCoronaFear *
  corona_fear_std, a ~ dnorm(0, 0.2), bSanitizing ~ dnorm(0,
  1), bGermLevel ~ dnorm(0, 1), bCoronaFear ~ dnorm(0, 1),
  sigma ~ dexp(1)), data = df)
```

```
precis(m10)
```

	mean	sd	5.5%	94.5%
## a	-1.241215e-07	0.02555019	-0.04083426	0.04083401
## bSanitizing	-5.256210e-02	0.20767386	-0.38446505	0.27934084
## bGermLevel	-6.351185e-01	0.07799783	-0.75977405	-0.51046285
## bCoronaFear	-9.294399e-03	0.16583924	-0.27433753	0.25574874
## sigma	8.146428e-01	0.01820475	0.78554806	0.84373746

```
plot(coeftab(m10), by.model = TRUE)
```



The plot above shows the posterior mean for our model's parameters σ , $b_{\text{CoronaFear}}$, $b_{\text{GermLevel}}$, $b_{\text{Sanitizing}}$ and our intercept, a . The posterior means are shown by the points and the 89% compatibility intervals are shown by the solid horizontal lines. We can interpret the posterior mean of the parameter $b_{\text{Sanitizing}}$ as: once we know a person's germ level and corona fear, there is no additional predictive power in also knowing the person's sanitizing level.

The posterior mean for sanitizing ($b_{\text{Sanitizing}}$) is -0.06 with very little probability of both sides of zero. This does not violate the implication that sanitizing becomes independent from energy level when stratifying by germ level and corona fear.

Model 11: `enrg // scl_ | grm_`

Testing the implication that the causal coefficient for social exposure becomes significantly close to zero when stratifying by germ level in a model that has energy level as outcome variable.

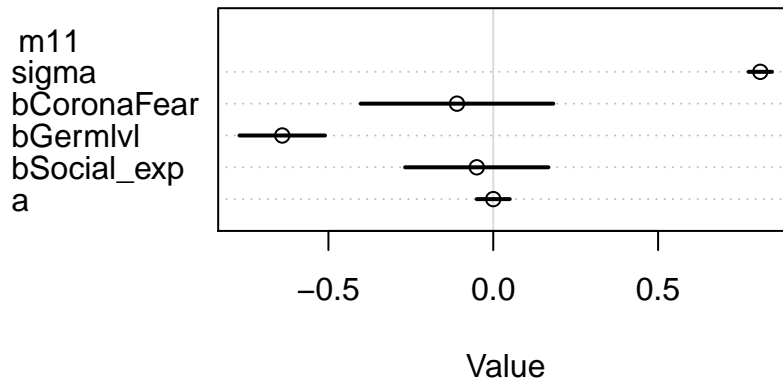
```
adjustmentSets(dag_corona, exposure = "social_exposure", outcome = "energylvl") # Adding { Corona Fear
```

```
## { sanitizing }
## { corona_fear }
```

```
m11 <- quap(alist(energylvl_std ~ dnorm(mu, sigma), mu <- a +
  bSocial_exp * social_exposure_std + bGermlvl * germ_std +
  bCoronaFear * corona_fear_std, a ~ dnorm(0, 0.2), bSocial_exp ~
  dnorm(0, 1), bGermlvl ~ dnorm(0, 1), bCoronaFear ~ dnorm(0,
  1), sigma ~ dexp(1)), data = df)
precis(m11)
```

	mean	sd	5.5%	94.5%
## a	4.735404e-07	0.02554846	-0.0408309	0.04083184
## bSocial_exp	-4.896199e-02	0.11091812	-0.2262306	0.12830659
## bGermlvl	-6.370576e-01	0.06619602	-0.7428516	-0.53126357
## bCoronaFear	-1.098451e-01	0.14906410	-0.3480783	0.12838815
## sigma	8.145867e-01	0.01820360	0.7854939	0.84367961

```
plot(coeftab(m11), by.model = TRUE)
```



The plot above shows the posterior mean for our model's parameters σ , $b_{\text{Social_exp}}$, b_{Germlvl} , $b_{\text{CoronaFear}}$ and our intercept, α (a). The posterior means are shown by the points and the 89% compatibility intervals are shown by the solid horizontal lines. We can interpret the posterior mean of the parameter $b_{\text{Social_exp}}$ as: once we know a person's germ level and corona fear, there is no additional predictive power in also knowing the person's social exposure level.

The posterior mean for social exposure ($b_{\text{Social_exp}}$) is 0.03 with very little probability of both sides of zero. This does not violate the implication that social exposure becomes independent from energy level when stratifying by germ level and corona fear.

Model 12: $\text{sntz} \parallel \text{scl_} \mid \text{crn_}$

Testing the implication that the causal coefficient for social exposure becomes significantly close to zero when stratifying by corona fear in a model that has sanitizing as outcome variable.

