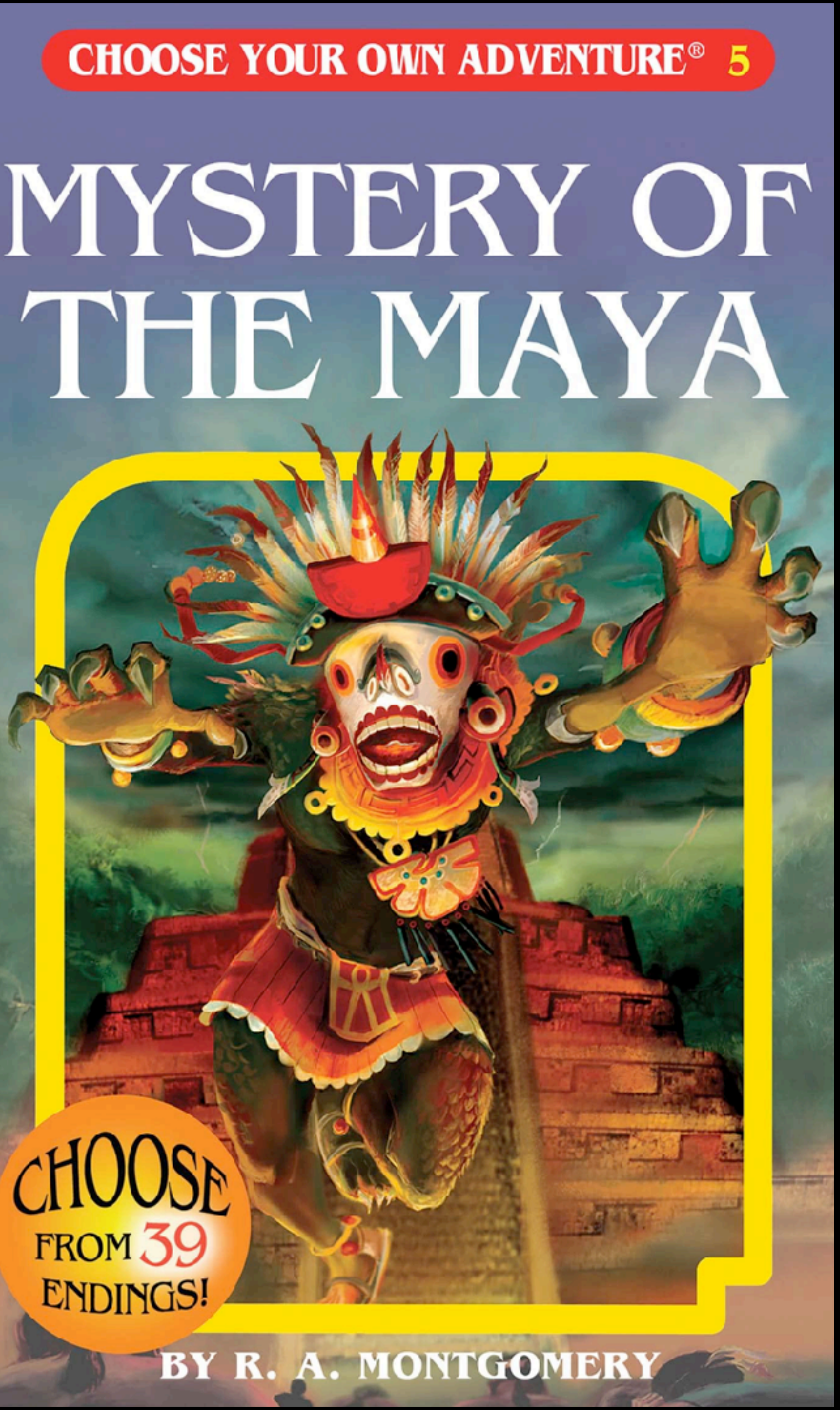
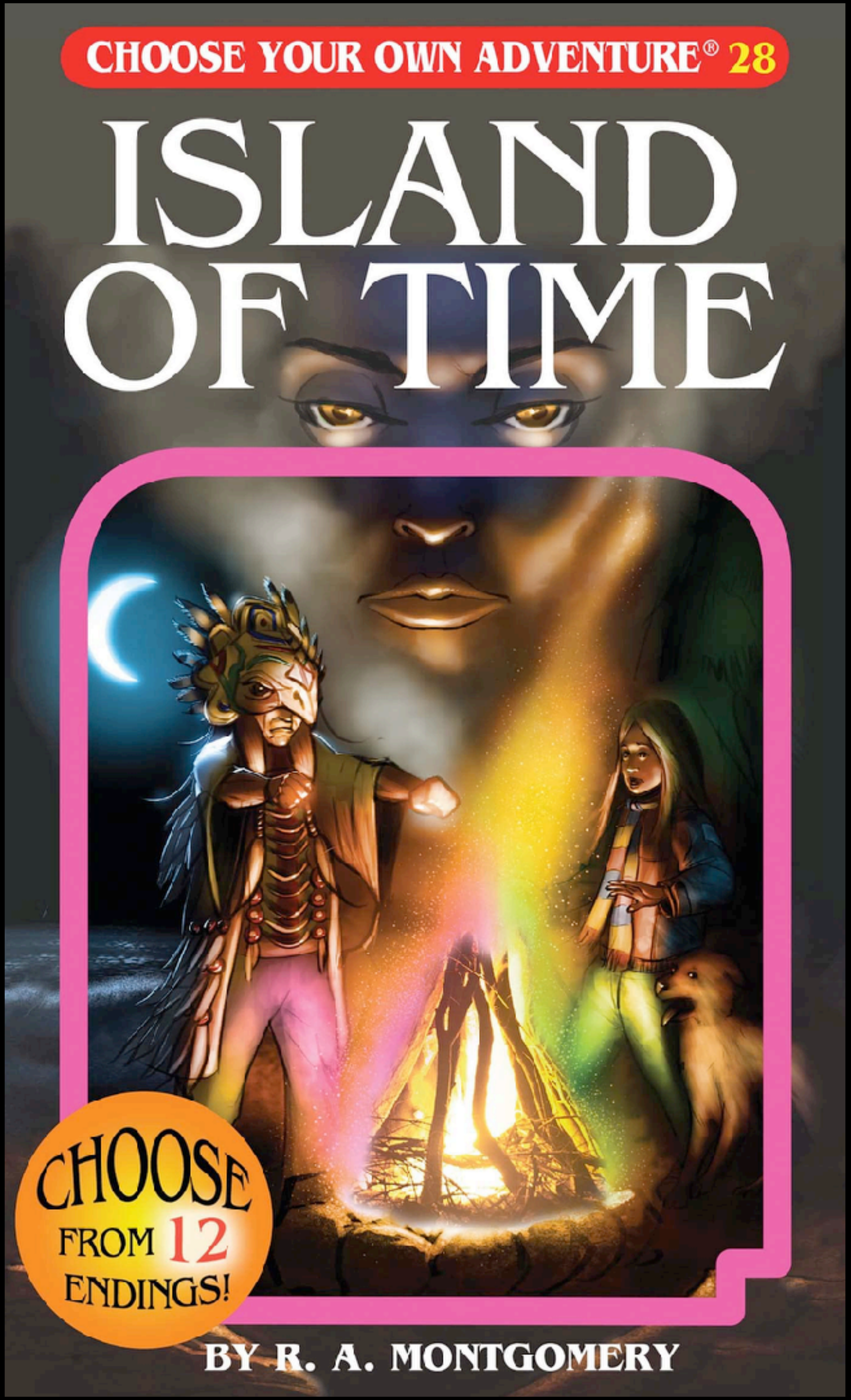
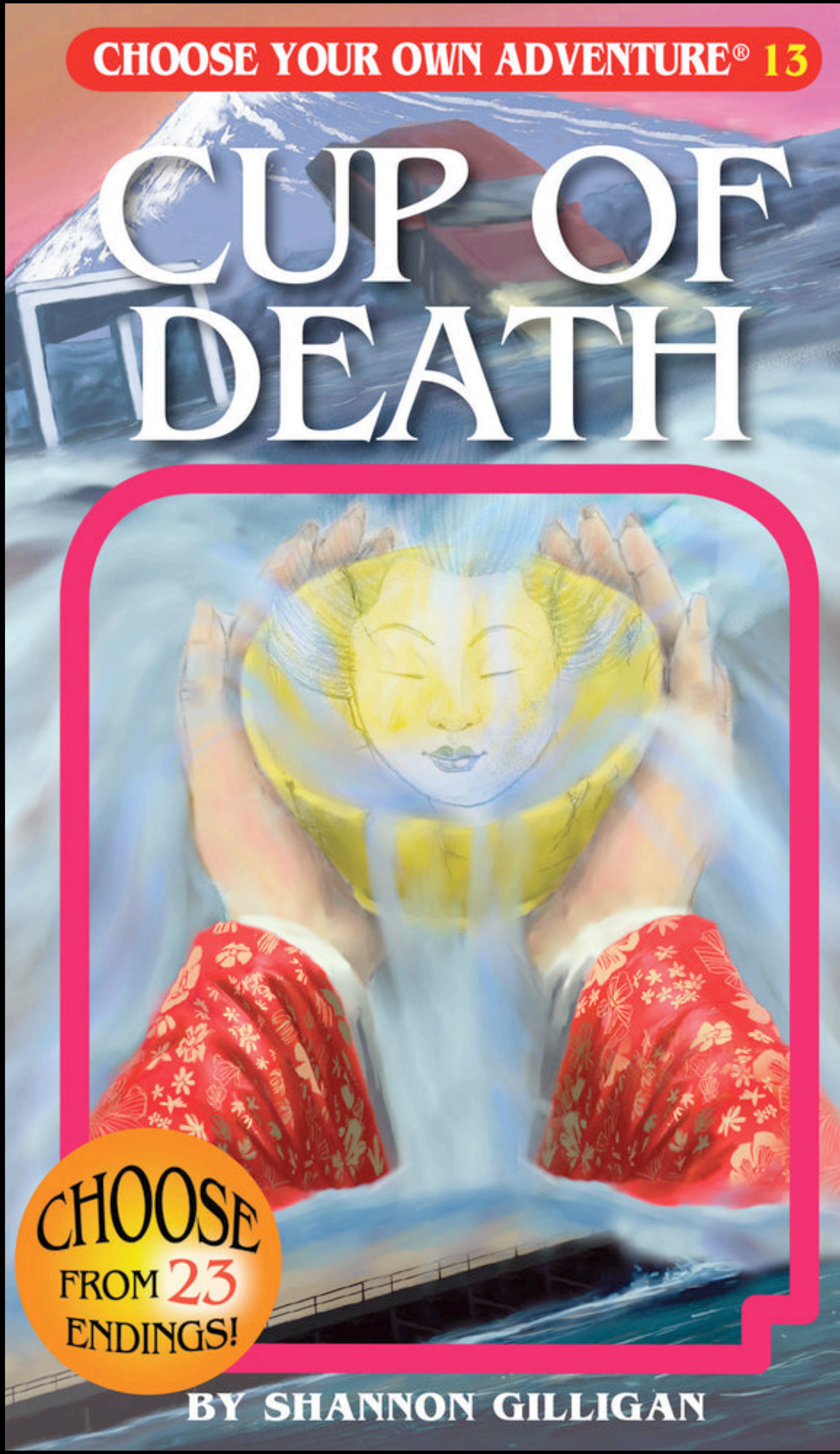
Statistical Rethinking

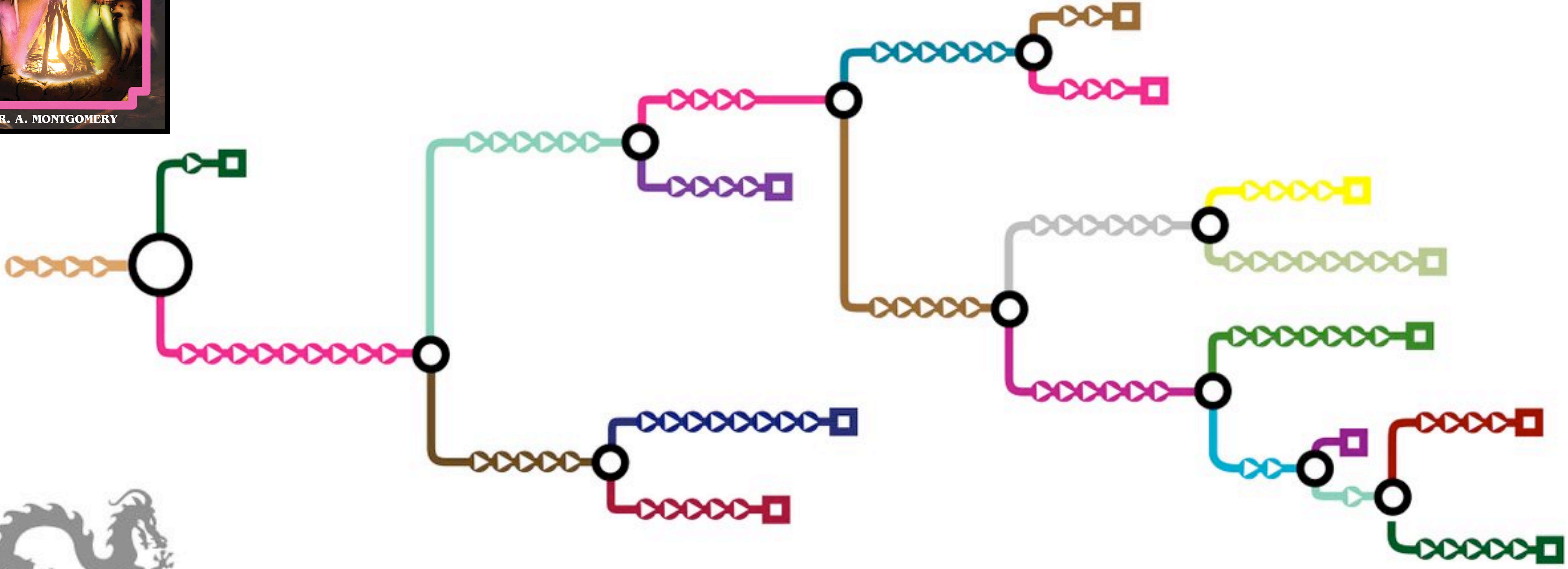
2023

13. Multilevel Adventures


CHOOSE YOUR OWN ADVENTURE® 5
MYSTERY OF THE MAYA
CHOOSE FROM 39 ENDINGS!
BY R. A. MONTGOMERY


CHOOSE YOUR OWN ADVENTURE® 28
ISLAND OF TIME
CHOOSE FROM 12 ENDINGS!
BY R. A. MONTGOMERY

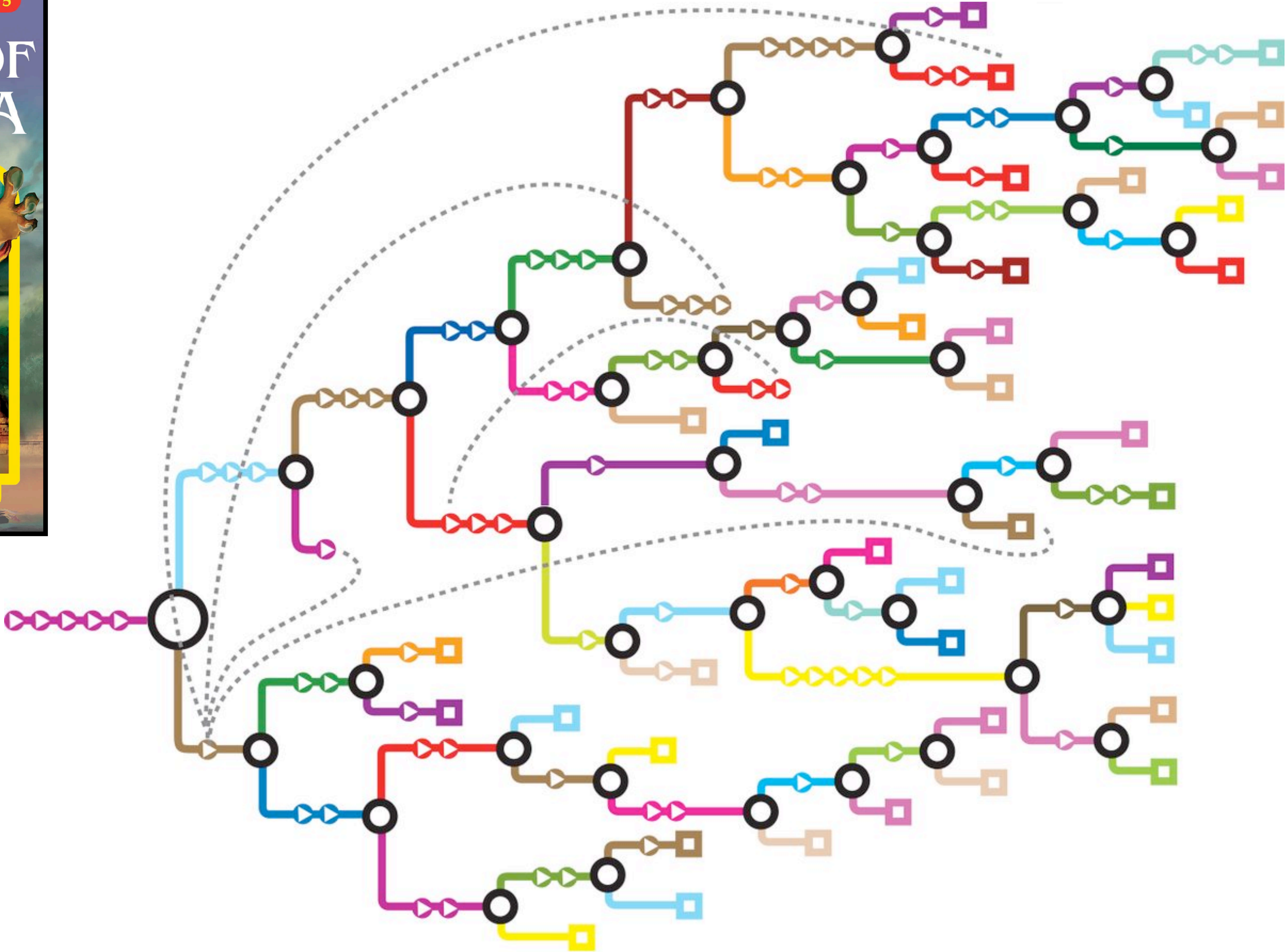

CHOOSE YOUR OWN ADVENTURE® 13
CUP OF DEATH
CHOOSE FROM 23 ENDINGS!
BY SHANNON GILLIGAN



CHOOSECO
Island of Time
Image © Chooseco LLC

https://www.atlasobscura.com/articles/cyoa-choose-your-own-adventure-maps


CHOOSE YOUR OWN ADVENTURE® 5
MYSTERY OF THE MAYA
CHOOSE FROM 39 ENDINGS!
BY R. A. MONTGOMERY



CHOOSECO

Drawing the Bayesian Owl

1. Theoretical estimand

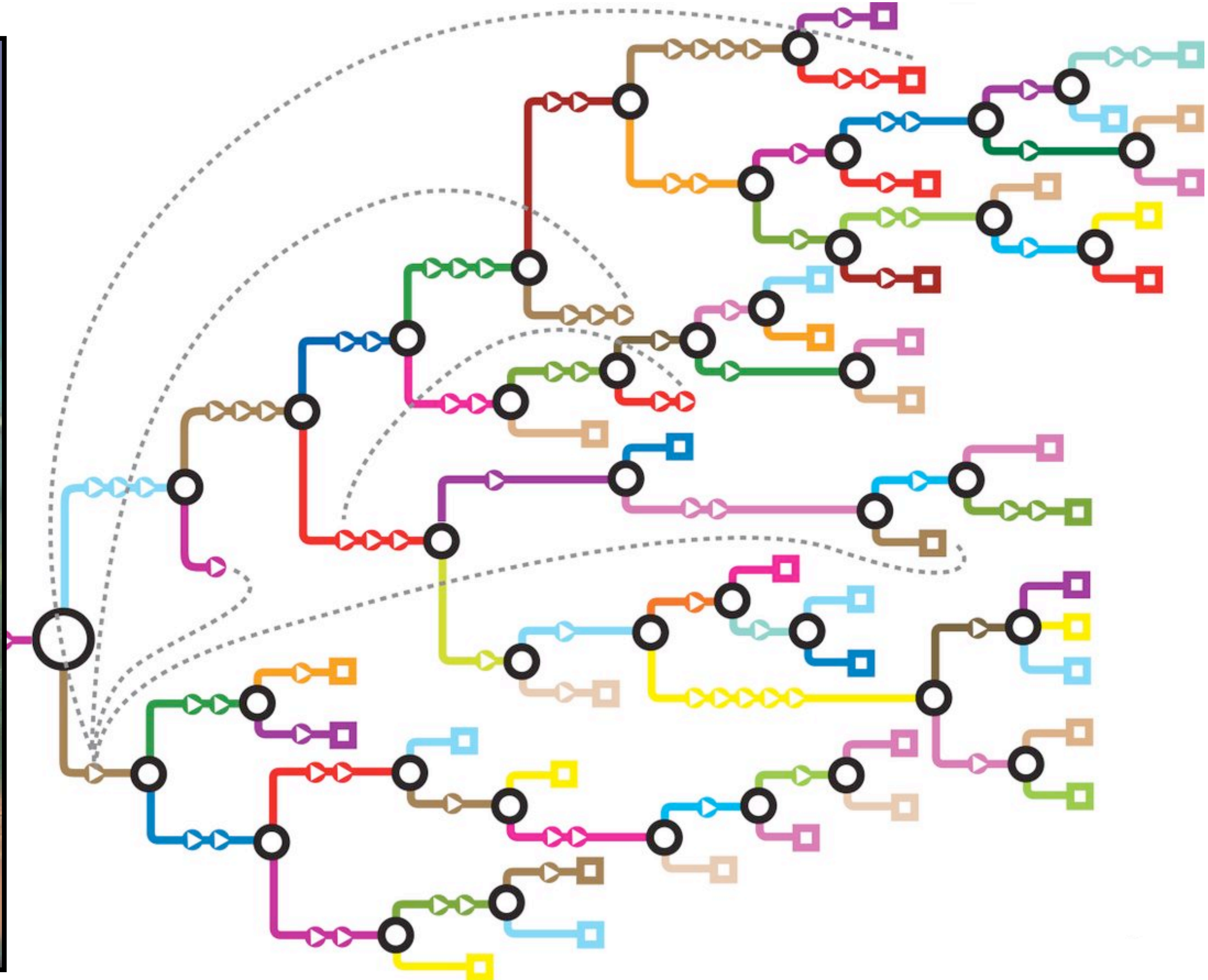
2. Scientific (causal) model(s)

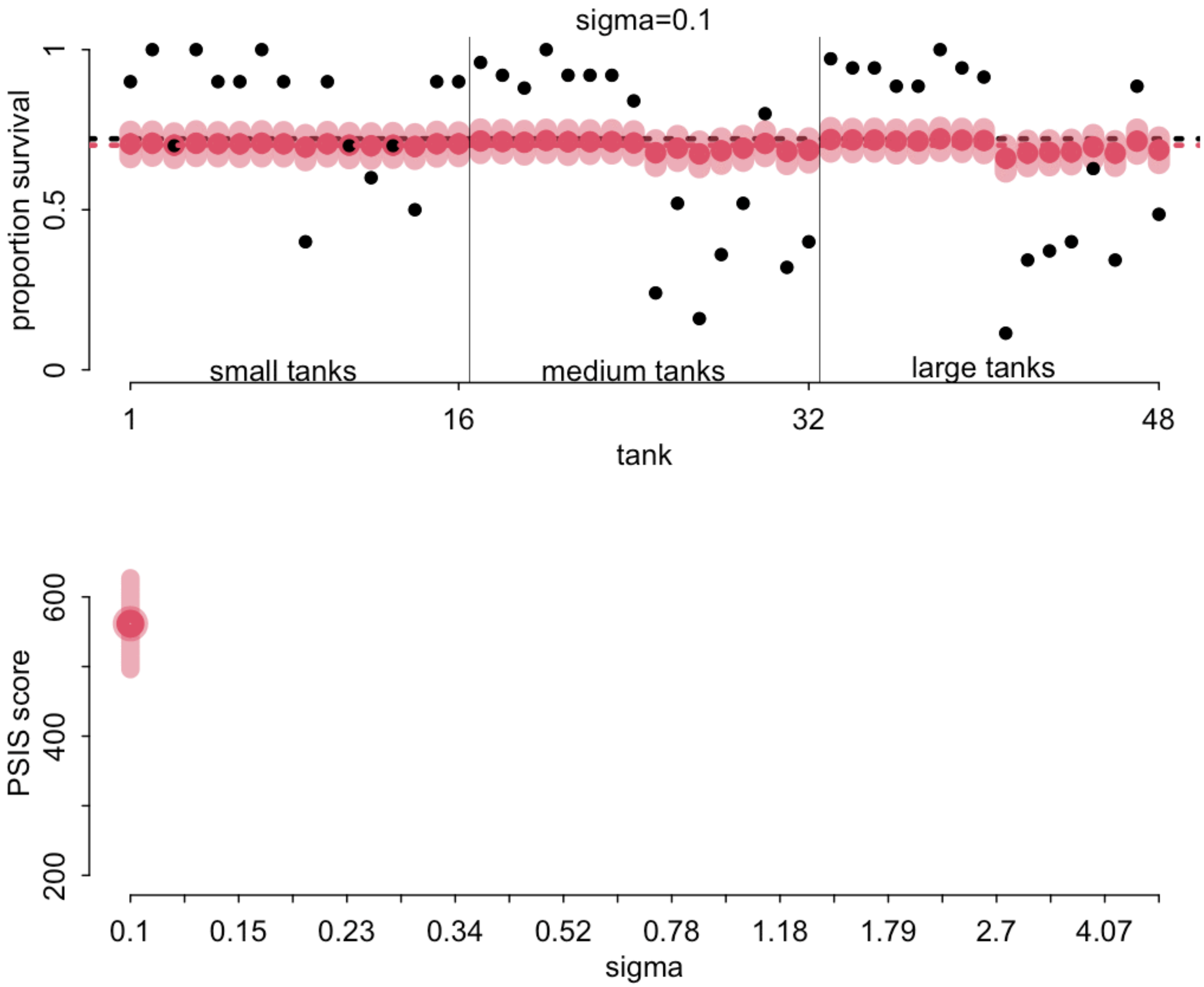
3. Use 1 & 2 to build statistical model(s)

