Statistical Rethinking

2022

14: Correlated Varying Effects

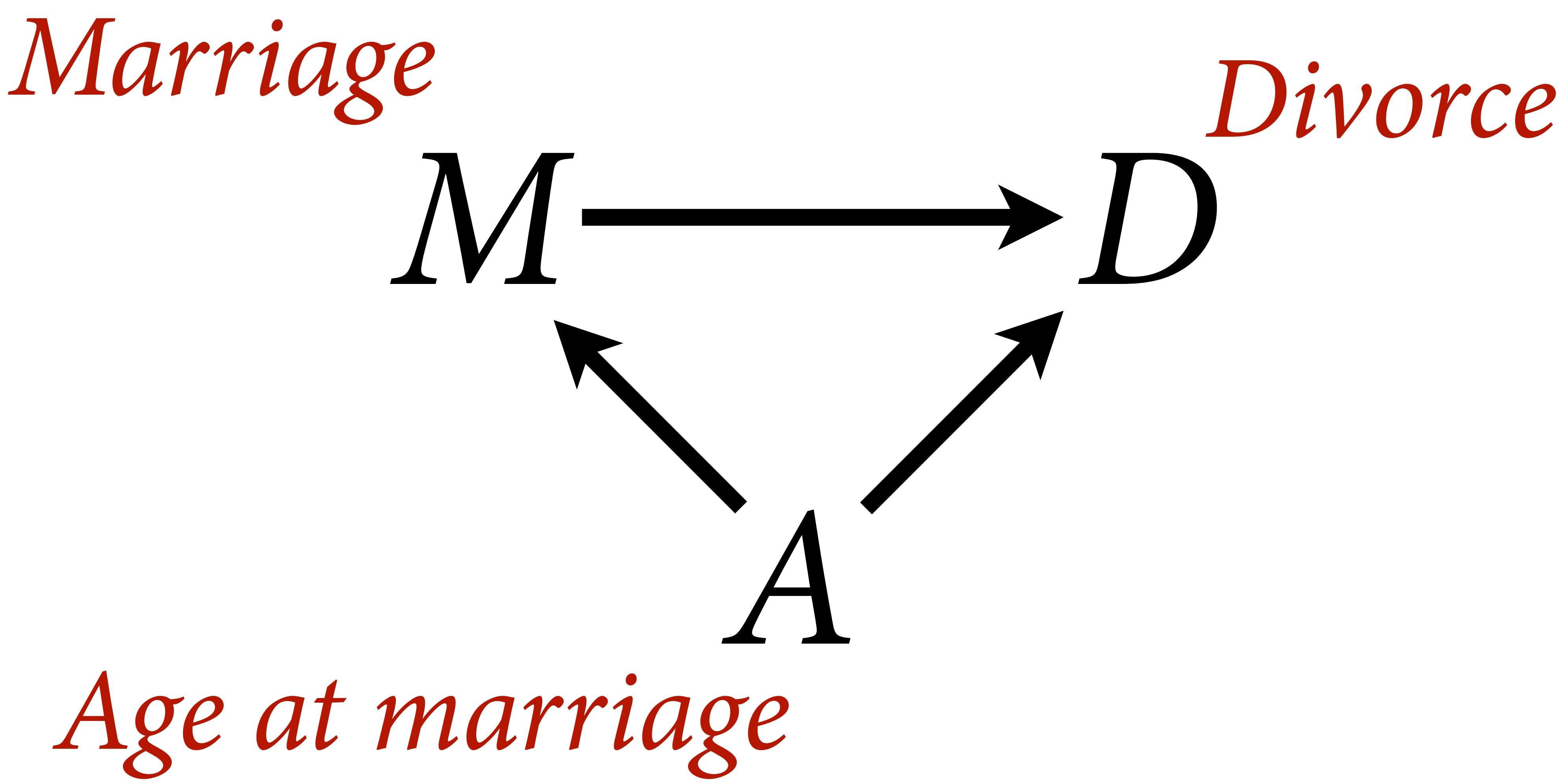

8,15

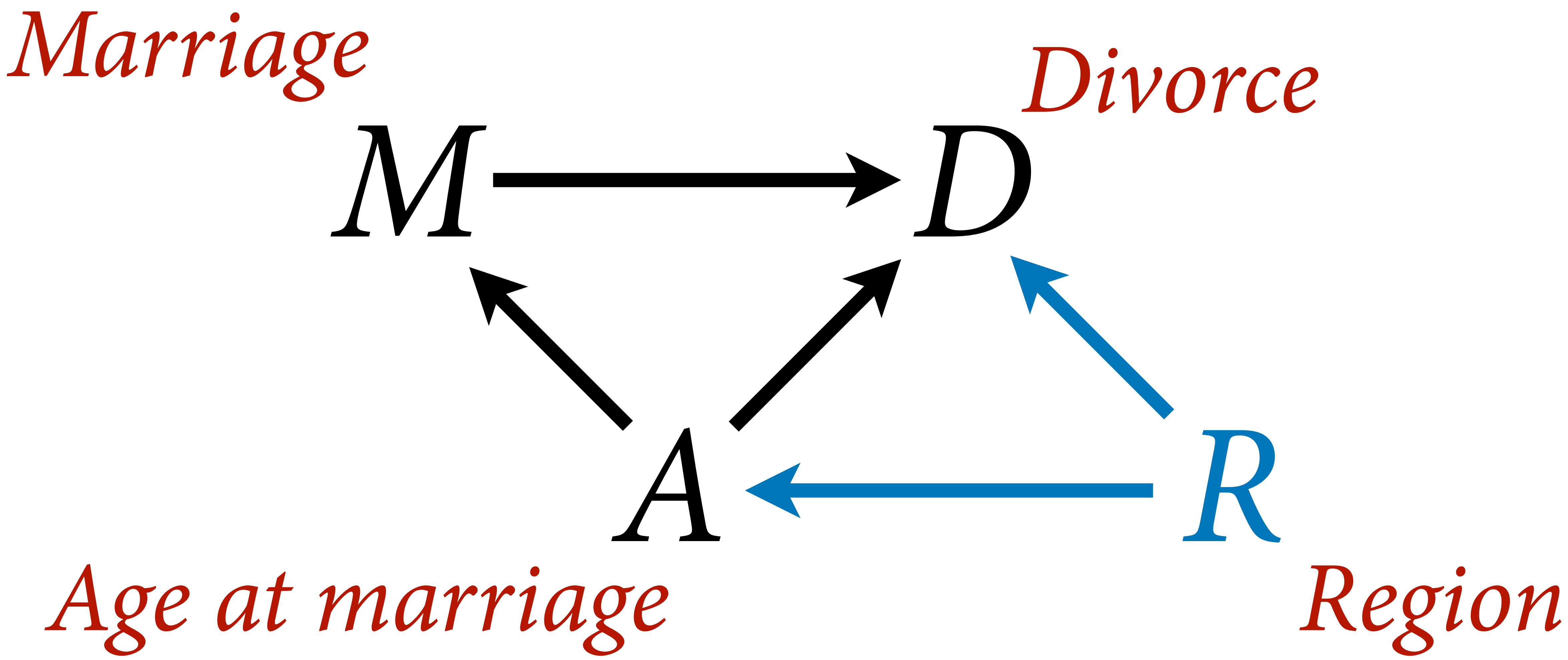
Varying effects as confounds

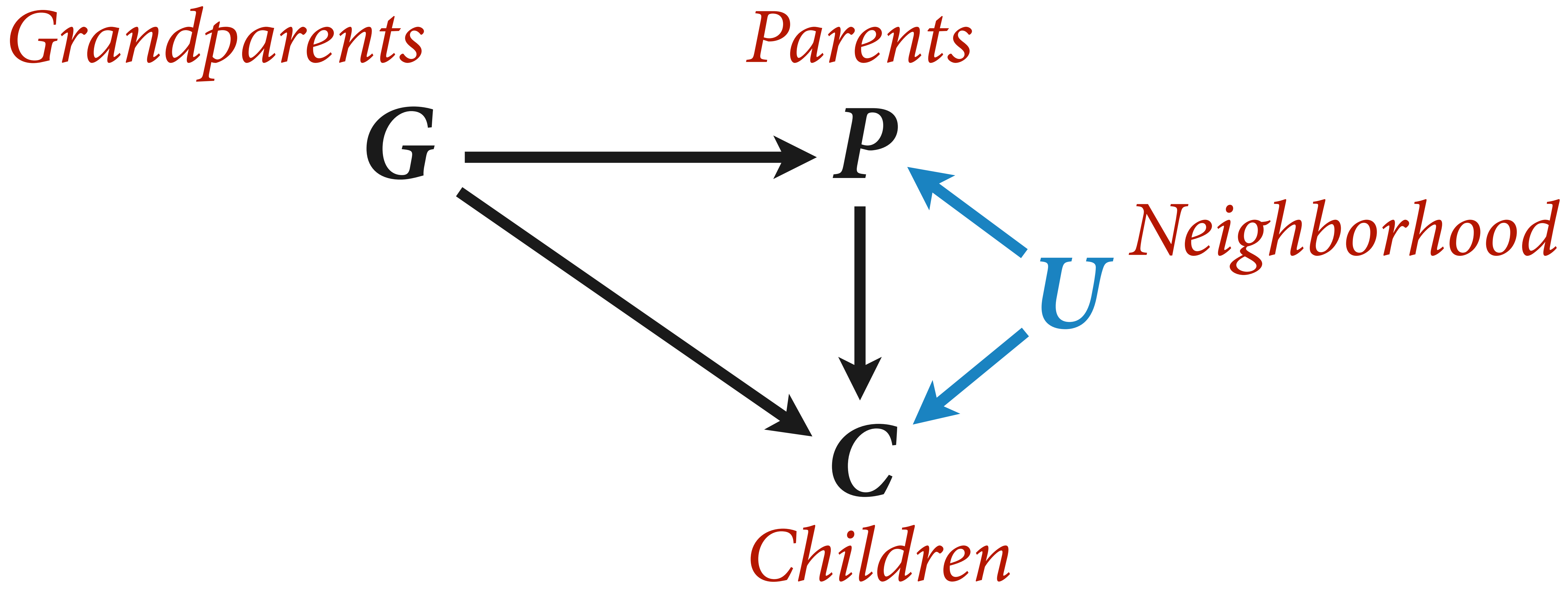
Varying effect strategy: Unmeasured features of clusters leave an imprint on the data that can be measured by (1) repeat observations of each cluster and (2) partial pooling among clusters

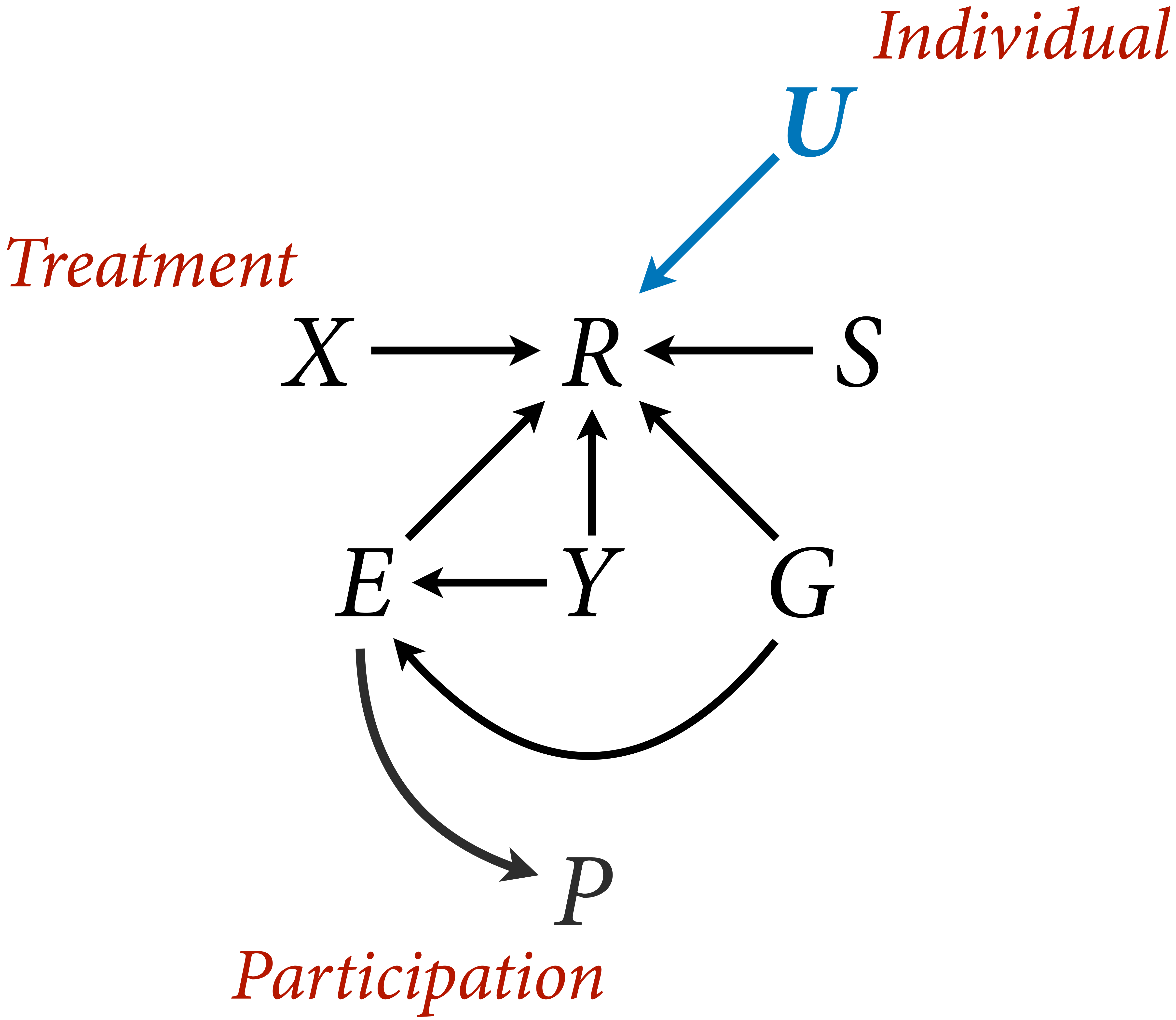
Predictive perspective: Important source of cluster-level variation, regularize

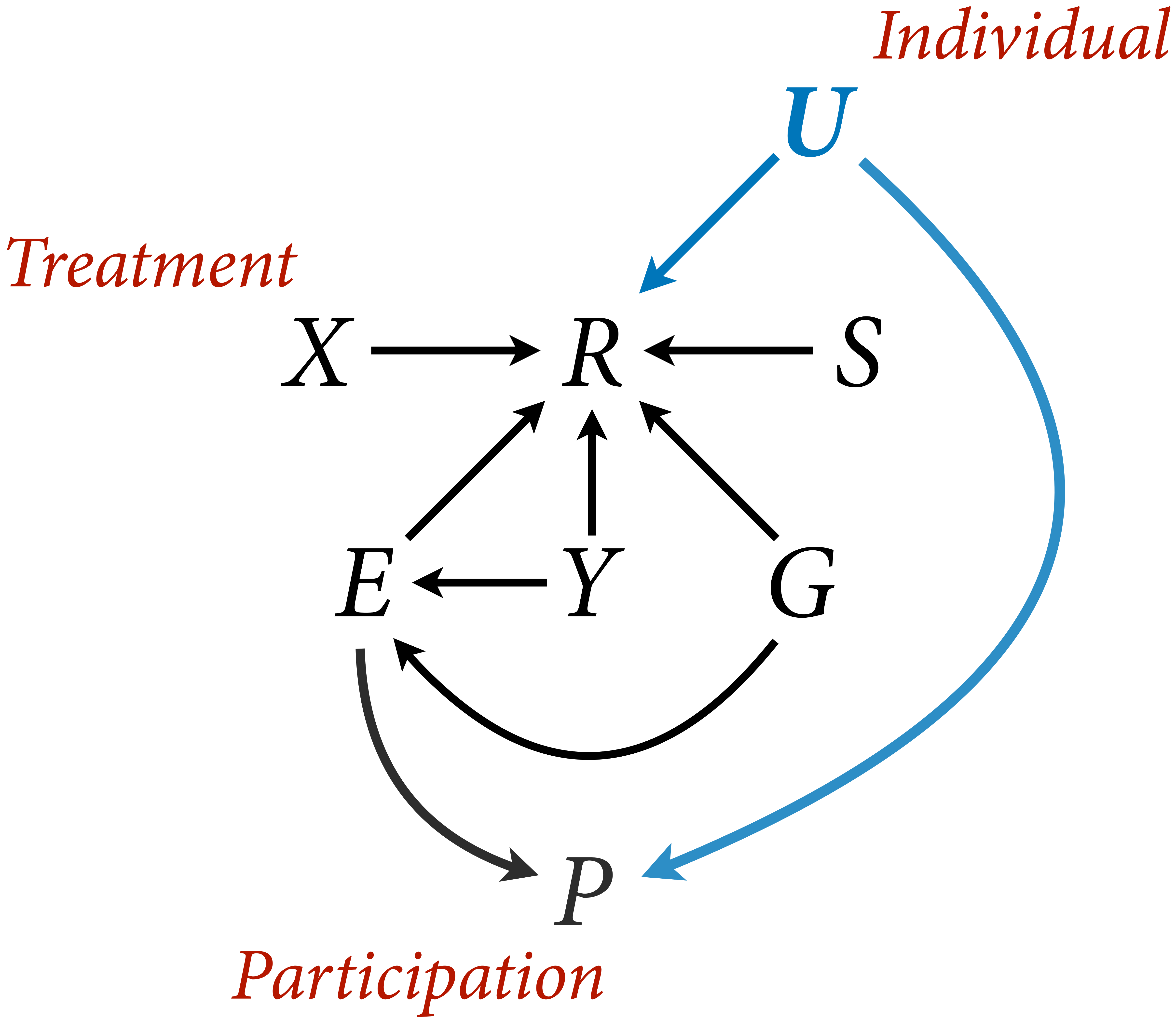
Causal perspective: Competing causes or unobserved confounds