```
adjustmentSets(dag_corona, exposure = "social_exposure", outcome = "sanitizing") # No additional to ad
```

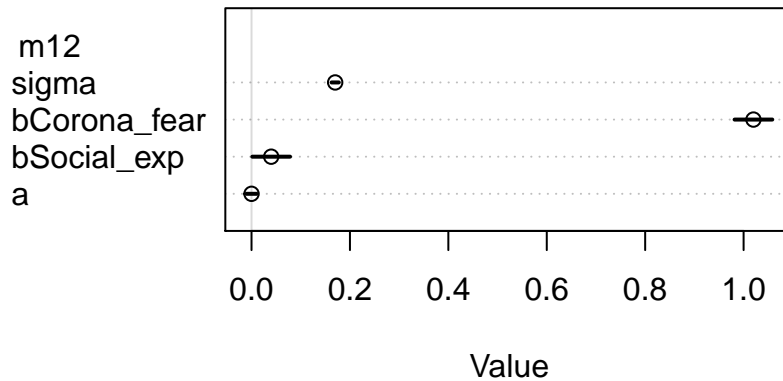
```
## { corona_fear }
```

```
m12 <- quap(alist(sanitizing_std ~ dnorm(mu, sigma), mu <- a +
  bSocial_exp * social_exposure_std + bCorona_fear * corona_fear_std,
  a ~ dnorm(0, 0.2), bSocial_exp ~ dnorm(0, 1), bCorona_fear ~
    dnorm(0, 1), sigma ~ dexp(1)), data = df)
```

```
precis(m12)
```

```
##               mean          sd      5.5%      94.5%
## a      3.168120e-07 0.005350296 -0.00855049 0.008551124
## bSocial_exp 4.014582e-02 0.019578405 0.00885575 0.071435894
## bCorona_fear 1.024078e+00 0.019578405 0.99278770 1.055367850
## sigma      1.692518e-01 0.003783559 0.16320495 0.175298660
```

```
plot(coeftab(m12), by.model = TRUE)
```



The plot above shows the posterior mean for our model's parameters *sigma*, *bSocial_exp*, *bCorona_fear* and our intercept, *alpha* (*a*). The posterior means are shown by the points and the 89% compatibility intervals are shown by the solid horizontal lines. We can interpret the posterior mean of the parameter *bSocial_exp* as: once we know a person's corona fear, there is no additional predictive power in also knowing the person's social exposure.

The posterior mean for social exposure (*bSocial_exp*) is 0.04 with a standard deviation of 0.02. This does not violate the implication that social exposure becomes independent from sanitizing when stratifying by corona fear.

Task 4: Messing it up

- Try and **deliberately have an open back door path** and see if you can get wrong inference.

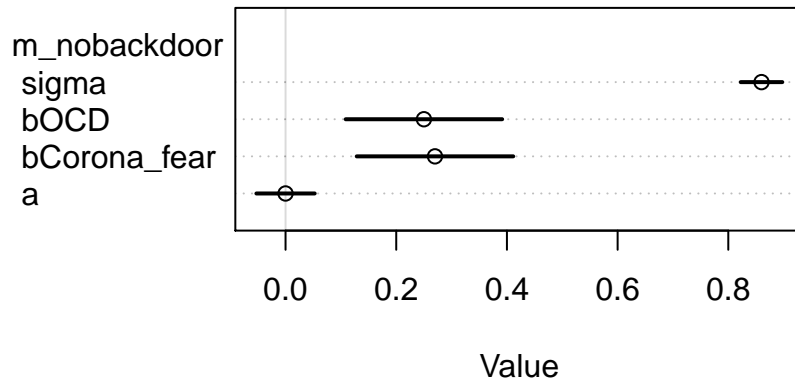
We know that we will get wrong inference if we don't include *OCD* in our model. Not including *OCD* in the model results in a different estimate of the level of corona fear's influence on energy level. Contrarily, including *OCD* in the model and thus closing the backdoor will allow us to estimate the total causal effect that corona fear has on energy level.