4. Simulate from 2 to validate 3 yields 1

5. Analyze real data


CHOOSE YOUR OWN ADVENTURE® 5
MYSTERY OF THE MODEL
CHOOSE FROM 39 ENDINGS!
BY R. A. MONTGOMERY




sigma=0.1
proportion survival
1
0.5
0
small tanks
medium tanks
large tanks
1
16
32
48
tank
PSIS score
600
400
200
0.1
0.15
0.23
0.34
0.52
0.78
1.18
1.79
2.7
4.07
sigma

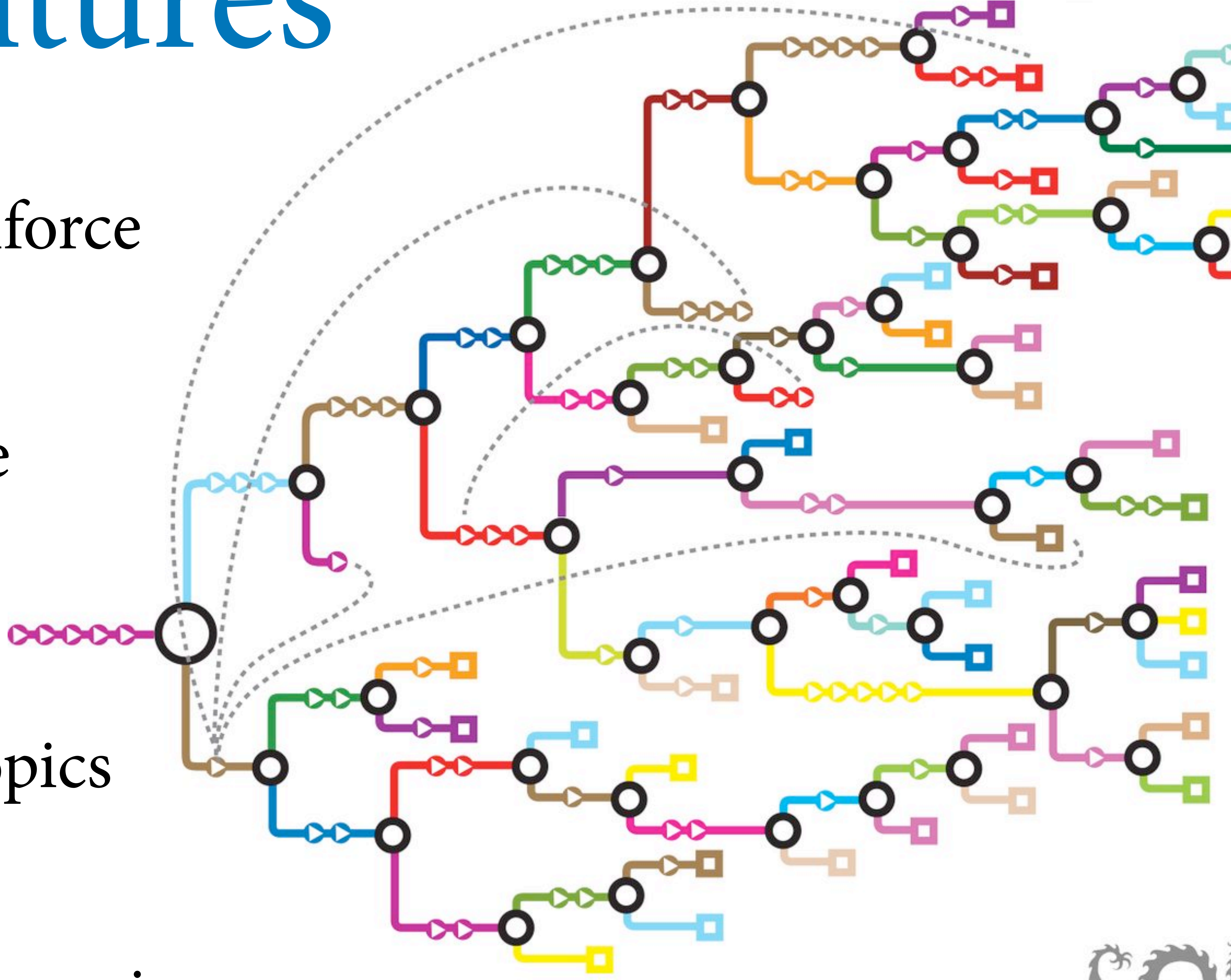
Multilevel Adventures

Return to the start: Start again, reinforce foundation

Skim & Index: Don’t try to learn the details; just acquaint yourself with possibilities

Pick & Choose: Engage only with topics that interest you

Bayesian Flow: Just enough to keep moving

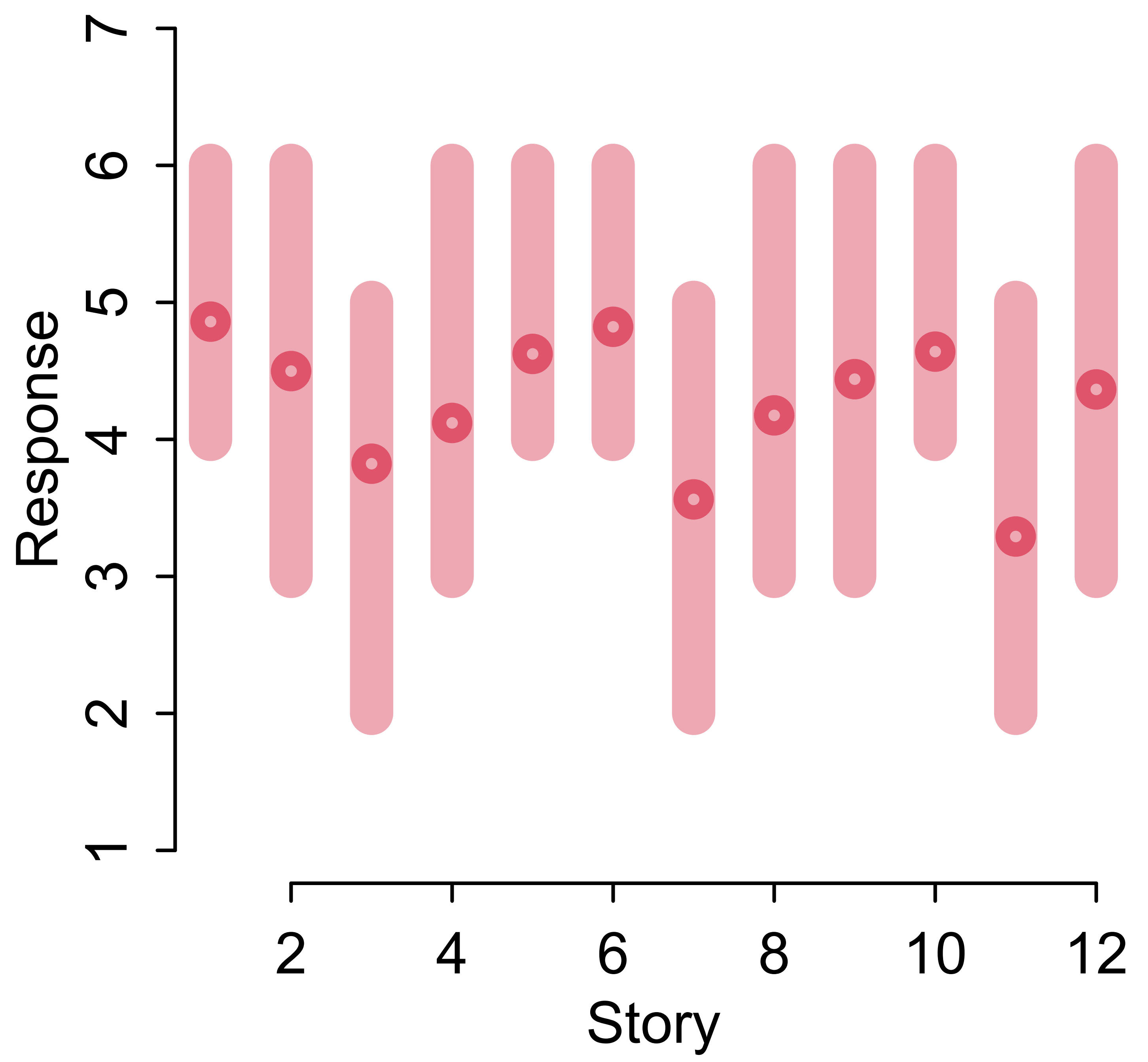


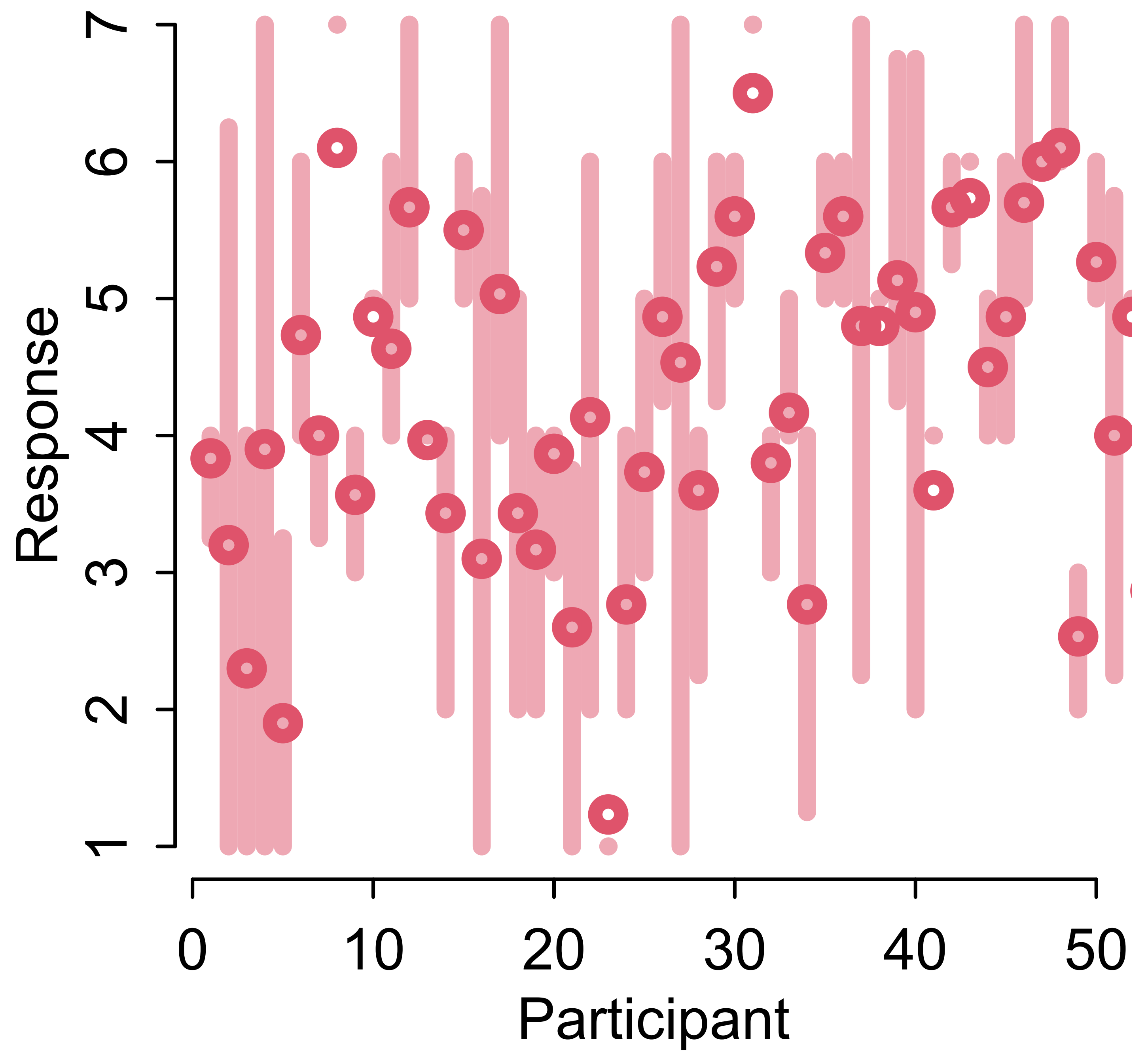
Multilevel Adventures

Clusters: Kinds of groups in the data

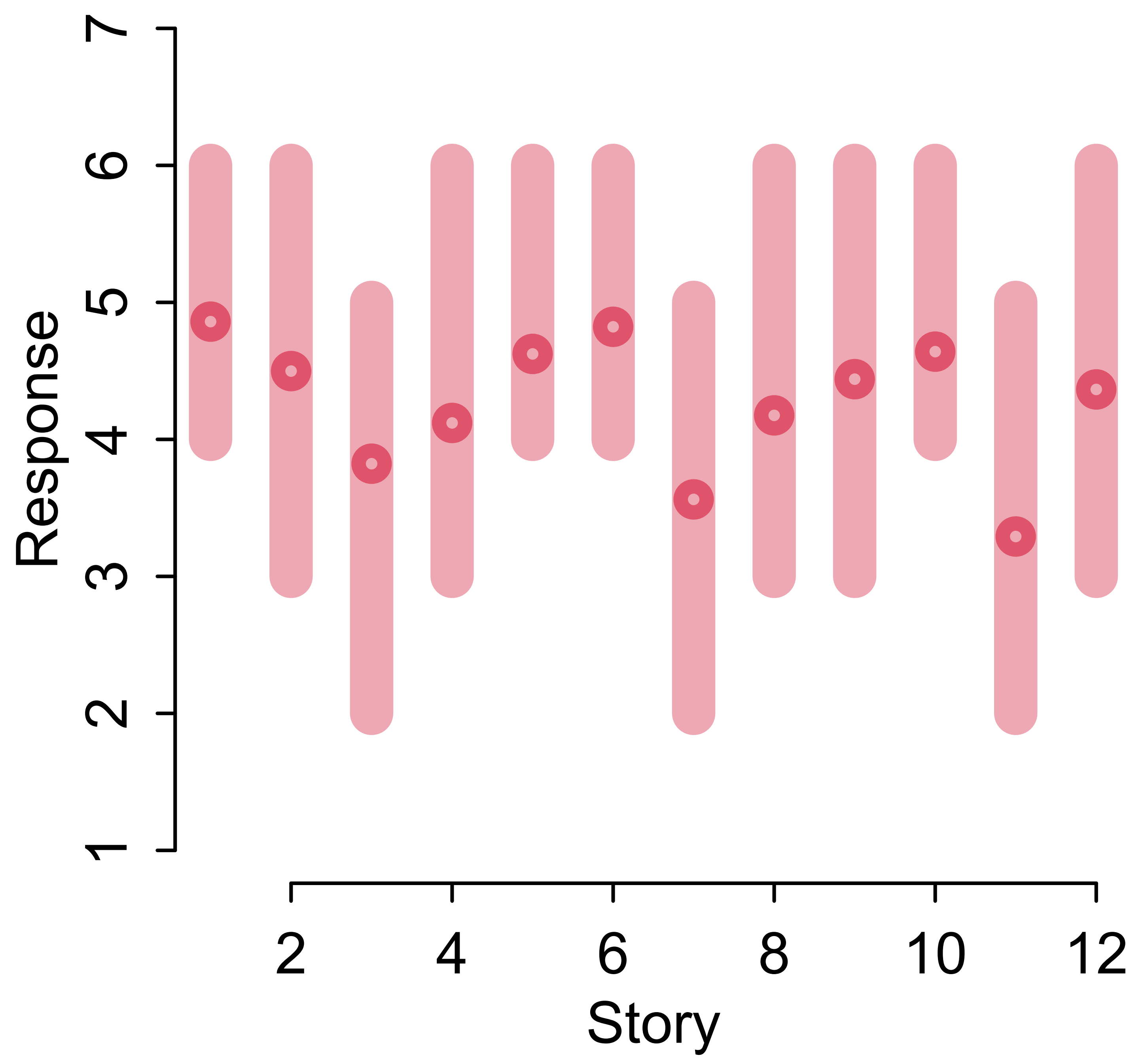
Features: Aspects of the model (parameters) that vary by cluster

Cluster		**Features**
tanks	⟶	survival
stories	⟶	treatment effect
individuals	⟶	average response
departments	⟶	admission rate, bias



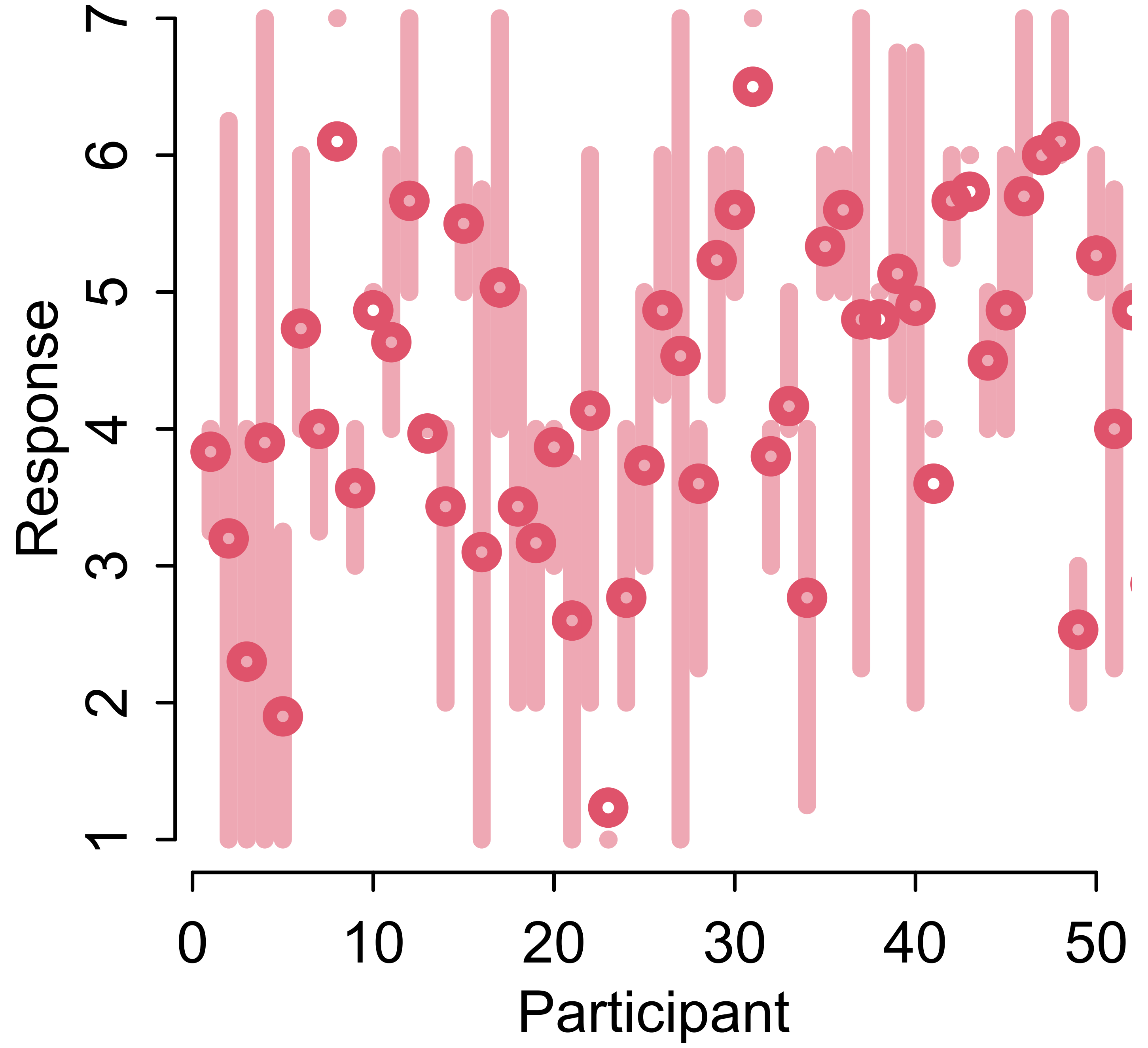


Cluster		Features
tanks	⟶	survival
stories	⟶	treatment effect
individuals	⟶	average response
departments	⟶	admission rate, bias



