

Many Variables (Part 2)

EES 5891-03

Bayesian Statistical Methods

Jonathan Gilligan

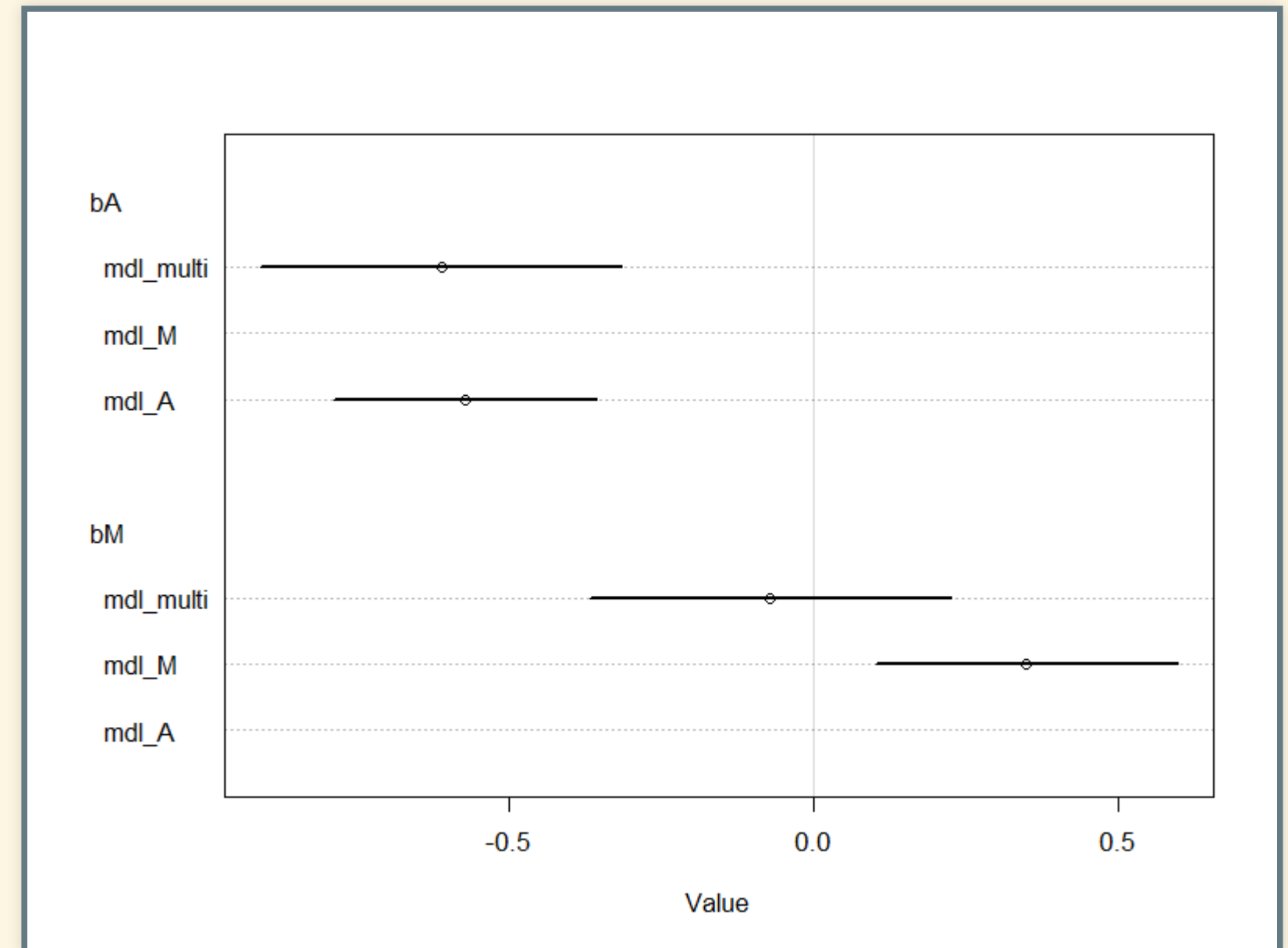
Class #7: Thursday, September 15 2022

Review

DAGs

- DAGs
 - DAGs are theoretical. They don't know about data
 - DAGs make predictions from hypotheses about causal relationships:
 - **Conditional independencies**
 - Test predictions, falsify incorrect hypotheses.
 - How to test predictions:
 - Compare posteriors of single- and multiple-regression models
- The graph shows information about the *posterior* of the model:
 - **Point:** maximum a-posterior estimate
 - **Line:** 95% percentile interval of posterior