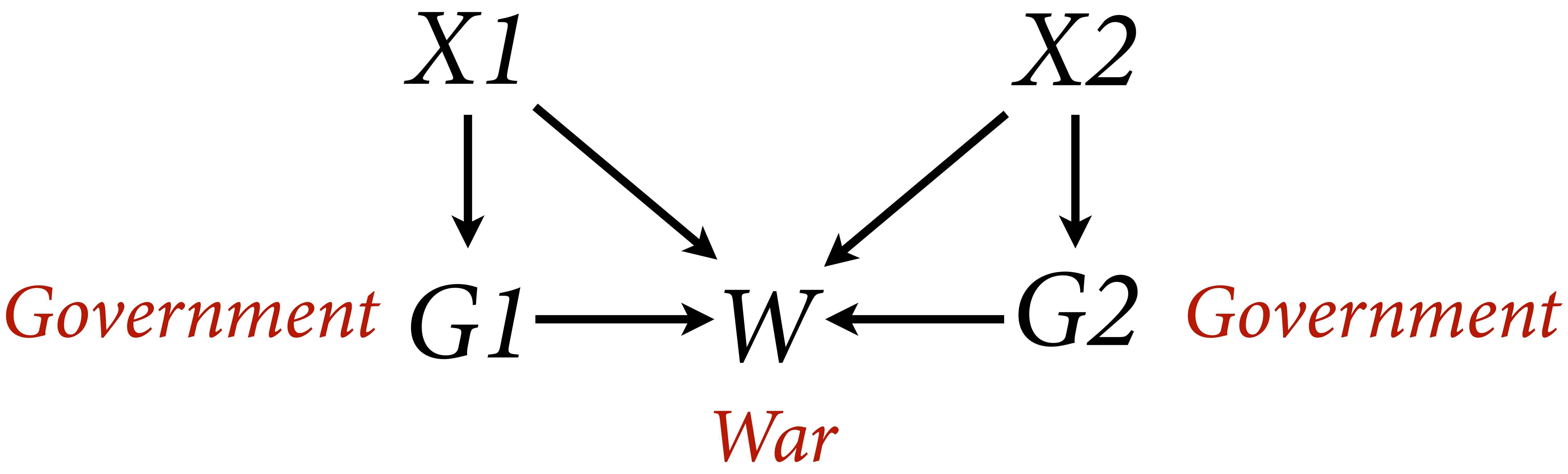

Marriage
M
D
Divorce
A
Age at marriage

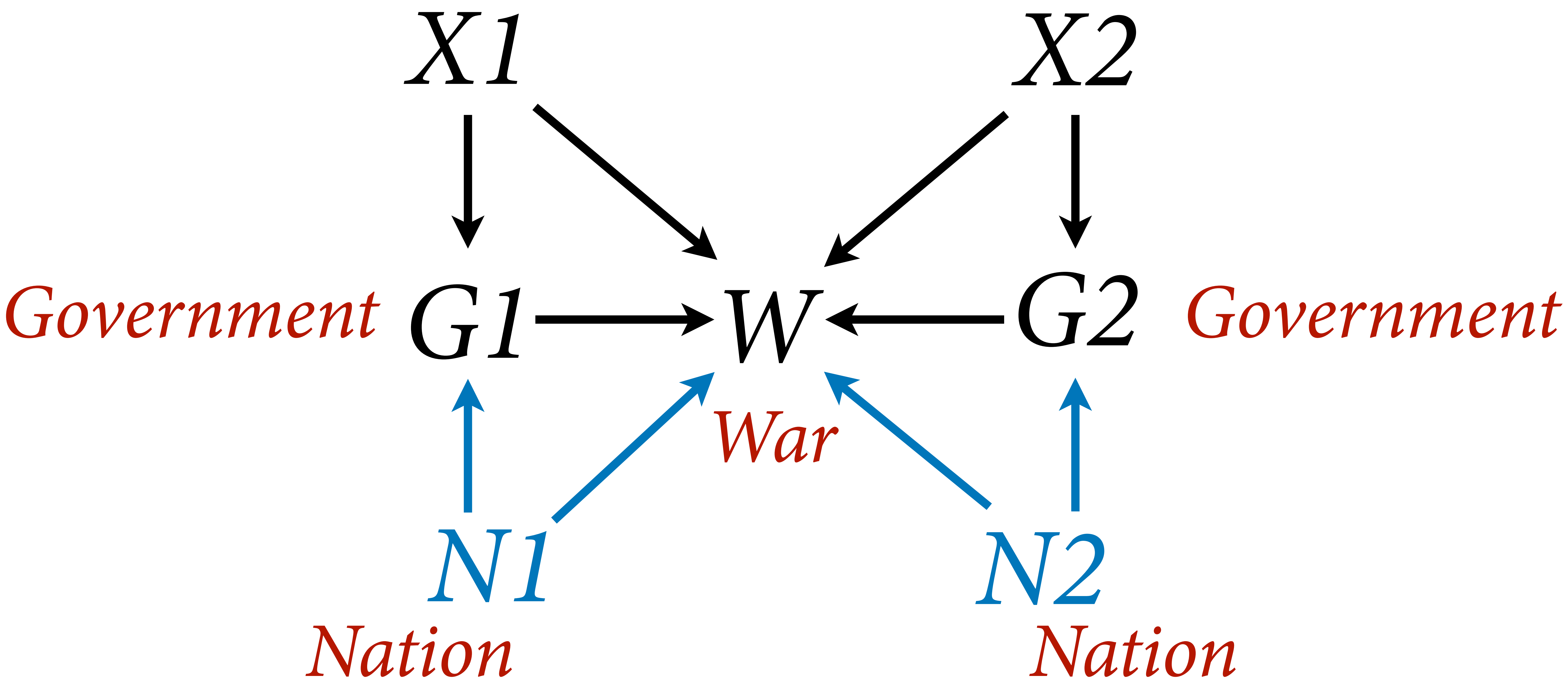

Marriage
Divorce
M
D
A
R
Age at marriage
Region


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
Government
G1
W
G2
Government
War
N1
N2
Nation
Nation

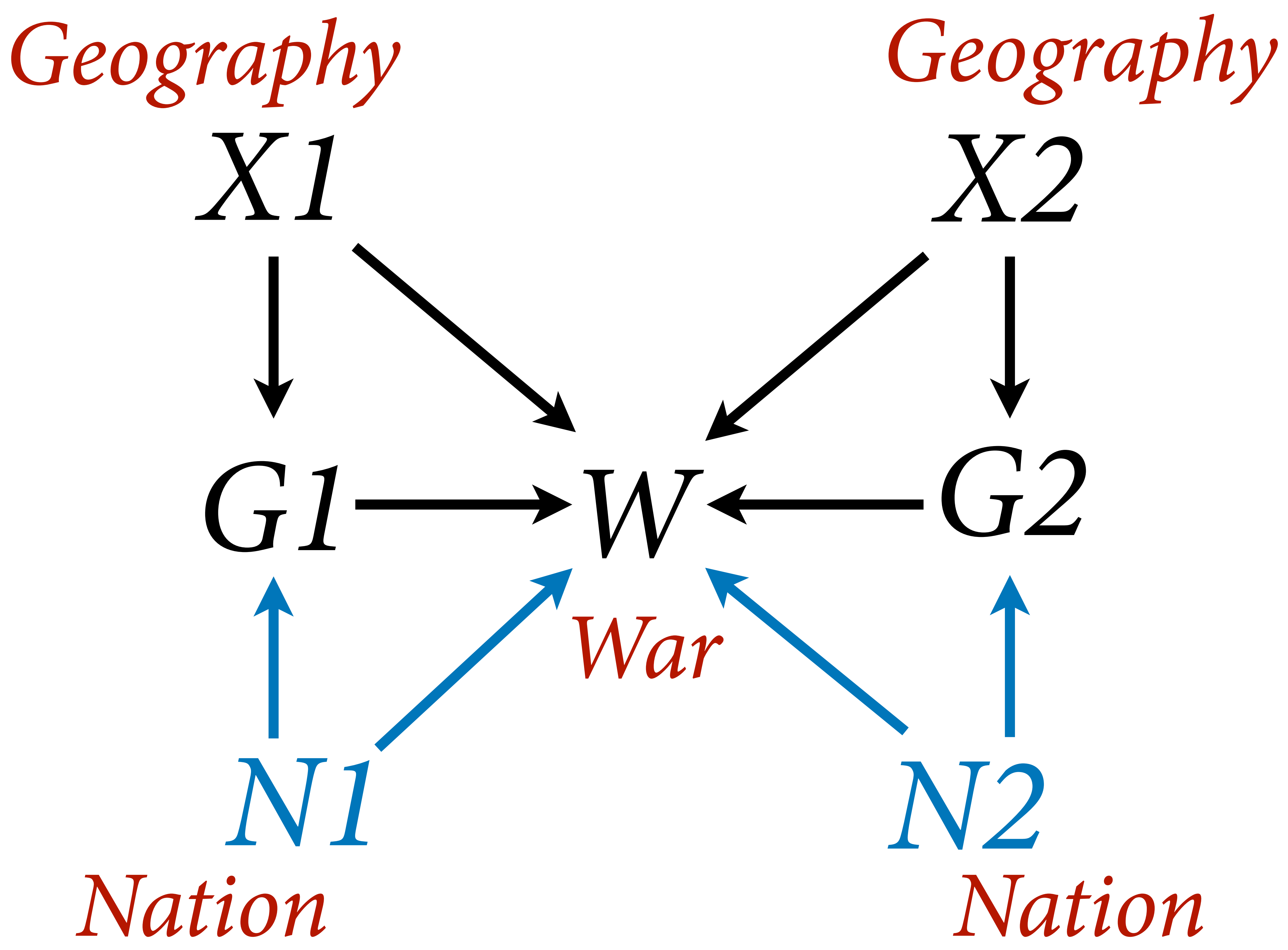
Varying effects as confounds

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



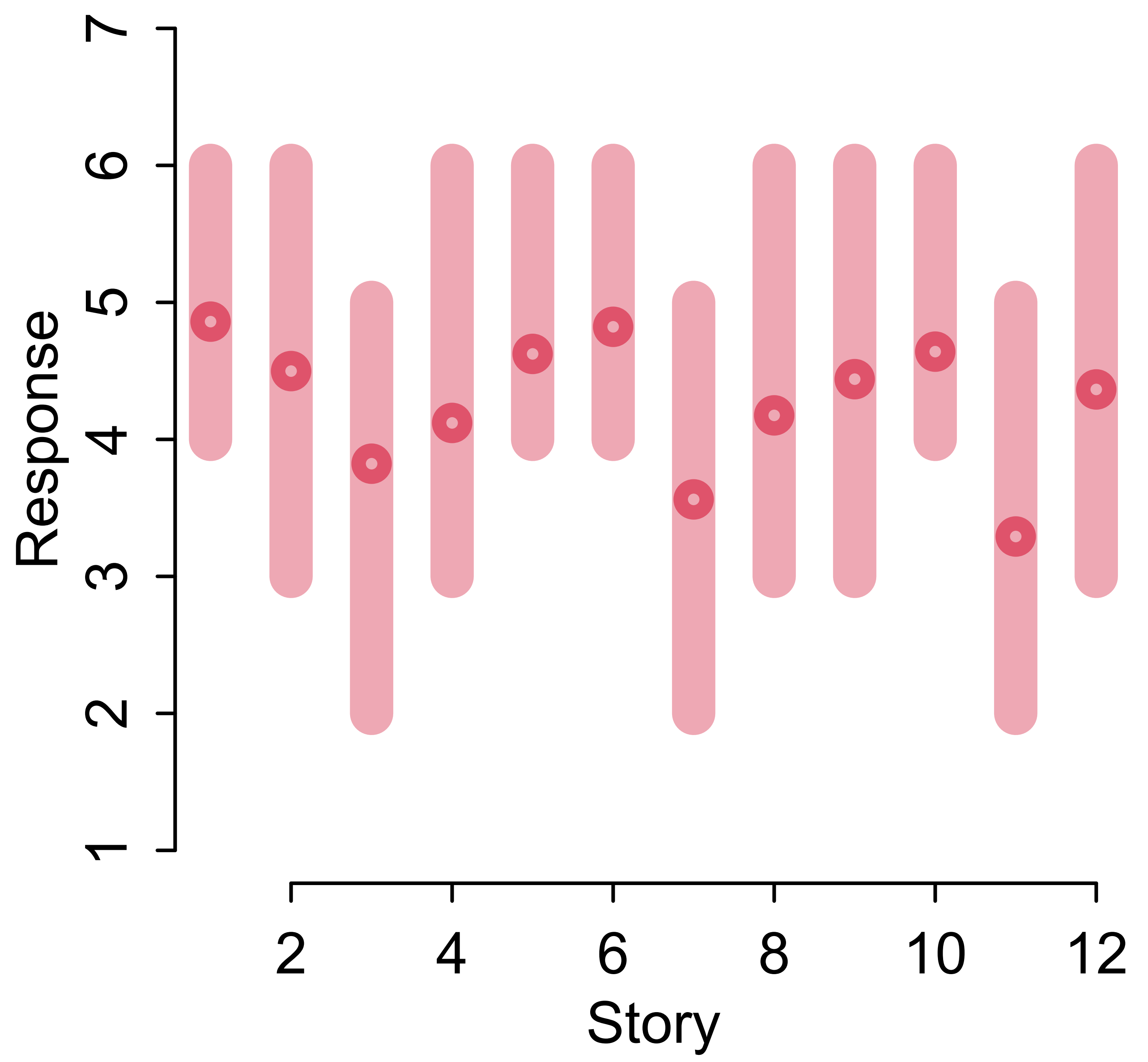
Cluster → **Features**

tanks ⟶ survival

stories ⟶ treatment effect

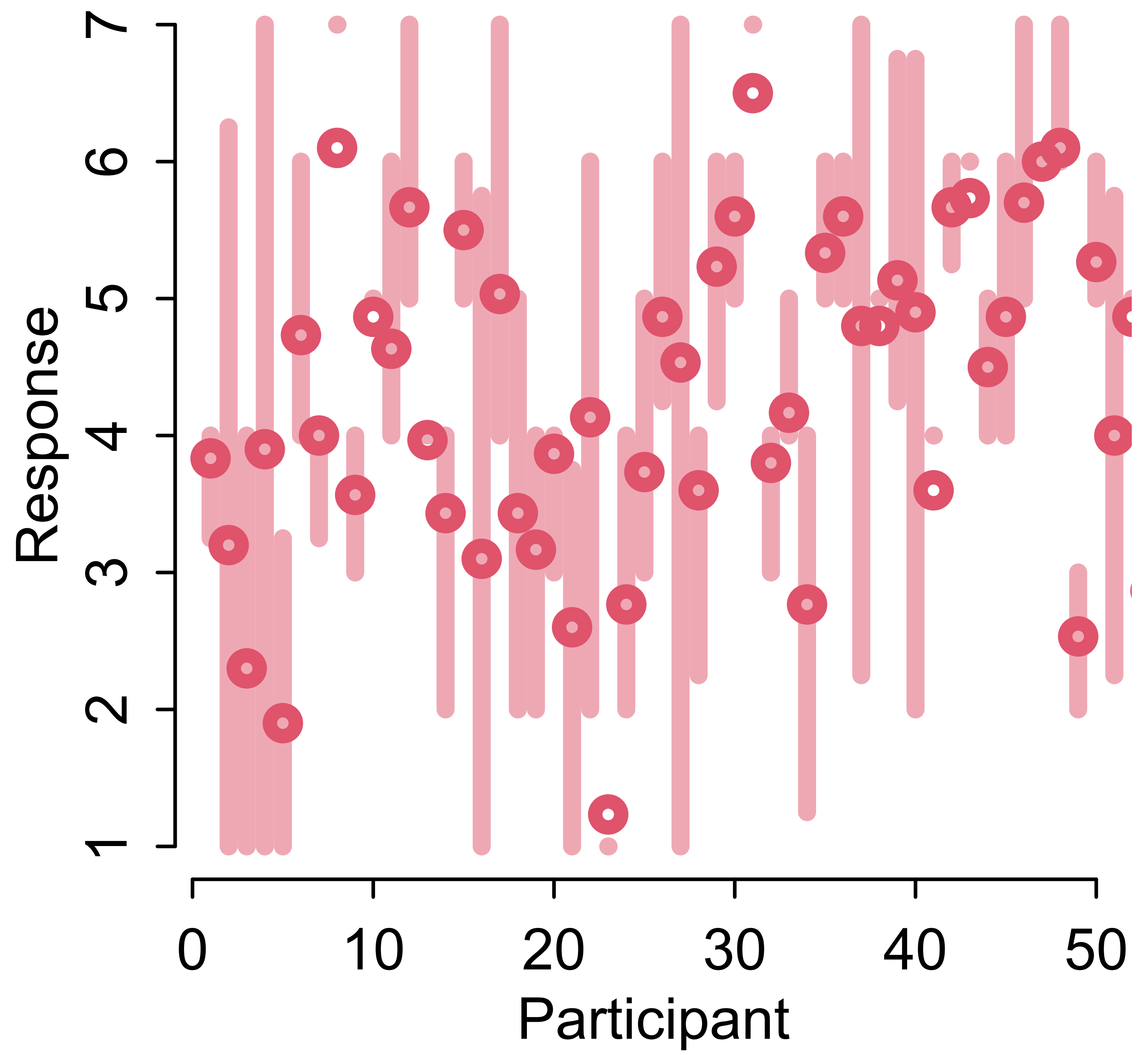
individuals ⟶ average response

departments ⟶ admission rate, bias



Add clusters: More index variables, more population priors (previous lecture)

Add features: More parameters, more dimensions *in each* population prior (this lecture)



Adding Features

One prior distribution for each cluster

One feature: One-dimensional distribution

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

Two features: Two-D distribution

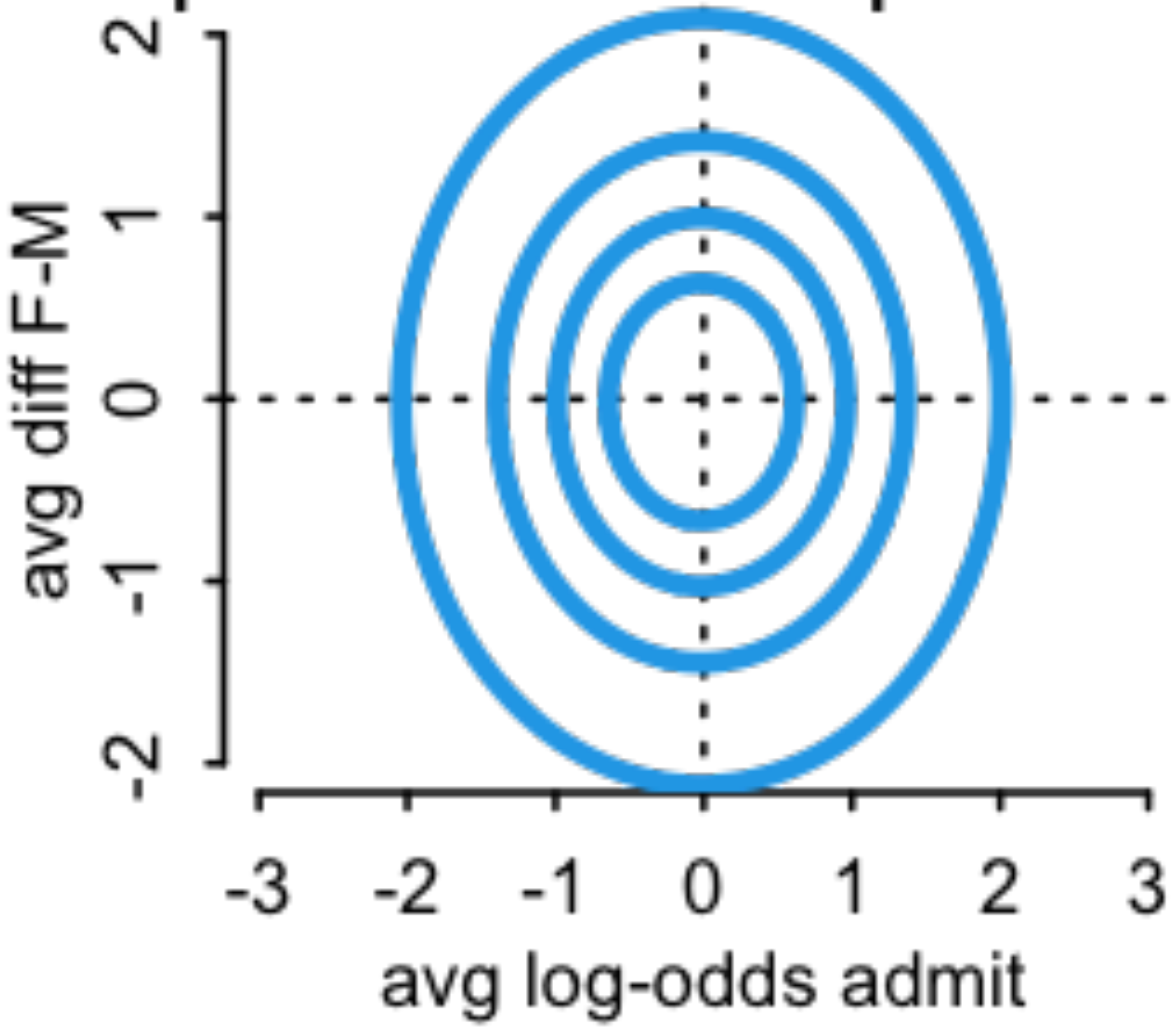
$$[\alpha_j, \beta_j] \sim \text{MVNormal}([\bar{\alpha}, \bar{\beta}], \Sigma)$$