Add clusters: More index variables, more population priors

Add features: More parameters, more dimensions *in each* population prior

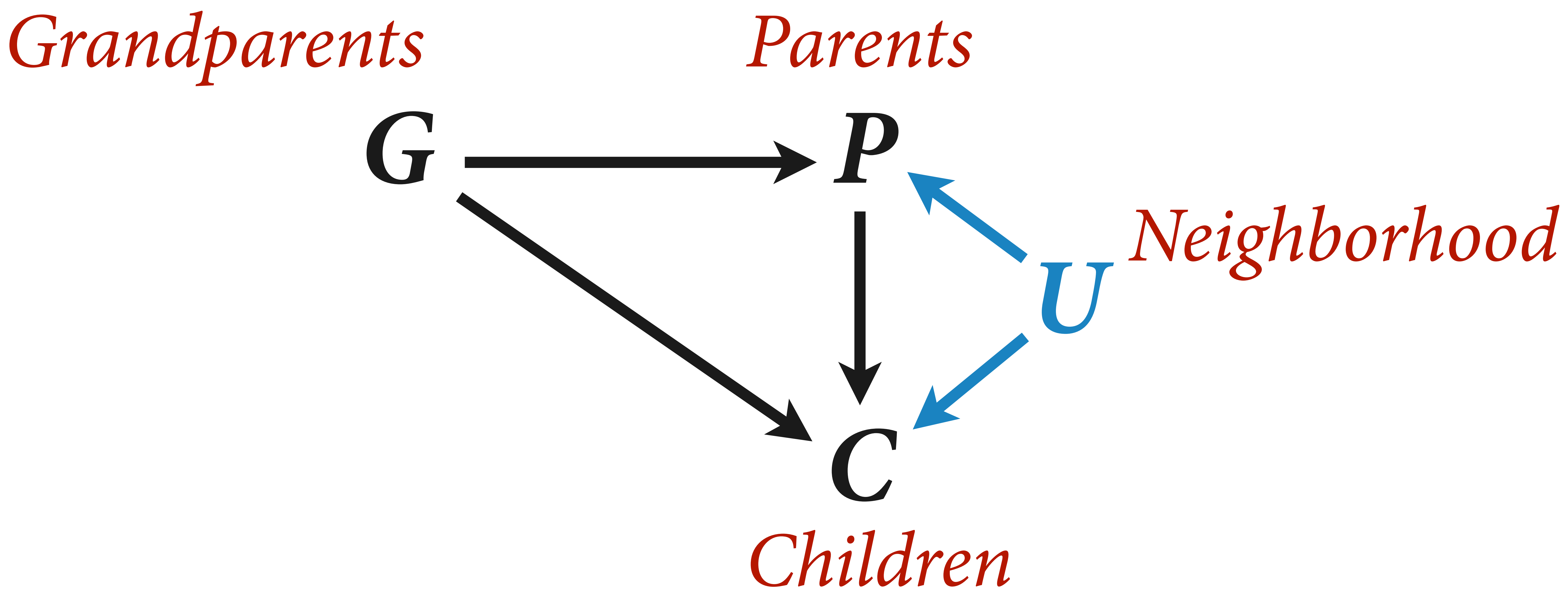


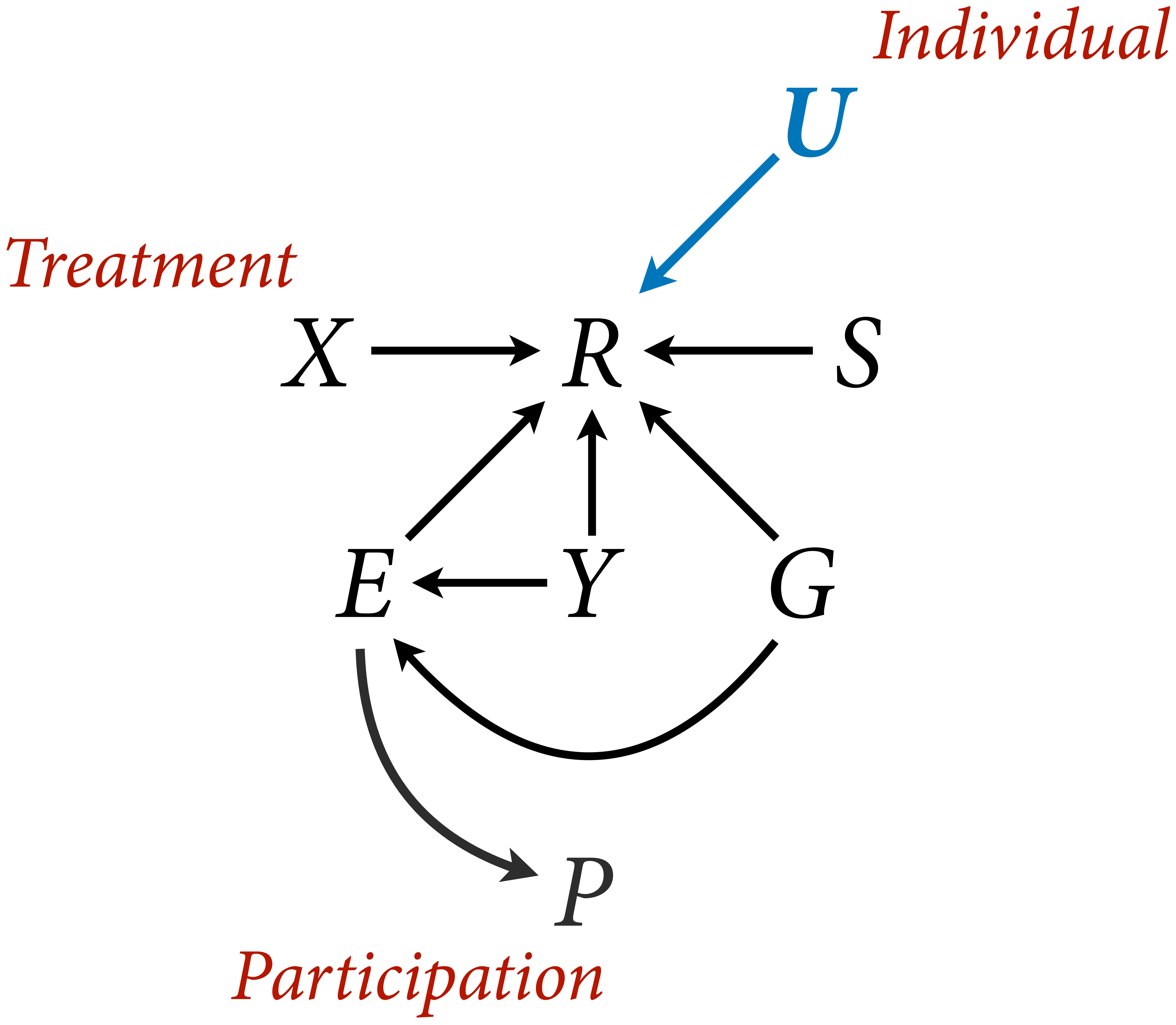
Varying effects as confounds

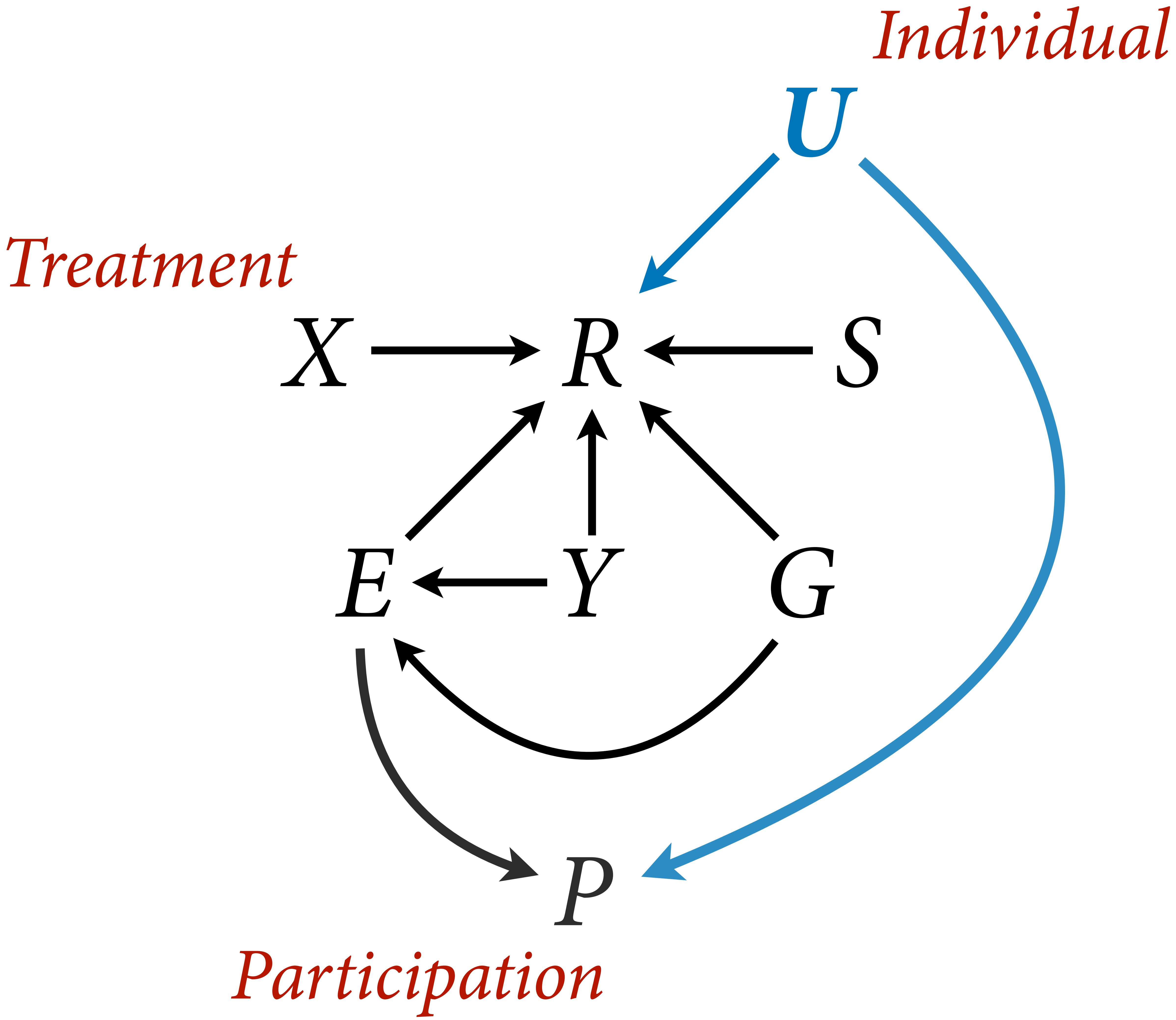
Varying effect strategy: Unmeasured features of clusters leave an imprint on the data that can be measured by (1) repeat observations of each cluster and (2) partial pooling among clusters

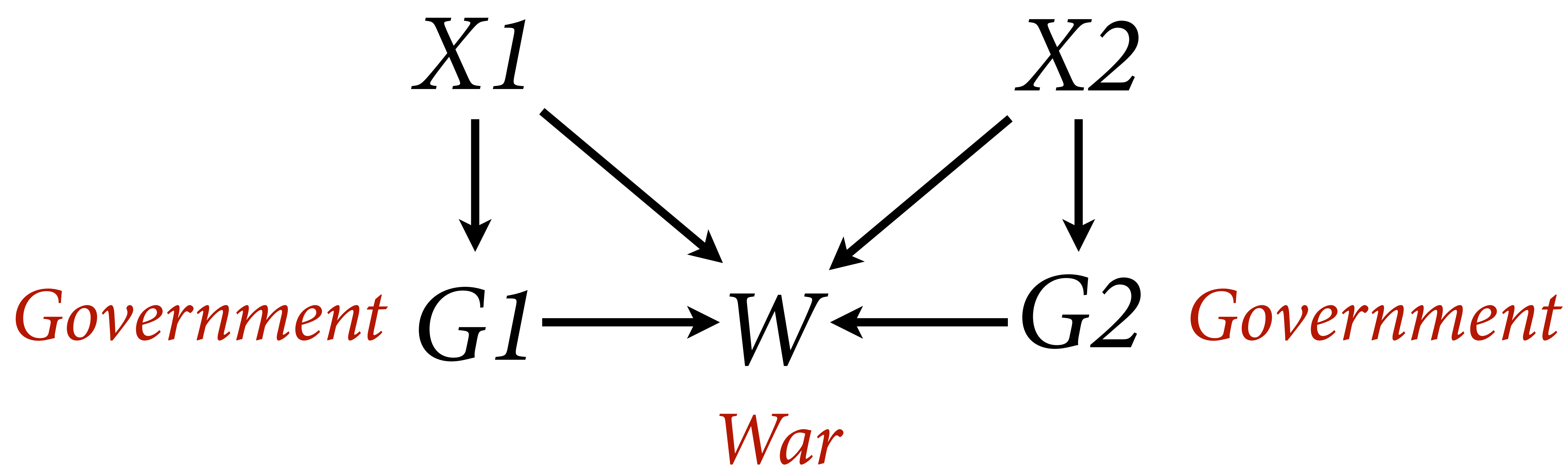
Predictive perspective: Important source of cluster-level variation, regularize

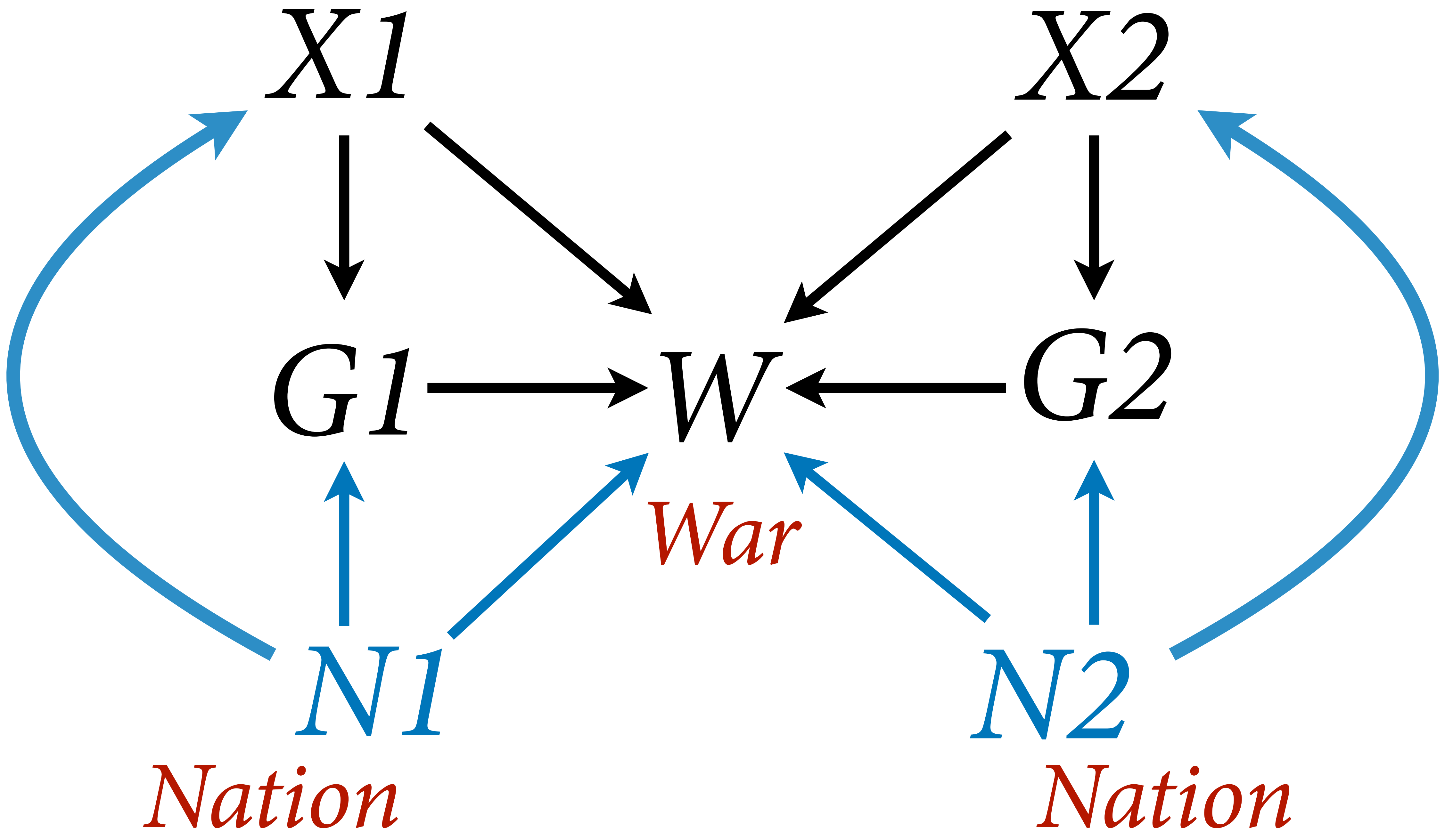
Causal perspective: Competing causes or unobserved confounds


Grandparents
Parents
G
P
Neighborhood
U
C
Children


Individual
U
Treatment
X
R
S
E
Y
G
P
Participation


Individual
U
Treatment
X
R
S
E
Y
G
P
Participation


X1
X2
Government
G1
W
G2
Government
War


X1
X2
G1
W
G2
War
N1
N2
Nation
Nation

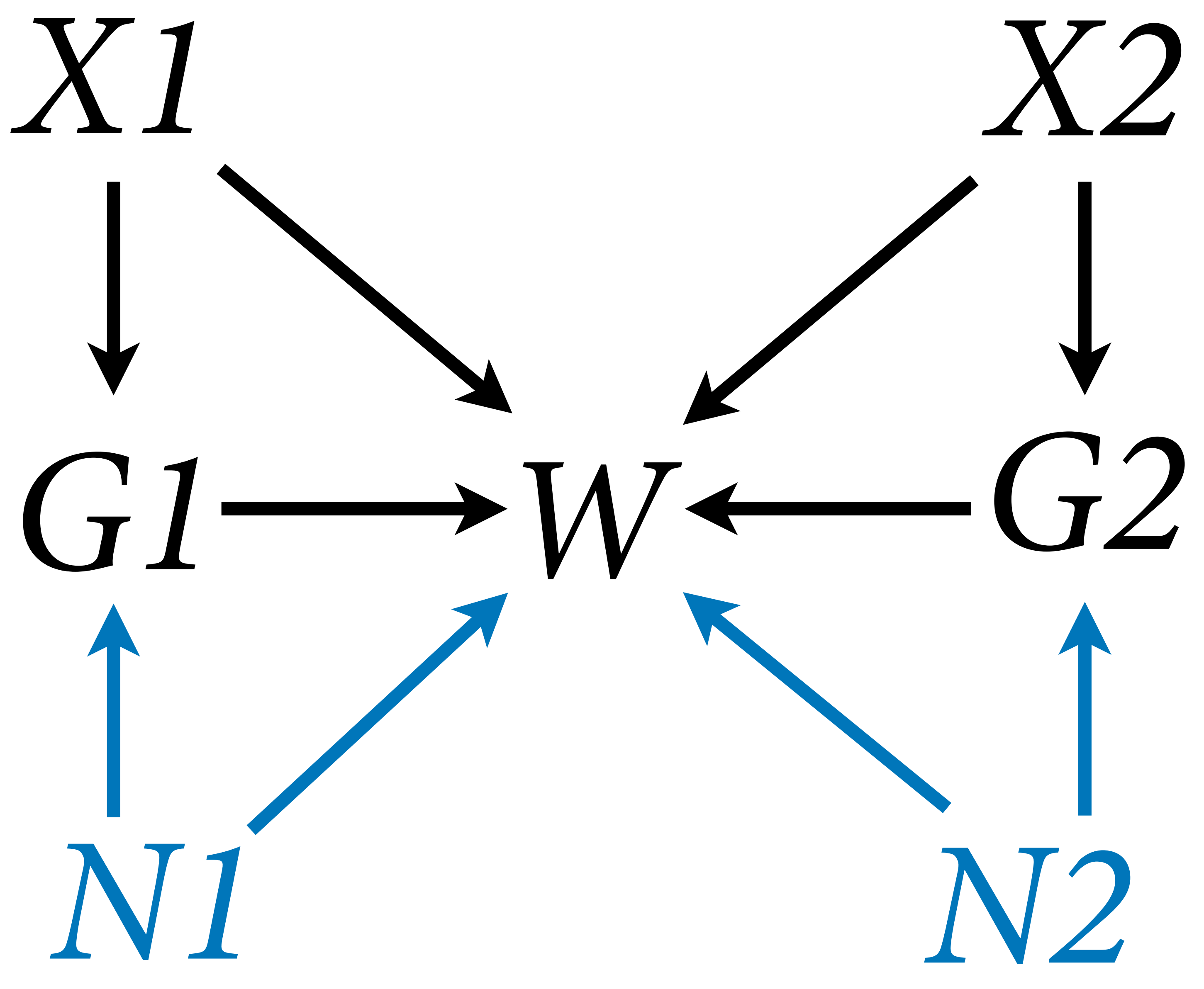
Varying effects as confounds

Causal perspective: Competing causes or actual confounds

Advantage over "fixed effect" approach: Can include other cluster-level (time invariant) causes

Fixed effects: Varying effects with variance fixed at infinity, no pooling

Don't panic: Make a generative model and draw the owl



Practical Difficulties

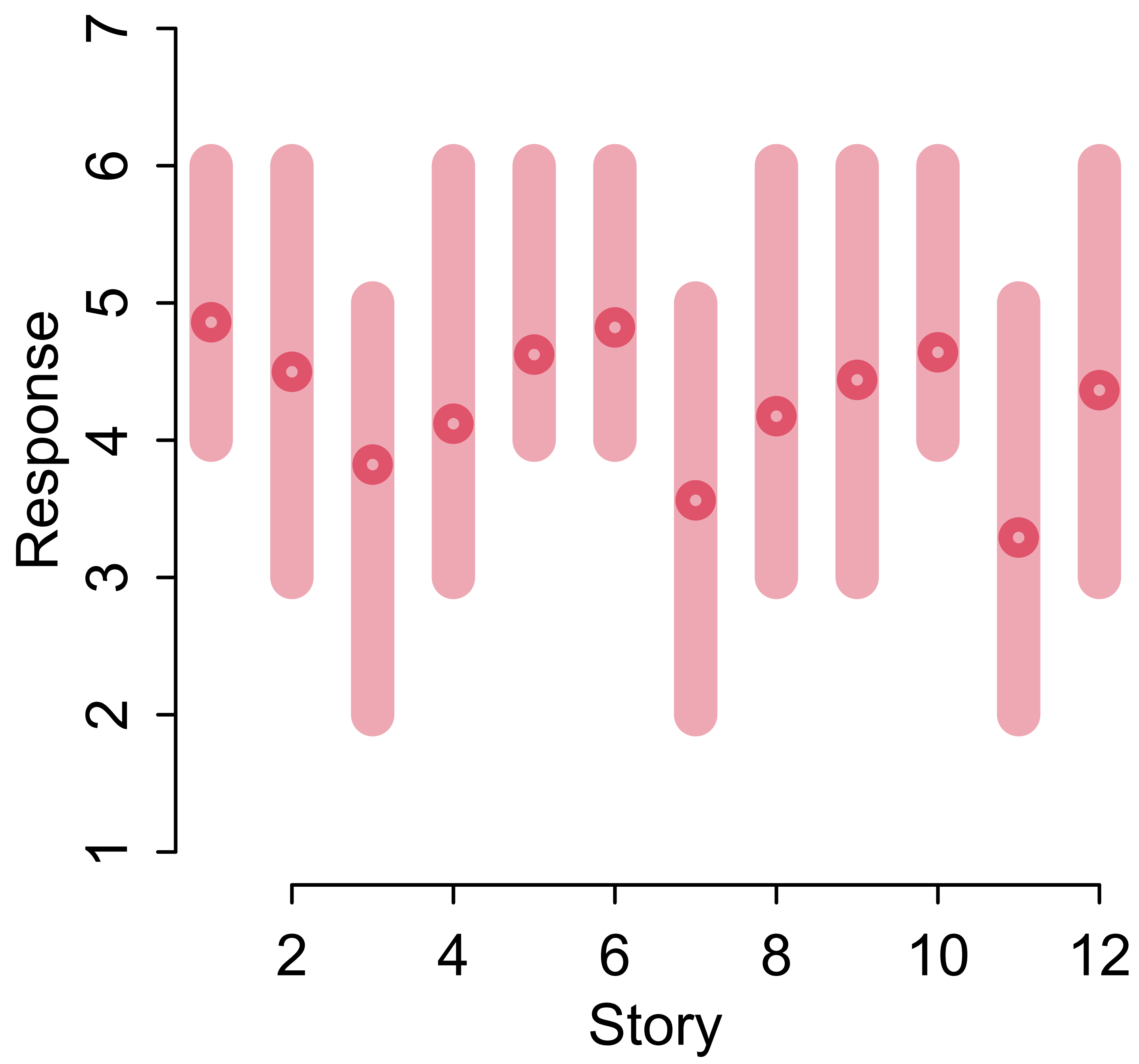
Varying effects are a good default, but…

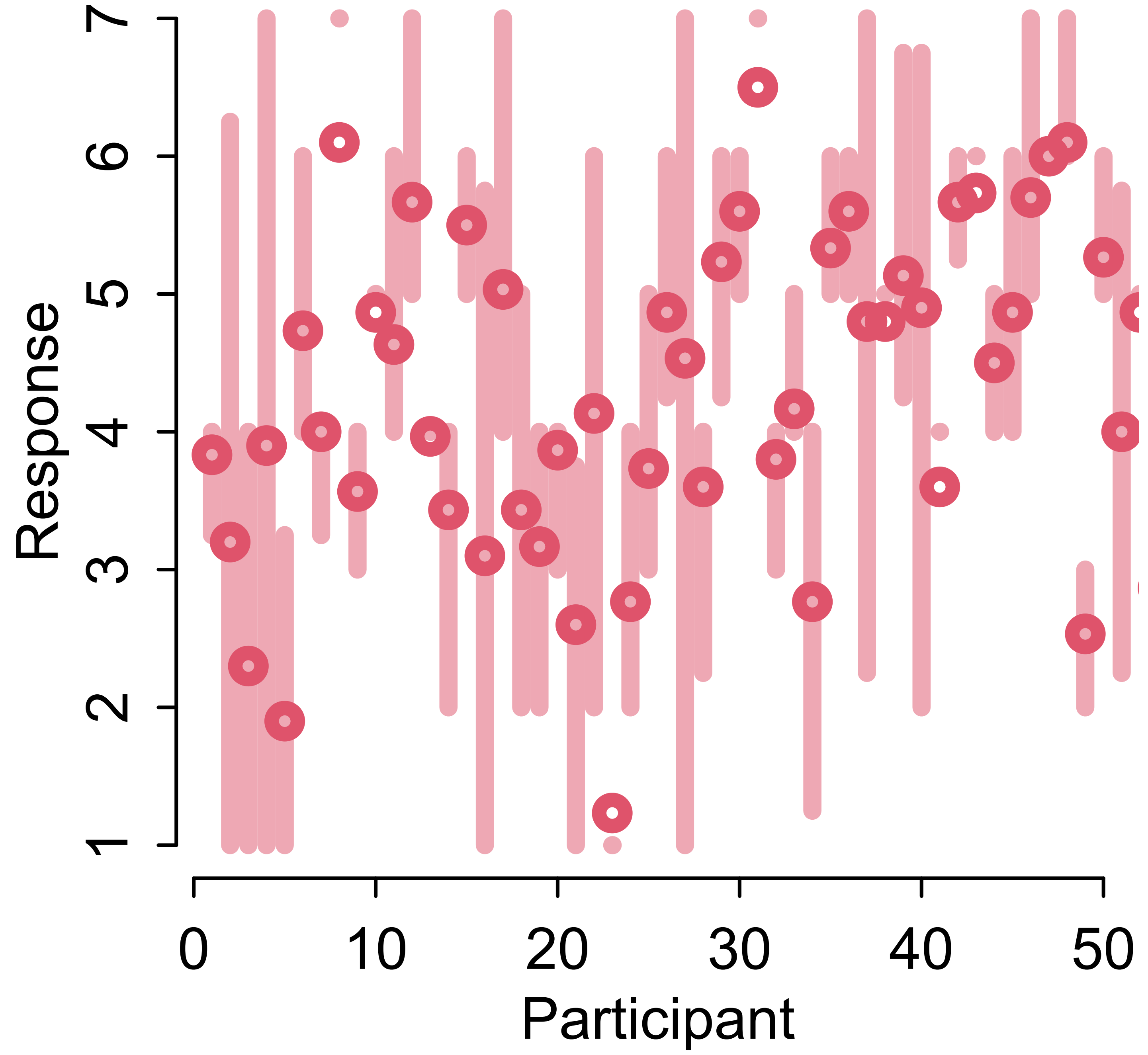
(1) How to use more than one cluster type at the same time?