```
# Model with closed back door
m_nobackdoor <- quap(alist(energylvl_std ~ dnorm(mu, sigma),
  mu <- a + bCorona_fear * corona_fear_std + bOCD * OCD_std,
  a ~ dnorm(0, 0.2), bCorona_fear ~ dnorm(0, 1), bOCD ~ dnorm(0,
    1), sigma ~ dexp(1)), data = df)

precis(m_nobackdoor)
```

	mean	sd	5.5%	94.5%
## a	7.253990e-07	0.02689977	-0.04299029	0.04299175
## bCorona_fear	2.662537e-01	0.07215300	0.15093928	0.38156814
## bOCD	2.546833e-01	0.07215300	0.13936888	0.36999774
## sigma	8.584453e-01	0.01918298	0.82778720	0.88910341


```
plot(coeftab(m_nobackdoor), by.model = TRUE) # bCorona_fear coef around 0.3
```

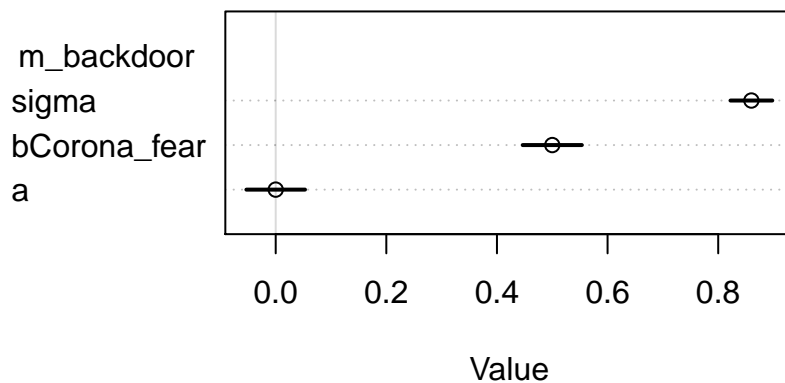


```
# Model with open back door
m_backdoor <- quap(alist(energylvl_std ~ dnorm(mu, sigma), mu <- a +
  bCorona_fear * corona_fear_std, a ~ dnorm(0, 0.2), bCorona_fear ~
  dnorm(0, 1), sigma ~ dexp(1)), data = df)
```

```
precis(m_backdoor)
```

```
##               mean          sd      5.5%      94.5%
## a      1.345068e-07 0.02706203 -0.04325022 0.04325049
## bCorona_fear 5.022153e-01 0.02731670 0.45855797 0.54587268
## sigma      8.637199e-01 0.01930077 0.83287355 0.89456626
```

```
plot(coeftab(m_backdoor), by.model = TRUE) # bCorona_fear coef around 0.5
```



When comparing the plots of the two models above, it becomes clear that having an open back door path makes the causal effect that corona fear has on energy level appear stronger than it actually is. Closing the backdoor path results in an estimated posterior mean of the corona fear coefficient around 0.3. Having an open backdoor path results in an estimated posterior mean of apx. 0.5, thus making it look as if the effect is bigger than it actually is. This is due to OCD being a confound that affects energy level through a front door path via corona fear, but also as a back door path via sanitizing.

- Try and deliberately **simulate some data that doesn't fit the DAG**, or **create a new DAG that doesn't fit the data**.

We'll now simulate some data that doesn't fit our DAG. Then we'll run the exact same tests of conditional independencies, now using the wrong data. We expect that several of the implications of our DAG will be violated as a result of our data simply not fitting the DAG. We'll run the exact same models as above, but only include the code output of those violated.

- Use the same approach as above to show that the DAG is wrong (by showing that conditional independencies don't exist in the data, for example).