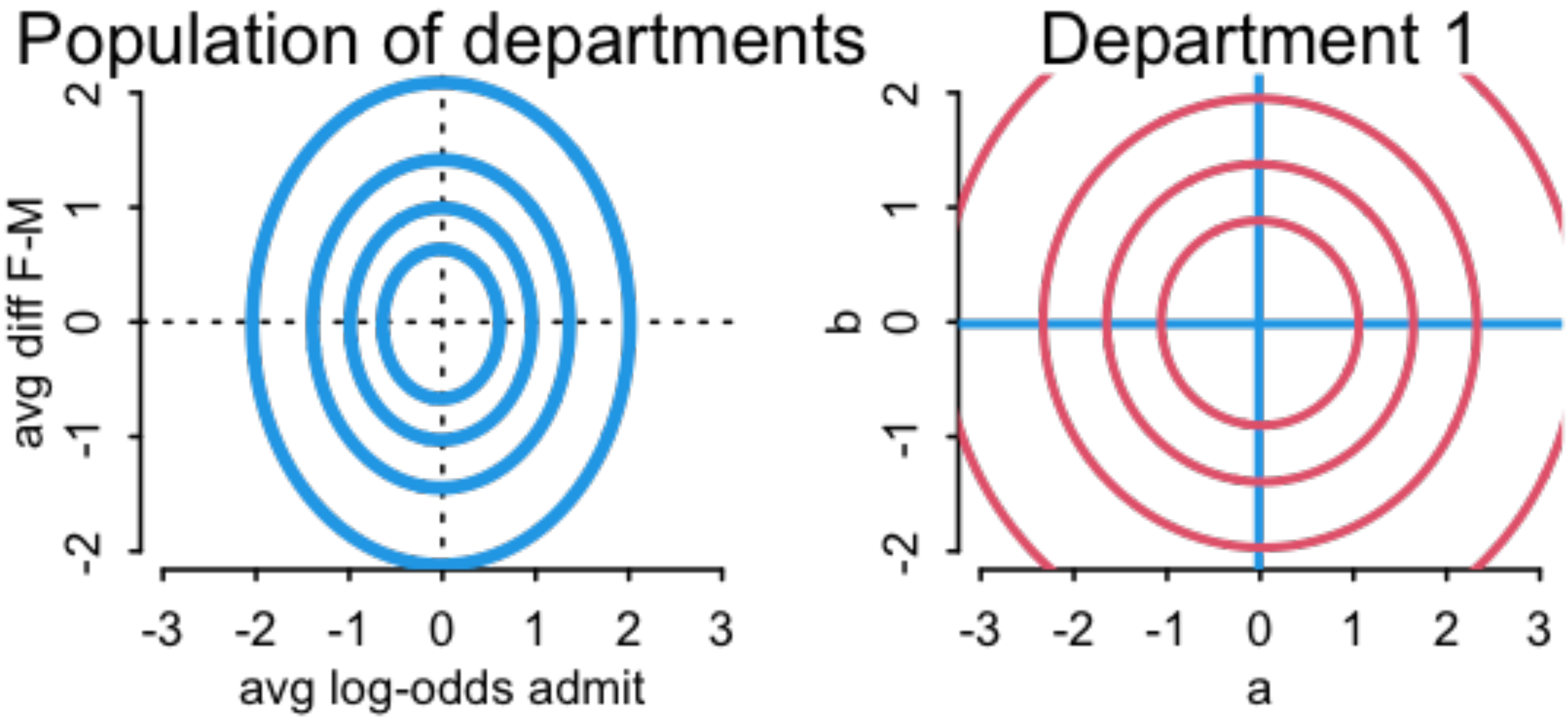
N features: *N*-dimensional distribution

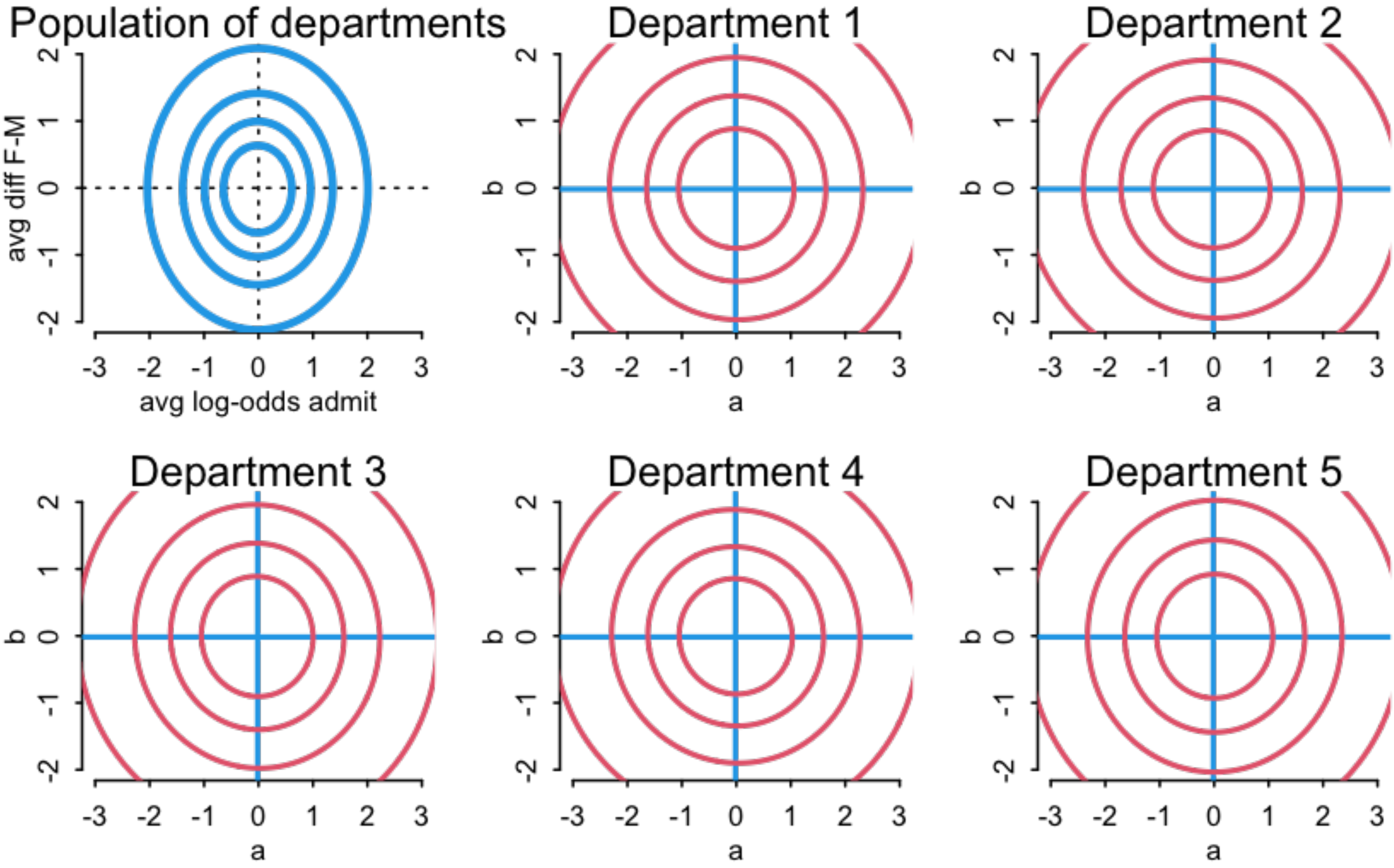
$$\alpha_{j,1..N} \sim \text{MVNormal}(\bar{\alpha}, \Sigma)$$

Hard part: Learning associations

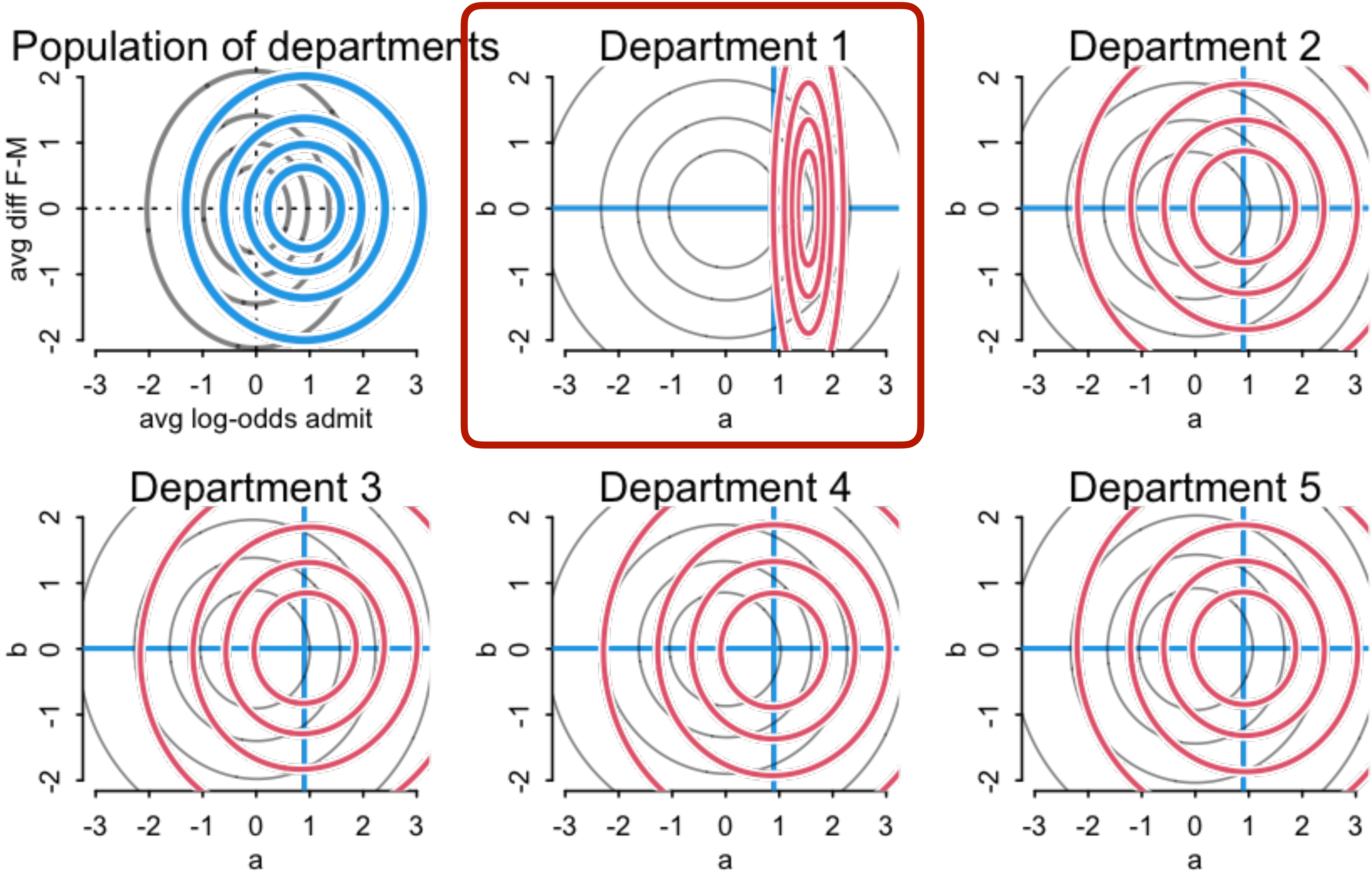
Population of departments

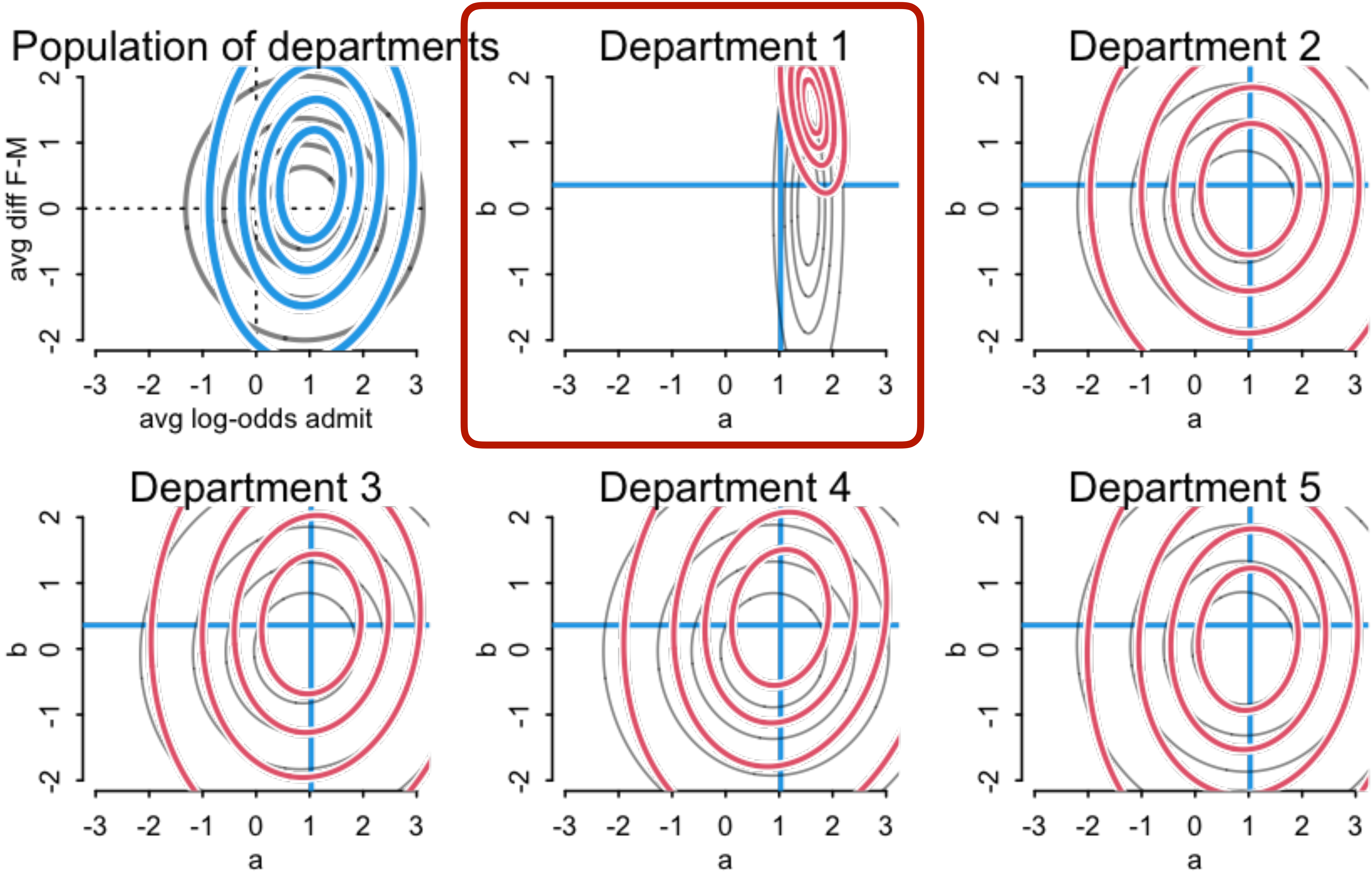



Population of departments
avg diff F-M
avg log-odds admit
Department 1
b
a
-3
-2
-1
0
1
2
3

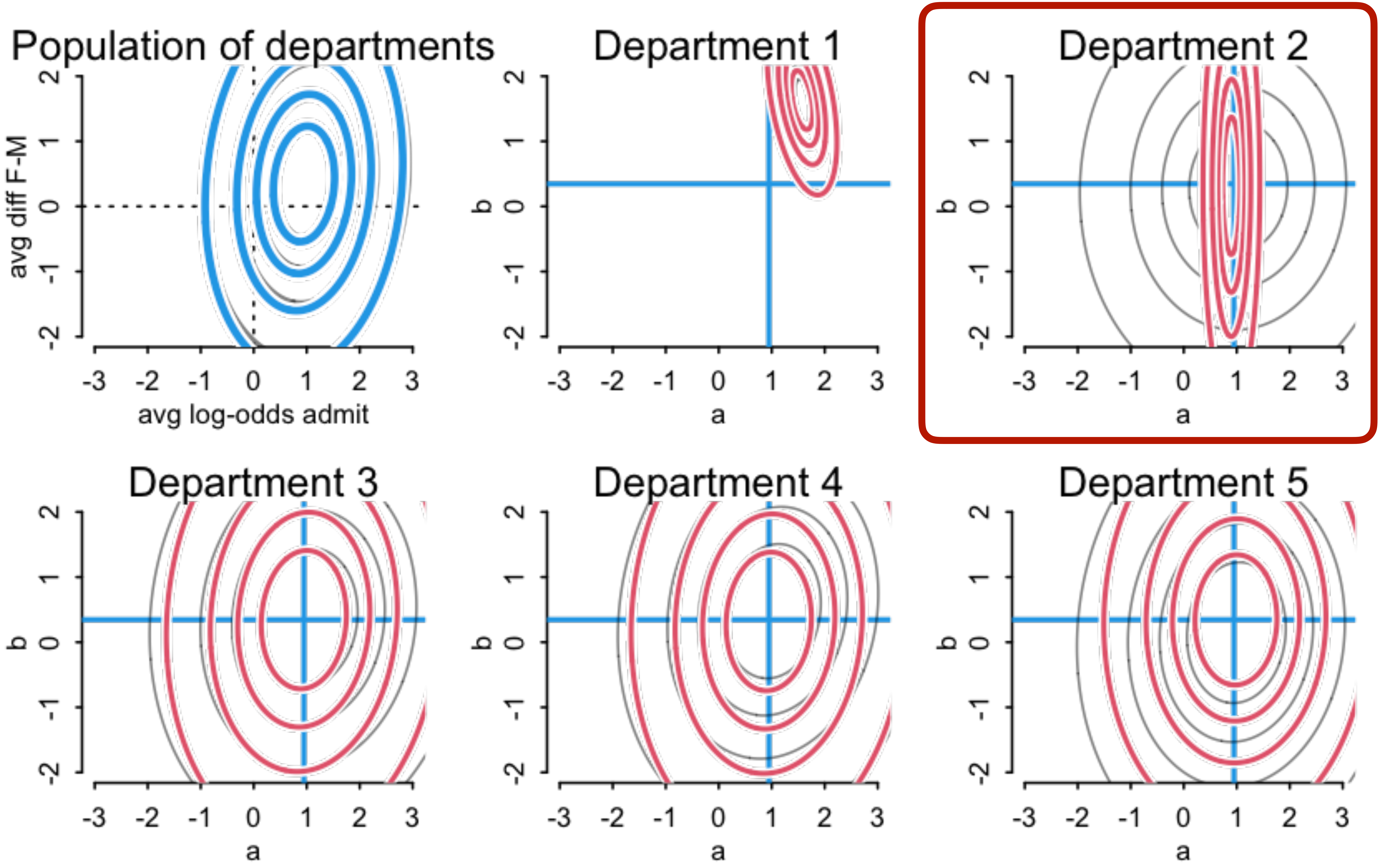


Population of departments
avg diff F-M
avg log-odds admit
Department 1
Department 2
Department 3
Department 4
Department 5
a
b
-3 -2 -1 0 1 2 3
-2 -1 0 1 2