(2) How to calculate predictions

(3) How to sample chains efficiently

(4) Group-level confounding





Fertility & behavior

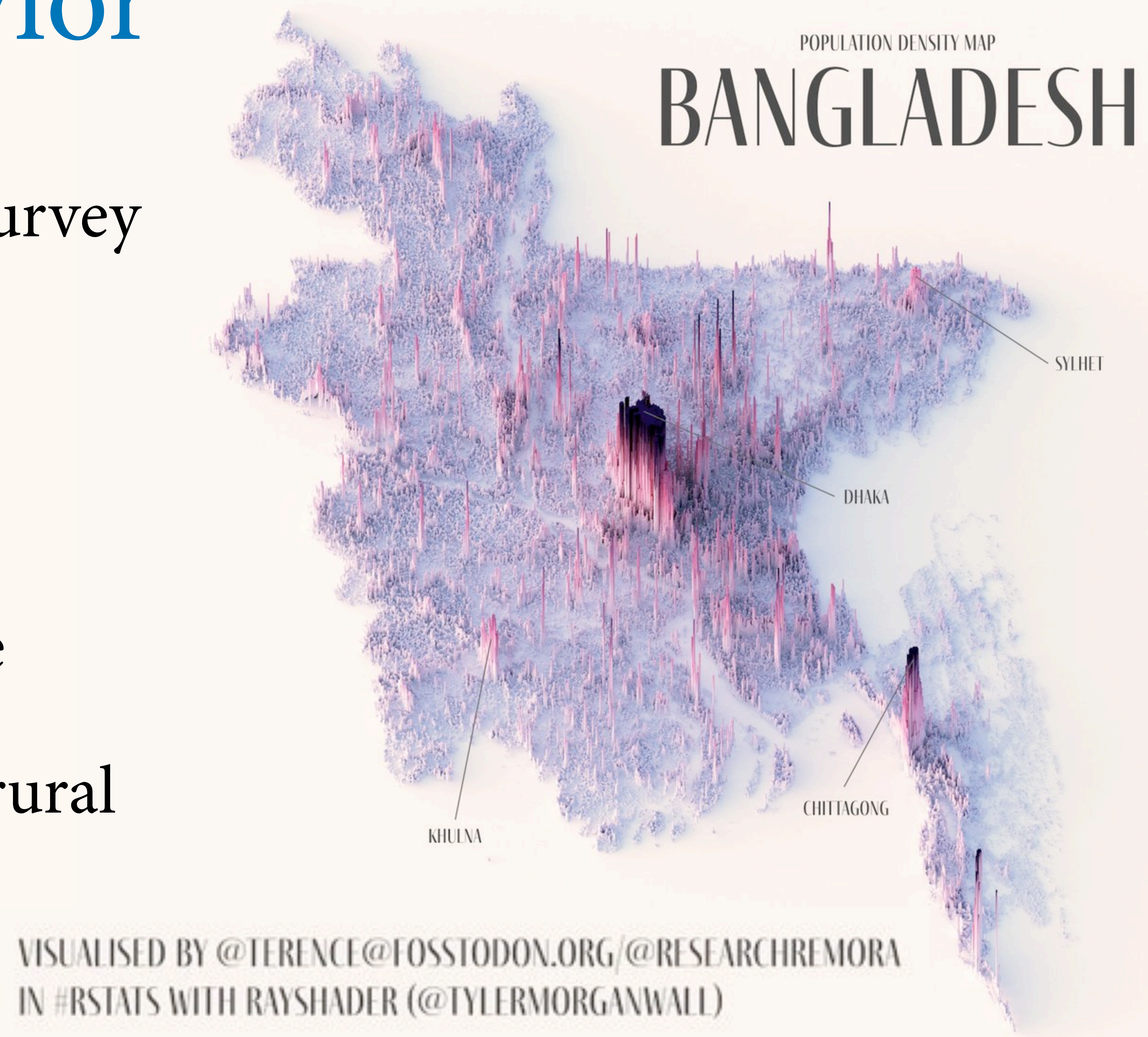
1989 Bangladesh Fertility Survey

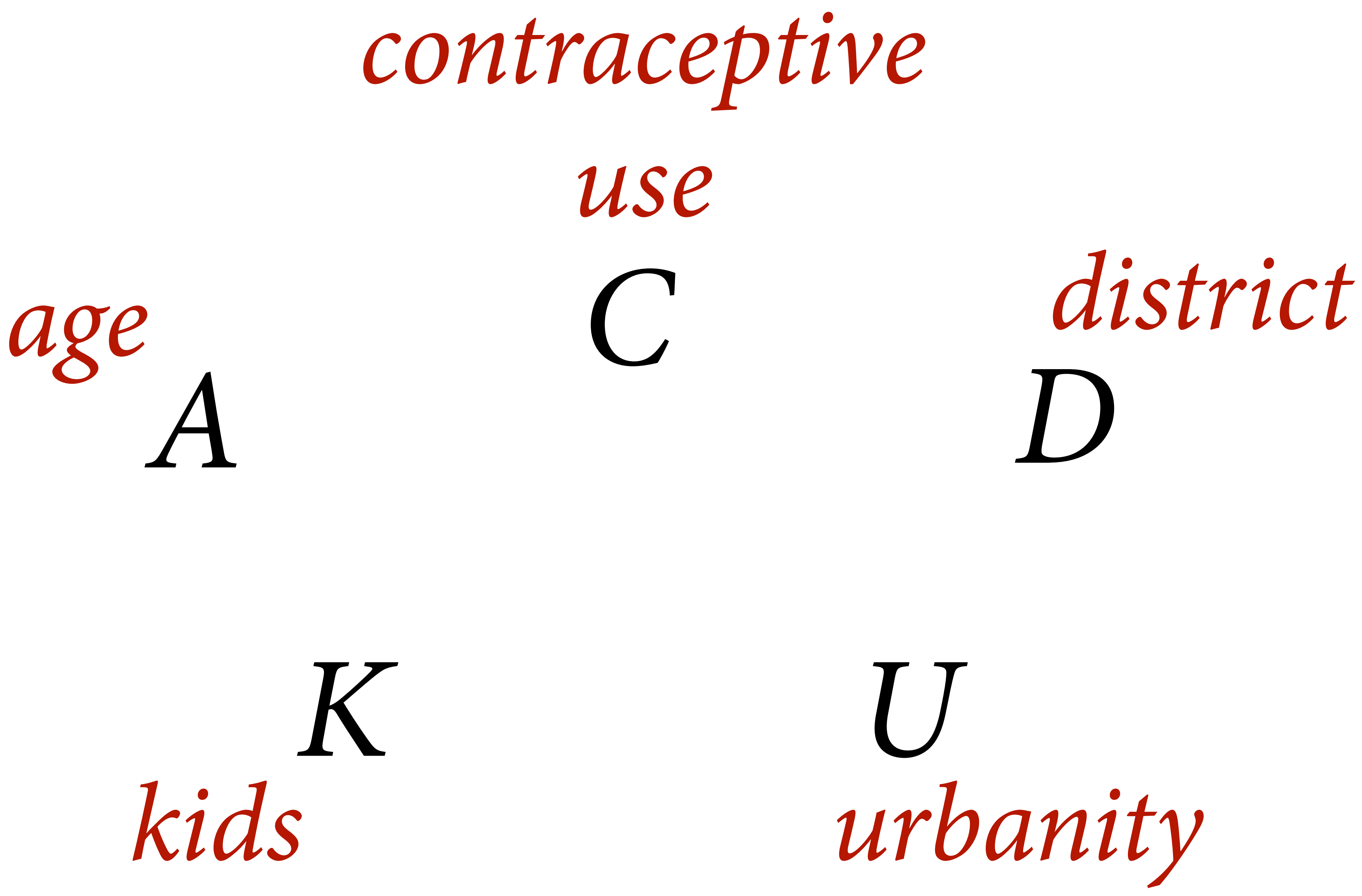
data(bangladesh)