```
# Simulating variables:
N <- 10000

# OCD
OCD_w <- rnorm(N, mean = 5, sd = 0.001)

# Corona fear
corona_fear_w <- rnorm(N, mean = 50, sd = 10)

# Social exposure
social_exposure_w <- rnorm(N, mean = (35 + (energylvl * 3)),
  sd = 1)

# Sanitizing
sanitizing_w <- rnorm(N, mean = (social_exposure_w/15), sd = 2)

# Energy level
energylvl_w <- rnorm(N, mean = 50 - ((social_exposure_w/100)),
  sd = 1)

# Germ level
germ_level_w <- rnorm(N, mean = 4600 - ((energylvl_w * 15)),
  sd = 100)

# creating df
df_w <- tibble(germ_level_w, sanitizing_w, energylvl_w, social_exposure_w,
  corona_fear_w, OCD_w)
```

Standardizing variables.

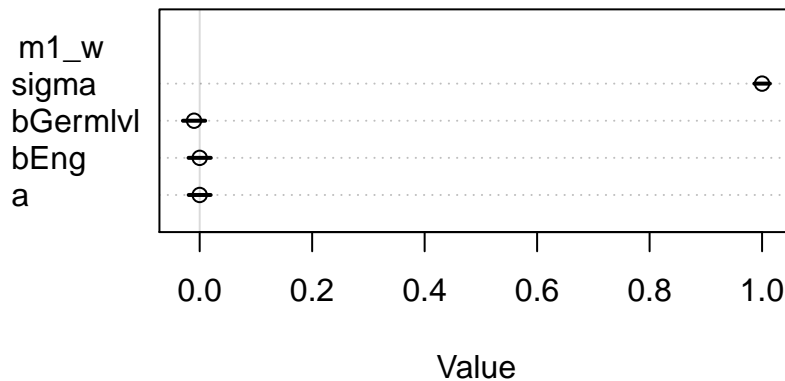
```
df_w <- df_w %>%
  mutate(germ_std_w = scale(germ_level_w), sanitizing_std_w = scale(sanitizing_w),
    energylvl_std_w = scale(energylvl_w), social_exposure_std_w = scale(social_exposure_w),
    corona_fear_std_w = scale(corona_fear_w), OCD_std_w = scale(OCD_w))
```

Wrong model 1: OCD // enrg | germ

```
adjustmentSets(dag_corona, exposure = "OCD", outcome = "energylvl")

m1_w <- quap(alist(OCD_std_w ~ dnorm(mu, sigma), mu <- a + bEng *
  energylvl_std_w + bGermlvl * germ_std_w, a ~ dnorm(0, 0.2),
  bEng ~ dnorm(0, 1), bGermlvl ~ dnorm(0, 1), sigma ~ dexp(1)),
  data = df_w)

plot(coeftab(m1_w), by.model = TRUE)
```

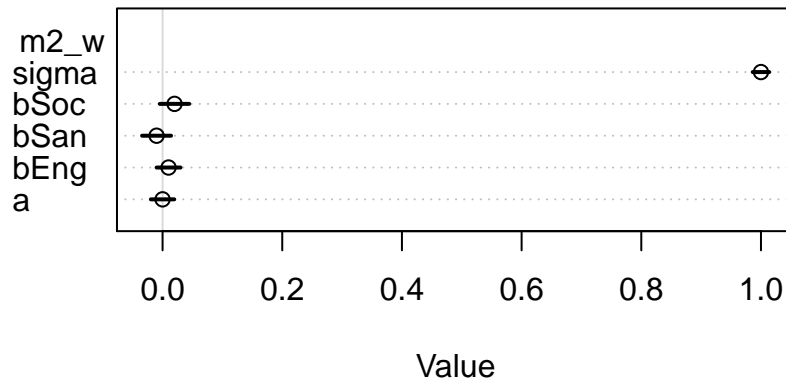


Wrong model 2: OCD // enrg | sntz, scl__

```
adjustmentSets(dag_corona, exposure = "energylvl", outcome = "OCD") # choosing { sanitizing, social_exposure }

m2_w <- quap(alist(OCD_std_w ~ dnorm(mu, sigma), mu <- a + bEng *
  energylvl_std_w + bSan * sanitizing_std_w + bSoc * social_exposure_std_w,
  a ~ dnorm(0, 0.2), bEng ~ dnorm(0, 1), bSan ~ dnorm(0, 1),
  bSoc ~ dnorm(0, 1), sigma ~ dexp(1) # OBS : why exp? It doesn't change anything changing it to dnorm(0, 1)
),
  data = df_w)