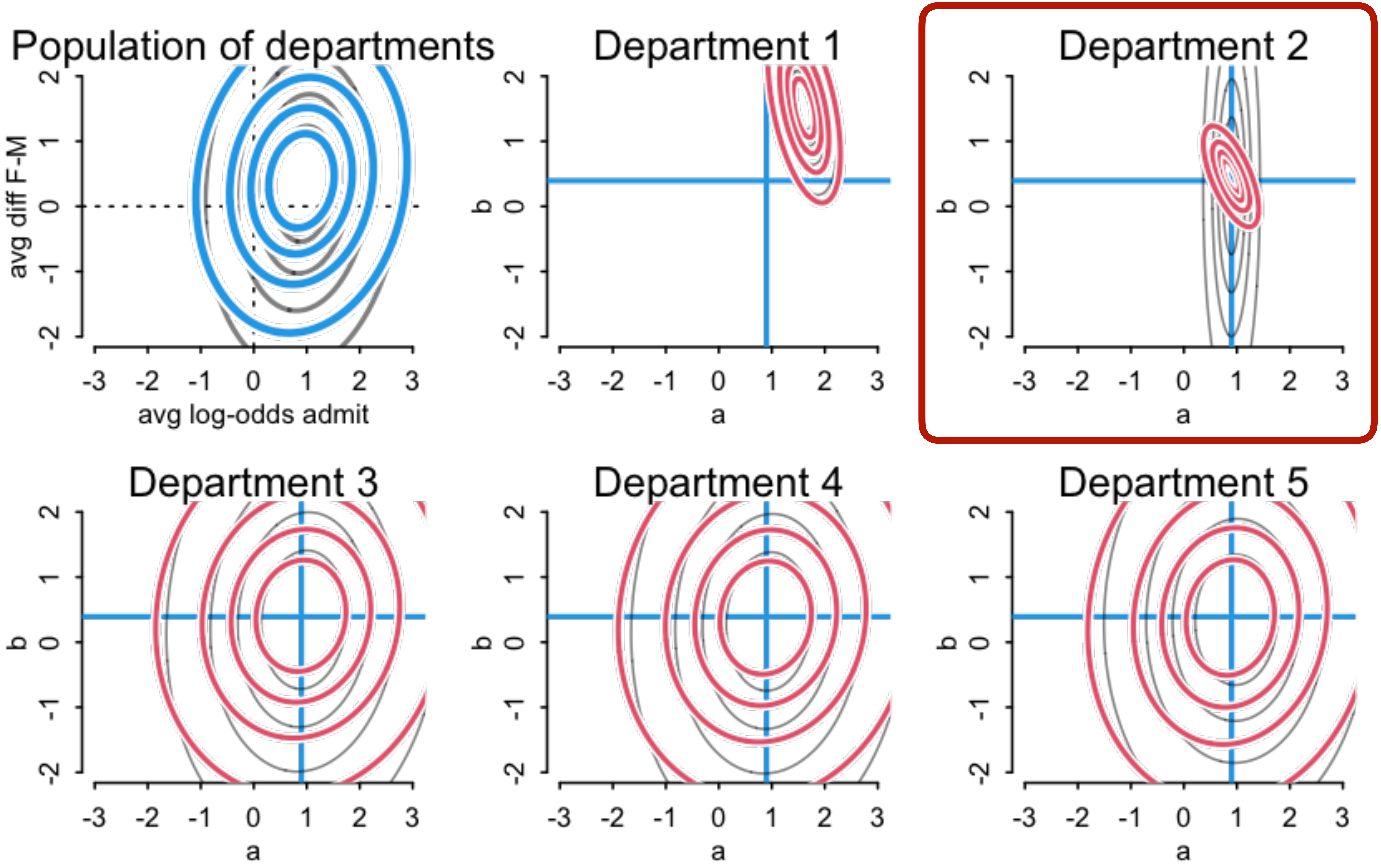

Population of departments
avg diff F-M
avg log-odds admit
Department 1
Department 2
Department 3
Department 4
Department 5
a
b
-3 -2 -1 0 1 2 3
-2 -1 0 1 2



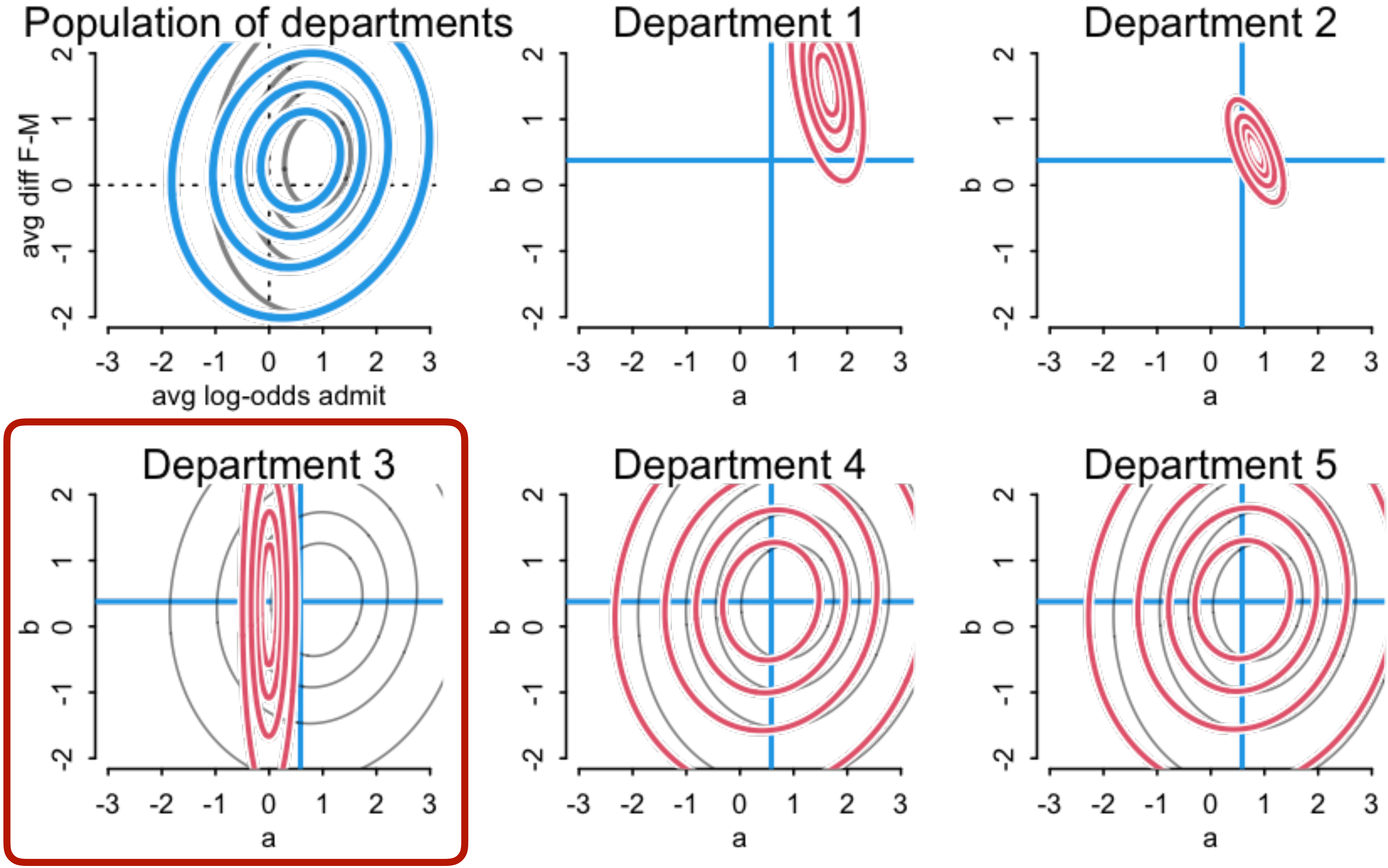
Population of departments
avg diff F-M
avg log-odds admit
Department 1
Department 2
Department 3
Department 4
Department 5
a
b
-3 -2 -1 0 1 2 3
-2 -1 0 1 2



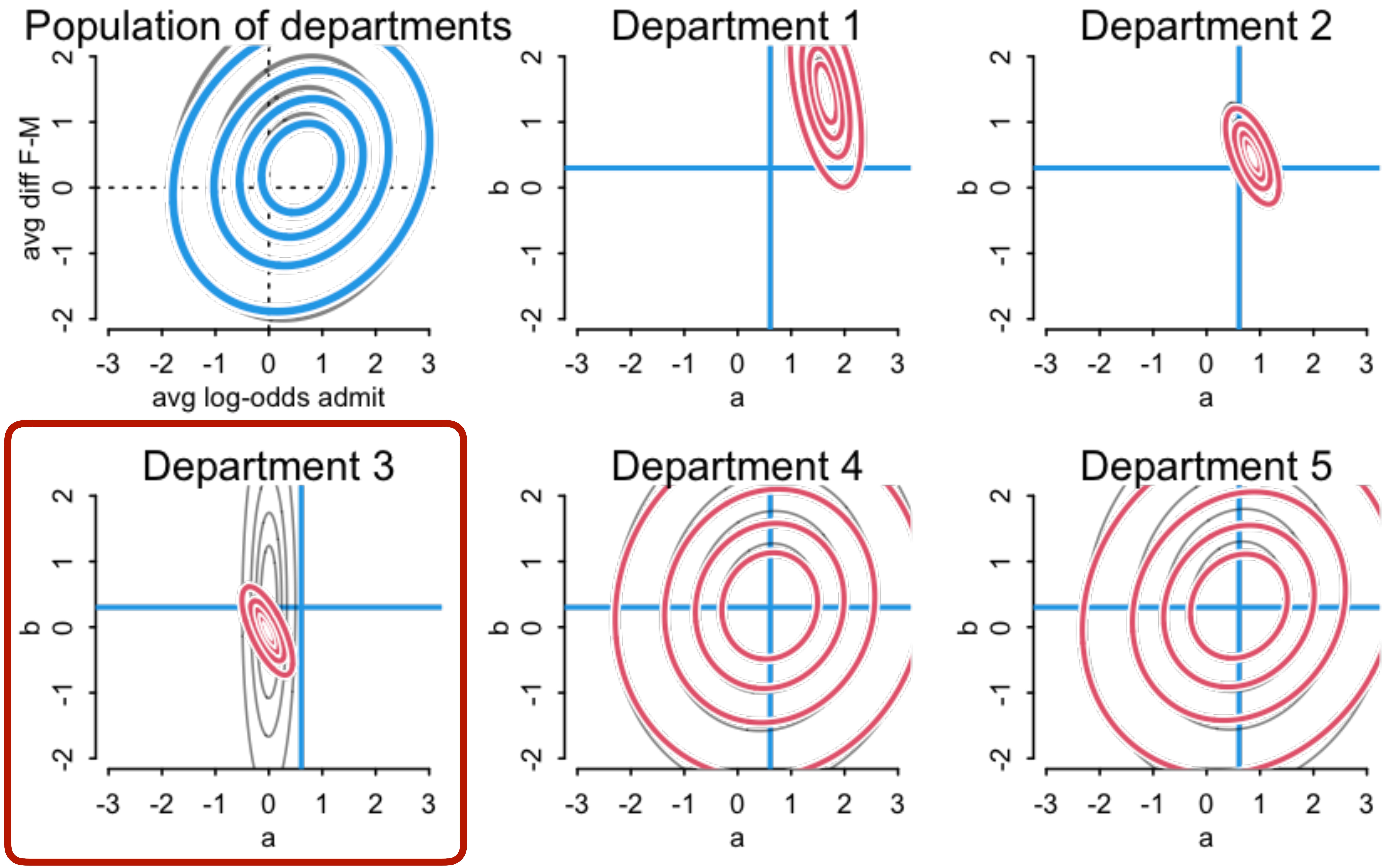
Population of departments
avg diff F-M
avg log-odds admit
Department 1
Department 2
Department 3
Department 4
Department 5
a
b
-3 -2 -1 0 1 2 3
-2 -1 0 1 2



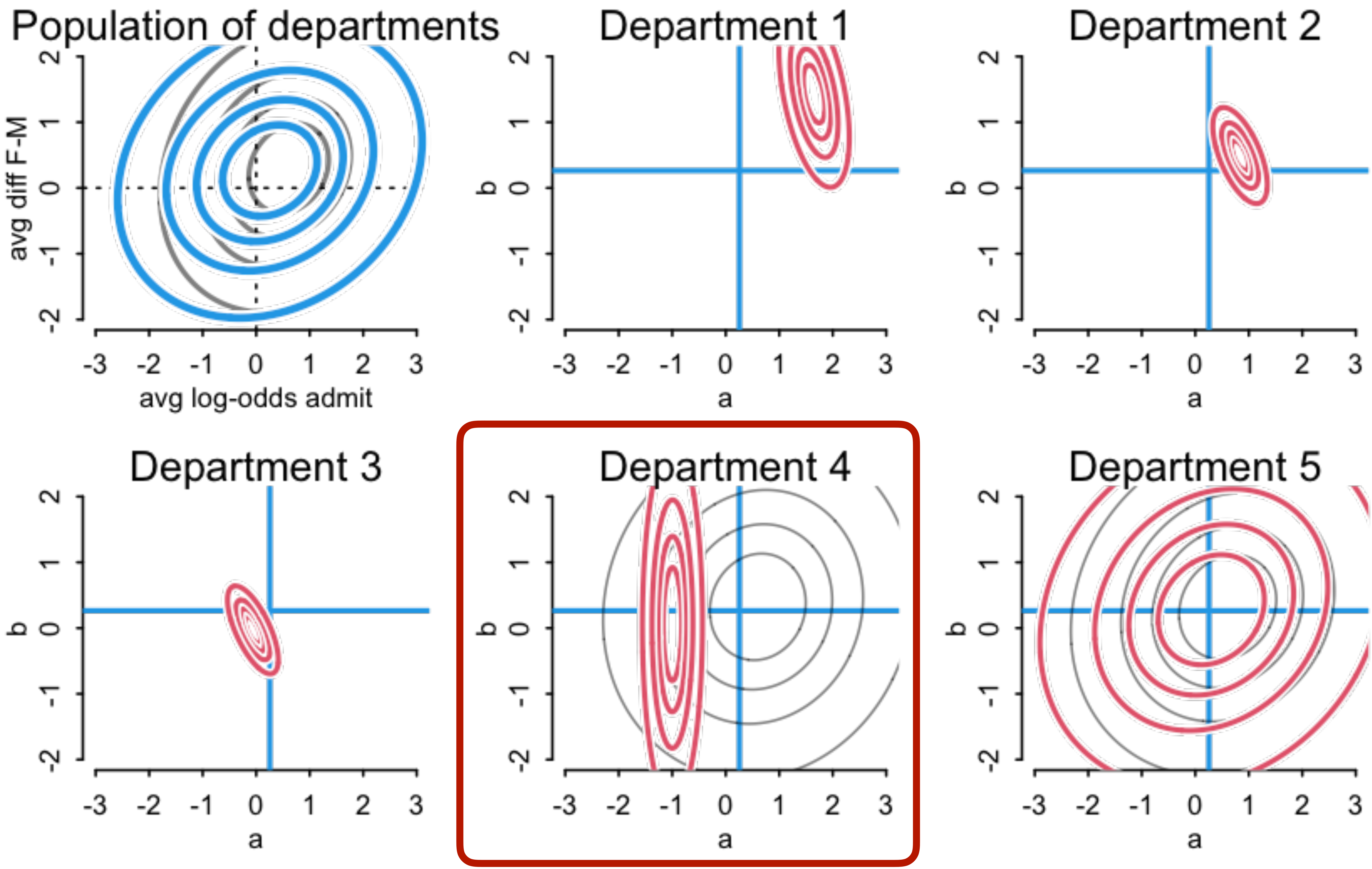
Population of departments
avg diff F-M
avg log-odds admit
Department 1
Department 2
Department 3
Department 4
Department 5
a
b
-3
-2
-1
0
1
2
3