1934 women, 61 districts

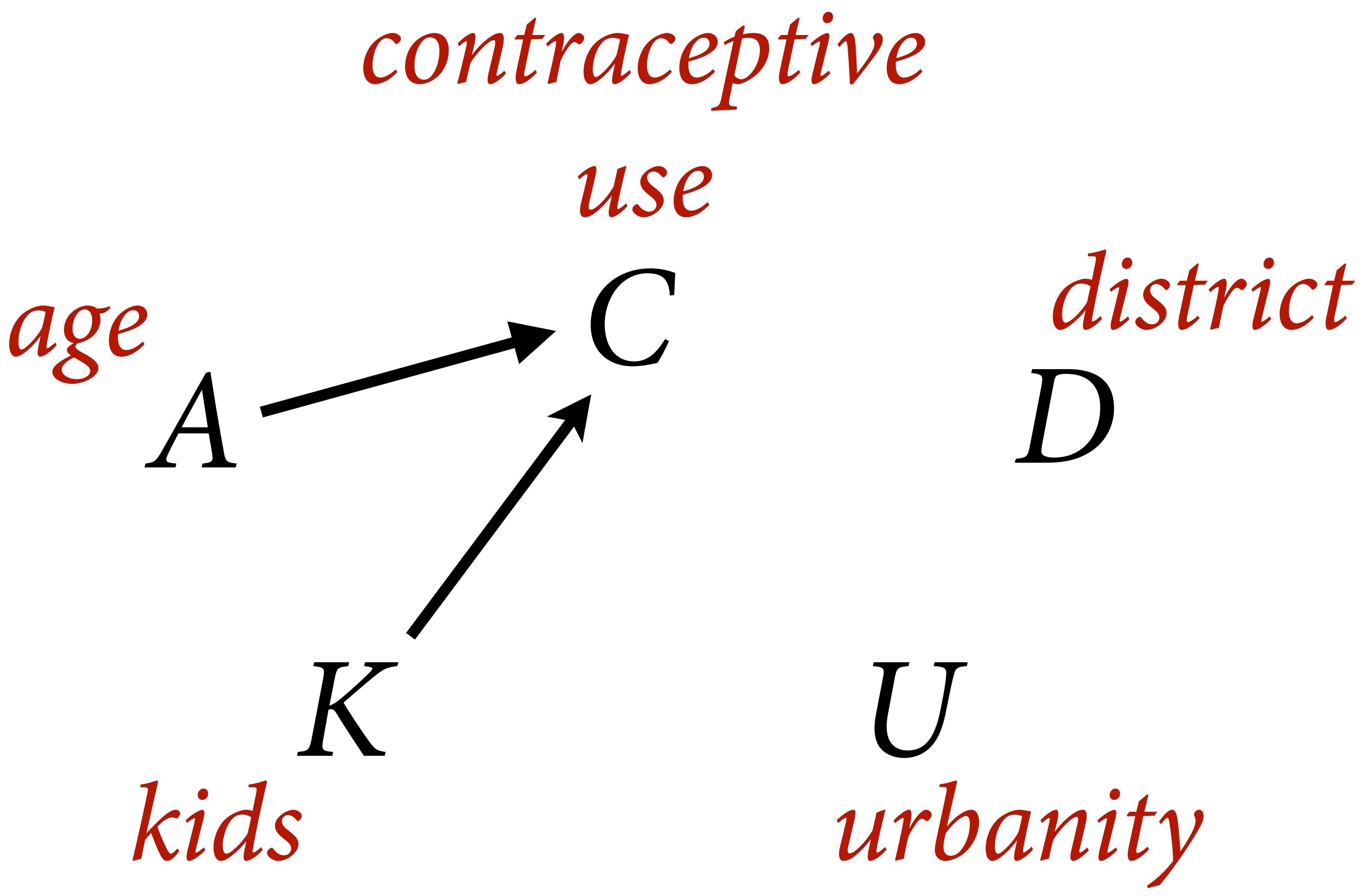
Outcome: contraceptive use

age, living children, urban/rural

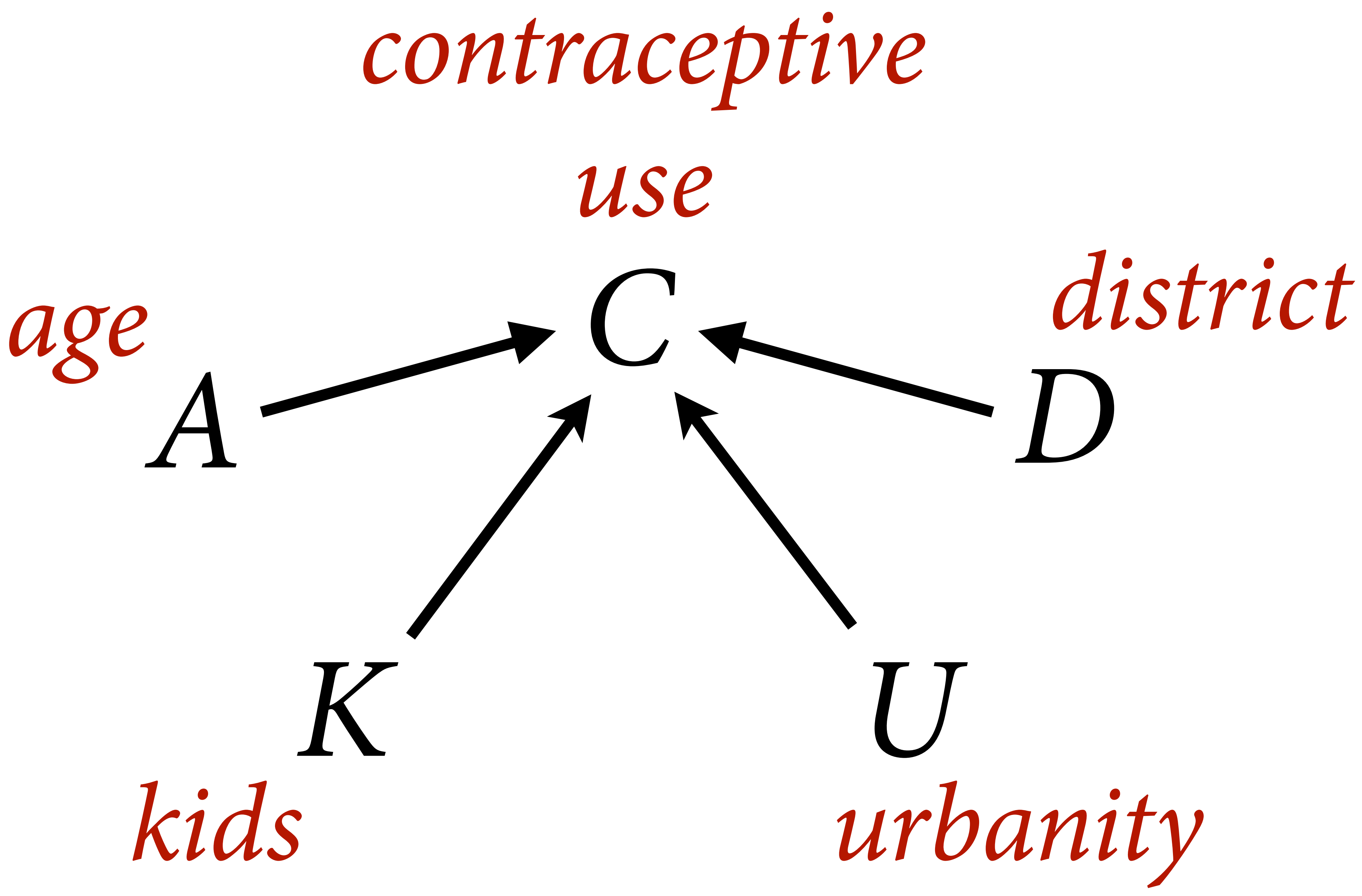



contraceptive
use
C
age
A
district
D
K
kids
U
urbanity

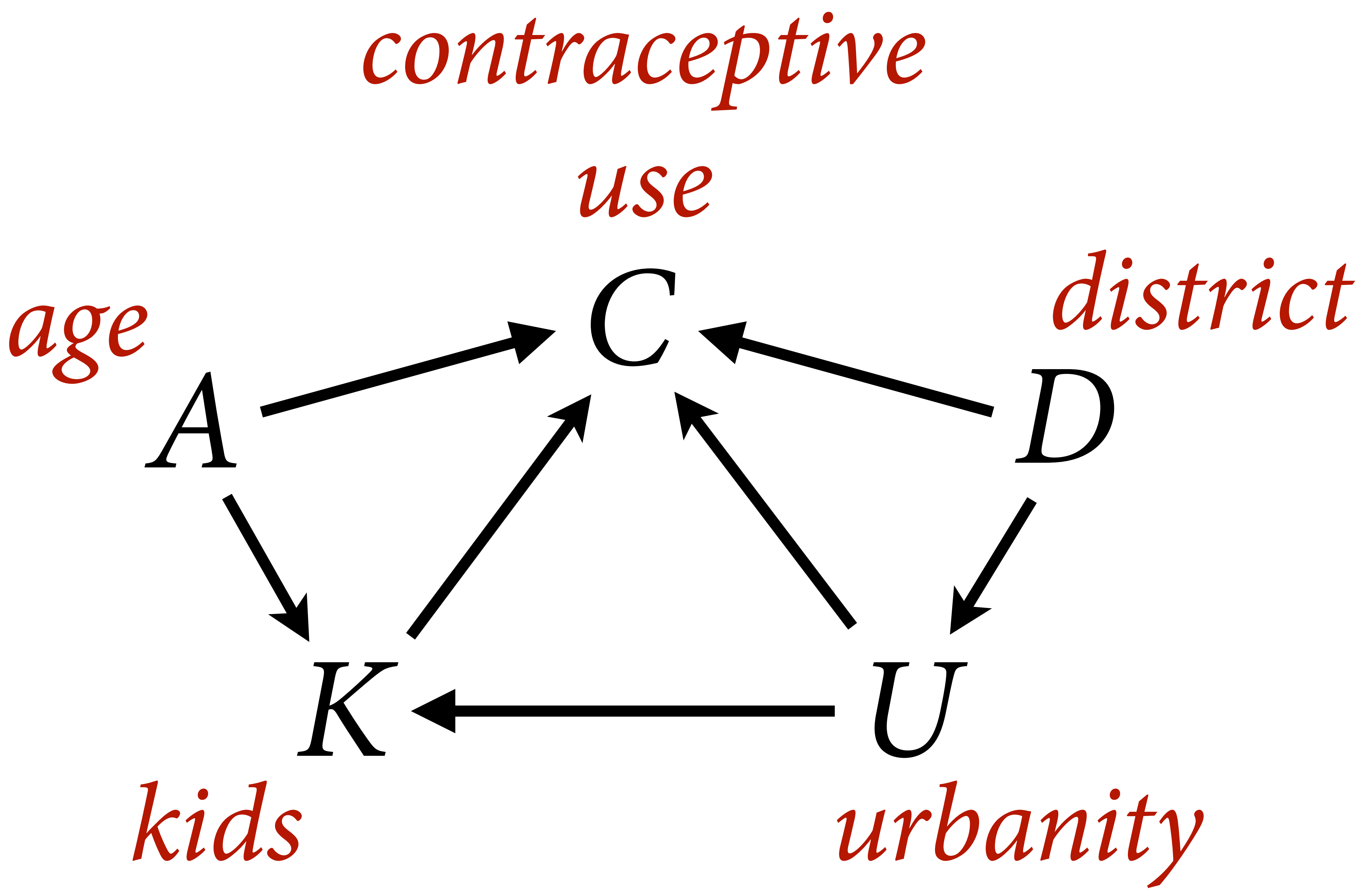
1. Causes of interest



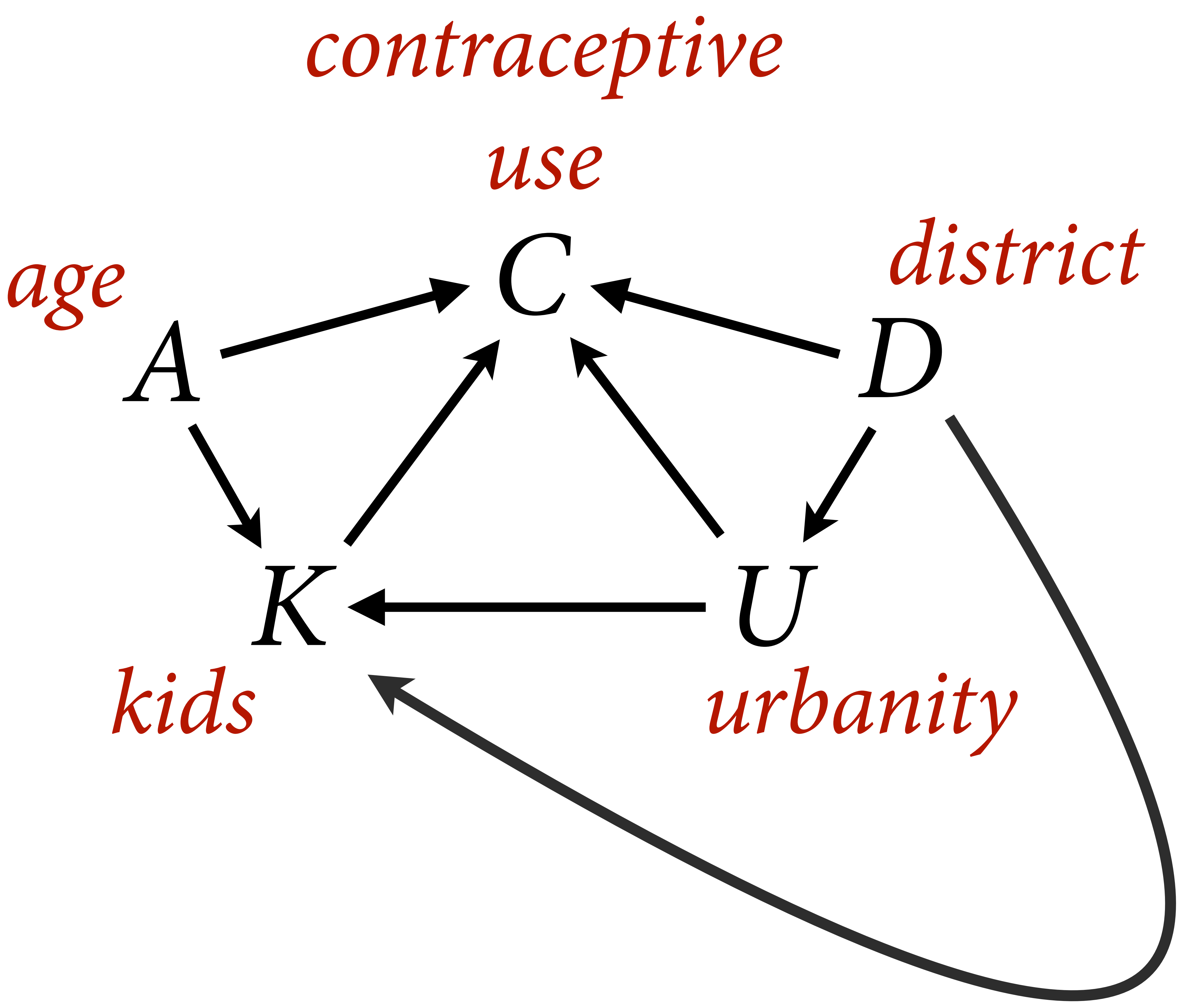
2. Competing causes



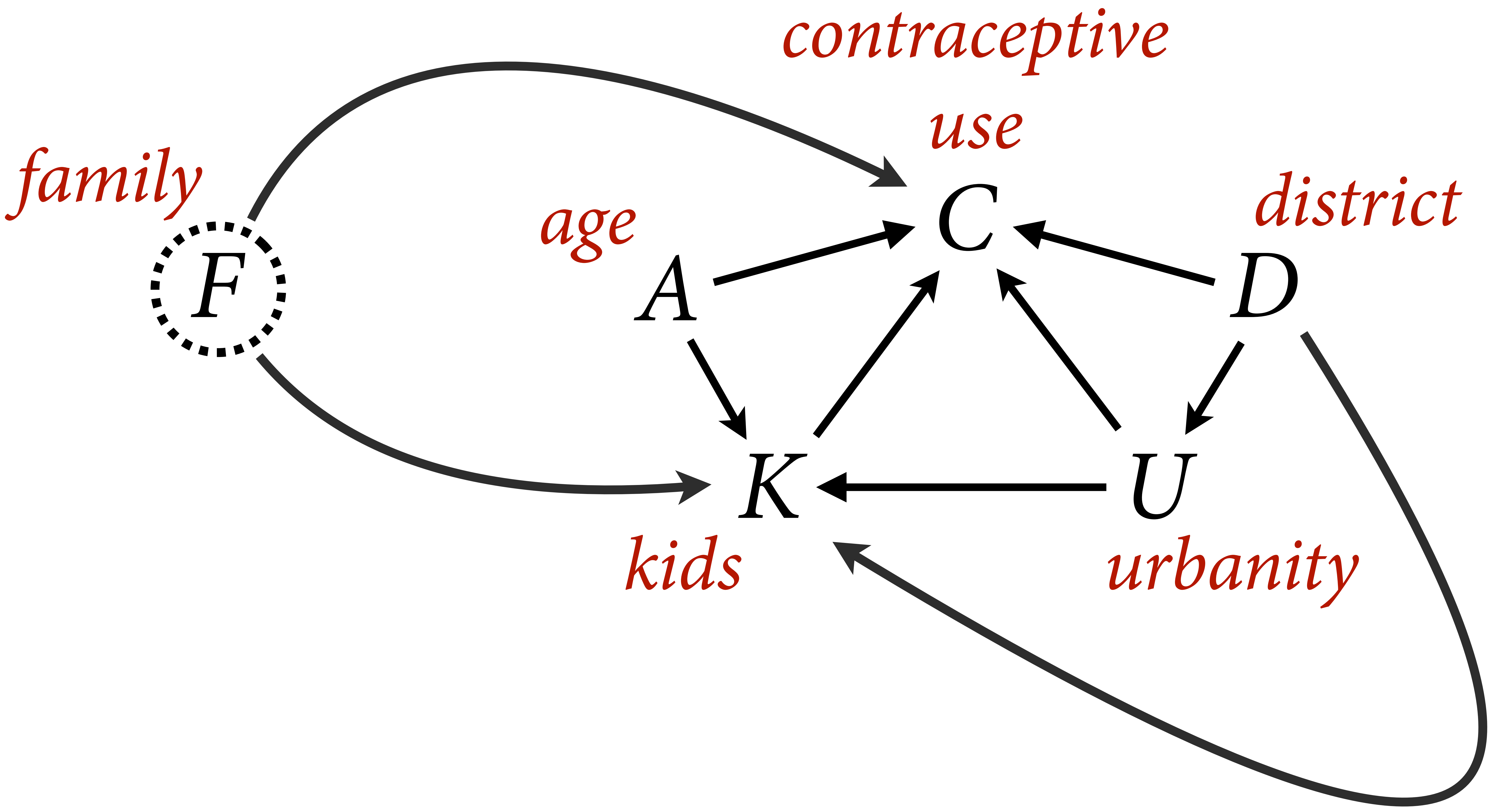
3. Relationships among causes



4. Unfortunate relationships among causes



5. A series of unfortunate relationships among causes

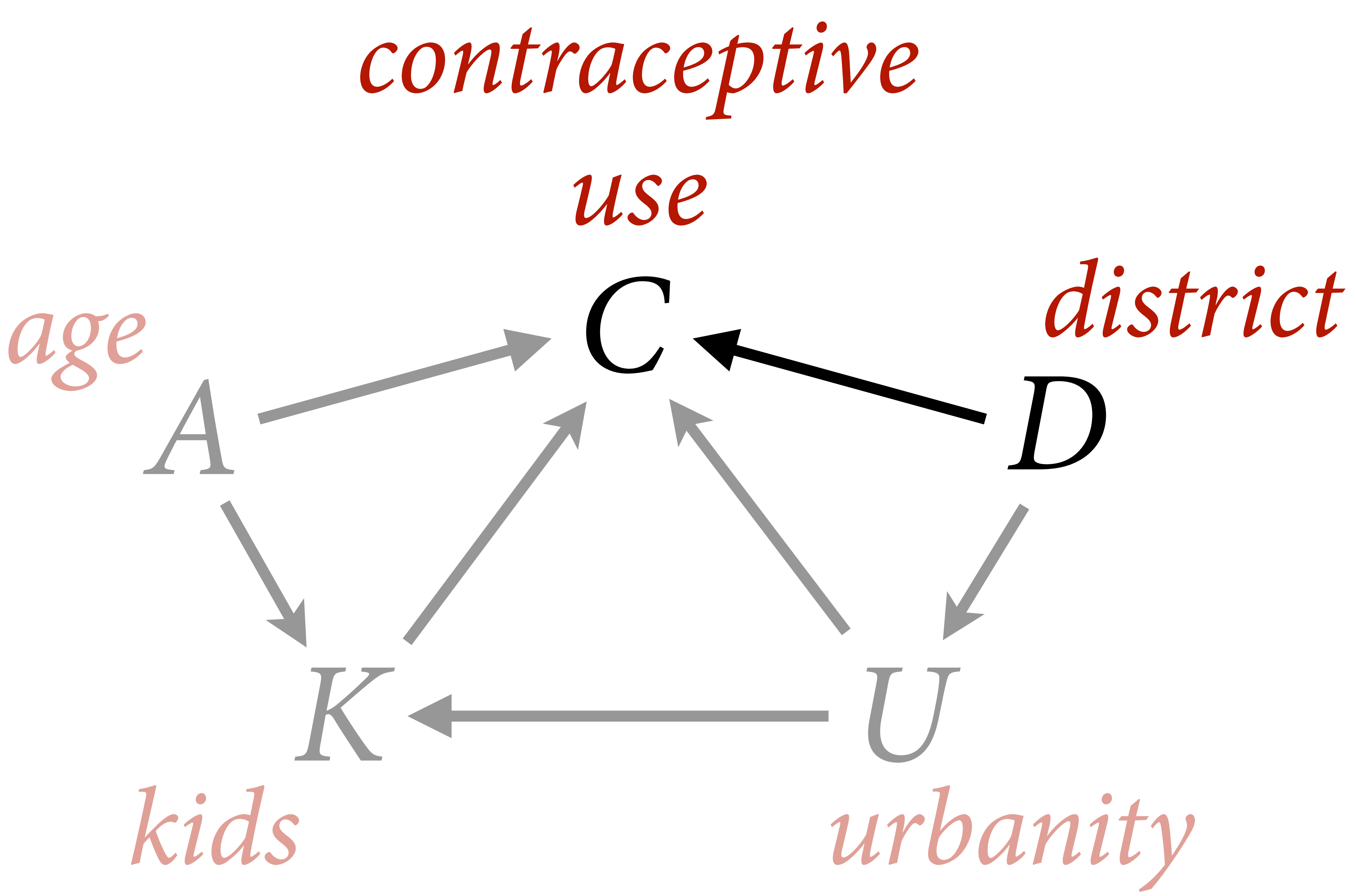


Varying districts

Estimand: C in each district, partially pooled

Varying intercept on each district

Another chance to understand partial pooling



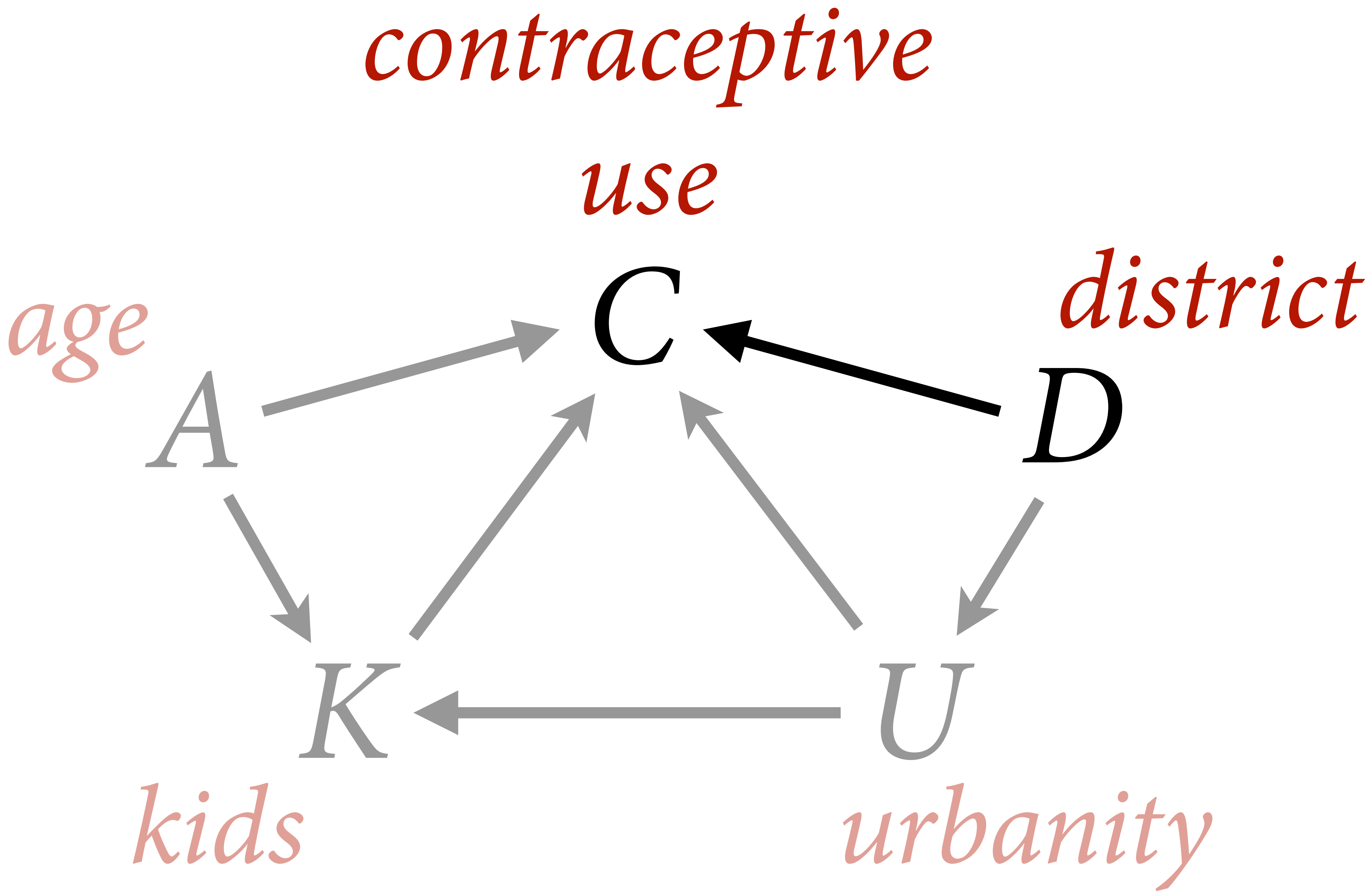
$$C_i \sim \text{Bernoulli}(p_i)$$

$$\text{logit}(p_i) = \alpha_{D[i]}$$

$$\alpha_j \sim \text{Normal}(\bar{\alpha}, \sigma)$$

$$\bar{\alpha} \sim \text{Normal}(0,1)$$

$$\sigma \sim \text{Exponential}(1)$$



$C_i \sim \text{Bernoulli}(p_i)$	Bernoulli because 0/1 outcome
$\text{logit}(p_i) = \alpha_{D[i]}$	log-odds of C=1 in each district D
$\alpha_j \sim \text{Normal}(\bar{\alpha}, \sigma)$	Regularizing prior for districts
$\bar{\alpha} \sim \text{Normal}(0,1)$	Average district
$\sigma \sim \text{Exponential}(1)$	Standard deviation among districts

```r
# simple varying intercepts model
library(rethinking)
data(bangladesh)
d <- bangladesh

dat <- list(
    C = d$use.contraception,
    D = as.integer(d$district) )

mCD <- ulam(
    alist(
        C ~ bernoulli(p),
        logit(p) <- a[D],
        vector[61]:a ~ normal(abar,sigma),
        abar ~ normal(0,1),
        sigma ~ exponential(1)
    ) , data=dat , chains=4 , cores=4 )
```

$$C_i \sim \text{Bernoulli}(p_i)$$

$$\text{logit}(p_i) = \alpha_{D[i]}$$

$$\alpha_j \sim \text{Normal}(\bar{\alpha}, \sigma)$$

$$\bar{\alpha} \sim \text{Normal}(0,1)$$

$$\sigma \sim \text{Exponential}(1)$$

```r
# simple varying intercepts model
library(rethinking)
data(bangladesh)
d <- bangladesh

dat <- list(
    C = d$use.contraception,
    D = as.integer(d$district) )

mCD <- ulam(
    alist(
        C ~ bernoulli(p),
        logit(p) <- a[D],
        vector[61]:a ~ normal(abar,sigma),
        abar ~ normal(0,1),
        sigma ~ exponential(1)
    ) , data=dat , chains=4 , cores=4 )
```

$$C_i \sim \text{Bernoulli}(p_i)$$

$$\text{logit}(p_i) = \alpha_{D[i]}$$

$$\alpha_j \sim \text{Normal}(\bar{\alpha}, \sigma)$$

$$\bar{\alpha} \sim \text{Normal}(0,1)$$

$$\sigma \sim \text{Exponential}(1)$$

```r
# simple varying intercepts model
library(rethinking)
data(bangladesh)
d <- bangladesh

dat <- list(
    C = d$use.contraception,
    D = as.integer(d$district) )

mCD <- ulam(
    alist(
        C ~ bernoulli(p),
        logit(p) <- a[D],
        vector[61]:a ~ normal(abar,sigma),
        abar ~ normal(0,1),
        sigma ~ exponential(1)
    ) , data=dat , chains=4 , cores=4 )
```

$$C_i \sim \text{Bernoulli}(p_i)$$

$$\text{logit}(p_i) = \alpha_{D[i]}$$

$$\alpha_j \sim \text{Normal}(\bar{\alpha}, \sigma)$$

$$\bar{\alpha} \sim \text{Normal}(0,1)$$

$$\sigma \sim \text{Exponential}(1)$$

```r
# simple varying intercepts model
library(rethinking)
data(bangladesh)
d <- bangladesh

dat <- list(
    C = d$use.contraception,
    D = as.integer(d$district) )

mCD <- ulam(
    alist(
        C ~ bernoulli(p),
        logit(p) <- a[D],
        vector[61]:a ~ normal(abar,sigma),
        abar ~ normal(0,1),
        sigma ~ exponential(1)
    ) , data=dat , chains=4 , cores=4 )
```

$$C_i \sim \text{Bernoulli}(p_i)$$

$$\text{logit}(p_i) = \alpha_{D[i]}$$

$$\alpha_j \sim \text{Normal}(\bar{\alpha}, \sigma)$$

$$\bar{\alpha} \sim \text{Normal}(0,1)$$

$$\sigma \sim \text{Exponential}(1)$$

```r
# simple varying intercepts model
library(rethinking)
data(bangladesh)
d <- bangladesh

dat <- list(
    C = d$use.contraception,
    D = as.integer(d$district) )

mCD <- ulam(
    alist(
        C ~ bernoulli(p),
        logit(p) <- a[D],
        vector[61]:a ~ normal(abar,sigma),
        abar ~ normal(0,1),
        sigma ~ exponential(1)
    ) , data=dat , chains=4 , cores=4 )
```

$$C_i \sim \text{Bernoulli}(p_i)$$

$$\text{logit}(p_i) = \alpha_{D[i]}$$

$$\alpha_j \sim \text{Normal}(\bar{\alpha}, \sigma)$$

$$\bar{\alpha} \sim \text{Normal}(0,1)$$

$$\sigma \sim \text{Exponential}(1)$$

```r
# simple varying intercepts model
library(rethinking)
data(bangladesh)
d <- bangladesh

dat <- list(
    C = d$use.contraception,
    D = as.integer(d$district) )

mCD <- ulam(
    alist(
        C ~ bernoulli(p),
        logit(p) <- a[D],
        vector[61]:a ~ normal(abar,sigma),
        abar ~ normal(0,1),
        sigma ~ exponential(1)
    ) , data=dat , chains=4 , cores=4 )
```

$$C_i \sim \text{Bernoulli}(p_i)$$

$$\text{logit}(p_i) = \alpha_{D[i]}$$

$$\alpha_j \sim \text{Normal}(\bar{\alpha}, \sigma)$$

$$\bar{\alpha} \sim \text{Normal}(0,1)$$

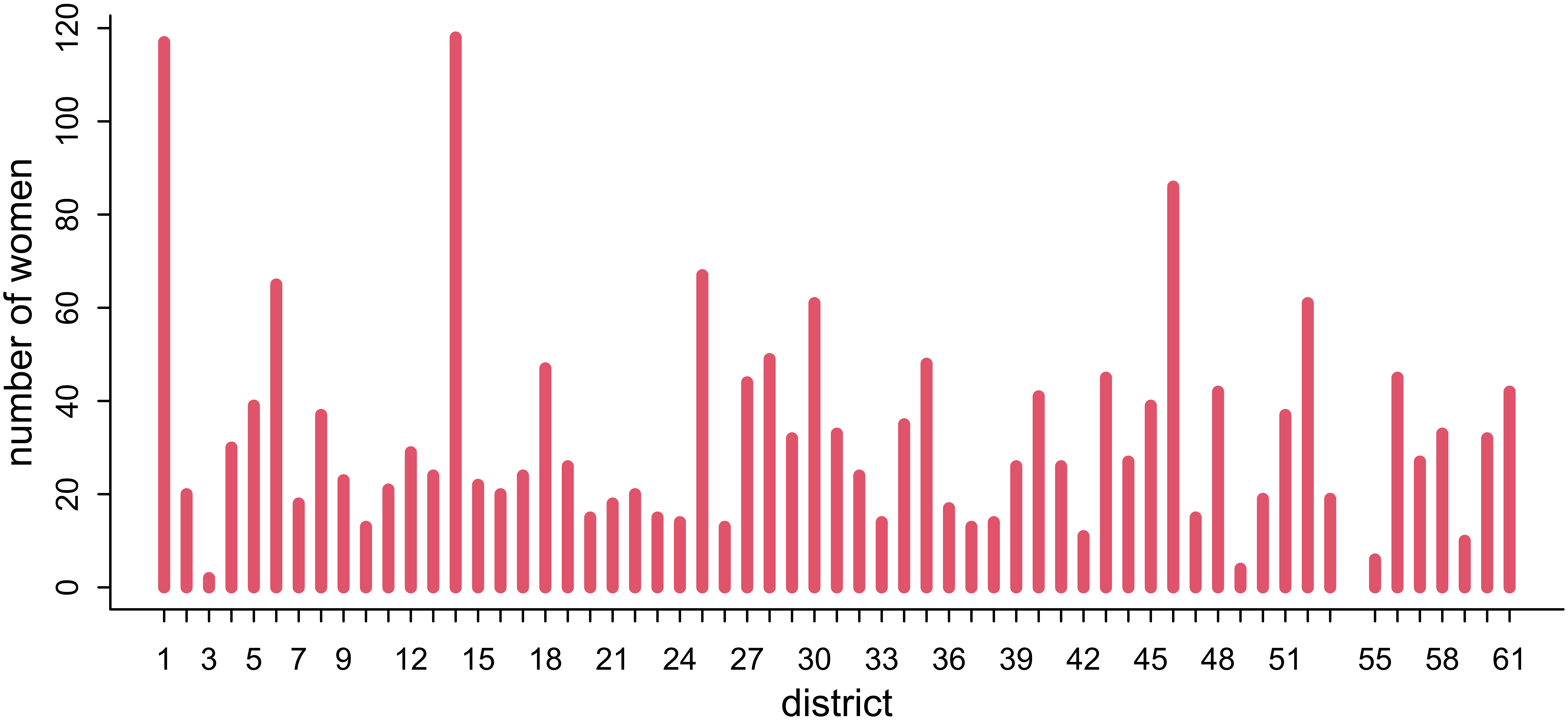
$$\sigma \sim \text{Exponential}(1)$$

```r
# simple varying intercepts model
library(rethinking)
data(bangladesh)
d <- bangladesh

dat <- list(
    C = d$use.contraception,
    D = as.integer(d$district) )

mCD <- ulam(
    alist(
        C ~ bernoulli(p),
        logit(p) <- a[D],
        vector[61]:a ~ normal(abar,sigma),
        abar ~ normal(0,1),
        sigma ~ exponential(1)
    ) , data=dat , chains=4 , cores=4 )
```

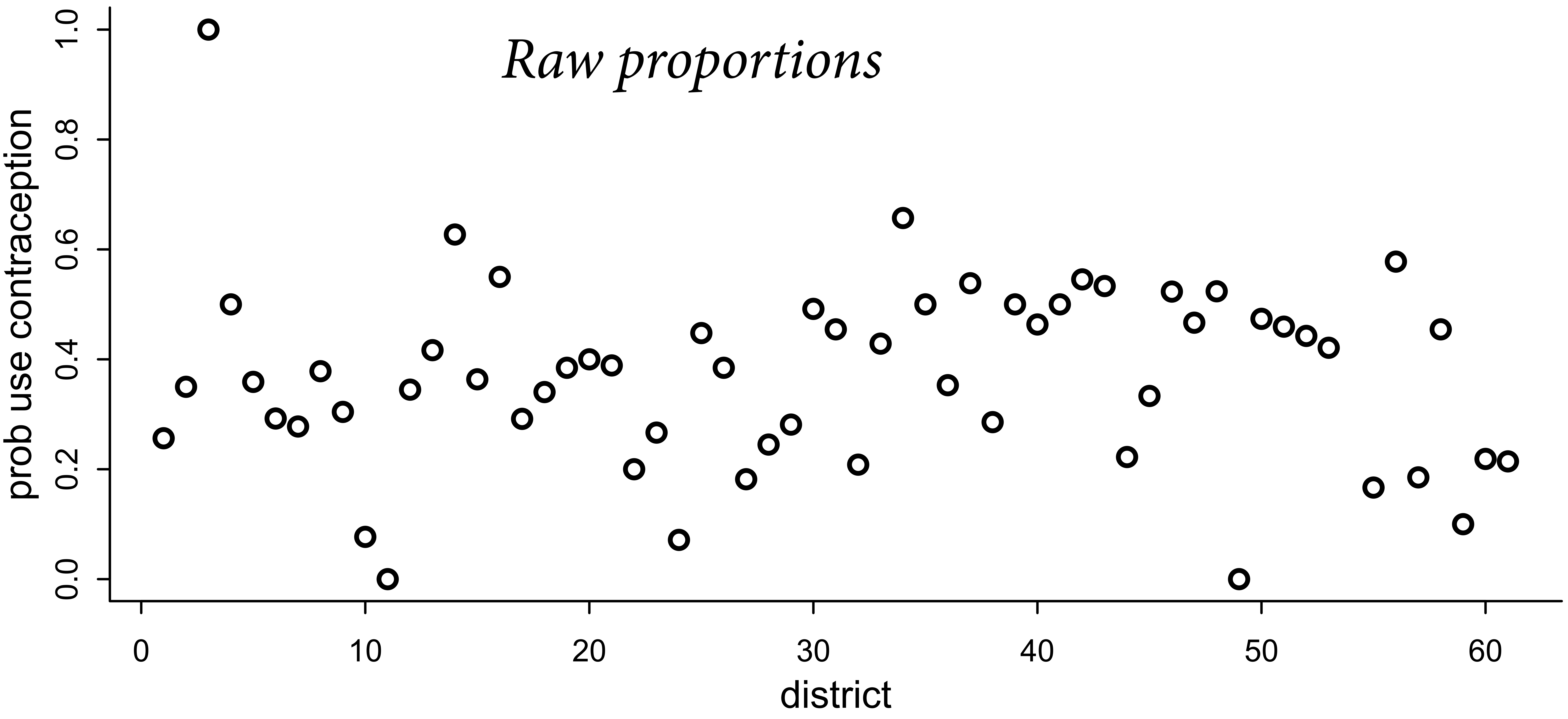
```
> precis(mCD,2)
       mean   sd  5.5% 94.5% n_eff Rhat4
a[1]  -0.99 0.20 -1.31 -0.67  3376  1.00
a[2]  -0.59 0.34 -1.16 -0.04  3847  1.00
a[3]  -0.22 0.50 -0.99  0.59  2918  1.00
a[4]  -0.18 0.30 -0.64  0.30  4063  1.00
a[5]  -0.58 0.28 -1.04 -0.16  3931  1.00
a[6]  -0.81 0.24 -1.21 -0.43  4513  1.00
a[7]  -0.76 0.38 -1.36 -0.15  3028  1.00
a[8]  -0.51 0.30 -0.98 -0.02  4220  1.00
a[9]  -0.71 0.34 -1.27 -0.18  3619  1.00
a[10] -1.14 0.43 -1.87 -0.48  2293  1.00
a[11] -1.54 0.43 -2.26 -0.88  2041  1.00
a[12] -0.61 0.32 -1.12 -0.11  4222  1.00
a[13] -0.43 0.33 -0.94  0.10  2602  1.00
a[14]  0.39 0.18  0.10  0.69  3264  1.00
a[15] -0.56 0.34 -1.11 -0.03  3745  1.00
a[16] -0.12 0.36 -0.68  0.46  3857  1.00
a[17] -0.75 0.34 -1.30 -0.20  4186  1.00
a[18] -0.64 0.25 -1.04 -0.26  3471  1.00
a[19] -0.50 0.33 -1.03  0.00  3715  1.00
a[20] -0.47 0.38 -1.09  0.13  4256  1.00
a[21] -0.50 0.36 -1.08  0.05  4264  1.00
a[22] -0.96 0.38 -1.59 -0.38  3273  1.00
a[23] -0.76 0.39 -1.38 -0.15  3241  1.00
a[24] -1.18 0.43 -1.91 -0.54  2182  1.00
a[25] -0.28 0.24 -0.66  0.09  3609  1.00
```

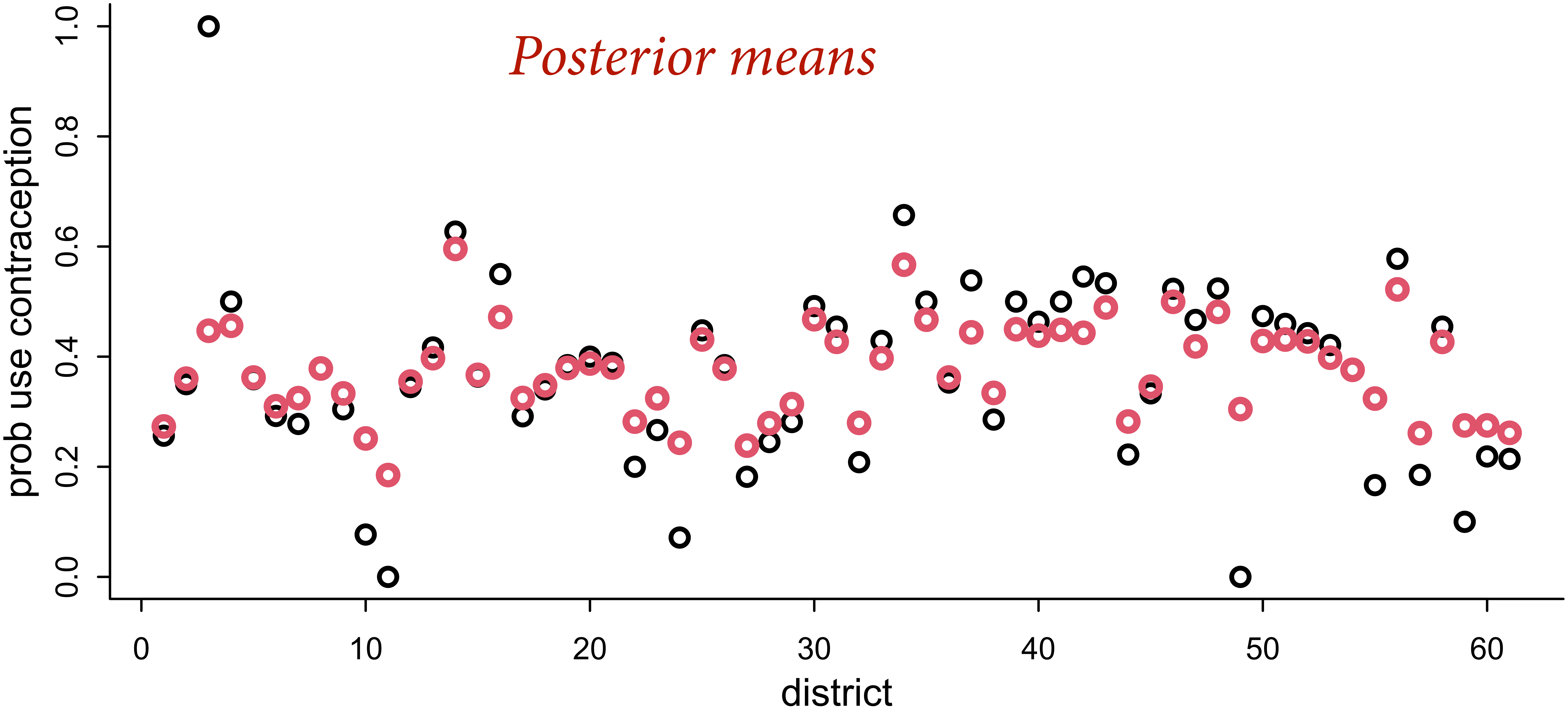
```
# simple varying intercepts model
library(rethinking)
data(bangladesh)
d <- bangladesh

dat <- list(
    C = d$use.contraception,
    D = as.integer(d$district) )

mCD <- ulam(
    alist(
        C ~ bernoulli(p),
        logit(p) <- a[D],
        vector[61]:a ~ normal(abar,sigma),
        abar ~ normal(0,1),
        sigma ~ exponential(1)
    ) , data=dat , chains=4 , cores=4 )
```

```
a[36] -0.58 0.36 -1.16 -0.01  3809  1.00
a[37] -0.23 0.38 -0.82  0.40  3096  1.00
a[38] -0.71 0.38 -1.33 -0.11  3188  1.00
a[39] -0.21 0.30 -0.69  0.27  3537  1.00
a[40] -0.25 0.25 -0.66  0.16  3996  1.00
a[41] -0.21 0.31 -0.71  0.28  3886  1.00
a[42] -0.24 0.40 -0.87  0.42  4422  1.00
a[43] -0.04 0.26 -0.46  0.36  4067  1.00
a[44] -0.96 0.33 -1.50 -0.44  3195  1.00
a[45] -0.65 0.28 -1.09 -0.21  4171  1.00
a[46]  0.00 0.20 -0.32  0.31  3739  1.00
a[47] -0.34 0.36 -0.91  0.24  3204  1.00
a[48] -0.08 0.27 -0.50  0.35  3590  1.00
a[49] -0.87 0.48 -1.63 -0.14  2809  1.00
a[50] -0.30 0.34 -0.82  0.24  4255  1.00
a[51] -0.28 0.29 -0.74  0.17  3392  1.00
a[52] -0.30 0.23 -0.67  0.08  3494  1.00
a[53] -0.43 0.35 -0.98  0.14  3814  1.00
a[54] -0.54 0.52 -1.40  0.28  3546  1.00
a[55] -0.77 0.46 -1.56 -0.07  3145  1.00
a[56]  0.09 0.27 -0.34  0.53  2499  1.00
a[57] -1.07 0.35 -1.67 -0.53  2797  1.00
a[58] -0.30 0.30 -0.75  0.17  3378  1.00
a[59] -1.01 0.43 -1.74 -0.35  2425  1.00
a[60] -0.99 0.33 -1.53 -0.48  3479  1.00
a[61] -1.06 0.30 -1.54 -0.59  3251  1.00
abar  -0.54 0.09 -0.68 -0.40  1734  1.00
sigma  0.52 0.09  0.39  0.67   626  1.01
```