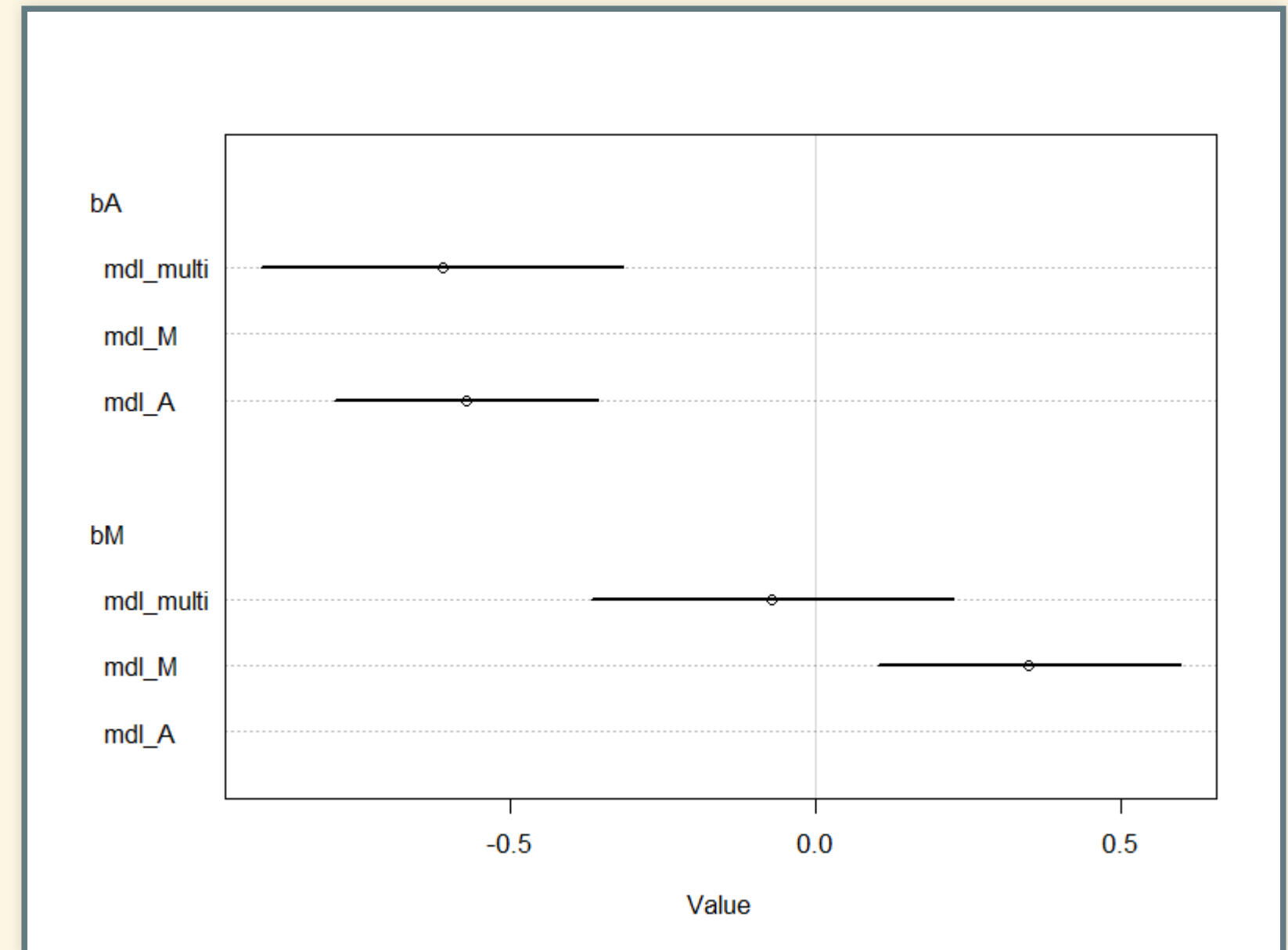
- `mdl_A`: $\mu = \alpha + \beta_A A$
- `mdl_M`: $\mu = \alpha + \beta_M M$
- `mdl_multi`: $\mu = \alpha + \beta_A A + \beta_M M$



Testing Causality Hypotheses

- The graph shows information about the *posterior* of the model:
 - **Point:** maximum a-posterior estimate
 - **Line:** 95% percentile interval of posterior
- **Conditional independencies**
 - Look at the parameter for one predictor.
 - When you add information about the other predictor, does the parameter change?
 - If it becomes consistent with zero, then the *dependent* variable (outcome) is *conditionally independent* of it.
 - It doesn't *cause* the dependent variable to change.

- `mdl_A`: $\mu = \alpha + \beta_A A$
- `mdl_M`: $\mu = \alpha + \beta_M M$
- `mdl_multi`: $\mu = \alpha + \beta_A A + \beta_M M$



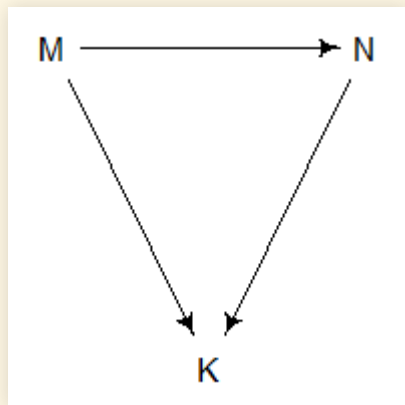
Diagnostic Plots

- Predictor Residual Plots
 - Provide information about the structure of a model
 - Don't address causality or quality of fit.
- Posterior Predictive Plots
 - Provide information about how well the structure of the model matches the data
 - Patterns in the mismatches may give insight into improving the model.
 - Doesn't say anything about causality
- Counterfactual plots
 - Tell us about causality