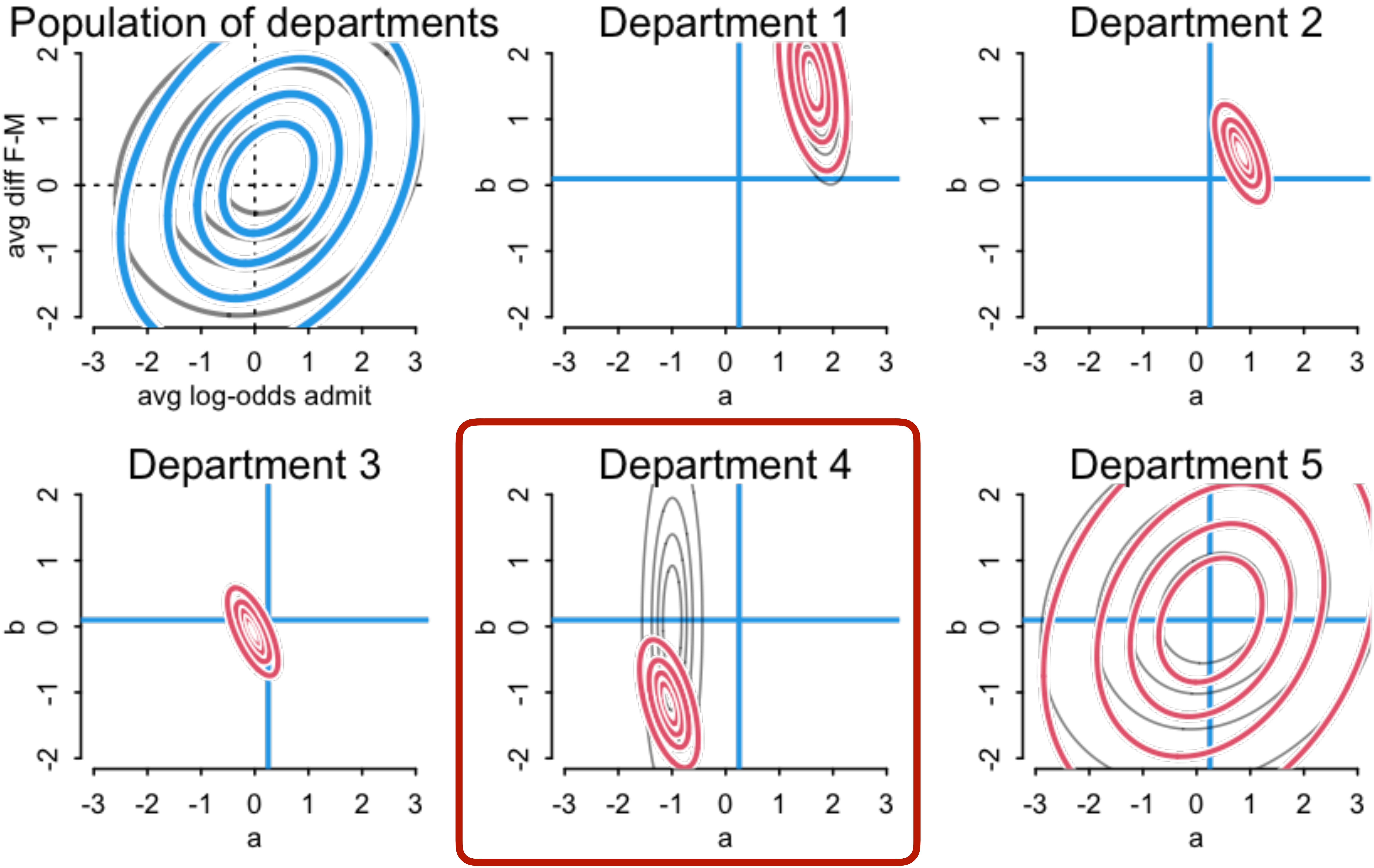


Population of departments
avg diff F-M
avg log-odds admit
Department 1
Department 2
Department 3
Department 4
Department 5
a
b
-3 -2 -1 0 1 2 3
-2 -1 0 1 2

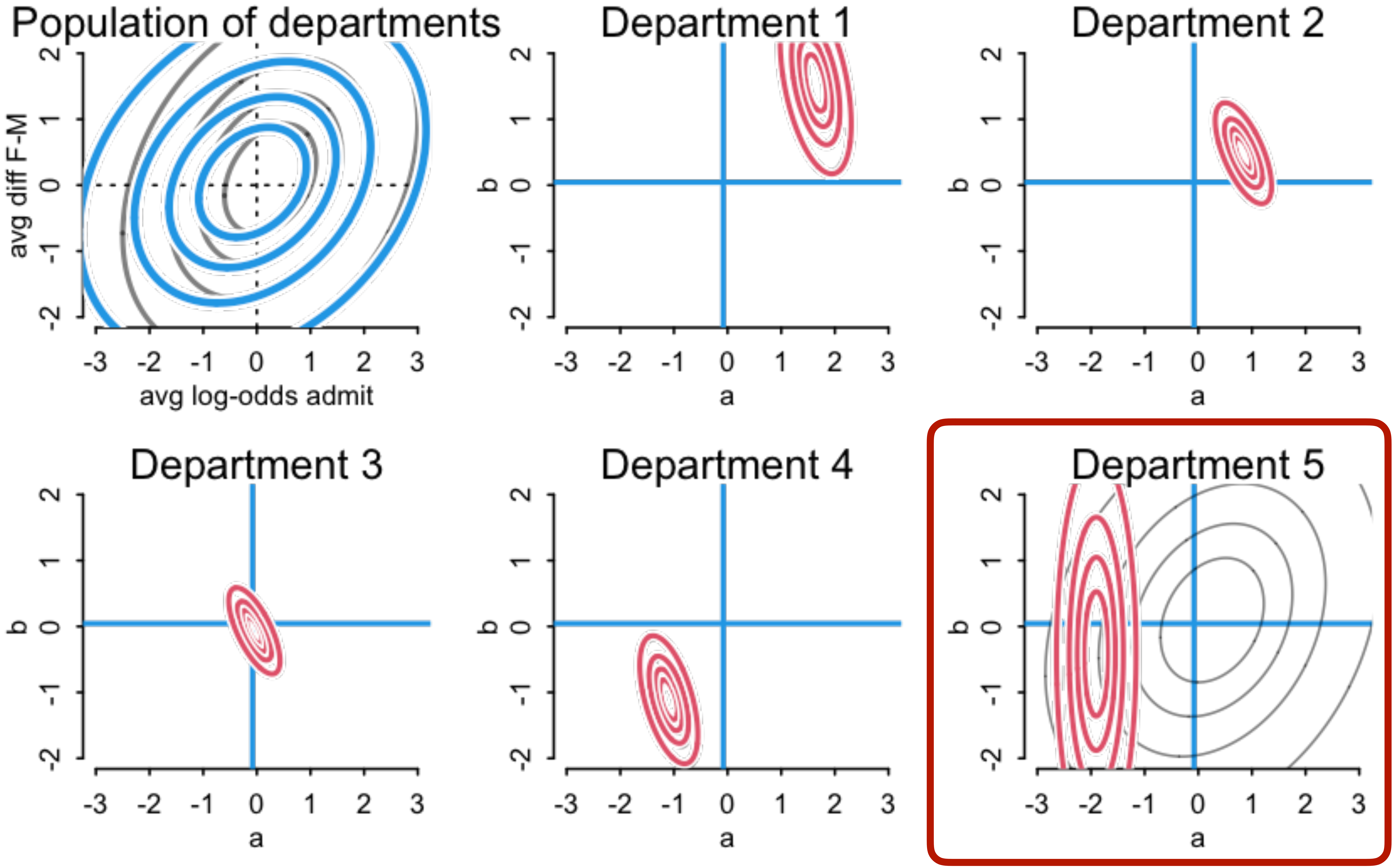


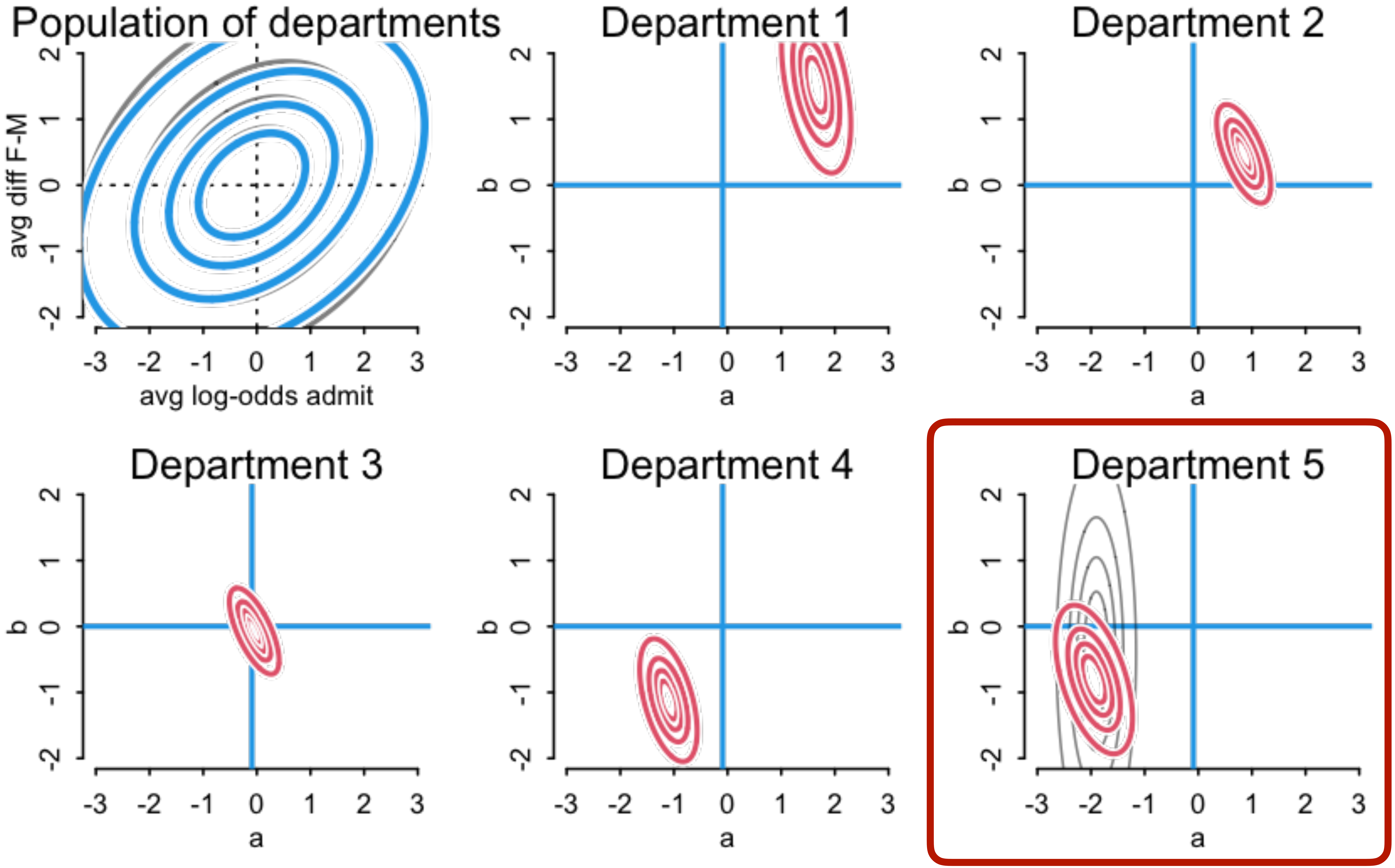
Population of departments
avg diff F-M
avg log-odds admit
Department 1
Department 2
Department 3
Department 4
Department 5
a
b
-3 -2 -1 0 1 2 3
-2 -1 0 1 2


Population of departments
avg diff F-M
avg log-odds admit
Department 1
Department 2
Department 3
Department 4
Department 5
a
b
-3
-2
-1
0
1
2
3



Population of departments
avg diff F-M
avg log-odds admit
Department 1
Department 2
Department 3
Department 4
Department 5
a
b
-3 -2 -1 0 1 2 3
-2 -1 0 1 2


Population of departments
avg diff F-M
avg log-odds admit
Department 1
Department 2
Department 3
Department 4
Department 5
a
b
-3 -2 -1 0 1 2 3
-2 -1 0 1 2



Population of departments
avg diff F-M
avg log-odds admit
Department 1
Department 2
Department 3
Department 4
Department 5
a
b
-3 -2 -1 0 1 2 3
-2 -1 0 1 2

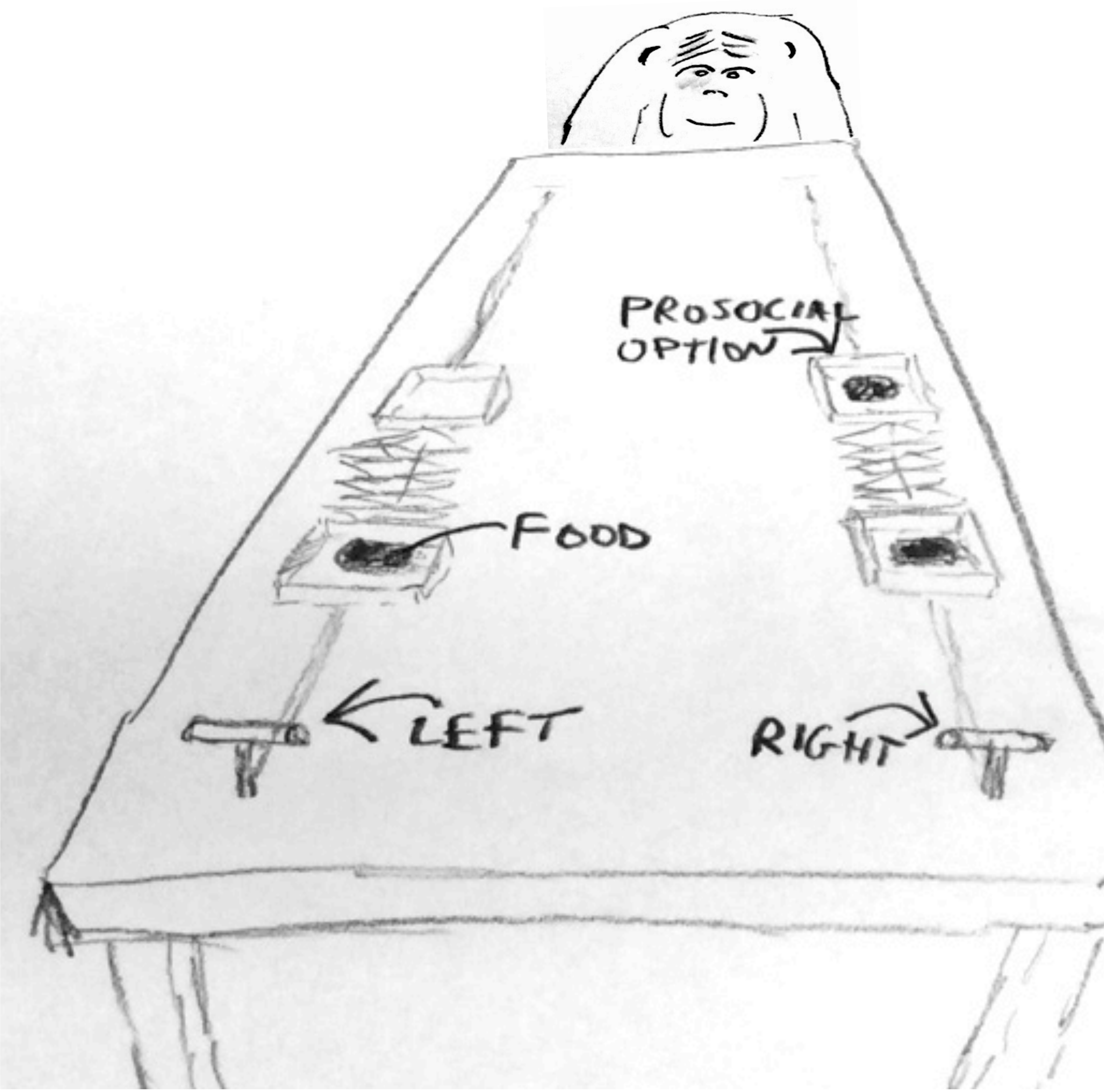
Prosocial chimpanzees

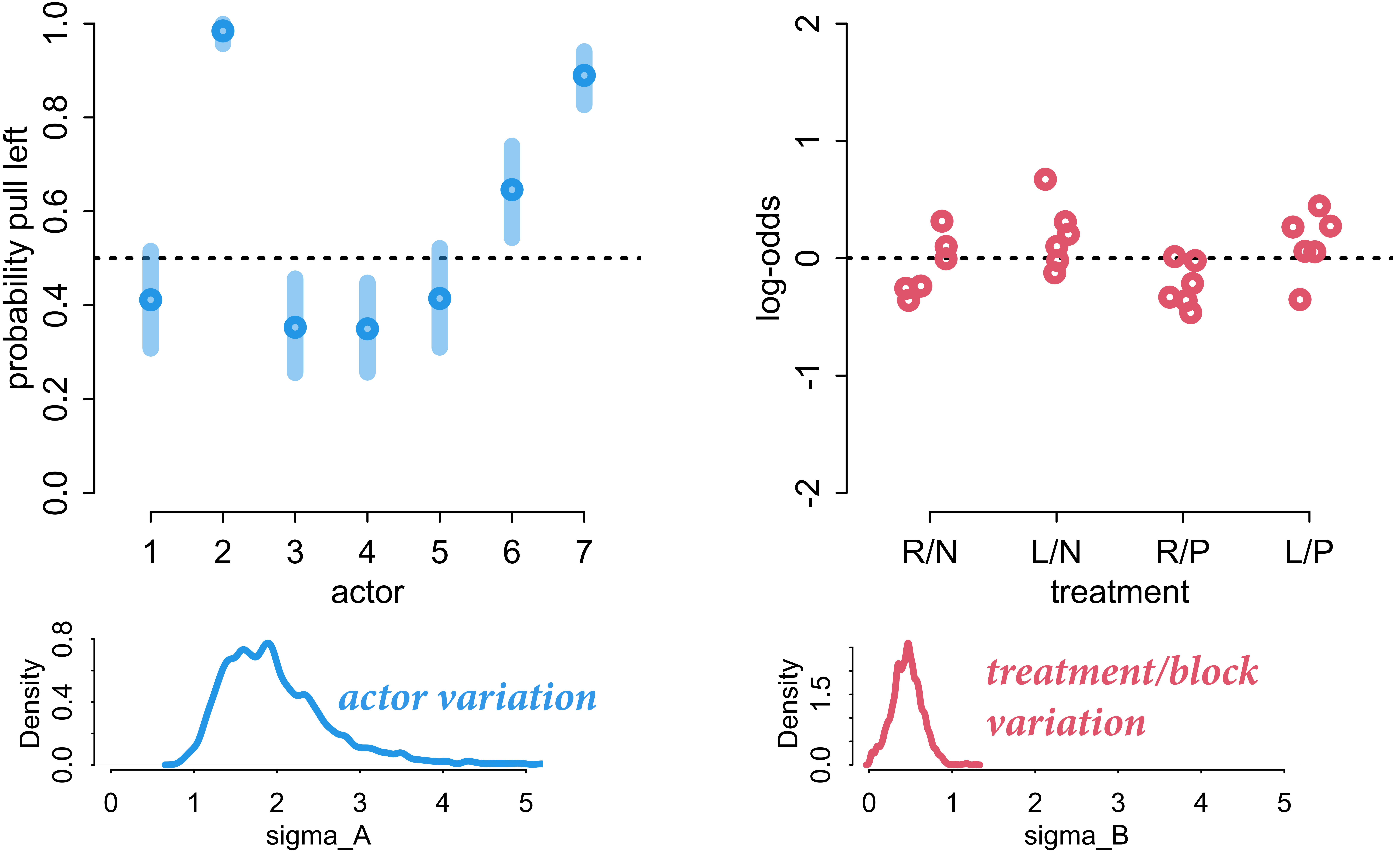
data(chimpanzees)

504 trials, 7 actors, 6 blocks

4 treatments:

(1) right, no partner
(2) left, no partner
(3) right, partner
(4) left, partner



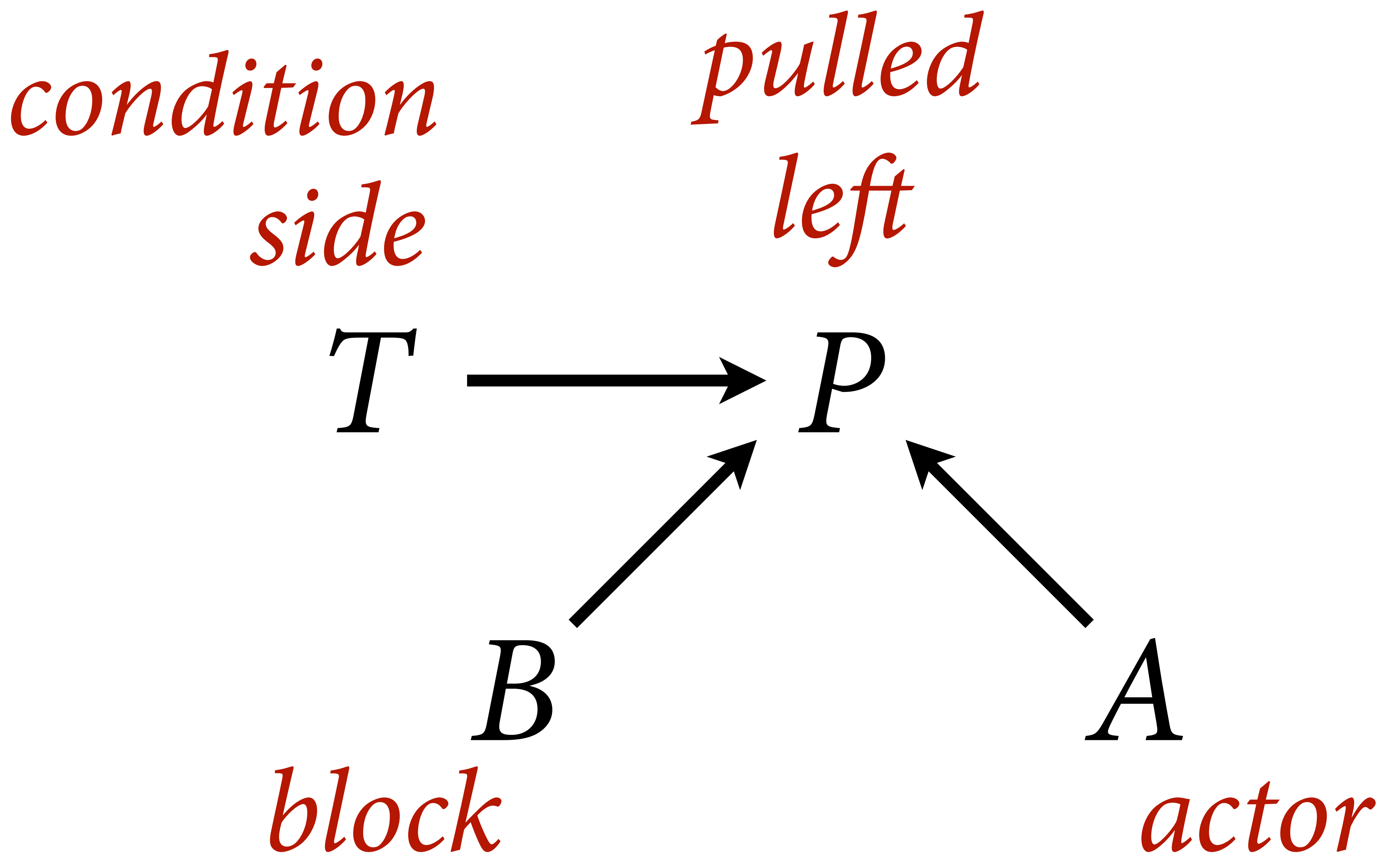


probability pull left
0.0
0.2
0.4
0.6
0.8
1.0
1
2
3
4
5
6
7
actor
log-odds
-2
-1
0
1
2
R/N
L/N
R/P
L/P
treatment
Density
0.0
0.4
0.8
0
1
2
3
4
5
sigma_A
actor variation
Density
0.0
1.5
sigma_B
treatment/block
variation

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

$$\text{logit}(p_i) = \bar{\alpha}_{A[i]} + \alpha_{A[i],T[i]} + \bar{\beta}_{B[i]} + \beta_{B[i],T[i]}$$

mean + Actor-treatment + Block-treatment



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

mean for each actor

actor offset for each treatment

$$\text{logit}(p_i) = \bar{\alpha}_{A[i]} + \alpha_{A[i],T[i]} + \bar{\beta}_{B[i]} + \beta_{B[i],T[i]}$$

mean for each block

block offset for each treatment

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

$$\text{logit}(p_i) = \bar{\alpha}_{A[i]} + \alpha_{A[i],T[i]} + \bar{\beta}_{B[i]} + \beta_{B[i],T[i]}$$

mean + Actor-treatment + Block-treatment

$$\alpha_j \sim \text{MVNormal}([0,0,0,0], \mathbf{R}_A, \mathbf{S}_A)$$
$$\text{for } j \in 1..7$$

prior for covarying actor-treatment effects
a vector of 4 parameters for each actor j

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

$$\text{logit}(p_i) = \bar{\alpha}_{A[i]} + \alpha_{A[i],T[i]} + \bar{\beta}_{B[i]} + \beta_{B[i],T[i]}$$

mean + Actor-treatment + Block-treatment

$$\alpha_j \sim \text{MVNormal}([0,0,0,0], \mathbf{R}_A, \mathbf{S}_A) \quad \text{for } j \in 1..7$$

prior for covarying actor-treatment effects
a vector of 4 parameters for each actor j

$$\beta_k \sim \text{MVNormal}([0,0,0,0], \mathbf{R}_B, \mathbf{S}_B) \quad \text{for } k \in 1..6$$

prior for covarying block-treatment effects
a vector of 4 parameters for each block k

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

$$\text{logit}(p_i) = \bar{\alpha}_{A[i]} + \alpha_{A[i],T[i]} + \bar{\beta}_{B[i]} + \beta_{B[i],T[i]}$$

mean + Actor-treatment + Block-treatment

$$\alpha_j \sim \text{MVNormal}([0,0,0,0], \mathbf{R}_A, \mathbf{S}_A)$$
$$\text{for } j \in 1..7$$

prior for covarying actor-treatment effects
a vector of 4 parameters for each actor j

$$\beta_k \sim \text{MVNormal}([0,0,0,0], \mathbf{R}_B, \mathbf{S}_B)$$
$$\text{for } k \in 1..6$$

prior for covarying block-treatment effects
a vector of 4 parameters for each block k

$$\bar{\alpha}_j \sim \text{Normal}(0, \tau_A)$$
$$\bar{\beta}_k \sim \text{Normal}(0, \tau_B)$$

priors for actor and treatment means

$$\text{S}_{A,j}, \text{S}_{B,j}, \tau_A, \tau_B \sim \text{Exponential}(1)$$
$$\text{for } j \in 1..4$$

each standard deviation gets same prior
one standard deviation for each treatment

$$\mathbf{R}_A, \mathbf{R}_B \sim \text{LKJcorr}(4)$$

correlation matrix prior

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

$$\text{logit}(p_i) = \bar{\alpha}_{A[i]} + \alpha_{A[i],T[i]} + \bar{\beta}_{B[i]} + \beta_{B[i],T[i]}$$

$$\alpha_j \sim \text{MVNormal}([0,0,0,0], \mathbf{R}_A, \mathbf{S}_A) \quad \text{for } j \in 1..7$$

$$\beta_k \sim \text{MVNormal}([0,0,0,0], \mathbf{R}_B, \mathbf{S}_B) \quad \text{for } k \in 1..6$$

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

$$\bar{\beta}_k \sim \text{Normal}(0, \tau_B)$$

$$\text{S}_{A,j}, \text{S}_{B,j}, \tau_A, \tau_B \sim \text{Exponential}(1) \quad \text{for } j \in 1..4$$

$$\mathbf{R}_A, \mathbf{R}_B \sim \text{LKJcorr}(4)$$

```r
data(chimpanzees)
d <- chimpanzees
dat <- list(
    P = d$pulled_left,
    A = as.integer(d$actor),
    B = as.integer(d$block),
    T = 1L + d$prosoc_left + 2L*d$condition)

m14.2 <- ulam(
alist(
    P ~ bernoulli(p),
    logit(p) <- abar[A] + a[A,T] + bbar[B] + b[B,T],

    # adaptive priors
    vector[4]:a[A] ~ multi_normal(0,Rho_A,sigma_A),
    vector[4]:b[B] ~ multi_normal(0,Rho_B,sigma_B),
    abar[A] ~ normal(0,tau_A),
    bbar[B] ~ normal(0,tau_B),

    # fixed priors
    c(tau_A,tau_B) ~ exponential(1),
    sigma_A ~ exponential(1),
    Rho_A ~ dlkjcorr(4),
    sigma_B ~ exponential(1),
    Rho_B ~ dlkjcorr(4)
) , data=dat , chains=4 , cores=4 )
```

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

$$\text{logit}(p_i) = \bar{\alpha}_{A[i]} + \alpha_{A[i],T[i]} + \bar{\beta}_{B[i]} + \beta$$

$$\alpha_j \sim \text{MVNormal}(0, \mathbf{R}_A, \mathbf{S}_A)$$

$$\beta_k \sim \text{MVNormal}(0, \mathbf{R}_B, \mathbf{S}_B)$$

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

$$\bar{\beta}_k \sim \text{Normal}(0, \tau_B)$$

$$\text{S}_{A,j}, \text{S}_{B,j}, \tau_A, \tau_B \sim \text{Exponential}(1)$$

$$\text{R}_A, \text{R}_B \sim \text{LKJcorr}(4)$$

```r
data(chimpanzees)
d <- chimpanzees
dat <- list(
    P = d$pulled_left,
    A = as.integer(d$actor),
    B = as.integer(d$block),
    T = 1L + d$prosoc_left + 2L*d$condition)

m14.2 <- ulam(
alist(
    P ~ bernoulli(p),
    logit(p) <- abar[A] + a[A,T] + bbar[B] + b[B,T],

    # adaptive priors
    vector[4]:a[A] ~ multi_normal(0,Rho_A,sigma_A),
    vector[4]:b[B] ~ multi_normal(0,Rho_B,sigma_B),
    abar[A] ~ normal(0,tau_A),
    bbar[B] ~ normal(0,tau_B),

    # fixed priors
    c(tau_A,tau_B) ~ exponential(1),
    sigma_A ~ exponential(1),
    Rho_A ~ dlkjcorr(4),
    sigma_B ~ exponential(1),
    Rho_B ~ dlkjcorr(4)
) , data=dat , chains=4 , cores=4 )
```

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

$$\text{logit}(p_i) = \bar{\alpha}_{A[i]} + \alpha_{A[i],T[i]} + \bar{\beta}_{B[i]} + \beta$$

$$\alpha_j \sim \text{MVNormal}(0, \mathbf{R}_A, \mathbf{S}_A)$$

$$\beta_k \sim \text{MVNormal}(0, \mathbf{R}_B, \mathbf{S}_B)$$

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

$$\bar{\beta}_k \sim \text{Normal}(0, \tau_B)$$

$$\text{S}_{A,j}, \text{S}_{B,j}, \tau_A, \tau_B \sim \text{Exponential}(1)$$

$$\text{R}_A, \text{R}_B \sim \text{LKJcorr}(4)$$

```r
data(chimpanzees)
d <- chimpanzees
dat <- list(
    P = d$pulled_left,
    A = as.integer(d$actor),
    B = as.integer(d$block),
    T = 1L + d$prosoc_left + 2L*d$condition)

m14.2 <- ulam(
alist(
    P ~ bernoulli(p),
    logit(p) <- abar[A] + a[A,T] + bbar[B] + b[B,T],

    # adaptive priors
    vector[4]:a[A] ~ multi_normal(0,Rho_A,sigma_A),
    vector[4]:b[B] ~ multi_normal(0,Rho_B,sigma_B),
    abar[A] ~ normal(0,tau_A),
    bbar[B] ~ normal(0,tau_B),

    # fixed priors
    c(tau_A,tau_B) ~ exponential(1),
    sigma_A ~ exponential(1),
    Rho_A ~ dlkjcorr(4),
    sigma_B ~ exponential(1),
    Rho_B ~ dlkjcorr(4)
) , data=dat , chains=4 , cores=4 )
```

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

$$\text{logit}(p_i) = \bar{\alpha}_{A[i]} + \alpha_{A[i],T[i]} + \bar{\beta}_{B[i]} + \beta$$

$$\alpha_j \sim \text{MVNormal}(0, \mathbf{R}_A, \mathbf{S}_A)$$

$$\beta_k \sim \text{MVNormal}(0, \mathbf{R}_B, \mathbf{S}_B)$$

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

$$\bar{\beta}_k \sim \text{Normal}(0, \tau_B)$$

$$\text{S}_{A,j}, \text{S}_{B,j}, \tau_A, \tau_B \sim \text{Exponential}(1)$$

$$\text{R}_A, \text{R}_B \sim \text{LKJcorr}(4)$$

```r
data(chimpanzees)
d <- chimpanzees
dat <- list(
    P = d$pulled_left,
    A = as.integer(d$actor),
    B = as.integer(d$block),
    T = 1L + d$prosoc_left + 2L*d$condition)

m14.2 <- ulam(
alist(
    P ~ bernoulli(p),
    logit(p) <- abar[A] + a[A,T] + bbar[B] + b[B,T],

    # adaptive priors
    vector[4]:a[A] ~ multi_normal(0,Rho_A,sigma_A),
    vector[4]:b[B] ~ multi_normal(0,Rho_B,sigma_B),
    abar[A] ~ normal(0,tau_A),
    bbar[B] ~ normal(0,tau_B),

    # fixed priors
    c(tau_A,tau_B) ~ exponential(1),
    sigma_A ~ exponential(1),
    Rho_A ~ dlkjcorr(4),
    sigma_B ~ exponential(1),
    Rho_B ~ dlkjcorr(4)
) , data=dat , chains=4 , cores=4 )
```

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

$$\text{logit}(p_i) = \bar{\alpha}_{A[i]} + \alpha_{A[i],T[i]} + \bar{\beta}_{B[i]} + \beta$$

$$\alpha_j \sim \text{MVNormal}(0, \mathbf{R}_A, \mathbf{S}_A)$$

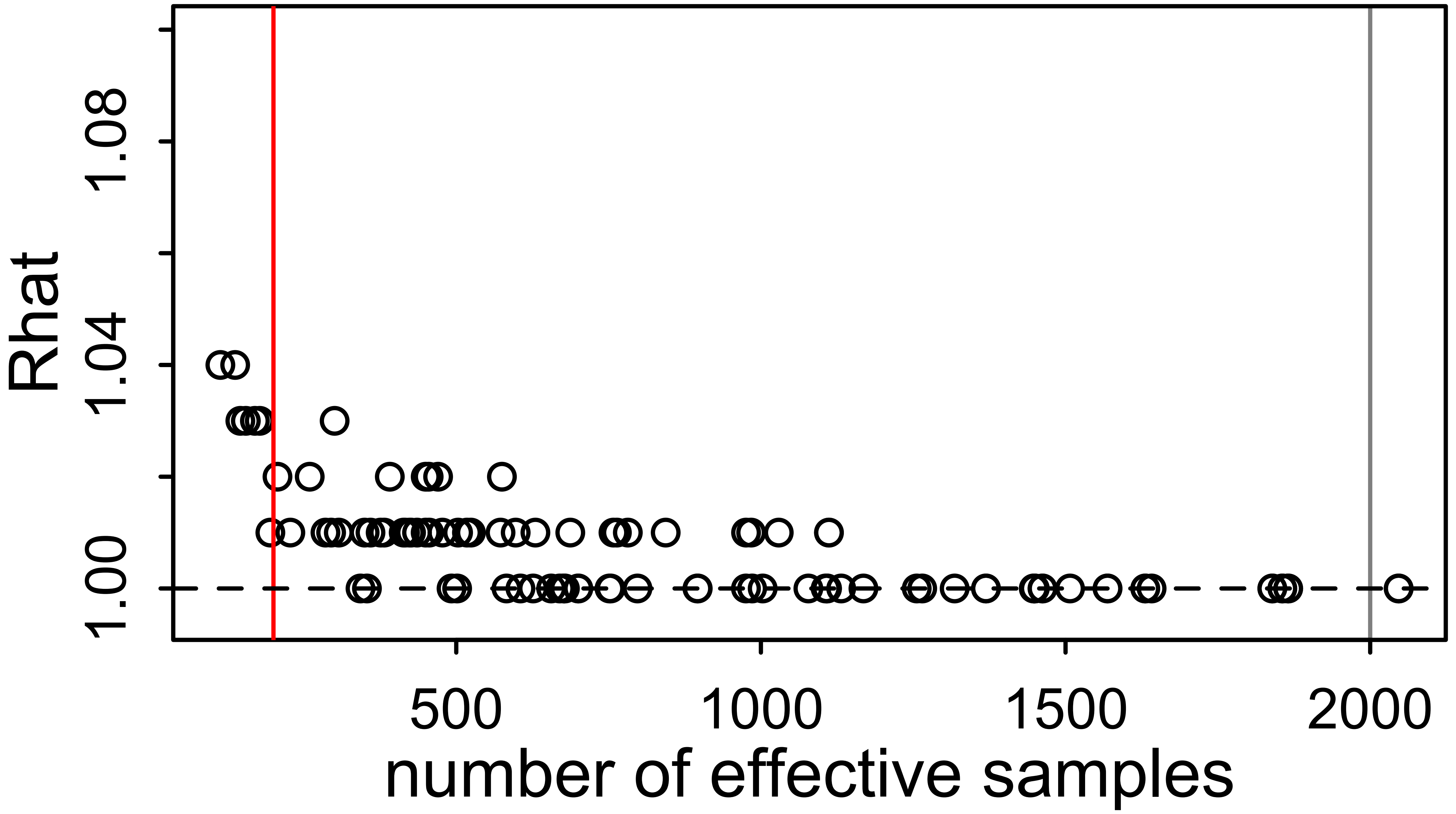
$$\beta_k \sim \text{MVNormal}(0, \mathbf{R}_B, \mathbf{S}_B)$$