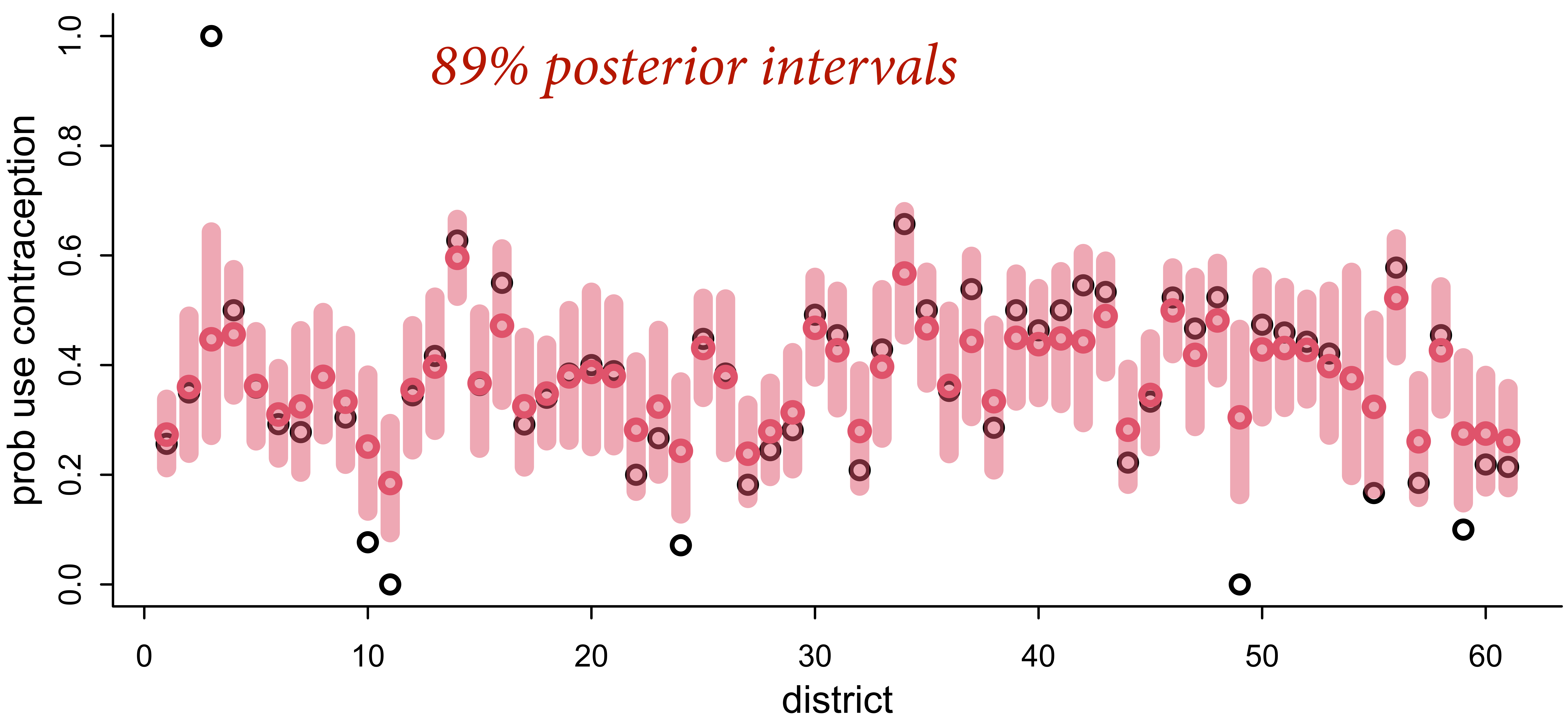


number of women
0
20
40
60
80
100
120
1
3
5
7
9
12
15
18
21
24
27
30
33
36
39
42
45
48
51
55
58
61
district

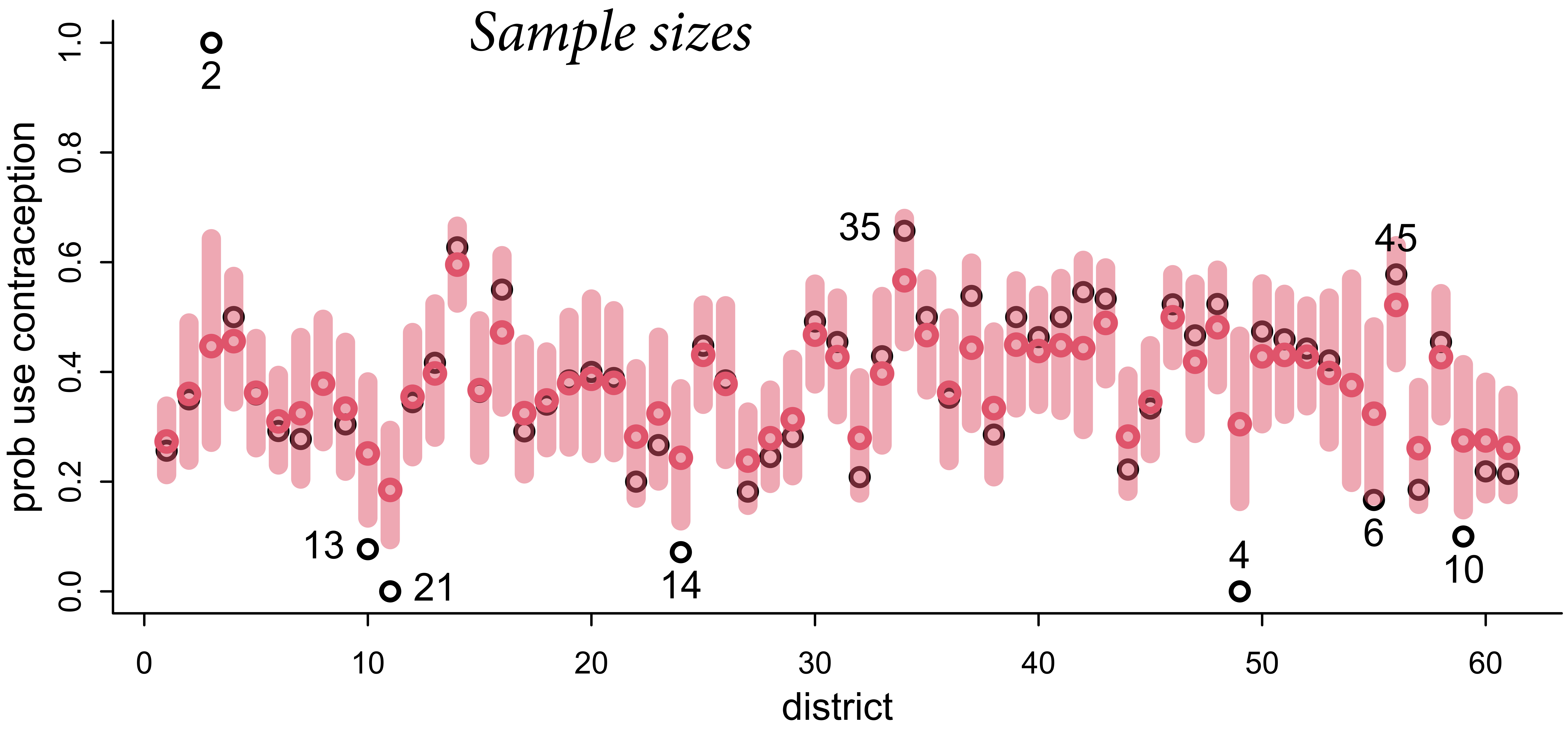

Raw proportions
prob use contraception
0.0
0.2
0.4
0.6
0.8
1.0
0
10
20
30
40
50
60
district



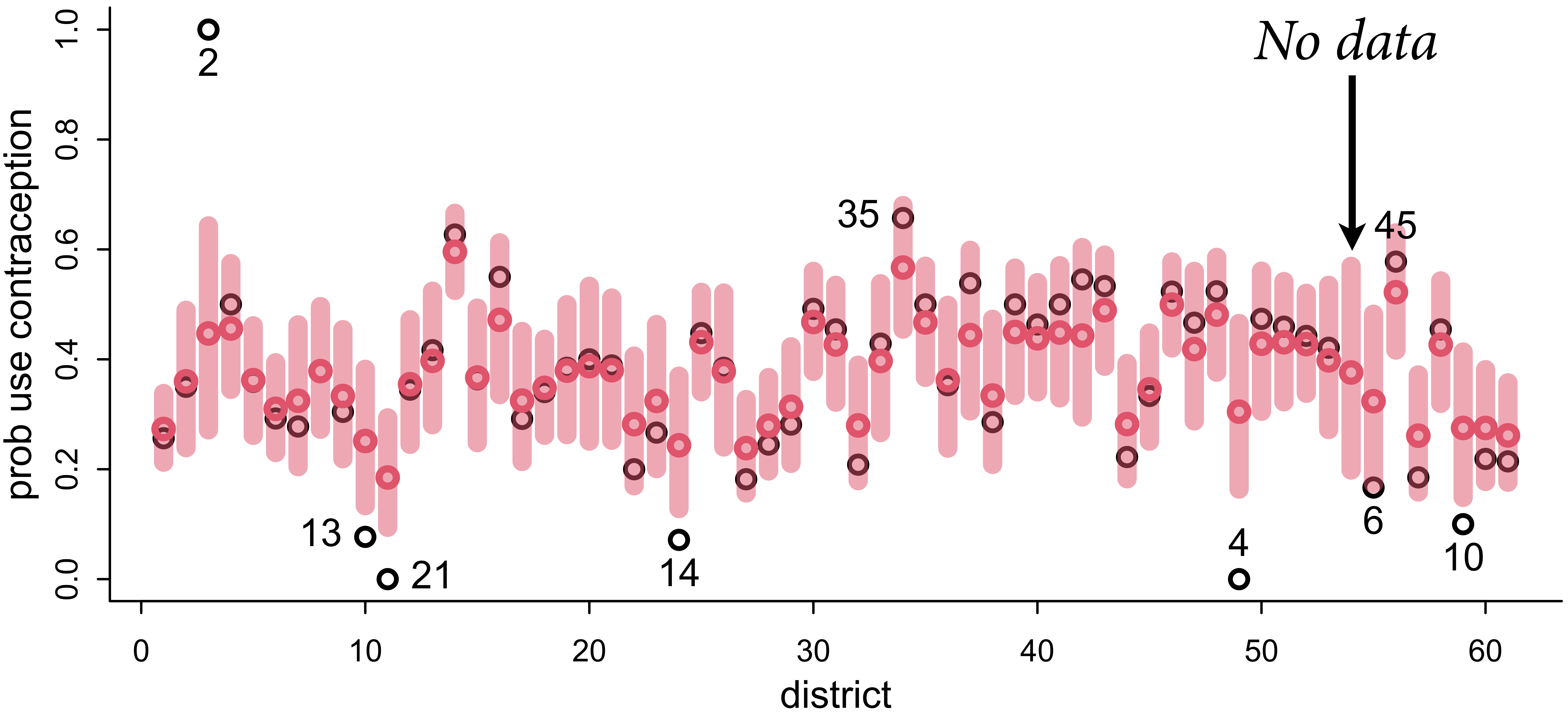
Posterior means
prob use contraception
0.0
0.2
0.4
0.6
0.8
1.0
district
0
10
20
30
40
50
60



89% posterior intervals
prob use contraception
0.0
0.2
0.4
0.6
0.8
1.0
0
10
20
30
40
50
60
district

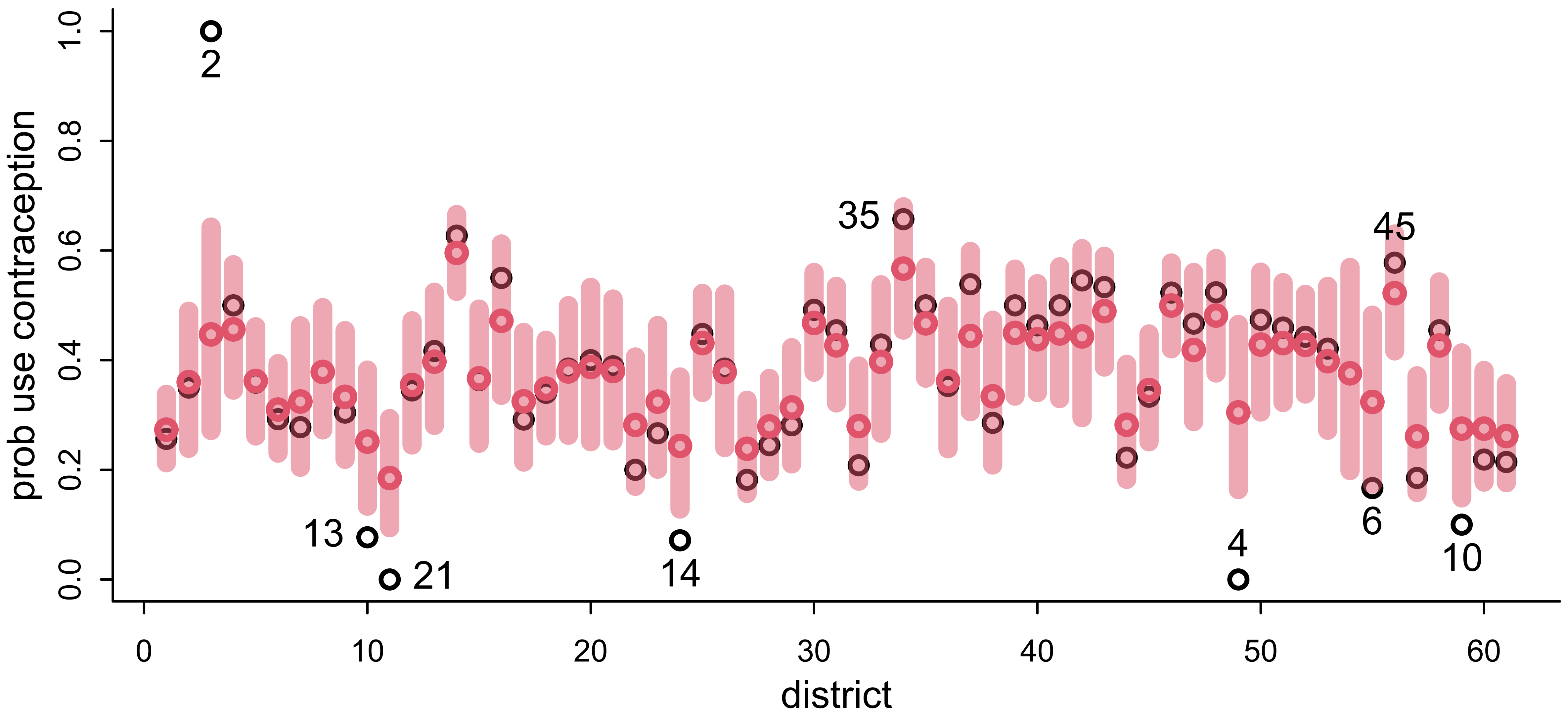


Sample sizes
prob use contraception
district
0.0
0.2
0.4
0.6
0.8
1.0
0
10
20
30
40
50
60
2
13
21
14
35
45
4
6
10



No data
prob use contraception
district
2
35
45
13
21
14
4
6
10
0.0
0.2
0.4
0.6
0.8
1.0
0
10
20
30
40
50
60

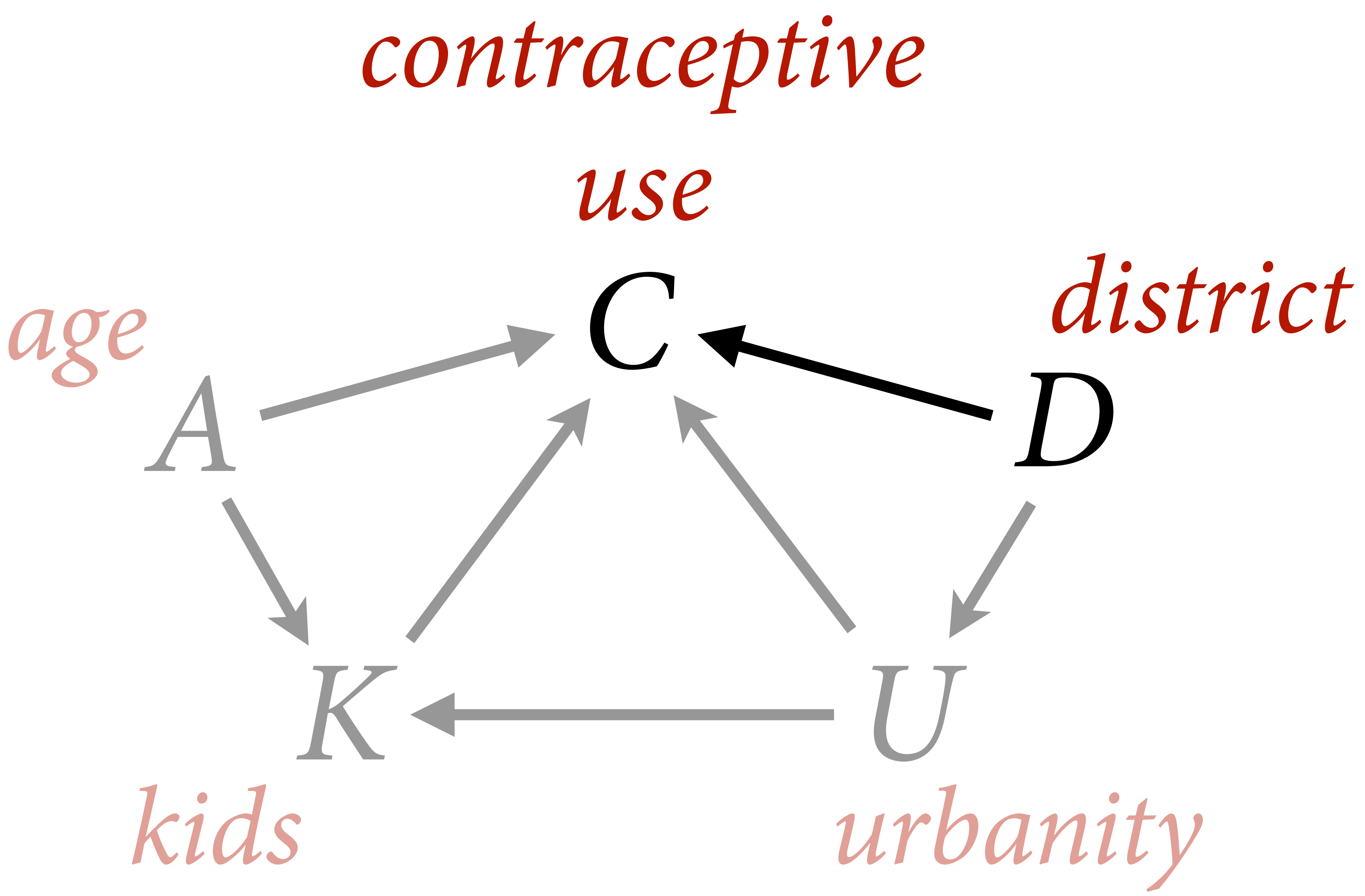
Varying districts



Partial pooling shrinks districts with low sampling towards mean

Better predictions

No inference yet



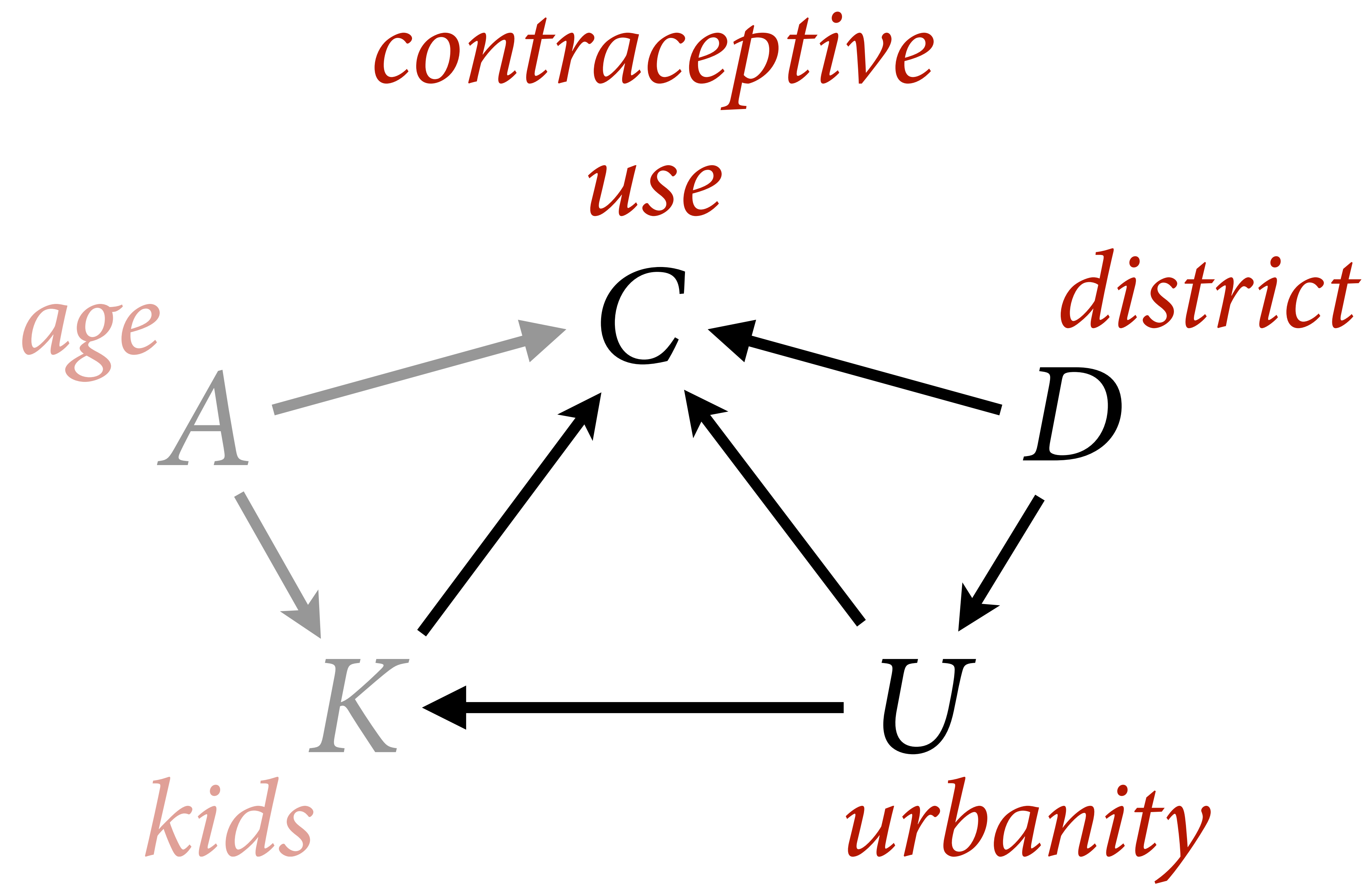
Varying districts + urban

What is the effect of urban living?