Counterfactual Models

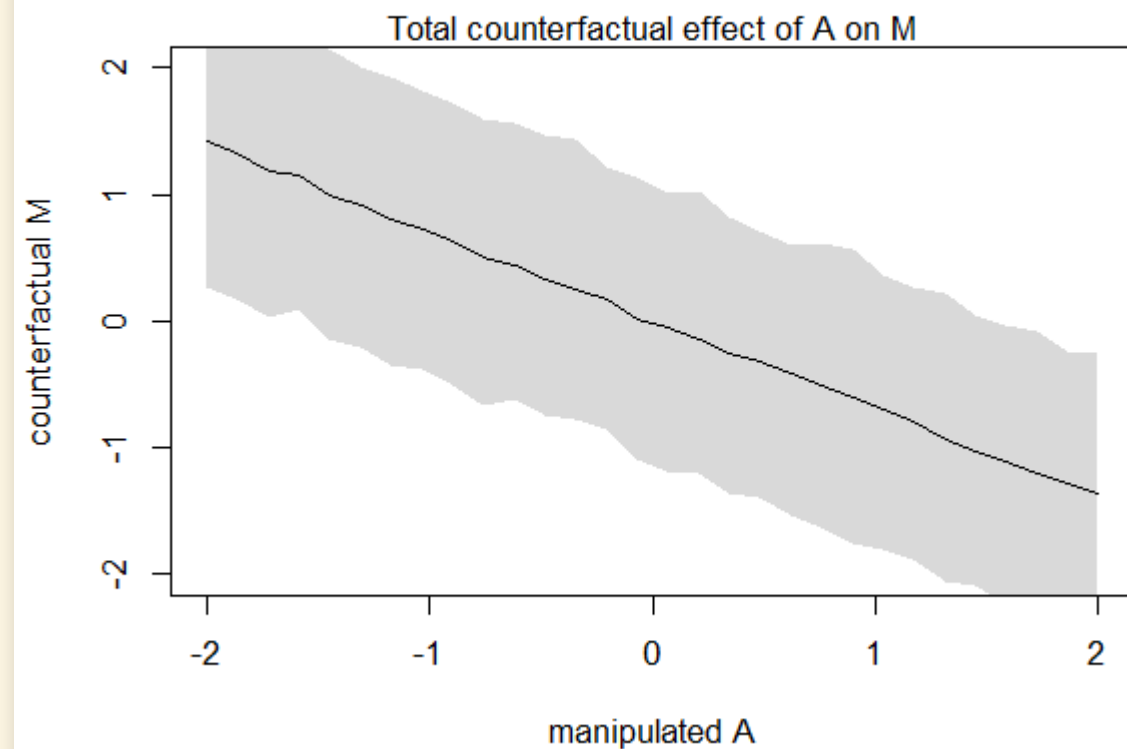
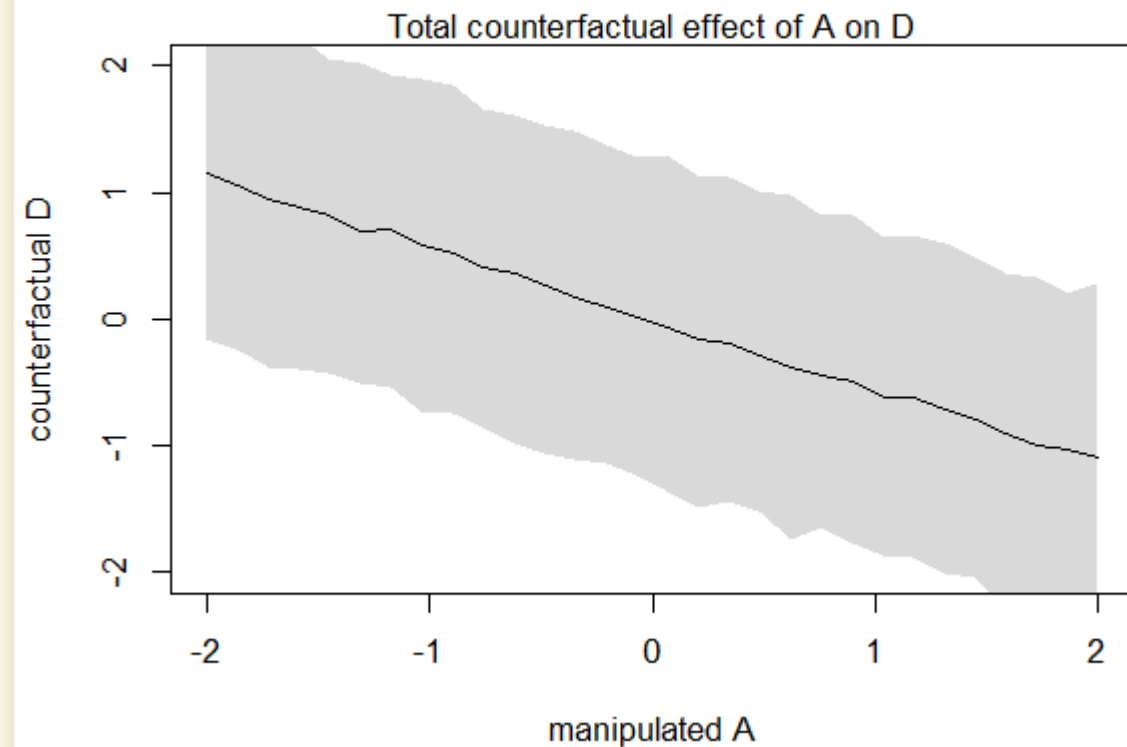
Counterfactual Models

- Counterfactual models:
 - Explore what the model would look like with data very different to what you actually observe.
 - What happens if no one gets married before they're 35?
- Process:
 1. Pick a variable to manipulate (the *intervention variable*)
 2. Define a range of values for it
 3. For each value, use the causal model to simulate values for other variables.
- Simulate the causal model where A influences M and D and M also influences D .



```
mdl_cf <- quap(  
  alist(  
    D ~ dnorm(mu, sigma),  
    mu <- a + bM * M + bA * A,  
    a ~ dnorm(0, 0.2),  
    bM ~ dnorm(0, 0.5),  
    bA ~ dnorm(0, 0.5),  
    sigma ~ dexp(1),  
  
    M ~ dnorm(mu_M, sigma_M),  
    mu_M <- aM + baM * A,  
    aM ~ dnorm(0, 0.2),  
    baM ~ dnorm(0, 0.5),  
    sigma_M ~ dexp(1)  
  ), data = d  
)
```

Manipulating the Counterfactual Model



- Prepare counterfactual

```
sim_dat <- data.frame(A = seq(-2, 2, length.out = 30))  
s <- sim mdl_cf, data = sim_dat, vars = c("M", "D"))
```