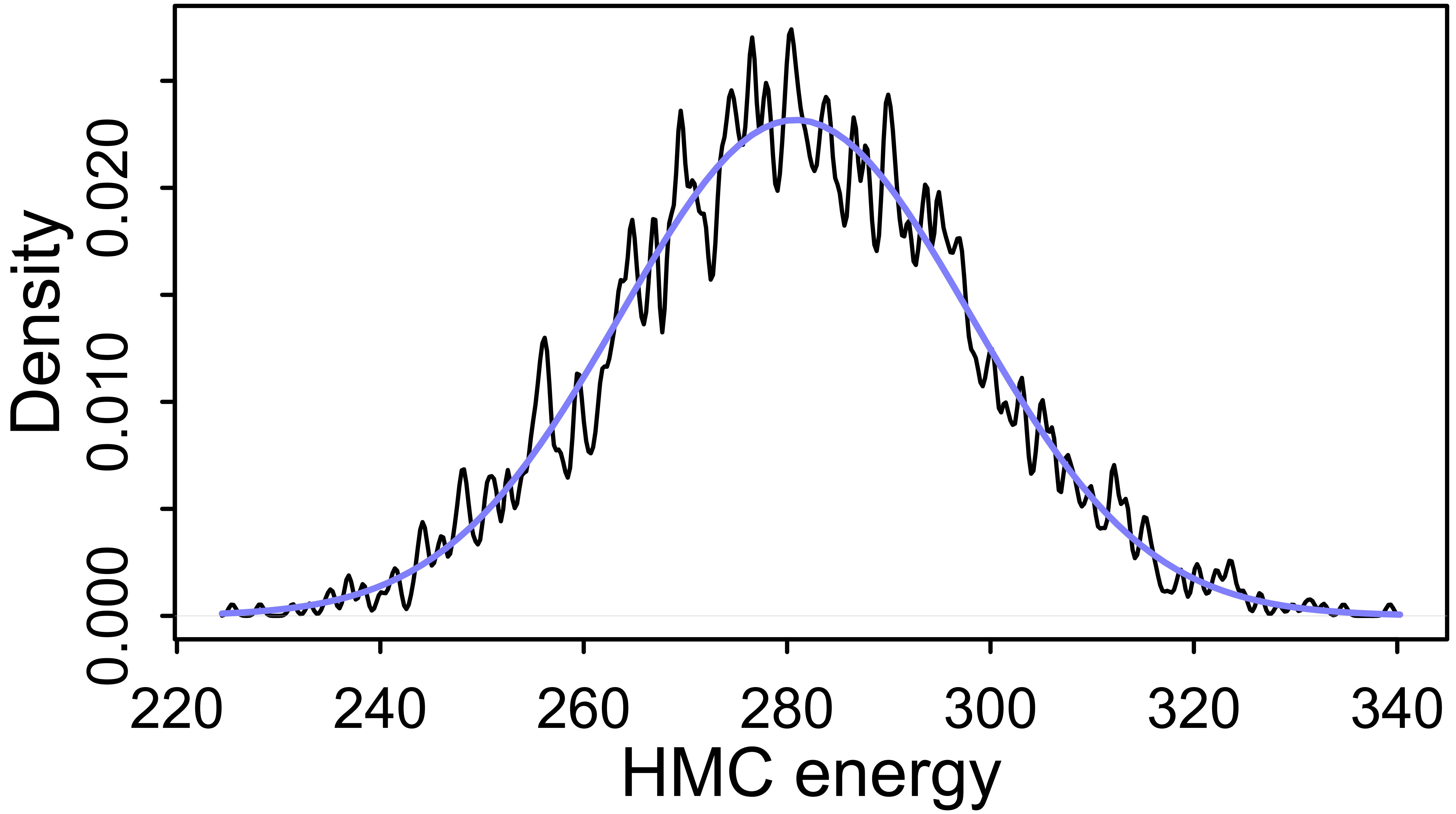
$$\bar{\alpha}_j \sim \text{Normal}(0, \tau_A)$$

$$\bar{\beta}_k \sim \text{Normal}(0, \tau_B)$$

$$\mathrm{S}_{A,j}, \mathrm{S}_{B,j}, \tau_A, \tau_B \sim \text{Exponential}(1)$$

$$\mathrm{R}_A, \mathrm{R}_B \sim \text{LKJcorr}(4)$$

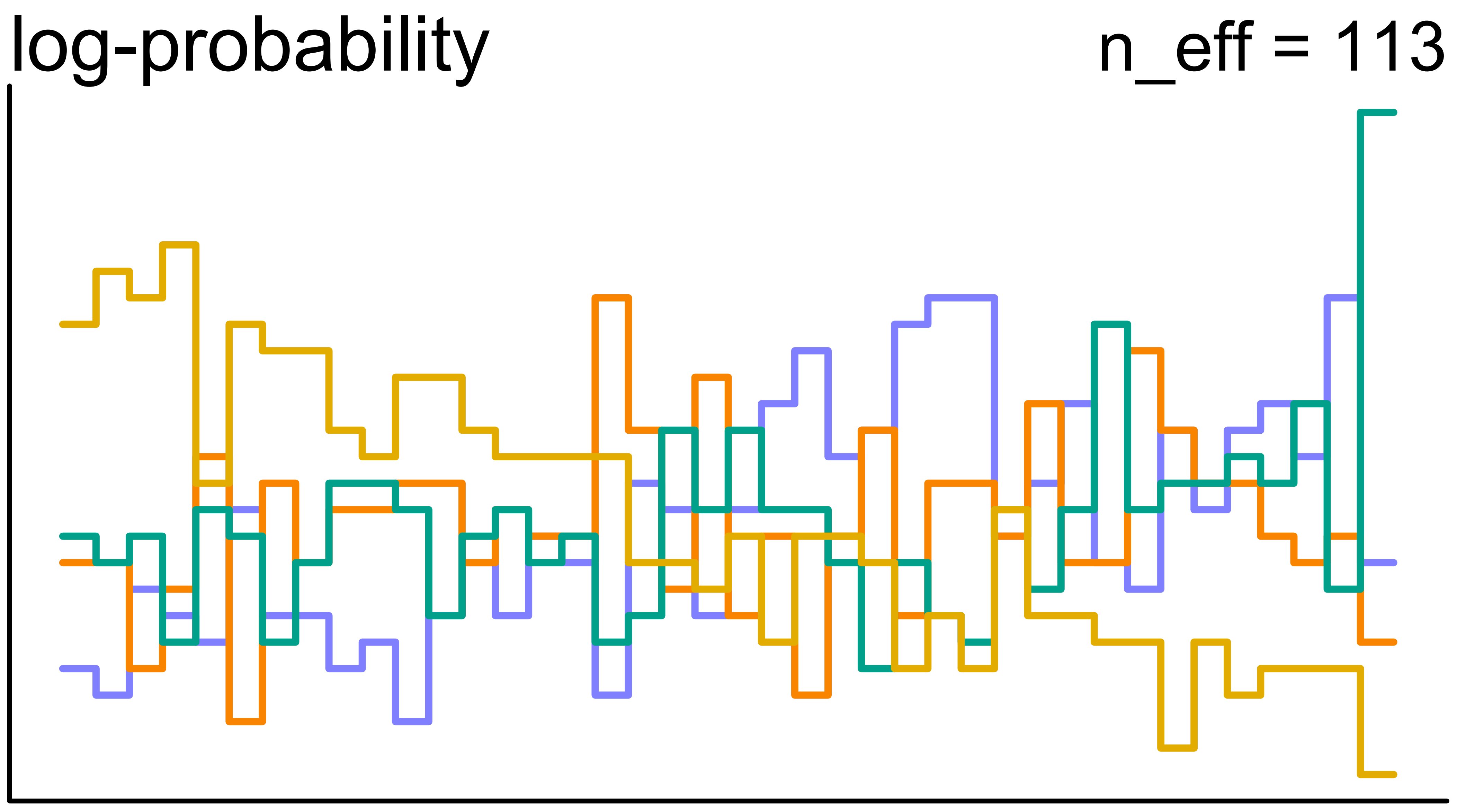




37

Divergent transitions

Check yourself before
you wreck yourself

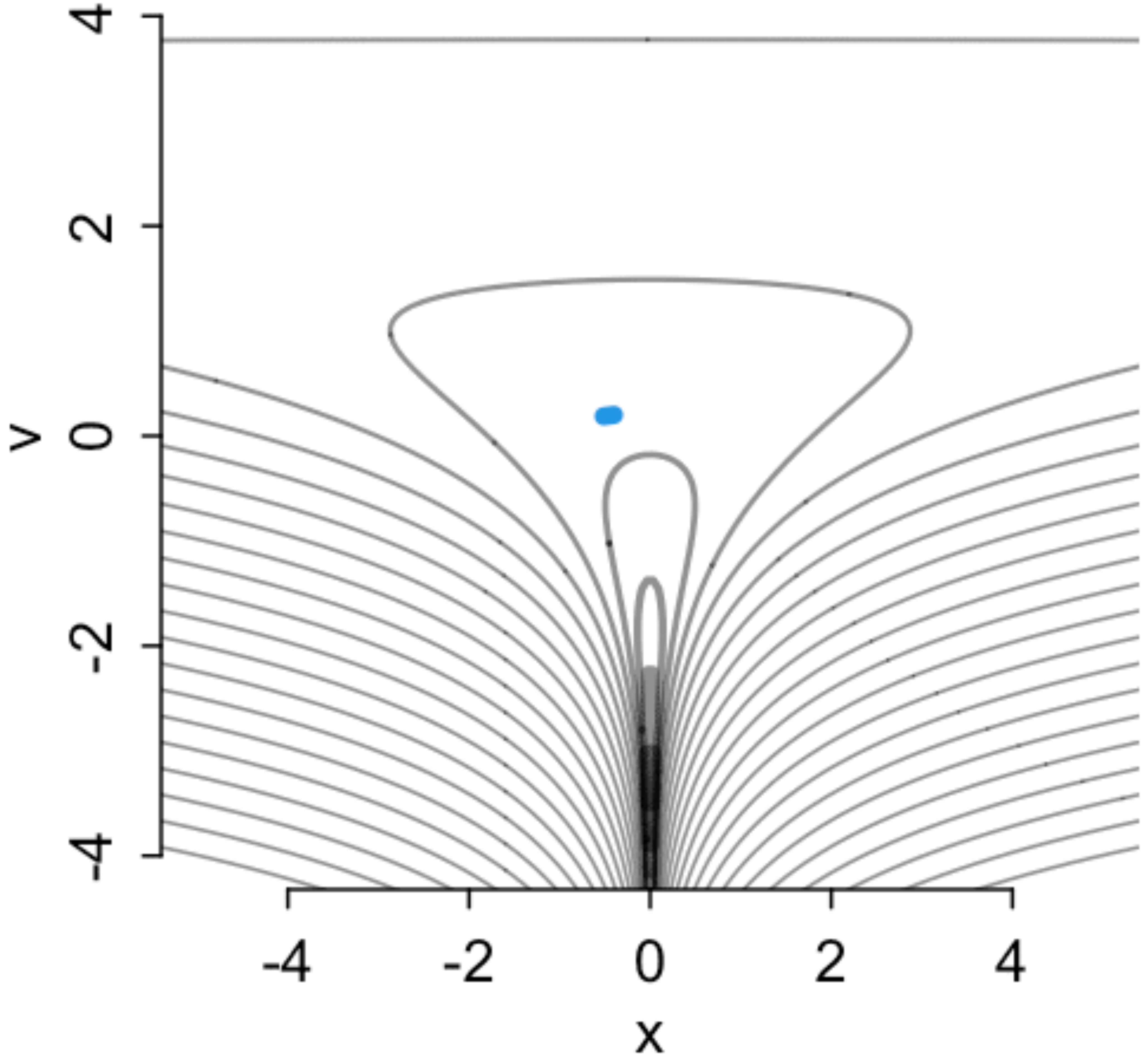


PAUSE

"Centered"

$v \sim \text{Normal}(0,3)$

$x \sim \text{Normal}(0, \exp(v))$

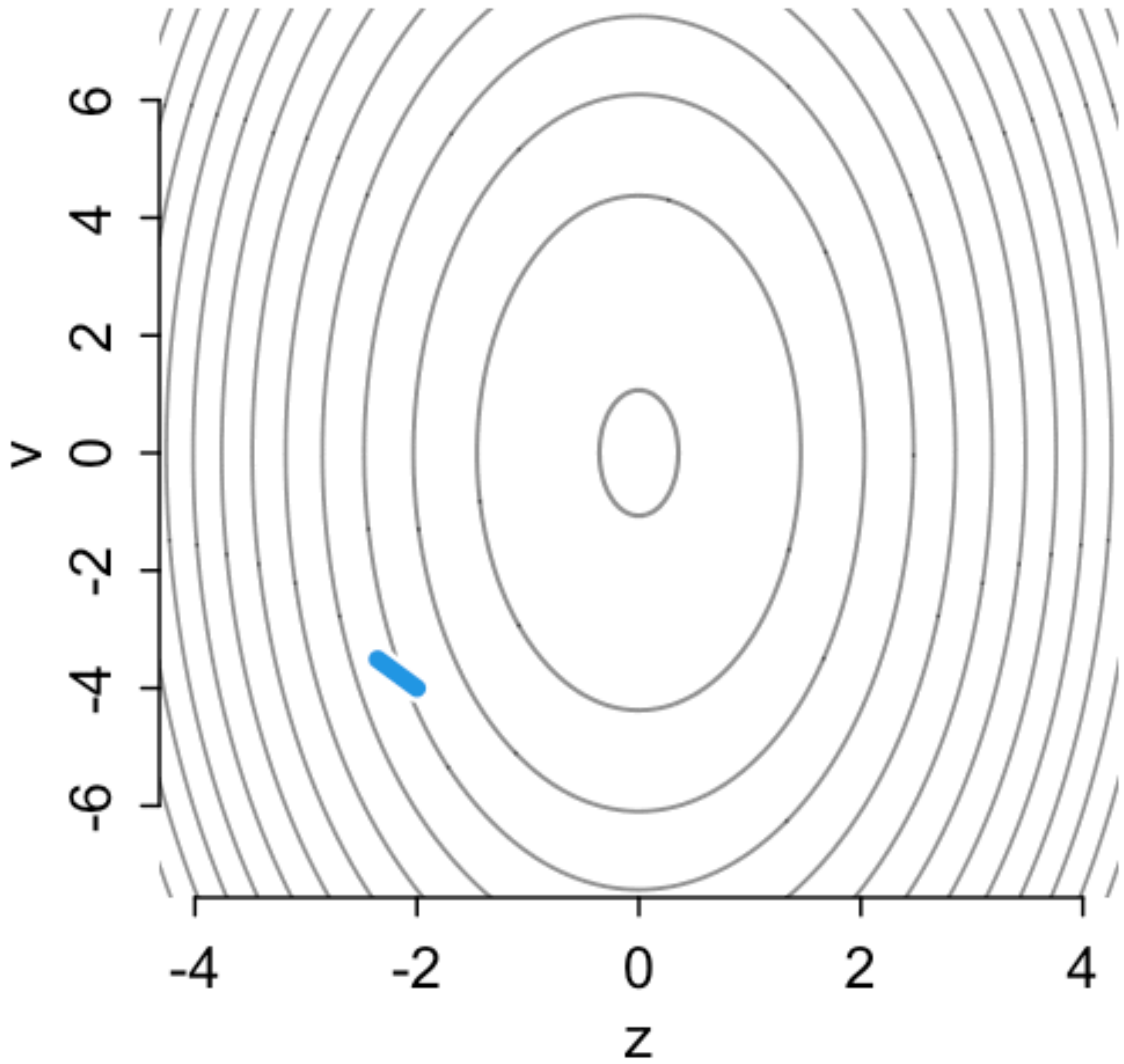


"Non-centered"

$v \sim \text{Normal}(0,3)$

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

$x = z \exp(v)$



Centered

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

$$\text{logit}(p_i) = \beta_{T[i],B[i]} + \alpha_{A[i]}$$

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

$$\beta_{j,k} \sim \text{Normal}(0, \sigma_B)$$

$$\sigma_A, \sigma_B \sim \text{Exponential}(1)$$

$$\bar{\alpha} \sim \text{Normal}(0, 1.5)$$

Centered

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

$$\text{logit}(p_i) = \beta_{T[i],B[i]} + \alpha_{A[i]}$$

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

$$\beta_{j,k} \sim \text{Normal}(0, \sigma_B)$$

$$\sigma_A, \sigma_B \sim \text{Exponential}(1)$$

$$\bar{\alpha} \sim \text{Normal}(0, 1.5)$$

Non-centered

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

$$\text{logit}(p_i) = \bar{\alpha} + (z_{\alpha,A[i]})\sigma_A + (z_{\beta,T[i],B[i]})\sigma_B$$

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

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

$$\sigma_A, \sigma_B \sim \text{Exponential}(1)$$

$$\bar{\alpha} \sim \text{Normal}(0, 1.5)$$

Centered

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

$$\text{logit}(p_i) = \beta_{T[i],B[i]} + \alpha_{A[i]}$$

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

$$\beta_{j,k} \sim \text{Normal}(0, \sigma_B)$$

$$\sigma_A, \sigma_B \sim \text{Exponential}(1)$$

$$\bar{\alpha} \sim \text{Normal}(0, 1.5)$$

Non-centered

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

$$\text{logit}(p_i) = \bar{\alpha} + \boxed{(z_{\alpha,A[i]})\sigma_A + (z_{\beta,T[i],B[i]})\sigma_B}$$

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

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

$$\sigma_A, \sigma_B \sim \text{Exponential}(1)$$

$$\bar{\alpha} \sim \text{Normal}(0, 1.5)$$

Centered

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

$$\text{logit}(p_i) = \bar{\alpha}_{A[i]} + \alpha_{A[i],T[i]} + \bar{\beta}_{B[i]} + \beta_{B[i],T[i]}$$

$$\alpha_j \sim \text{MVNormal}([0,0,0,0], \mathbf{R}_A, \mathbf{S}_A)$$

$$\beta_k \sim \text{MVNormal}([0,0,0,0], \mathbf{R}_B, \mathbf{S}_B)$$

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

$$\bar{\beta}_k \sim \text{Normal}(0, \tau_B)$$

$$\text{S}_{A,j}, \text{S}_{B,j}, \tau_A, \tau_B \sim \text{Exponential}(1)$$

$$\mathbf{R}_A, \mathbf{R}_B \sim \text{LKJcorr}(4)$$

How can we factor R and S out of the priors?

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

$$\text{logit}(p_i) = \bar{\alpha}_{A[i]} + \alpha_{A[i],T[i]} + \bar{\beta}_{B[i]} + \beta_{B[i],T[i]}$$

a 7-by-4 matrix

$$\alpha = \left(\underbrace{\begin{bmatrix} S_{A,1} & 0 & 0 & 0 \\ 0 & S_{A,2} & 0 & 0 \\ 0 & 0 & S_{A,3} & 0 \\ 0 & 0 & 0 & S_{A,4} \end{bmatrix}}_{\text{diagonal matrix of standard deviations}} \mathbf{L}_A \mathbf{Z}_{T,A} \right)^{\intercal}$$

transpose!
flips rows and columns

matrix of treatment-actor z-scores

diagonal matrix of standard deviations

Cholesky factor of correlation matrix across treatments

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

$$\text{logit}(p_i) = \bar{\alpha}_{A[i]} + \alpha_{A[i],T[i]} + \bar{\beta}_{B[i]} + \beta_{B[i],T[i]}$$

$$\alpha = \left(\begin{bmatrix} S_{A,1} & 0 & 0 & 0 \\ 0 & S_{A,2} & 0 & 0 \\ 0 & 0 & S_{A,3} & 0 \\ 0 & 0 & 0 & S_{A,4} \end{bmatrix} \mathbf{L}_A \mathbf{Z}_{T,A} \right)^\intercal$$

$$\alpha = \left(\text{diag}(\mathbf{S}_A)\mathbf{L}_A \mathbf{Z}_{T,A}\right)^\intercal$$

Cholesky factor of correlation matrix across treatments

I. — NOTICES SCIENTIFIQUES

Commandant BENOIT[1].