plot(coeftab(m2_w), by.model = TRUE)
```



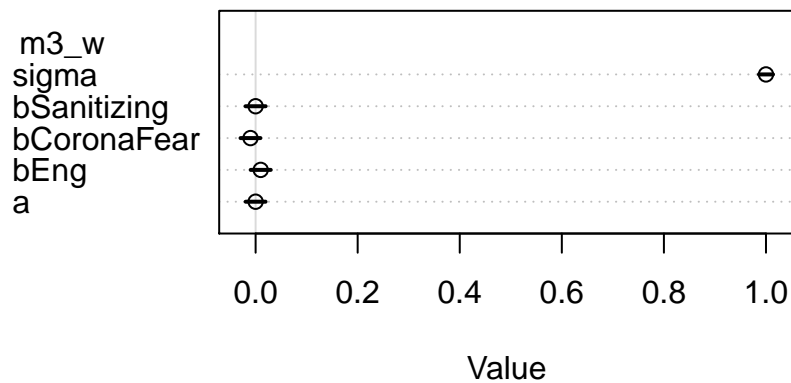
```
precis(m2_w)
```

Wrong model 3: OCD // enrg | crn_, sntz

```
adjustmentSets(dag_corona, exposure = "energylvl", outcome = "OCD") # { corona_fear, sanitizing }

m3_w <- quap(alist(OCD_std_w ~ dnorm(mu, sigma), mu <- a + bEng *
  energylvl_std_w + bCoronaFear * corona_fear_std_w + bSanitizing *
  sanitizing_std_w, a ~ dnorm(0, 0.2), bEng ~ dnorm(0, 1),
  bCoronaFear ~ dnorm(0, 1), bSanitizing ~ dnorm(0, 1), sigma ~
    dexp(1)), data = df_w)

plot(coeftab(m3_w), by.model = TRUE)
```



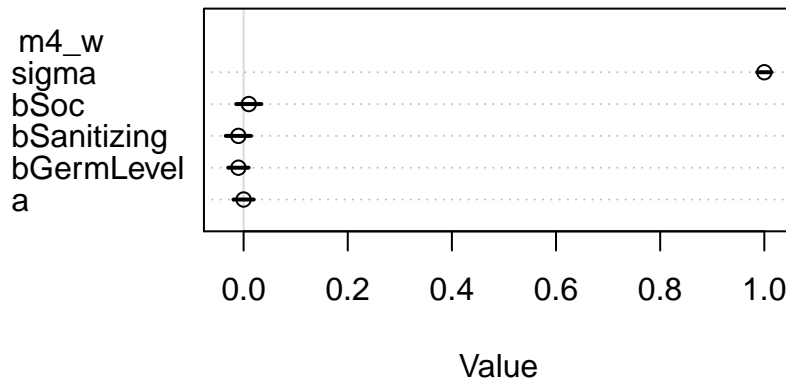
Wrong model 4: OCD // grm_ | sntz, scl_

```

m4_w <- quap(alist(OCD_std_w ~ dnorm(mu, sigma), mu <- a + bGermLevel *
  germ_std_w + bSanitizing * sanitizing_std_w + bSoc * social_exposure_std_w,
  a ~ dnorm(0, 0.2), bGermLevel ~ dnorm(0, 1), bSanitizing ~
    dnorm(0, 1), bSoc ~ dnorm(0, 1), sigma ~ dexp(1)), data = df_w)

plot(coeftab(m4_w), by.model = TRUE)

```



Wrong Model 5: OCD || grm_ | crn_, sntz

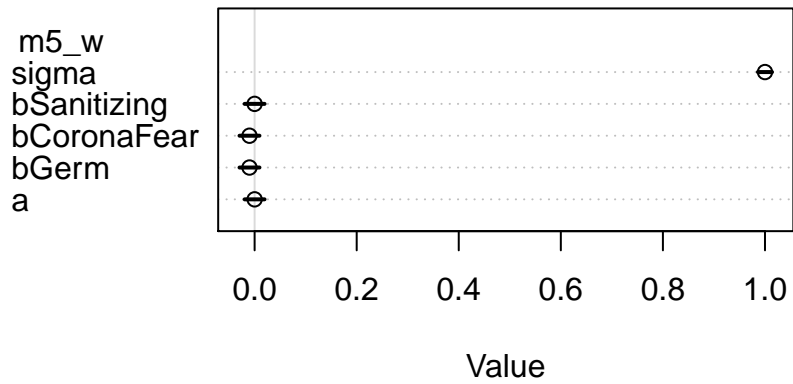
```

adjustmentSets(dag_corona, exposure = "germ_level", outcome = "OCD") # OCD _||_ grm_ | crn_, sntz