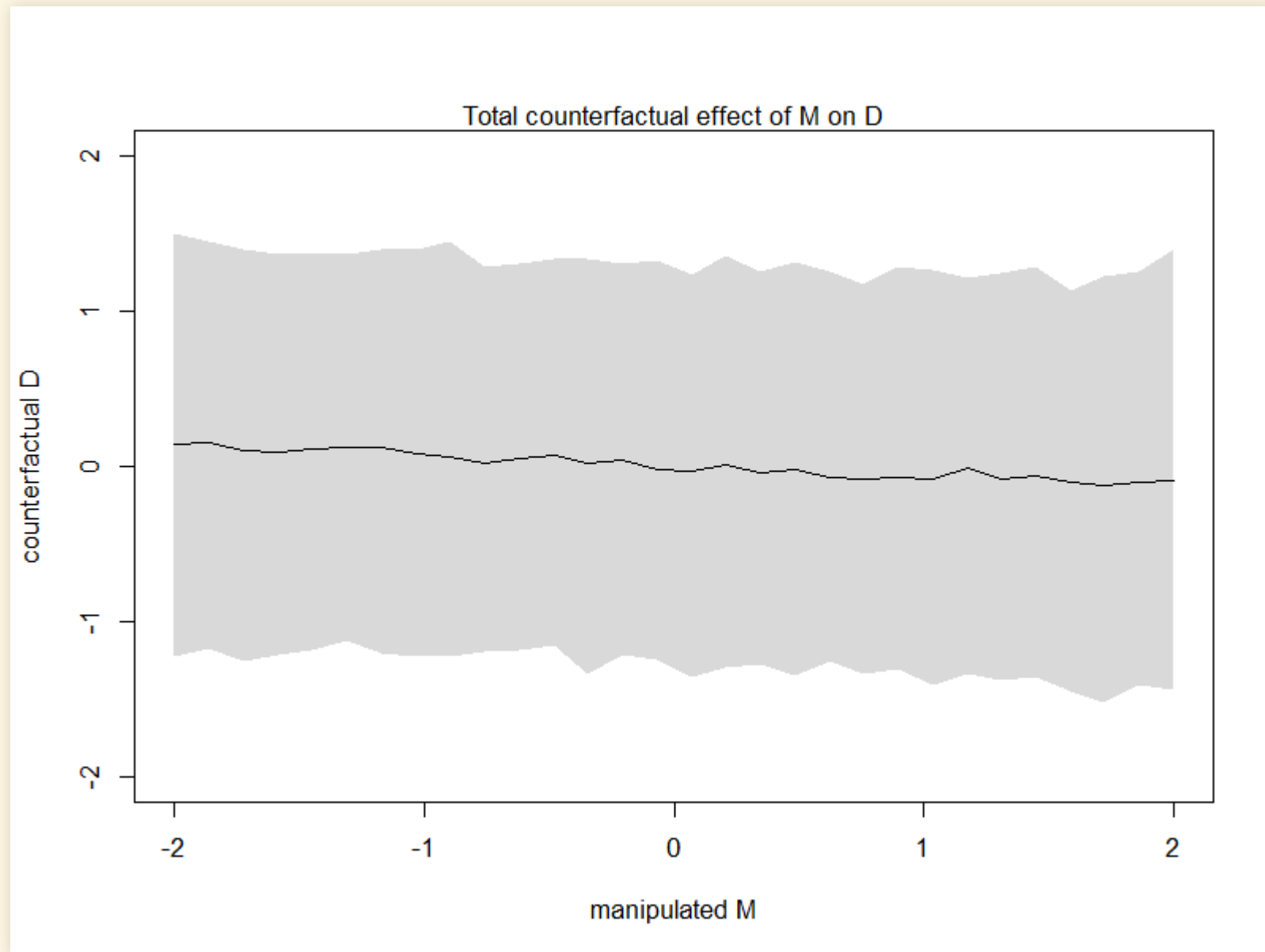
- Plot effect of A on D , including indirect effect through M

```
plot(sim_dat$A, colMeans(s$D), ylim = c(-2,2), type =  
      "l",  
      xlab = "manipulated A", ylab = "counterfactual D")  
shade(apply(s$D, 2, PI), sim_dat$A)  
mtext( "Total counterfactual effect of A on D")
```

- Plot effect of A on M

```
plot(sim_dat$A, colMeans(s$M), ylim = c(-2,2), type =  
      "l",  
      xlab = "manipulated A", ylab = "counterfactual M")  
shade(apply(s$M, 2, PI), sim_dat$A)  
mtext( "Total counterfactual effect of A on M")
```

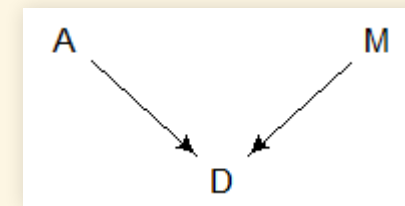

More Counterfactual Manipulations



- Set A to 0 and calculate the direct effect of M on D :

```
sim_dat <- data.frame(M = seq(-2, 2, length.out = 30),  
                      A = 0)  
  
s <- sim mdl_cf, data = sim_dat, vars = c("D"))
```

- Manipulating M directly breaks the causal link with A :



- Plot the results:

```
plot(sim_dat$M, colMeans(s), ylim = c(-2,2), type =  
      "l",  
      xlab = "manipulated M", ylab = "counterfactual D")  
shade(apply(s, 2, PI), sim_dat$M)  
mtext( "Total counterfactual effect of M on D")
```

- When we break the causality between A and M , the direct effect of M on D is very small.

Masked Relationships

Masked Relationships

- Hypothesis:
 - Primates with larger brains produce higher-calorie milk so infant brains grow faster.
- Data:
 - Load data on characteristics of milk in different primate species.
 - `kcal.per.g`: Kilocalories energy per kg milk
 - `mass`: female body mass (kg)
 - `neocortex.perc`: percent of total brain-mass that is neocortex
- Standardize data
 - Note: we standardize the log of mass.
- Clean data
 - *complete-case analysis*: Remove rows with missing values