District features are potential group-level confounds

Total effect of *U* passes through *K*

Do not stratify by *K*!



$$C_i \sim \text{Bernoulli}(p_i)$$

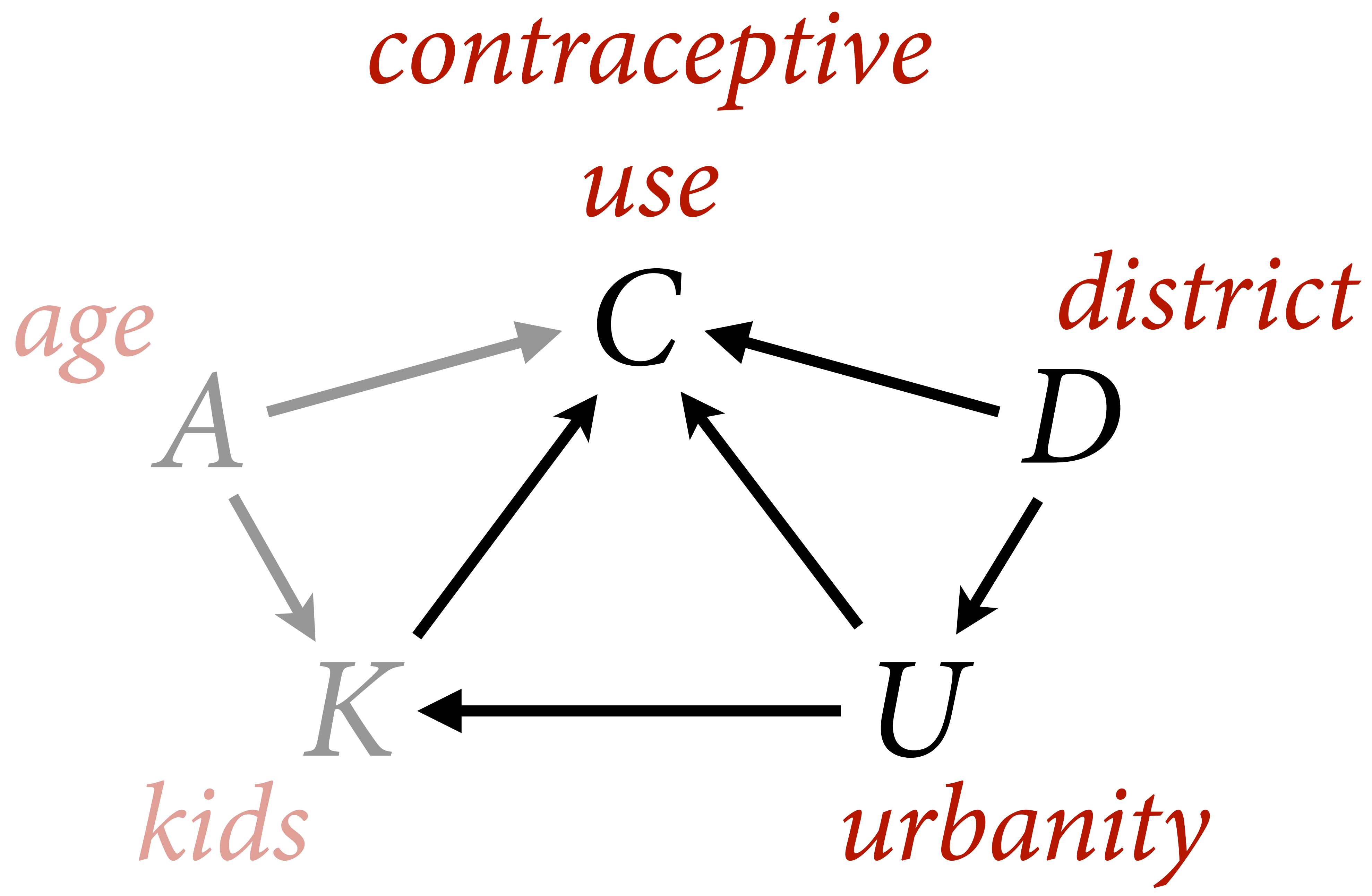
$$\text{logit}(p_i) = \alpha_{D[i]} + \beta_{D[i]} U_i$$

$$\alpha_j \sim \text{Normal}(\bar{\alpha}, \sigma)$$

$$\beta_j \sim \text{Normal}(\bar{\beta}, \tau)$$

$$\bar{\alpha}, \bar{\beta} \sim \text{Normal}(0,1)$$

$$\sigma, \tau \sim \text{Exponential}(1)$$



$$C_i \sim \text{Bernoulli}(p_i)$$

$$\text{logit}(p_i) = \alpha_{D[i]} + \beta_{D[i]} U_i$$

$$\alpha_j \sim \text{Normal}(\bar{\alpha}, \sigma)$$ Regularizing prior for rural

$$\beta_j \sim \text{Normal}(\bar{\beta}, \tau)$$ Regularizing prior for urban effect

$$\bar{\alpha}, \bar{\beta} \sim \text{Normal}(0,1)$$ Averages

$$\sigma, \tau \sim \text{Exponential}(1)$$ Standard deviations

```r
dat <- list(
    C = d$use.contraception,
    D = as.integer(d$district),
    U = ifelse(d$urban==1,1,0) )

# total U
mCDU <- ulam(
    alist(
        C ~ bernoulli(p),
        logit(p) <- a[D] + b[D]*U,
        vector[61]:a ~ normal(abar,sigma),
        vector[61]:b ~ normal(bbar,tau),
        c(abar,bbar) ~ normal(0,1),
        c(sigma,tau) ~ exponential(1)
    ) , data=dat , chains=4 , cores=4 )
```

$$C_i \sim \text{Bernoulli}(p_i)$$

$$\text{logit}(p_i) = \alpha_{D[i]} + \beta_{D[i]} U_i$$

$$\alpha_j \sim \text{Normal}(\bar{\alpha}, \sigma)$$

$$\beta_j \sim \text{Normal}(\bar{\beta}, \tau)$$

$$\bar{\alpha}, \bar{\beta} \sim \text{Normal}(0,1)$$

$$\sigma, \tau \sim \text{Exponential}(1)$$

```r
dat <- list(
    C = d$use.contraception,
    D = as.integer(d$district),
    U = ifelse(d$urban==1,1,0) )

# total U
mCDU <- ulam(
    alist(
        C ~ bernoulli(p
        logit(p) <- a[D
        vector[61]:a ~
        vector[61]:b ~
        c(abar,bbar) ~
        c(sigma,tau) ~
    ) , data=dat , chains=4 , cores=4 )
```

```
All 4 chains finished successfully.
Mean chain execution time: 4.3 seconds.
Total execution time: 4.7 seconds.

Warning: 4 of 2000 (0.0%) transitions ended with a divergence.
See https://mc-stan.org/misc/warnings for details.

Warning: 3 of 4 chains had an E-BFMI less than 0.2.
See https://mc-stan.org/misc/warnings for details.
```

$$C_i \sim \text{Bernoulli}(p_i)$$

$$\text{logit}(p_i) = \alpha_{D[i]} + \beta_{D[i]} U_i$$

$$\sigma, \tau \sim \text{Exponential}(1)$$

```
dat <- list(
    C = d$use.contraception,
    D = as.in
    U = ifels

# total U
mCDU <- ulam(
    alist(
        C ~ b
        logit(p) <- a[D
        vector[61]:a ~
        vector[61]:b ~
        c(abar,bbar) ~
        c(sigma,tau) ~
    ) , data=dat , chains=4 , cores=4 )
```

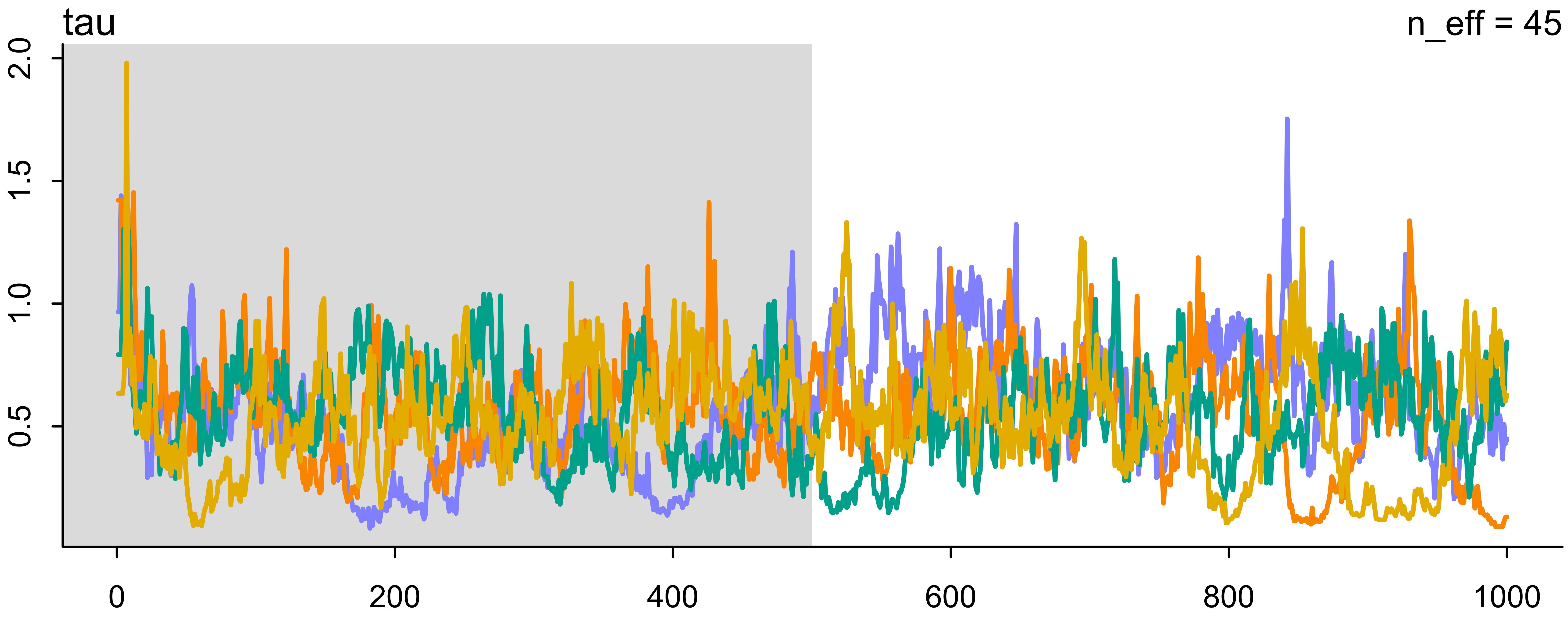
	mean	sd	5.5%	94.5%	n_eff	Rhat4
bbar	0.62	0.16	0.37	0.87	678	1.01
abar	-0.70	0.09	-0.85	-0.56	1123	1.00
tau	0.57	0.24	0.17	0.94	45	1.11
sigma	0.49	0.09	0.34	0.64	330	1.02

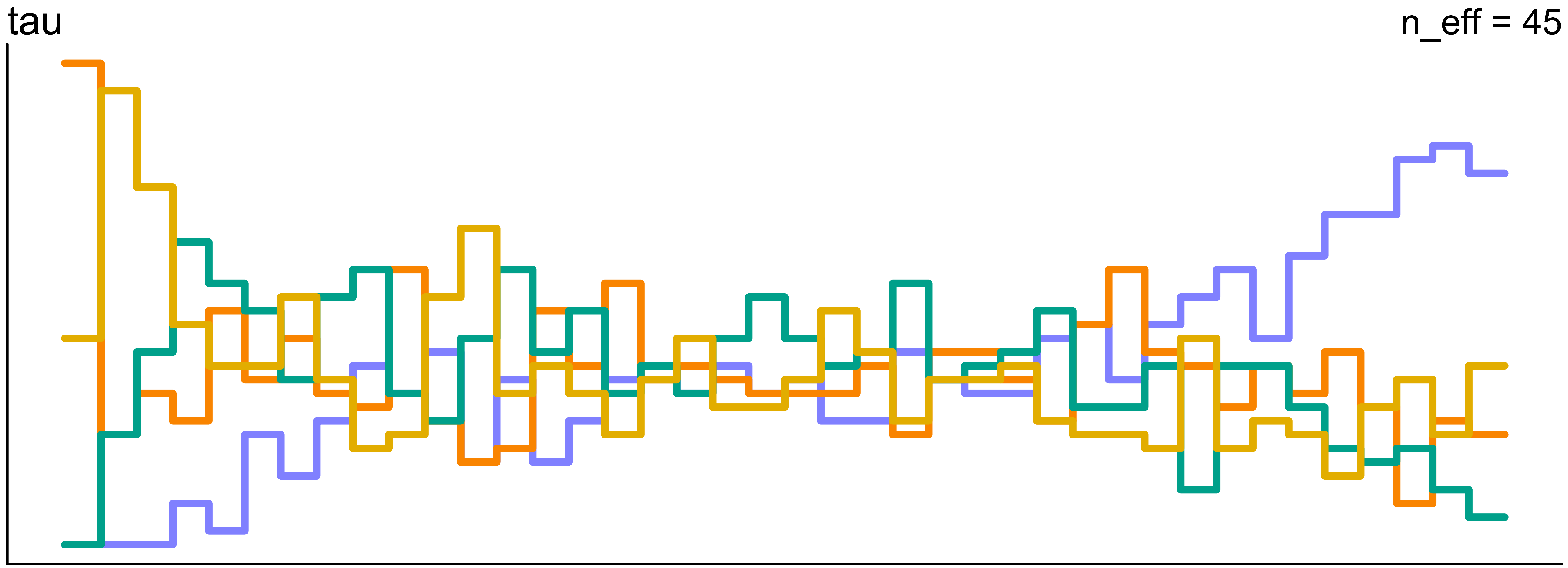
```
Warning: 4 of 2000 (0.0%) transitions ended with a divergence.
See https://mc-stan.org/misc/warnings for details.

Warning: 3 of 4 chains had an E-BFMI less than 0.2.
See https://mc-stan.org/misc/warnings for details.
```

$$C_i \sim \text{Bernoulli}(p_i)$$

$$= \alpha_{D[i]} + \beta_{D[i]} U_i$$

$$\sigma, \tau \sim \text{Exponential}(1)$$


tau
n_eff = 45
2.0
1.5
1.0
0.5
0
200
400
600
800
1000


tau
n_eff = 45

PAUSE

More priors, more problems

$$C_i \sim \text{Bernoulli}(p_i)$$

$$\text{logit}(p_i) = \alpha_{D[i]} + \beta_{D[i]} U_i$$

$$\alpha_j \sim \text{Normal}(\bar{\alpha}, \sigma)$$

$$\beta_j \sim \text{Normal}(\bar{\beta}, \tau)$$

$$\bar{\alpha}, \bar{\beta} \sim \text{Normal}(0,1)$$

$$\sigma, \tau \sim \text{Exponential}(1)$$

Priors inside priors: "centered"

More priors, more problems

$$C_i \sim \text{Bernoulli}(p_i)$$

$$\text{logit}(p_i) = \alpha_{D[i]} + \beta_{D[i]} U_i$$

$$\alpha_j \sim \text{Normal}(\bar{\alpha}, \sigma)$$

$$\beta_j \sim \text{Normal}(\bar{\beta}, \tau)$$

$$\bar{\alpha}, \bar{\beta} \sim \text{Normal}(0,1)$$

$$\sigma, \tau \sim \text{Exponential}(1)$$

$$z_j = (\alpha_j - \bar{\alpha})/\sigma$$

More priors, more problems

$$C_i \sim \text{Bernoulli}(p_i)$$

$$\text{logit}(p_i) = \alpha_{D[i]} + \beta_{D[i]} U_i$$

$$\alpha_j \sim \text{Normal}(\bar{\alpha}, \sigma)$$

$$\beta_j \sim \text{Normal}(\bar{\beta}, \tau)$$

$$\bar{\alpha}, \bar{\beta} \sim \text{Normal}(0,1)$$

$$\sigma, \tau \sim \text{Exponential}(1)$$

$$z_j = (\alpha_j - \bar{\alpha})/\sigma$$

$$\alpha_j = \bar{\alpha} + z_{\alpha,j} \times \sigma$$

More priors, more problems

$$C_i \sim \text{Bernoulli}(p_i)$$

$$\text{logit}(p_i) = \alpha_{D[i]} + \beta_{D[i]} U_i$$

$$\alpha_j \sim \text{Normal}(\bar{\alpha}, \sigma)$$

$$\beta_j \sim \text{Normal}(\bar{\beta}, \tau)$$

$$\bar{\alpha}, \bar{\beta} \sim \text{Normal}(0,1)$$

$$\sigma, \tau \sim \text{Exponential}(1)$$

$$z_j = (\alpha_j - \bar{\alpha})/\sigma$$

$$\alpha_j = \bar{\alpha} + z_{\alpha,j} \times \sigma$$

$$z_{\alpha,j} \sim \text{Normal}(0,1)$$

Centered varying intercepts

$$C_i \sim \text{Bernoulli}(p_i)$$

$$\text{logit}(p_i) = \alpha_{D[i]} + \beta_{D[i]} U_i$$

$$\alpha_j \sim \text{Normal}(\bar{\alpha}, \sigma)$$

$$\beta_j \sim \text{Normal}(\bar{\beta}, \tau)$$

$$\bar{\alpha}, \bar{\beta} \sim \text{Normal}(0,1)$$

$$\sigma, \tau \sim \text{Exponential}(1)$$

Non-centered varying intercepts

$$C_i \sim \text{Bernoulli}(p_i)$$

$$\text{logit}(p_i) = \alpha_{D[i]} + \beta_{D[i]} U_i$$

$$\alpha_j = \bar{\alpha} + z_{\alpha,j} \times \sigma$$

$$\beta_j = \bar{\beta} + z_{\beta,j} \times \tau$$

$$z_{\alpha,j} \sim \text{Normal}(0,1)$$

$$z_{\beta,j} \sim \text{Normal}(0,1)$$

$$\bar{\alpha}, \bar{\beta} \sim \text{Normal}(0,1)$$

$$\sigma, \tau \sim \text{Exponential}(1)$$

```
mCDUnc <- ulam(
    alist(
        C ~ bernoulli(p),
        logit(p) <- a[D] + b[D]*U,
        # define effects using other parameters
        save> vector[61]:a <<- abar + za*sigma,
        save> vector[61]:b <<- bbar + zb*tau,
        # z-scored effects
        vector[61]:za ~ normal(0,1),
        vector[61]:zb ~ normal(0,1),
        # ye olde hyper-priors
        c(abar,bbar) ~ normal(0,1),
        c(sigma,tau) ~ exponential(1)
    ) , data=dat , chains=4 , cores=4 )
```

$$C_i \sim \text{Bernoulli}(p_i)$$

$$\text{logit}(p_i) = \alpha_{D[i]} + \beta_{D[i]} U_i$$

$$\alpha_j = \bar{\alpha} + z_{\alpha,j} \times \sigma$$

$$\beta_j = \bar{\beta} + z_{\beta,j} \times \tau$$

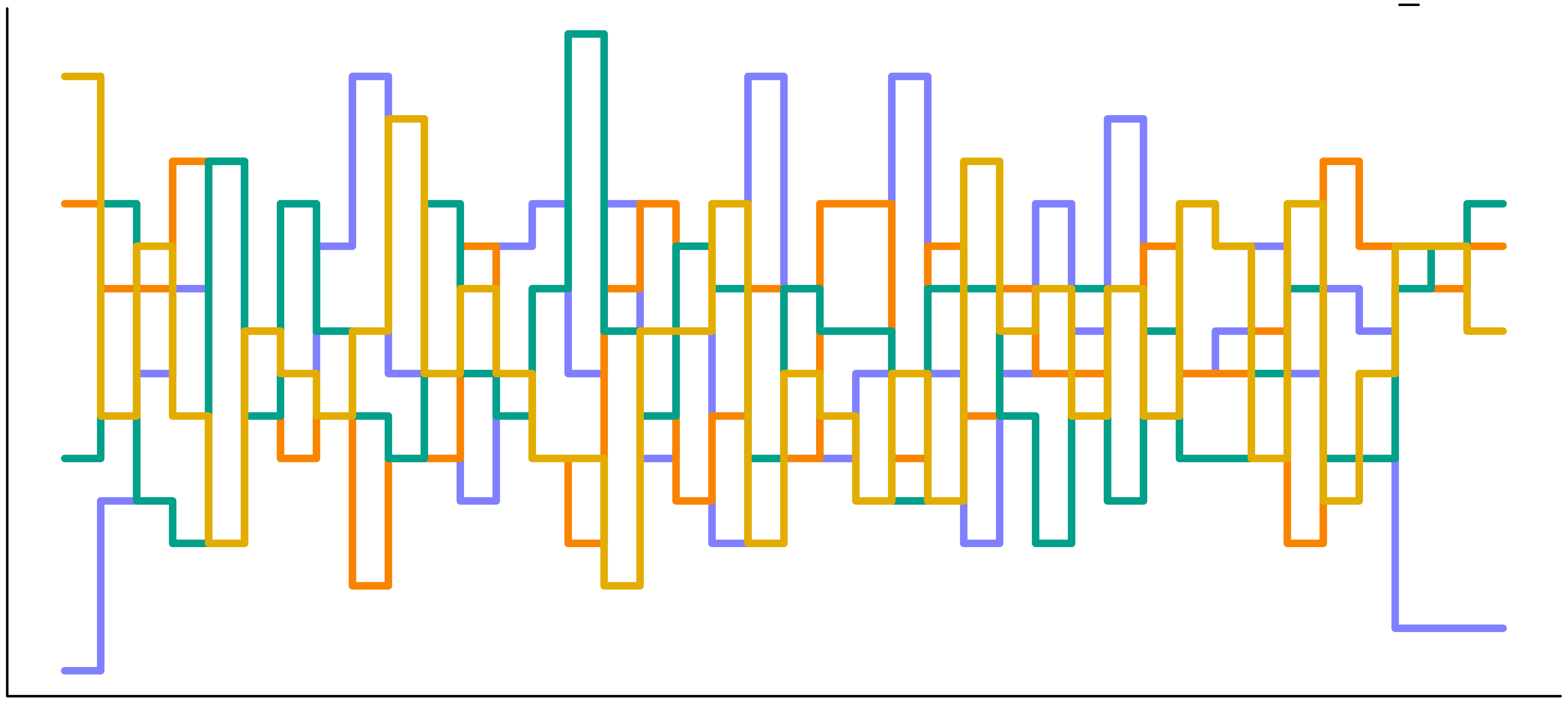
$$z_{\alpha,j} \sim \text{Normal}(0,1)$$

$$z_{\beta,j} \sim \text{Normal}(0,1)$$

$$\bar{\alpha}, \bar{\beta} \sim \text{Normal}(0,1)$$

$$\sigma, \tau \sim \text{Exponential}(1)$$

	mean	sd	5.5%	94.5%	n_eff	Rhat4
bbar	0.62	0.16	0.37	0.86	1513	1.00
abar	-0.70	0.09	-0.84	-0.56	1457	1.00
tau	0.55	0.23	0.17	0.92	368	1.01
sigma	0.49	0.09	0.36	0.64	753	1.00



$$C_i \sim \text{Bernoulli}(p_i)$$

$$\text{logit}(p_i) = \alpha_{D[i]} + \beta_{D[i]} U_i$$

$$\alpha_j = \bar{\alpha} + z_{\alpha,j} \times \sigma$$

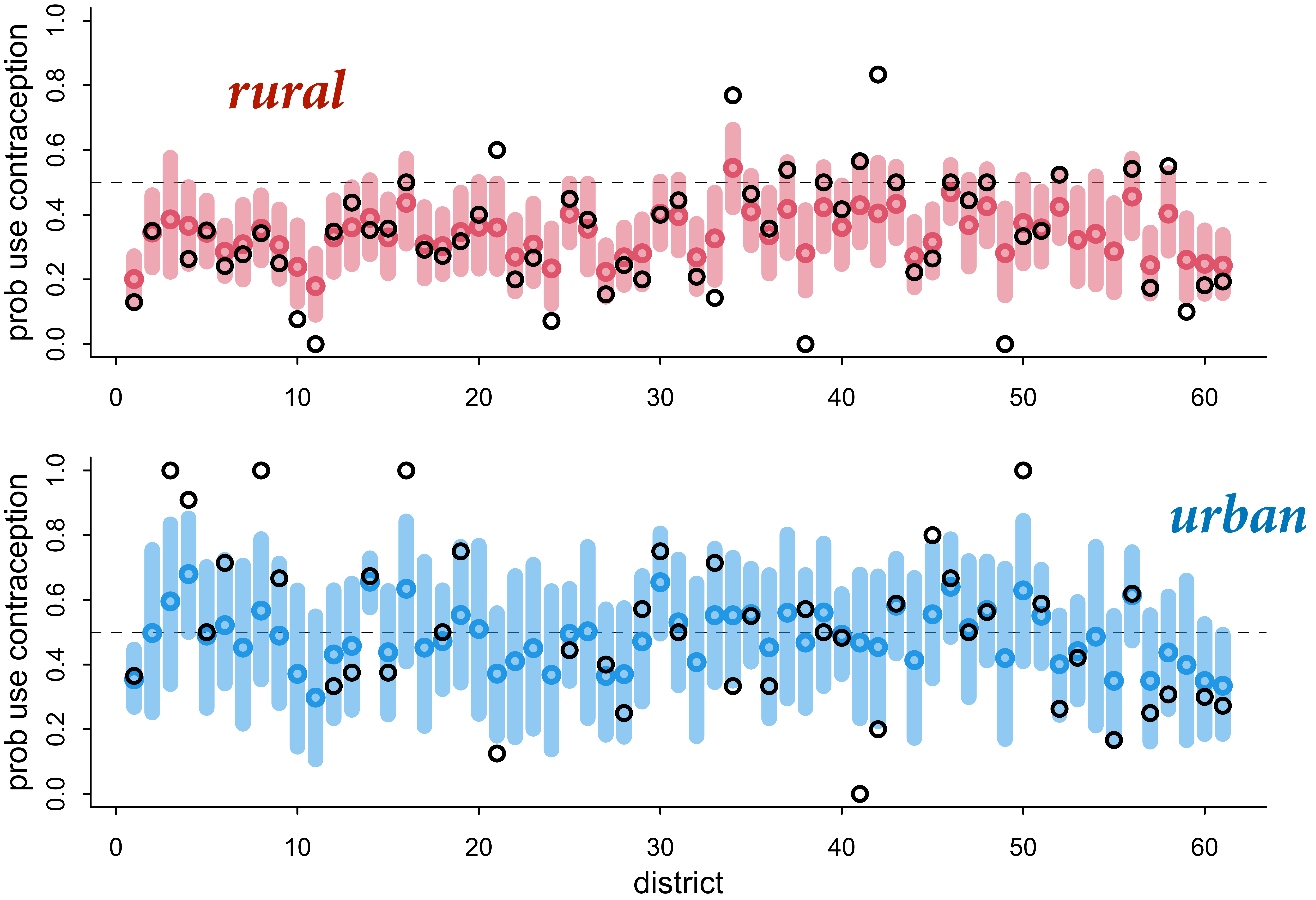
$$\beta_j = \bar{\beta} + z_{\beta,j} \times \tau$$