m5_w <- quap(alist(OCD_std_w ~ dnorm(mu, sigma), mu <- a + bGerm *
  germ_std_w + bCoronaFear * corona_fear_std_w + bSanitizing *
  sanitizing_std_w, a ~ dnorm(0, 0.2), bGerm ~ dnorm(0, 1),
  bCoronaFear ~ dnorm(0, 1), bSanitizing ~ dnorm(0, 1), sigma ~
    dexp(1)), data = df_w)

plot(coeftab(m5_w), by.model = TRUE)

```

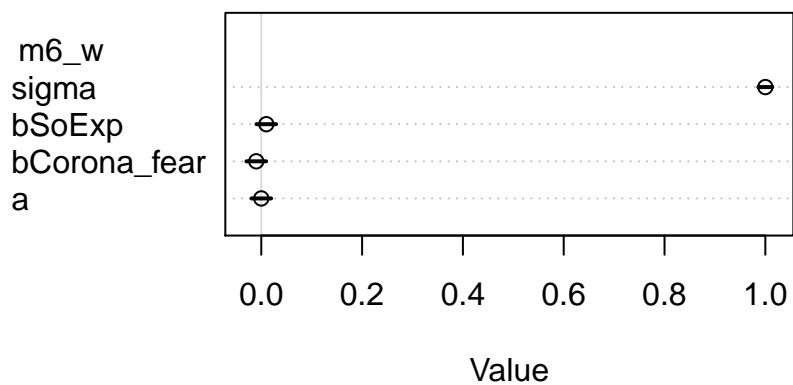


Wrong Model 6:OCD || scl_ | crn_

```
adjustmentSets(dag_corona, exposure = "social_exposure", outcome = "OCD") #OCD _||_ scl_ / crn_

m6_w <- quap(alist(OCD_std_w ~ dnorm(mu, sigma), mu <- a + bCorona_fear *
  corona_fear_std_w + bSoExp * social_exposure_std_w, a ~ dnorm(0,
  0.2), bCorona_fear ~ dnorm(0, 1), bSoExp ~ dnorm(0, 1), sigma ~
  dexp(1)), data = df_w)

plot(coeftab(m6_w), by.model = TRUE)
```



Wrong Model 7: crn_ || enrg | grm_

```
adjustmentSets(dag_corona, exposure = "energylvl", outcome = "corona_fear") #crn_ _||_ enrg / grm_

m7_w <- quap(alist(corona_fear_std_w ~ dnorm(mu, sigma), mu <- a +
```

```

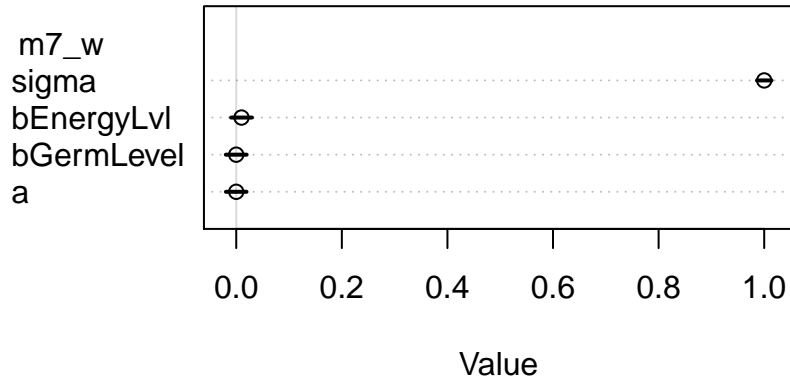
bGermLevel * germ_std_w + bEnergyLvl * energylvl_std_w, a ~
dnorm(0, 0.2), bGermLevel ~ dnorm(0, 1), bEnergyLvl ~ dnorm(0,
1), sigma ~ dexp(1)), data = df_w)

```

```

plot(coeftab(m7_w), by.model = TRUE)

```



WrongModel 8: $crn_ || \text{ enrg } | \text{ sntz, scl_}$

```

adjustmentSets(dag_corona, exposure = "energylvl", outcome = "corona_fear") #crn_ ||_ enrg | grm_

```