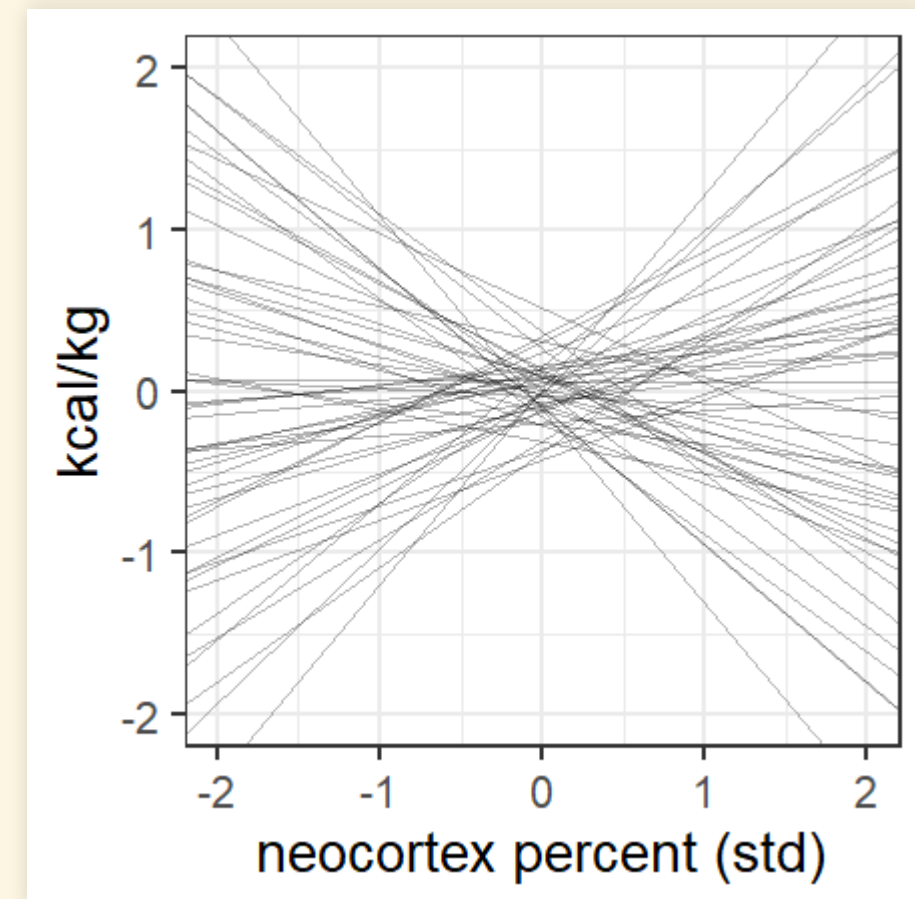
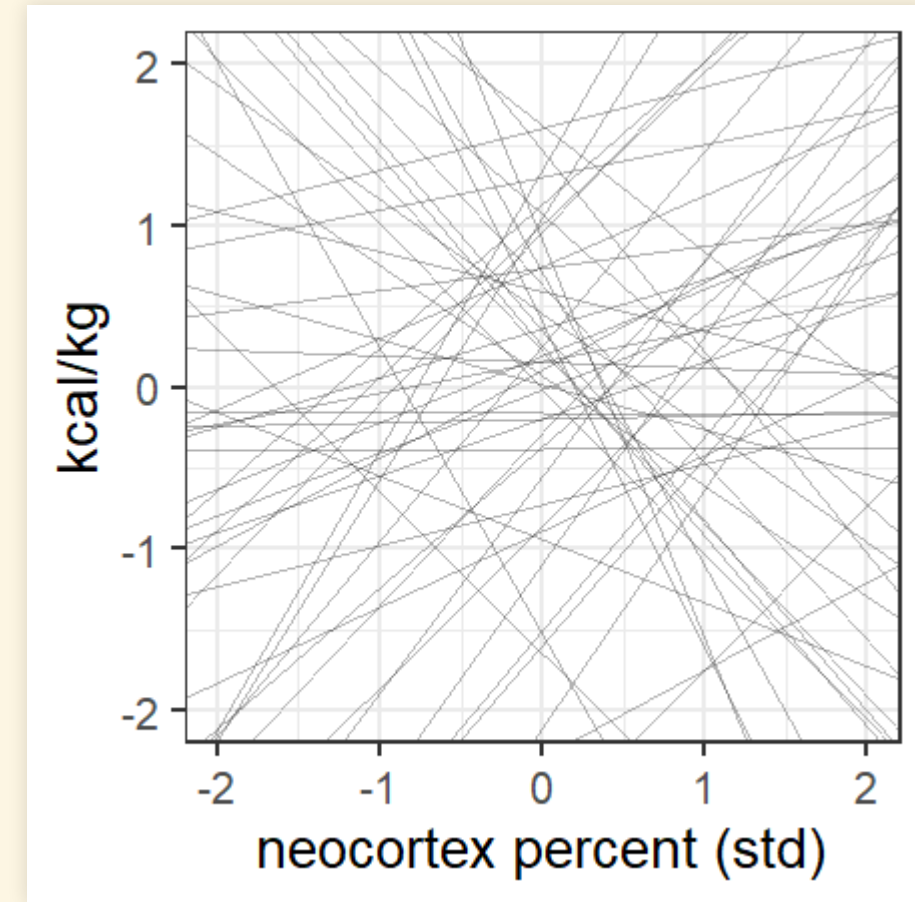
```
data(milk)
d <- milk
```

```
d <- d %>% mutate(
  K = standardize(kcal.per.g),
  N = standardize(neocortex.perc),
  M = standardize(log(mass))
)
```

```
dcc <- drop_na(d, K, N, M)
```

Simple Regression

- Model
$$K \sim \text{Normal}(\mu, \sigma) \quad \mu = \alpha + \beta_N N \quad \alpha \sim \text{Normal}(0, 1) \quad \beta_N \sim \text{Normal}(0, 1) \quad \sigma \sim \text{Exponential}(1)$$
- Plot prior predictive distributions for variables.
 - Prior predictions look absurd.
- Choose better priors:
$$\alpha \sim \text{Normal}(0, 0.2) \quad \beta_N \sim \text{Normal}(0, 0.5)$$