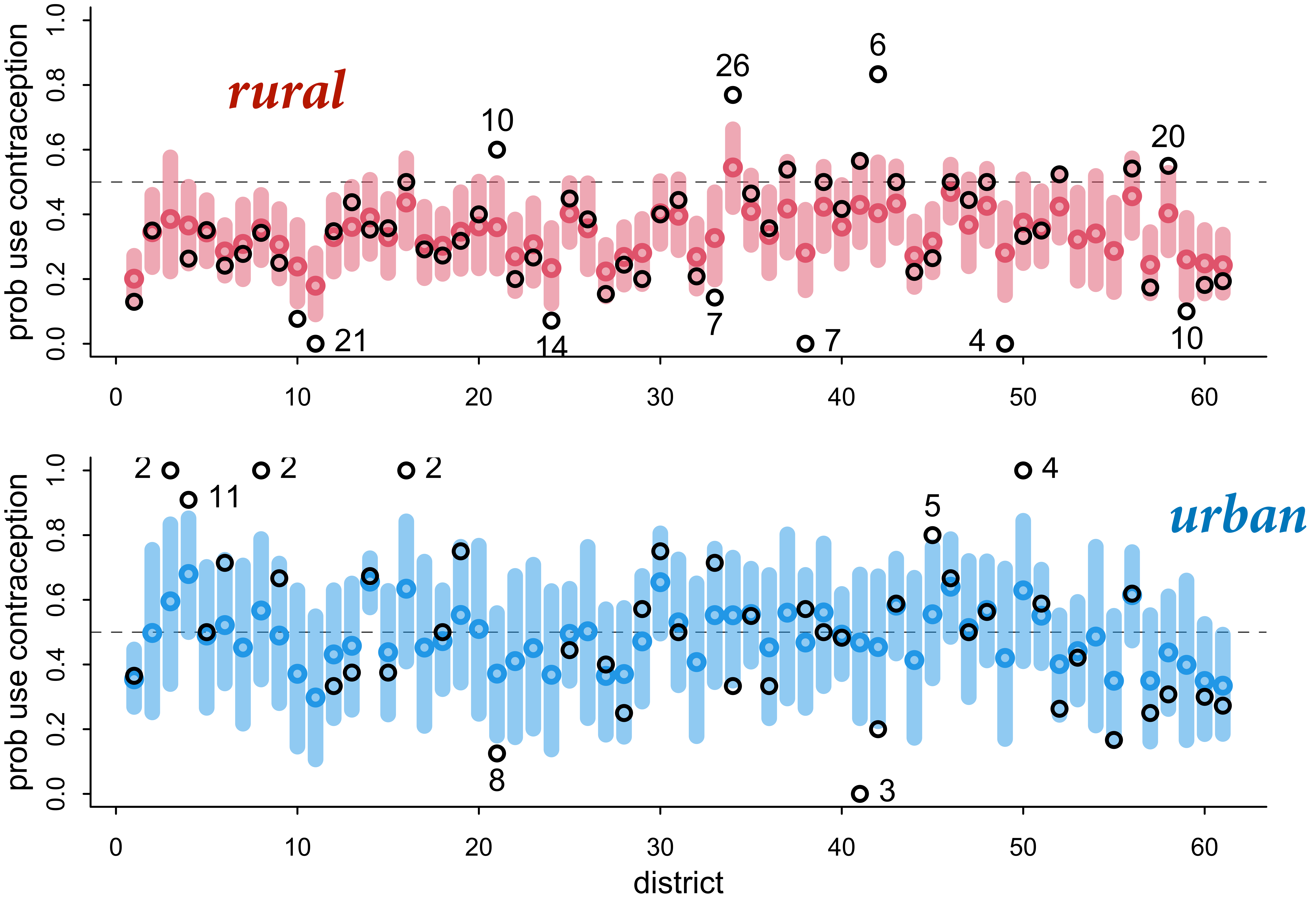
$$z_{\alpha,j} \sim \text{Normal}(0,1)$$

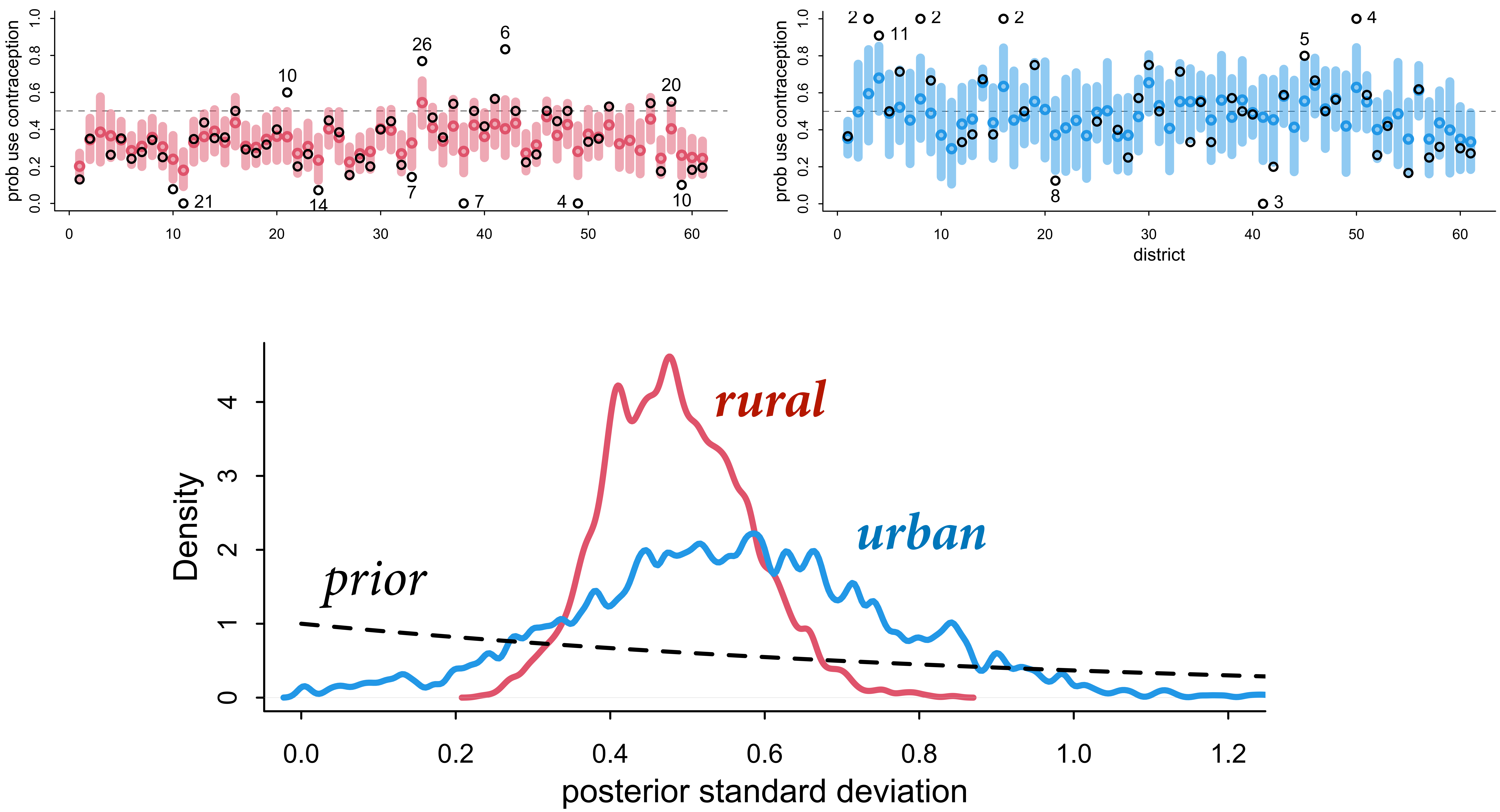
$$z_{\beta,j} \sim \text{Normal}(0,1)$$

$$\bar{\alpha}, \bar{\beta} \sim \text{Normal}(0,1)$$

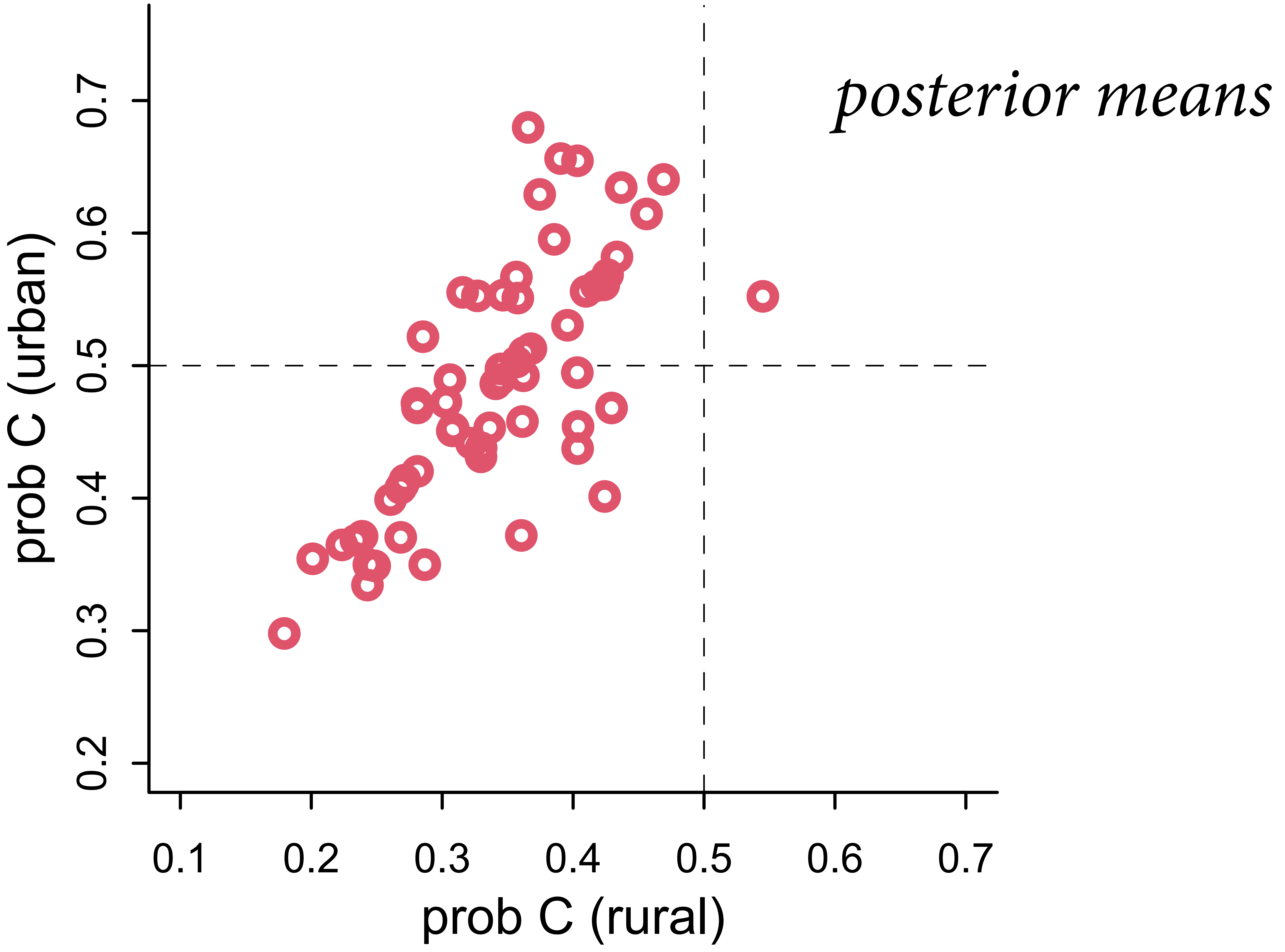
$$\sigma, \tau \sim \text{Exponential}(1)$$

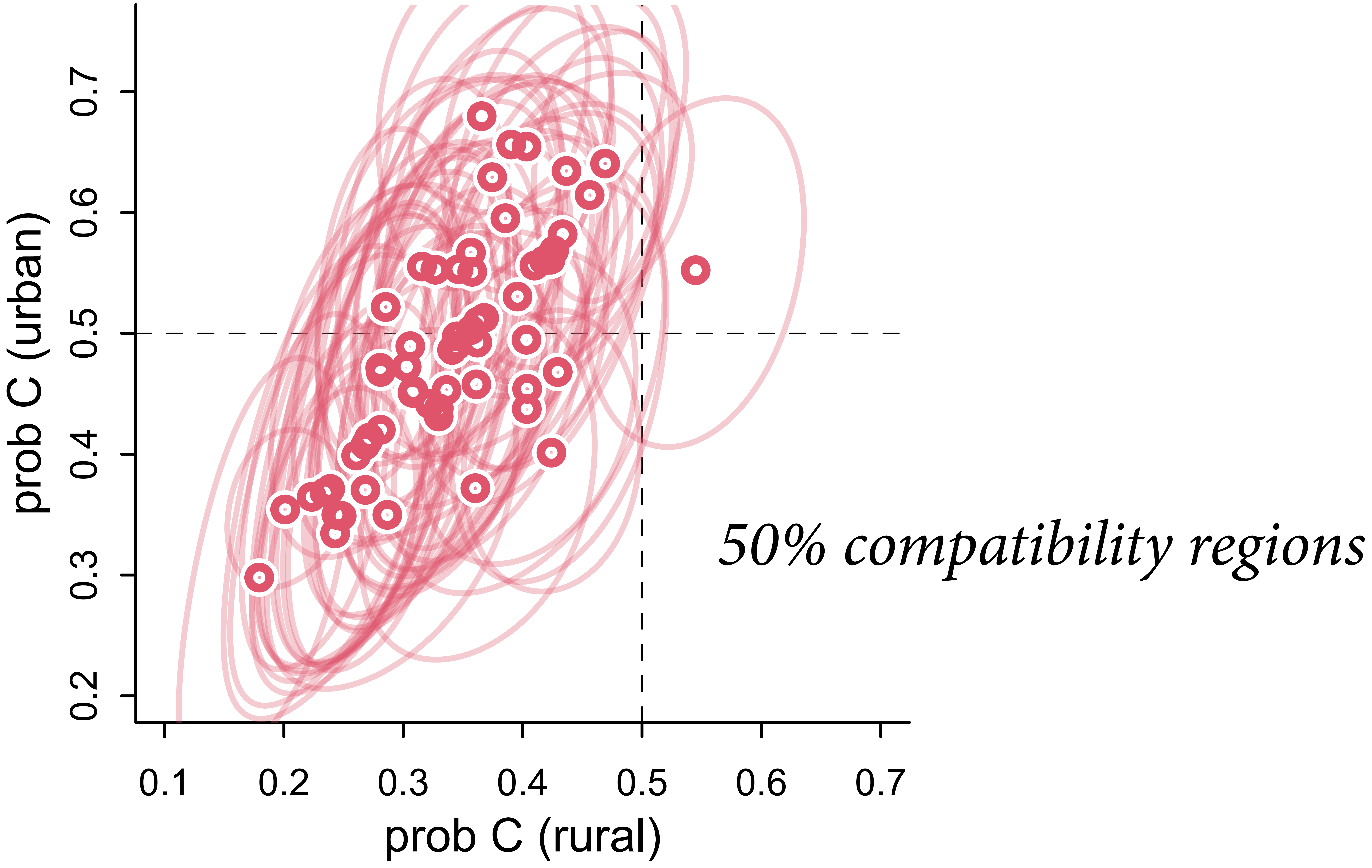

rural
prob use contraception
0.0
0.2
0.4
0.6
0.8
1.0
0
10
20
30
40
50
60
urban
prob use contraception
0.0
0.2
0.4
0.6
0.8
1.0
0
10
20
30
40
50
60
district


rural
prob use contraception
0.0
0.2
0.4
0.6
0.8
1.0
0
10
20
30
40
50
60
10
26
6
20
21
14
7
7
4
10
urban
prob use contraception
2
11
2
2
4
5
8
3
district

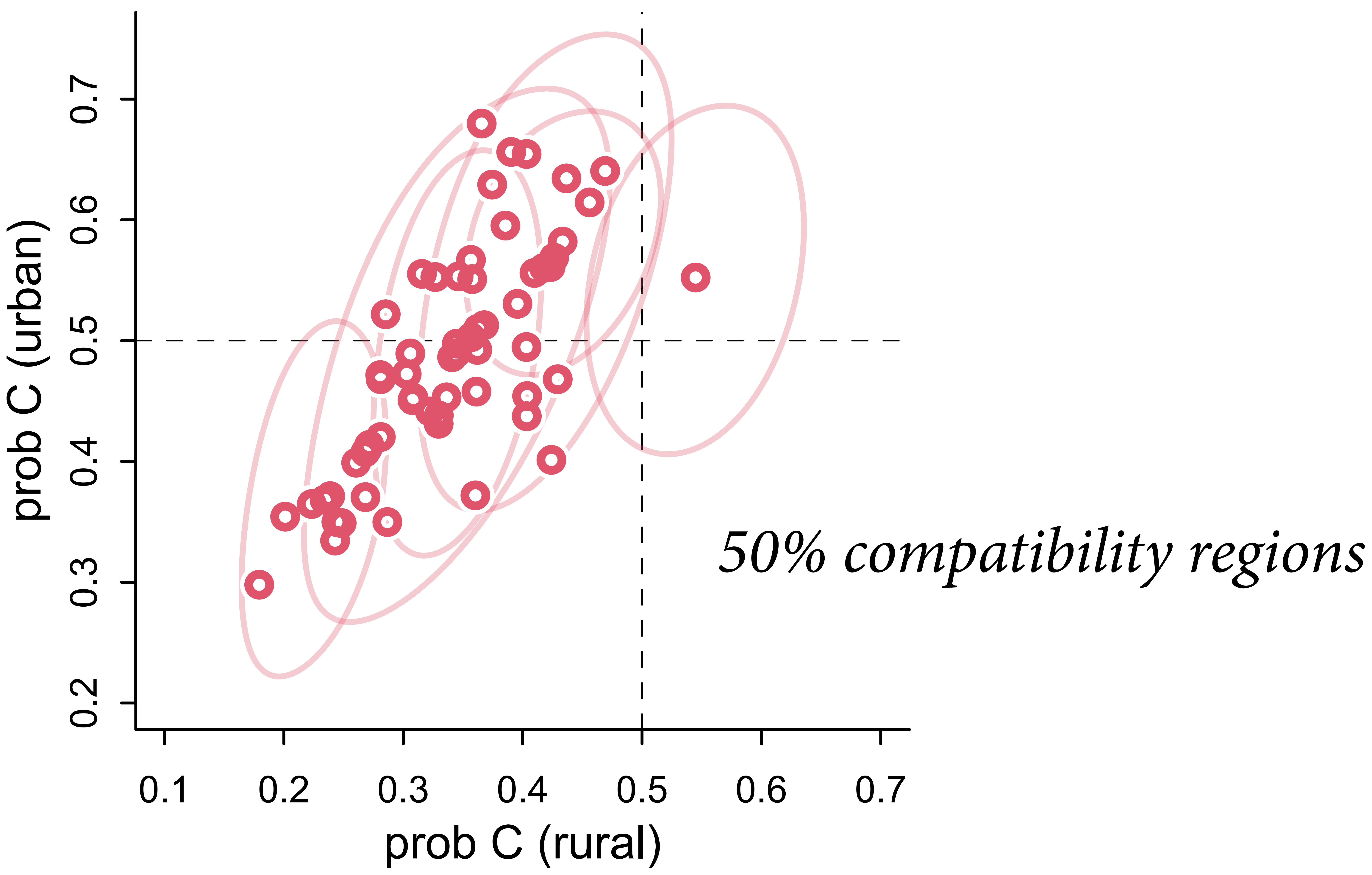


prob use contraception
0.0 0.2 0.4 0.6 0.8 1.0
0 10 20 30 40 50 60
26
6
10
20
21
14
7
7
4
10
prob use contraception
0.0 0.2 0.4 0.6 0.8 1.0
0 10 20 30 40 50 60
district
2
11
2
2
4
5
8
3
Density
0 1 2 3 4
prior
rural
urban
0.0 0.2 0.4 0.6 0.8 1.0 1.2
posterior standard deviation


posterior means
prob C (urban)
0.2
0.3
0.4
0.5
0.6
0.7
0.1
0.2
0.3
0.4
0.5
0.6
0.7
prob C (rural)



prob C (urban)
0.2
0.3
0.4
0.5
0.6
0.7
0.1
0.2
0.3
0.4
0.5
0.6
0.7
prob C (rural)
50% compatibility regions



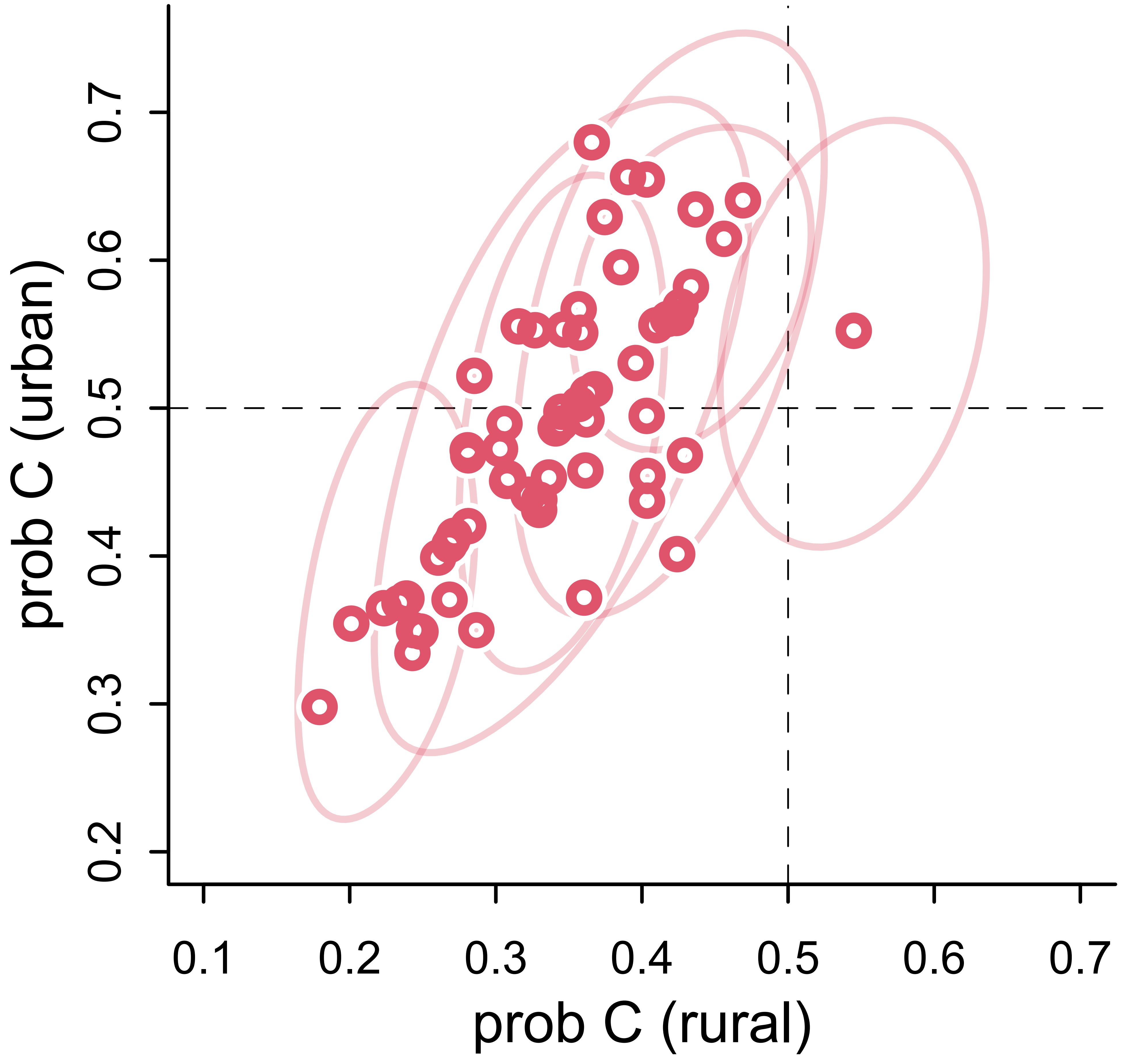
prob C (urban)
0.2
0.3
0.4
0.5
0.6
0.7
prob C (rural)
0.1
0.2
0.3
0.4
0.5
0.6
0.7
50% compatibility regions

Multilevel adventures

Clusters: Kinds of groups in the data (districts)

Features: Aspects of the model (parameters) that vary by cluster (rural, urban)

There is useful information to transfer across features



Course Schedule

Week 1	Bayesian inference	Chapters 1, 2, 3
Week 2	Linear models & Causal Inference	Chapter 4
Week 3	Causes, Confounds & Colliders	Chapters 5 & 6
Week 4	Overfitting / MCMC	Chapters 7, 8, 9
Week 5	Generalized Linear Models	Chapters 10, 11
Week 6	Ordered categories & Multilevel models	Chapters 12 & 13
Week 7	More Multilevel models	Chapters 13 & 14
Week 8	Multilevel models & Gaussian processes	Chapter 14
Week 9	Measurement & Missingness	Chapter 15
Week 10	Generalized Linear Madness	Chapter 16

`https://github.com/rmcelreath/stat_rethinking_2023`