```

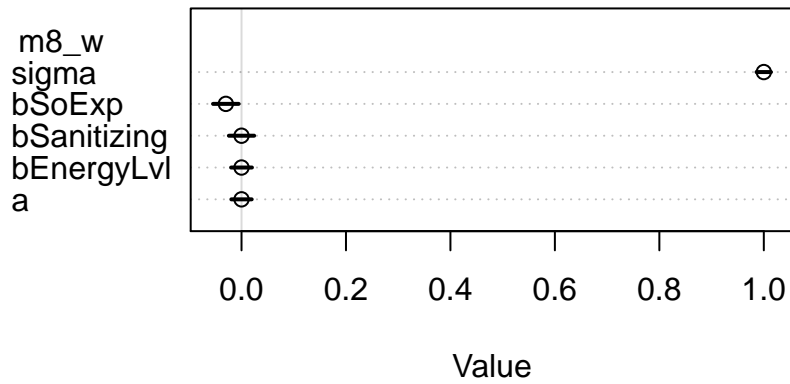
m8_w <- quap(alist(corona_fear_std_w ~ dnorm(mu, sigma), mu <- a +
bEnergyLvl * energylvl_std_w + bSanitizing * sanitizing_std_w +
bSoExp * social_exposure_std_w, a ~ dnorm(0, 0.2), bEnergyLvl ~
dnorm(0, 1), bSanitizing ~ dnorm(0, 1), bSoExp ~ dnorm(0,
1), sigma ~ dexp(1)), data = df_w)

```

```

plot(coeftab(m8_w), by.model = TRUE)

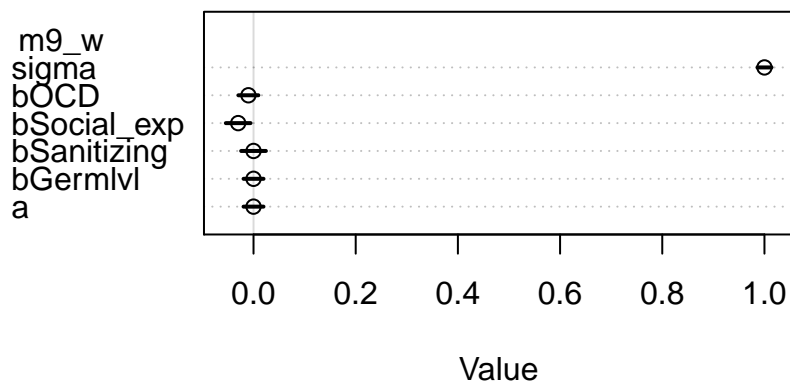
```



Wrong model 9: crn_ // grm_ | sntz, scl_

```
m9_w <- quap(alist(corona_fear_std_w ~ dnorm(mu, sigma), mu <- a +
  bGermlvl * germ_std_w + bSanitizing * sanitizing_std_w +
  bSocial_exp * social_exposure_std_w + bOCD * OCD_std_w, a ~
  dnorm(0, 0.2), bGermlvl ~ dnorm(0, 1), bSanitizing ~ dnorm(0,
  1), bSocial_exp ~ dnorm(0, 1), bOCD ~ dnorm(0, 1), sigma ~
  dexp(1)), data = df_w)

plot(coeftab(m9_w), by.model = TRUE)
```

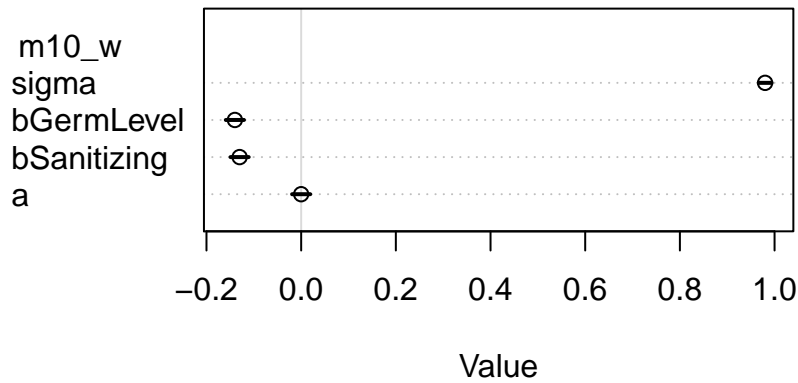


Wrong Model 10: enrg // sntz | grm_

Testing whether the causal coefficient for sanitizing becomes significantly close to zero when stratifying by germ level in a model that has energy level as outcome variable.