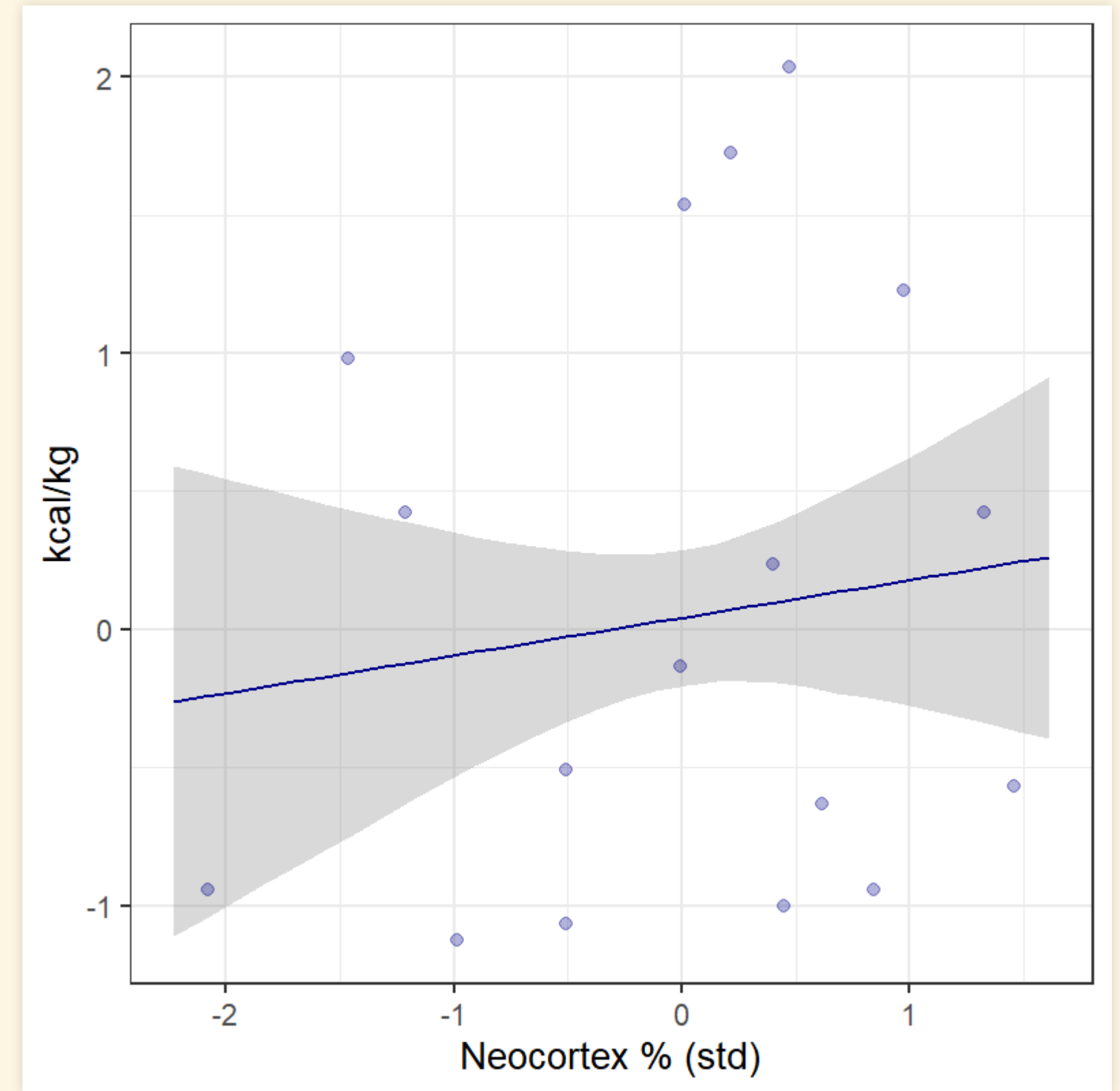


Examine Model

```
precis_show(precis mdl_milk, digits = 2))
```

```
##      mean    sd  5.5% 94.5%  
## a      0.04 0.15 -0.21  0.29  
## bN     0.13 0.22 -0.22  0.49  
## sigma 1.00 0.16  0.74  1.26
```

- Both a and bN are consistent with zero. There isn't a strong relationship between N and K .

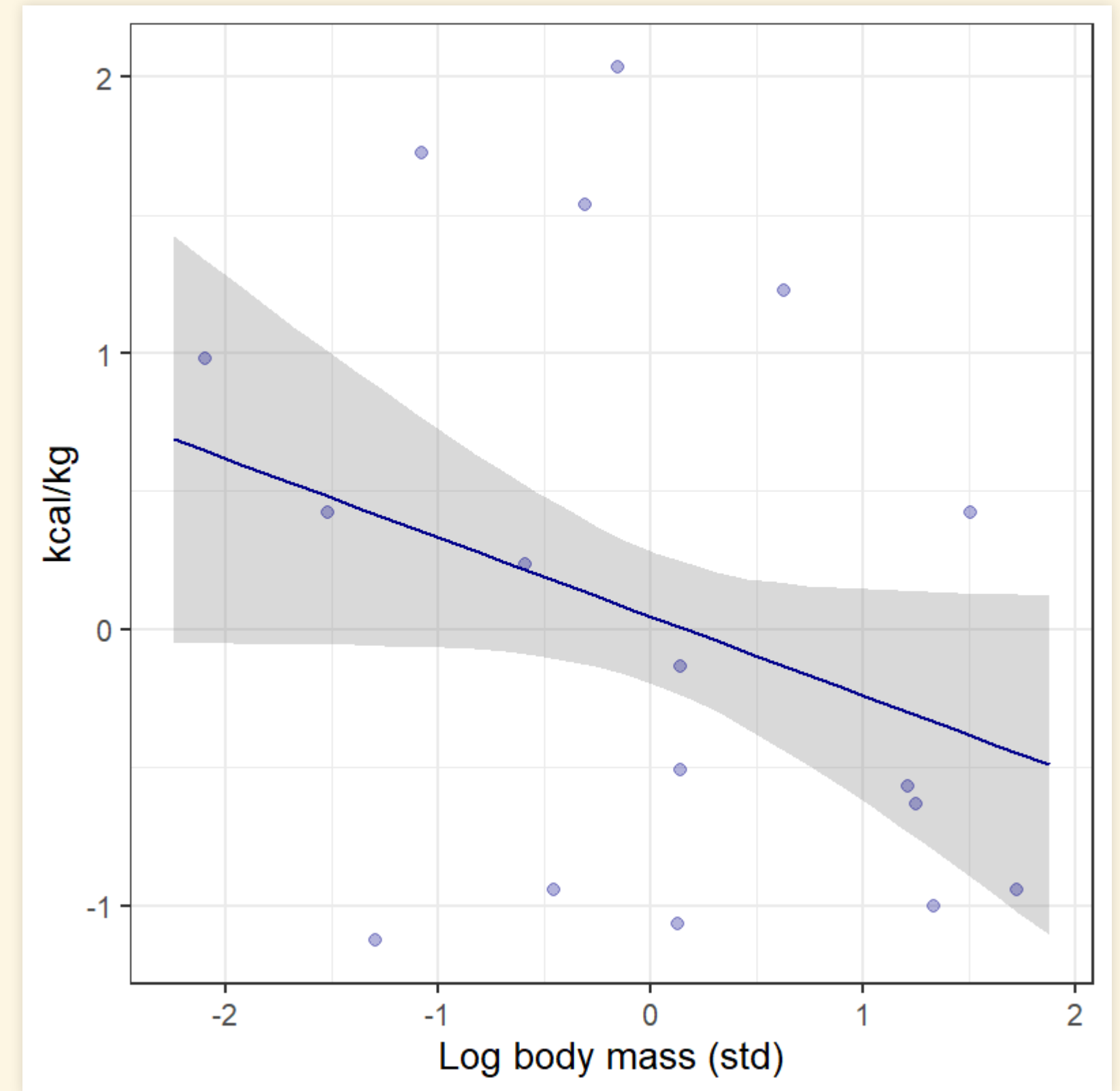


Try A Different Model

- Model
$$K \sim \text{Normal}(\mu, \sigma) \\ \mu = \alpha + \beta_M M \\ \alpha \sim \text{Normal}(0, 1) \\ \beta_M \sim \text{Normal}(0, 1) \\ \sigma \sim \text{Exponential}(1)$$

```
precis_show(precis(mdl_milk_2, digits = 2))
```

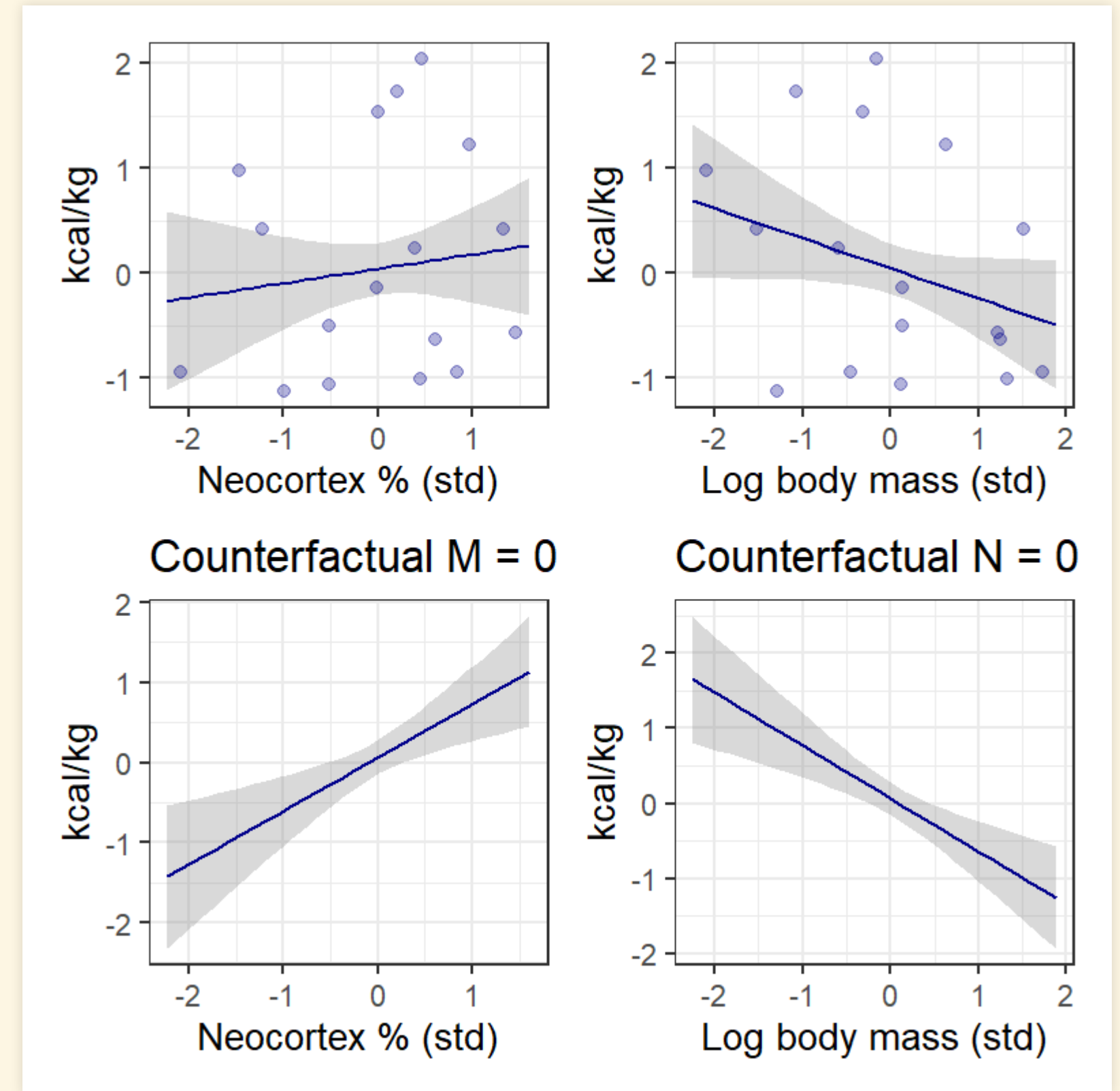
| ## | | mean | sd | 5.5% | 94.5% |
|----|-------|-------|------|-------|-------|
| ## | a | 0.05 | 0.15 | -0.20 | 0.29 |
| ## | bM | -0.28 | 0.19 | -0.59 | 0.03 |
| ## | sigma | 0.95 | 0.16 | 0.70 | 1.20 |



Compare Models using Counterfactuals

- Model
$$K \sim \text{Normal}(\mu, \sigma) \quad \mu = \alpha + \beta_M M + \beta_N N \quad \alpha \sim \text{Normal}(0, 1) \quad \beta_M \sim \text{Normal}(0, 1) \quad \beta_N \sim \text{Normal}(0, 1) \quad \sigma \sim \text{Exponential}(1)$$
- M and N have opposite effects, so they cancel out.