```
m10_w <- quap(alist(energylvl_std_w ~ dnorm(mu, sigma), mu <- a +
  bSanitizing * sanitizing_std_w + bGermLevel * germ_std_w,
  a ~ dnorm(0, 0.2), bSanitizing ~ dnorm(0, 1), bGermLevel ~
    dnorm(0, 1), sigma ~ dexp(1)), data = df_w)

plot(coeftab(m10_w), by.model = TRUE)
```



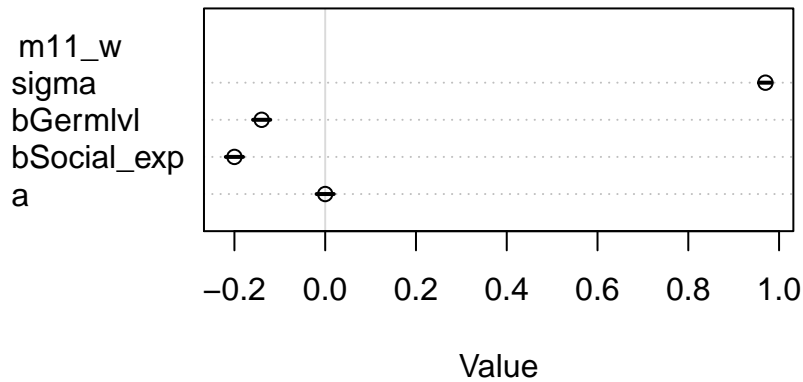
As seen in the above plot, *bSanitizing* does not become independent from energy level when stratifying by germ level. This violates one of the implications listed conditional independencies and tells us that our DAG is not compatible with the wrong data we just simulated.

Wrong Model 11: enrg // scl_ | grm_

Testing whether the causal coefficient for social exposure becomes significantly close to zero when stratifying by germ level in a model that has energy level as outcome variable.

```
m11_w <- quap(alist(energylvl_std_w ~ dnorm(mu, sigma), mu <- a +
  bSocial_exp * social_exposure_std_w + bGermlvl * germ_std_w,
  a ~ dnorm(0, 0.2), bSocial_exp ~ dnorm(0, 1), bGermlvl ~
    dnorm(0, 1), sigma ~ dexp(1)), data = df_w)

plot(coeftab(m11_w), by.model = TRUE)
```



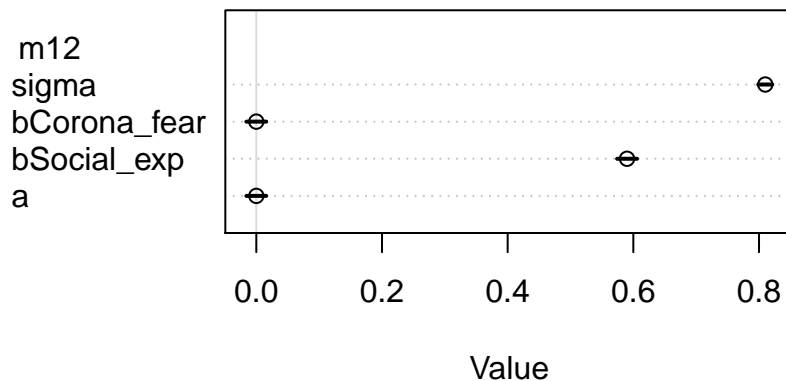
As seen in the above plot, `bSocial_exp` does not become independent from energy level when stratifying by germ level. This again violates one of the implications listed conditional independencies and tells us that our DAG is not compatible with the wrong data we just simulated.

Model 12: `sntz // scl_ | crn_`

Testing whether the causal coefficient for social exposure becomes significantly close to zero when stratifying by corona fear in a model that has sanitizing as outcome variable.

```
m12 <- quap(alist(sanitizing_std_w ~ dnorm(mu, sigma), mu <- a +
  bSocial_exp * social_exposure_std_w + bCorona_fear * corona_fear_std_w,
  a ~ dnorm(0, 0.2), bSocial_exp ~ dnorm(0, 1), bCorona_fear ~
    dnorm(0, 1), sigma ~ dexp(1)), data = df_w)

plot(coeftab(m12), by.model = TRUE)
```



As seen in the above plot, `bSocial_exp` does not become independent from sanitizing when stratifying by germ

level. This violates one of the implications listed conditional independencies and tells us that our DAG is not compatible with the wrong data we just simulated.

In summary, model 10, 11, and 12 does not fit the data we deliberately simulated to not fit our DAG.