■ Masking



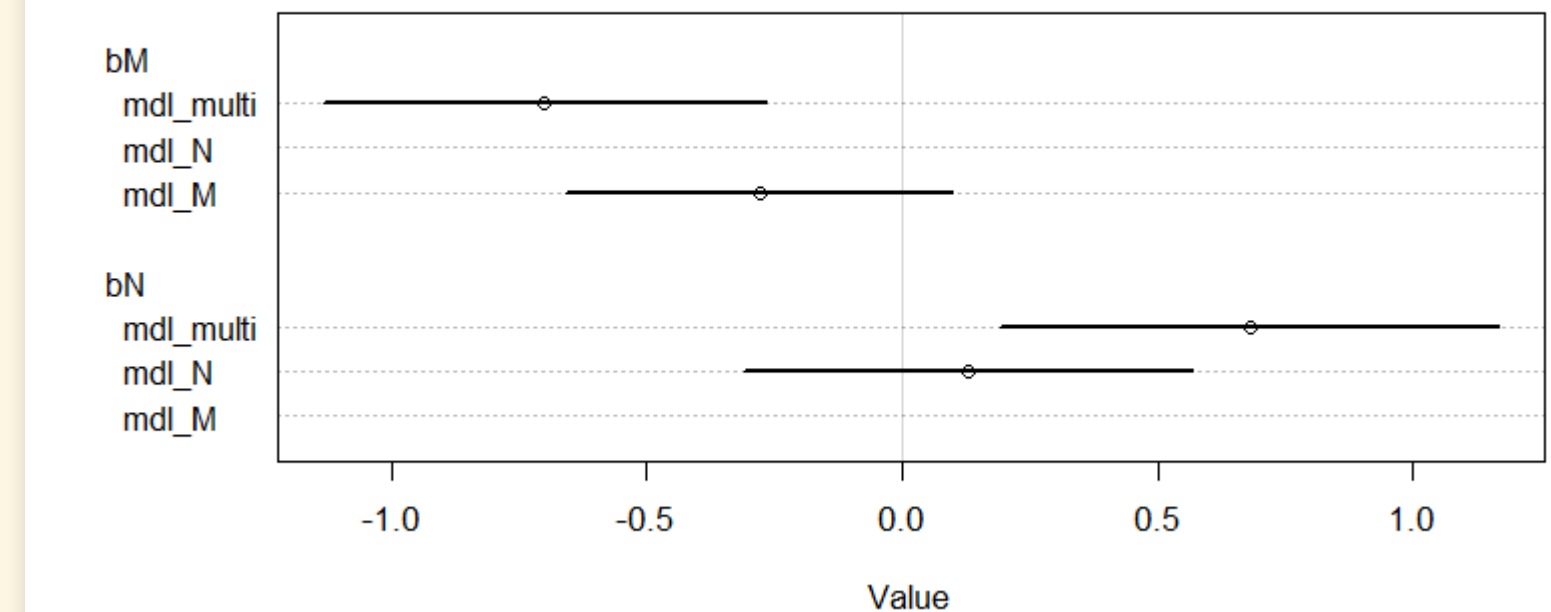
Multiple Regression Model

- Model
$$K \sim \text{Normal}(\mu, \sigma) \quad \mu = \alpha + \beta_M M + \beta_N N \quad \alpha \sim \text{Normal}(0, 1) \quad \beta_M \sim \text{Normal}(0, 1) \quad \beta_N \sim \text{Normal}(0, 1) \quad \sigma \sim \text{Exponential}(1)$$

```
precis_show(precis(mdl_milk_3, digits = 2))
```

```
##      mean    sd  5.5% 94.5%
## a      0.07 0.13 -0.15  0.28
## bM     -0.70 0.22 -1.06 -0.35
## bN      0.68 0.25  0.28  1.07
## sigma  0.74 0.13  0.53  0.95
```

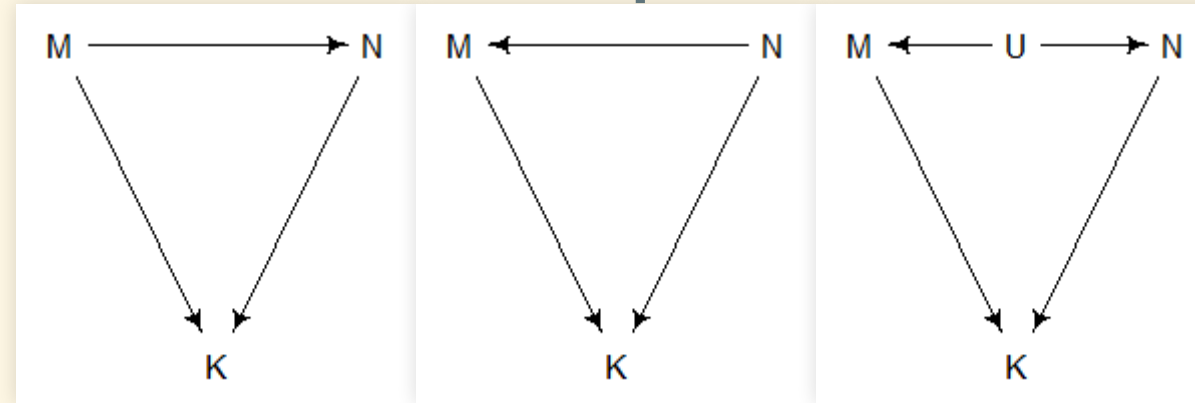
```
coefplot(coefplot(mdl_M, mdl_N, mdl_multi),
         pars = c("bM", "bN"))
```



- This is the opposite of what we saw for divorce rates.
- The parameters for each predictor are consistent with zero for the single-predictor models
- When we include both predictors, the association with each is stronger.

Interpreting Result

- There are 3 possible DAGs



- U is a variable that we didn't observe (sometimes called a **latent variable**)
 - We will look at latent variables in more detail in Chapter 6.
- Figuring out which diagram is the right one is **hard**.
 - All three have the same *conditional independencies*.
 - Data alone won't solve this.
 - Use scientific knowledge can rule out absurd possibilities.