NOTE SUR UNE MÉTHODE DE RÉSOLUTION DES ÉQUATIONS NORMALES PROVENANT DE L'APPLICATION DE LA MÉTHODE DES MOINDRES CARRÉS A UN SYSTÈME D'ÉQUATIONS LINÉAIRES EN NOMBRE INFÉRIEUR A CELUI DES INCONNUES. — APPLICATION DE LA MÉTHODE A LA RÉSOLUTION D'UN SYSTÈME DEFINI D'ÉQUATIONS LINÉAIRES.

(Procédé du Commandant CHOLESKY[2].)

Le Commandant d'Artillerie Cholesky, du Service géographique de l'Armée, tué pendant la grande guerre, a imaginé, au cours de recherches sur la compensation des réseaux géodésiques, un procédé très ingénieux de résolution des équations dites *normales*, obtenues par application de la méthode des moindres carrés à des équations linéaires en nombre inférieur à celui des inconnues. Il en a conclu une méthode générale de résolution des équations linéaires.

Nous suivrons, pour la démonstration de cette méthode, la progression même qui a servi au Commandant Cholesky pour l'imaginer.

André-Louis Cholesky (1875–1918)

I. — NOTICES SCIENTIFIQUES

Commandant BENOIT[1].

NOTE SUR UNE MÉTHODE DE RÉSOLUTION DES ÉQUATIONS NORMALES PROVENANT DE L'APPLICATION DE LA MÉTHODE DES MOINDRES CARRÉS A UN SYSTÈME D'ÉQUATIONS LINÉAIRES EN NOMBRE INFÉRIEUR A CELUI DES INCONNUES. — APPLICATION DE LA MÉTHODE A LA RÉSOLUTION D'UN SYSTÈME DEFINI D'ÉQUATIONS LINÉAIRES.

(Procédé du Commandant CHOLESKY[2].)

Le Commandant d'Artillerie Cholesky, du Service géographique de l'Armée, tué pendant la grande guerre, a imaginé, au cours de recherches sur la compensation des réseaux géodésiques, un procédé très ingénieux de résolution des équations dites *normales*, obtenues par application de la méthode des moindres carrés à des équations linéaires en nombre inférieur à celui des inconnues. Il en a conclu une méthode générale de résolution des équations linéaires.

André-Louis Cholesky (1875–1918)

I. — NOTICES SCIENTIFIQUES

Commandant BENOIT[1].

NOTE SUR UNE MÉTHODE DE RÉSOLUTION DES ÉQUATIONS NORMALES PROVENANT DE L'APPLICATION DE LA MÉTHODE DES MOINDRES CARRÉS A UN SYSTÈME D'ÉQUATIONS LINÉAIRES EN NOMBRE INFÉRIEUR A CELUI DES INCONNUES. — APPLICATION DE LA MÉTHODE A LA RÉSOLUTION D'UN SYSTÈME DEFINI D'ÉQUATIONS LINÉAIRES.

(Procédé du Commandant Cholesky[2].)

The artillery commander Cholesky, of the Geographical Service of the army, killed during the Great War, imagined, during research on the compensation of the geodesic networks, a very ingenious process of solving the equations known as normal, obtained by application of the method of least squares to linear equations in lower number than that of the unknowns.

résolution des équations linéaires.

André-Louis Cholesky (1875–1918)

ques, un procédé très ingénieux de résolution des équations dites *normales*, obtenues par application de la méthode des

```
# define 2D Gaussian with correlation 0.6
N <- 1e4
sigma1 <- 2
sigma2 <- 0.5
rho <- 0.6

# independent z-scores
z1 <- rnorm( N )
z2 <- rnorm( N )

# use Cholesky to blend in correlation
a1 <- z1 * sigma1
a2 <- ( rho*z1 + sqrt( 1-rho^2 )*z2 )*sigma2
```

ques, un procédé très ingénieux de résolution des équations dites *normales*, obtenues par application de la méthode des

```
# define 2D Gaussian with correlation 0.6
N <- 1e4
sigma1 <- 2
sigma2 <- 0.5
rho <- 0.6

# independent z-scores
z1 <- rnorm( N )
z2 <- rnorm( N )

# use Cholesky to blend in correlation
a1 <- z1 * sigma1
a2 <- ( rho*z1 + sqrt( 1-rho^2 )*z2 )*sigma2
```

ques, un procédé très ingénieux de résolution des équations dites *normales*, obtenues par application de la méthode des

```r
# define 2D Gaussian with correlation 0.6
N <- 1e4
sigma1 <- 2
sigma2 <- 0.5
rho <- 0.6

# independent z-scores
z1 <- rnorm( N )
z2 <- rnorm( N )

# use Cholesky to blend in correlation
a1 <- z1 * sigma1
a2 <- ( rho*z1 + sqrt( 1-rho^2 )*z2 )*sigma2
```

```
> cor(z1,z2)
[1] -0.0005542644
> cor(a1,a2)
[1] 0.5999334
> sd(a1)
[1] 1.997036
> sd(a2)
[1] 0.4989456
```

$$\alpha = \left(\operatorname{diag}(\mathbf{S}_A)\mathbf{L}_A\mathbf{Z}_{T,A}\right)^{\top}$$

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

$$\text{logit}(p_i) = \bar{\alpha}_{A[i]} + \alpha_{A[i],T[i]} + \bar{\beta}_{B[i]} + \beta_{B[i],T[i]}$$ *mean + Actor-treatment + Block-treatment*

$$\alpha = \left(\text{diag}(\mathbf{S}_A)\mathbf{L}_A\mathbf{Z}_{T,A}\right)^\intercal$$ *compute alpha from non-centered pieces*

$$\beta = \left(\text{diag}(\mathbf{S}_B)\mathbf{L}_B\mathbf{Z}_{T,B}\right)^\intercal$$ *compute beta from non-centered pieces*

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

$$\text{logit}(p_i) = \bar{\alpha}_{A[i]} + \alpha_{A[i],T[i]} + \bar{\beta}_{B[i]} + \beta_{B[i],T[i]}$$ *mean + Actor-treatment + Block-treatment*

$$\alpha = \left(\text{diag}(\mathbf{S}_A)\mathbf{L}_A\mathbf{Z}_{T,A}\right)^\intercal$$ *compute alpha from non-centered pieces*

$$\beta = \left(\text{diag}(\mathbf{S}_B)\mathbf{L}_B\mathbf{Z}_{T,B}\right)^\intercal$$ *compute beta from non-centered pieces*

$$\mathbf{Z}_{T,A} \sim \text{Normal}(0,1)$$ *matrix of treatment-actor z-scores*

$$\mathbf{Z}_{T,B} \sim \text{Normal}(0,1)$$ *matrix of treatment-block z-scores*

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

$$\text{logit}(p_i) = \bar{\alpha}_{A[i]} + \alpha_{A[i],T[i]} + \bar{\beta}_{B[i]} + \beta_{B[i],T[i]}$$ *mean + Actor-treatment + Block-treatment*

$$\alpha = \left(\text{diag}(\mathbf{S}_A)\mathbf{L}_A\mathbf{Z}_{T,A}\right)^{\intercal}$$ *compute alpha from non-centered pieces*

$$\beta = \left(\text{diag}(\mathbf{S}_B)\mathbf{L}_B\mathbf{Z}_{T,B}\right)^{\intercal}$$ *compute beta from non-centered pieces*

$$\mathbf{Z}_{T,A} \sim \text{Normal}(0,1)$$ *matrix of treatment-actor z-scores*

$$\mathbf{Z}_{T,B} \sim \text{Normal}(0,1)$$ *matrix of treatment-block z-scores*

$$z_{\bar{A},j}, z_{\bar{B},k} \sim \text{Normal}(0,1)$$ *mean actor and block z-scores*

$$\bar{\alpha} = z_{\bar{A}}\tau_A, \bar{\beta} = z_{\bar{B}}\tau_B$$ *compute mean effects*

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

$$\text{logit}(p_i) = \bar{\alpha}_{A[i]} + \alpha_{A[i],T[i]} + \bar{\beta}_{B[i]} + \beta_{B[i],T[i]}$$ *mean + Actor-treatment + Block-treatment*

$$\alpha = \left(\text{diag}(\mathbf{S}_A)\mathbf{L}_A\mathbf{Z}_{T,A}\right)^\intercal$$ *compute alpha from non-centered pieces*

$$\beta = \left(\text{diag}(\mathbf{S}_B)\mathbf{L}_B\mathbf{Z}_{T,B}\right)^\intercal$$ *compute beta from non-centered pieces*

$$\mathbf{Z}_{T,A} \sim \text{Normal}(0,1)$$ *matrix of treatment-actor z-scores*

$$\mathbf{Z}_{T,B} \sim \text{Normal}(0,1)$$ *matrix of treatment-block z-scores*

$$z_{\bar{A},j}, z_{\bar{B},k} \sim \text{Normal}(0,1)$$ *mean actor and block z-scores*

$$\bar{\alpha} = z_{\bar{A}}\tau_A, \bar{\beta} = z_{\bar{B}}\tau_B$$ *compute mean effects*

$$\text{S}_{A,j}, \text{S}_{B,j}, \tau_A, \tau_B \sim \text{Exponential}(1)$$ *each standard deviation gets same prior*

$$\mathbf{R}_A, \mathbf{R}_B \sim \text{LKJcorr}(4)$$ *correlation matrix prior*

```
m14.3 <- ulam(
alist(
    P ~ bernoulli(p),
    logit(p) <- abar[A] + a[A,T] + bbar[B] + b[B,T],

    # adaptive priors - non-centered
    transpars> matrix[A,4]:a <-
            compose_noncentered( sigma_A , L_Rho_A , zA ),
    transpars> matrix[B,4]:b <-
            compose_noncentered( sigma_B , L_Rho_B , zB ),
    matrix[4,A]:zA ~ normal( 0 , 1 ),
    matrix[4,B]:zB ~ normal( 0 , 1 ),
    zAbar[A] ~ normal(0,1),
    zBbar[B] ~ normal(0,1),
    transpars> vector[A]:abar <<- zAbar*tau_A,
    transpars> vector[B]:bbar <<- zBbar*tau_B,

    # fixed priors
    c(tau_A,tau_B) ~ exponential(1),
    vector[4]:sigma_A ~ exponential(1),
    cholesky_factor_corr[4]:L_Rho_A ~ lkj_corr_cholesky(4),
    vector[4]:sigma_B ~ exponential(1),
    cholesky_factor_corr[4]:L_Rho_B ~ lkj_corr_cholesky(4),

    # compute ordinary correlation matrixes
    gq> matrix[4,4]:Rho_A <<- Chol_to_Corr(L_Rho_A),
    gq> matrix[4,4]:Rho_B <<- Chol_to_Corr(L_Rho_B)
) , data=dat , chains=4 , cores=4 , log_lik=TRUE )
```

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

$$\text{logit}(p_i) = \bar{\alpha}_{A[i]} + \alpha_{A[i],T[i]} + \bar{\beta}_{B[i]} + \beta_{B[i],T[i]}$$

$$\alpha = \left(\text{diag}(\mathbf{S}_A)\mathbf{L}_A\mathbf{Z}_{T,A}\right)^\intercal$$

$$\beta = \left(\text{diag}(\mathbf{S}_B)\mathbf{L}_B\mathbf{Z}_{T,B}\right)^\intercal$$

$$\mathbf{Z}_{T,A} \sim \text{Normal}(0,1)$$

$$\mathbf{Z}_{T,B} \sim \text{Normal}(0,1)$$

$$z_{\bar{A},j}, z_{\bar{B},k} \sim \text{Normal}(0,1)$$

$$\bar{\alpha} = z_{\bar{A}}\tau_A, \bar{\beta} = z_{\bar{B}}\tau_B$$

$$\text{S}_{A,j}, \text{S}_{B,j}, \tau_A, \tau_B \sim \text{Exponential}(1)$$

$$\mathbf{R}_A, \mathbf{R}_B \sim \text{LKJcorr}(4)$$

```r
m14.3 <- ulam(
alist(
    P ~ bernoulli(p),
    logit(p) <- abar[A] + a[A,T] + bbar[B] + b[B,T],

    # adaptive priors - non-centered
    transpars> matrix[A,4]:a <-
            compose_noncentered( sigma_A , L_Rho_A , zA ),
    transpars> matrix[B,4]:b <-
            compose_noncentered( sigma_B , L_Rho_B , zB ),
    matrix[4,A]:zA ~ normal( 0 , 1 ),
    matrix[4,B]:zB ~ normal( 0 , 1 ),
    zAbar[A] ~ normal(0,1),
    zBbar[B] ~ normal(0,1),
    transpars> vector[A]:abar <<- zAbar*tau_A,
    transpars> vector[B]:bbar <<- zBbar*tau_B,

    # fixed priors
    c(tau_A,tau_B) ~ exponential(1),
    vector[4]:sigma_A ~ exponential(1),
    cholesky_factor_corr[4]:L_Rho_A ~ lkj_corr_cholesky(4),
    vector[4]:sigma_B ~ exponential(1),
    cholesky_factor_corr[4]:L_Rho_B ~ lkj_corr_cholesky(4),

    # compute ordinary correlation matrixes
    gq> matrix[4,4]:Rho_A <<- Chol_to_Corr(L_Rho_A),
    gq> matrix[4,4]:Rho_B <<- Chol_to_Corr(L_Rho_B)
) , data=dat , chains=4 , cores=4 , log_lik=TRUE )
```

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

$$\text{logit}(p_i) = \bar{\alpha}_{A[i]} + \alpha_{A[i],T[i]} + \bar{\beta}_{B[i]} + \beta_{B[i],T[i]}$$

$$\boxed{\begin{aligned} \alpha &= \left(\text{diag}(\mathbf{S}_A)\mathbf{L}_A\mathbf{Z}_{T,A}\right)^{\intercal} \\ \beta &= \left(\text{diag}(\mathbf{S}_B)\mathbf{L}_B\mathbf{Z}_{T,B}\right)^{\intercal} \end{aligned}}$$

$$\mathbf{Z}_{T,A} \sim \text{Normal}(0,1)$$

$$\mathbf{Z}_{T,B} \sim \text{Normal}(0,1)$$

$$z_{\bar{A},j}, z_{\bar{B},k} \sim \text{Normal}(0,1)$$

$$\bar{\alpha} = z_{\bar{A}}\tau_A, \bar{\beta} = z_{\bar{B}}\tau_B$$

$$\text{S}_{A,j}, \text{S}_{B,j}, \tau_A, \tau_B \sim \text{Exponential}(1)$$

$$\mathbf{R}_A, \mathbf{R}_B \sim \text{LKJcorr}(4)$$

```
m14.3 <- ulam(
alist(
    P ~ bernoulli(p),
    logit(p) <- abar[A] + a[A,T] + bbar[B] + b[B,T],

    # adaptive priors - non-centered
    transpars> matrix[A,4]:a <-
            compose_noncentered( sigma_A , L_Rho_A , zA ),
    transpars> matrix[B,4]:b <-
            compose_noncentered( sigma_B , L_Rho_B , zB ),
    matrix[4,A]:zA ~ normal( 0 , 1 ),
    matrix[4,B]:zB ~ normal( 0 , 1 ),
    zAbar[A] ~ normal(0,1),
    zBbar[B] ~ normal(0,1),
    transpars> vector[A]:abar <<- zAbar*tau_A,
    transpars> vector[B]:bbar <<- zBbar*tau_B,

    # fixed priors
    c(tau_A,tau_B) ~ exponential(1),
    vector[4]:sigma_A ~ exponential(1),
    cholesky_factor_corr[4]:L_Rho_A ~ lkj_corr_cholesky(4),
    vector[4]:sigma_B ~ exponential(1),
    cholesky_factor_corr[4]:L_Rho_B ~ lkj_corr_cholesky(4),

    # compute ordinary correlation matrixes
    gq> matrix[4,4]:Rho_A <<- Chol_to_Corr(L_Rho_A),
    gq> matrix[4,4]:Rho_B <<- Chol_to_Corr(L_Rho_B)
) , data=dat , chains=4 , cores=4 , log_lik=TRUE )
```

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

$$\text{logit}(p_i) = \bar{\alpha}_{A[i]} + \alpha_{A[i],T[i]} + \bar{\beta}_{B[i]} + \beta_{B[i],T[i]}$$

$$\alpha = \left(\text{diag}(\mathbf{S}_A)\mathbf{L}_A\mathbf{Z}_{T,A}\right)^\intercal$$

$$\beta = \left(\text{diag}(\mathbf{S}_B)\mathbf{L}_B\mathbf{Z}_{T,B}\right)^\intercal$$

$$\mathbf{Z}_{T,A} \sim \text{Normal}(0,1)$$

$$\mathbf{Z}_{T,B} \sim \text{Normal}(0,1)$$

$$z_{\bar{A},j}, z_{\bar{B},k} \sim \text{Normal}(0,1)$$

$$\bar{\alpha} = z_{\bar{A}}\tau_A, \bar{\beta} = z_{\bar{B}}\tau_B$$

$$\text{S}_{A,j}, \text{S}_{B,j}, \tau_A, \tau_B \sim \text{Exponential}(1)$$

$$\mathbf{R}_A, \mathbf{R}_B \sim \text{LKJcorr}(4)$$

```r
m14.3 <- ulam(
alist(
    P ~ bernoulli(p),
    logit(p) <- abar[A] + a[A,T] + bbar[B] + b[B,T],

    # adaptive priors - non-centered
    transpars> matrix[A,4]:a <-
            compose_noncentered( sigma_A , L_Rho_A , zA ),
    transpars> matrix[B,4]:b <-
            compose_noncentered( sigma_B , L_Rho_B , zB ),
    matrix[4,A]:zA ~ normal( 0 , 1 ),
    matrix[4,B]:zB ~ normal( 0 , 1 ),
    zAbar[A] ~ normal(0,1),
    zBbar[B] ~ normal(0,1),
    transpars> vector[A]:abar <<- zAbar*tau_A,
    transpars> vector[B]:bbar <<- zBbar*tau_B,

    # fixed priors
    c(tau_A,tau_B) ~ exponential(1),
    vector[4]:sigma_A ~ exponential(1),
    cholesky_factor_corr[4]:L_Rho_A ~ lkj_corr_cholesky(4),
    vector[4]:sigma_B ~ exponential(1),
    cholesky_factor_corr[4]:L_Rho_B ~ lkj_corr_cholesky(4),

    # compute ordinary correlation matrixes
    gq> matrix[4,4]:Rho_A <<- Chol_to_Corr(L_Rho_A),
    gq> matrix[4,4]:Rho_B <<- Chol_to_Corr(L_Rho_B)
) , data=dat , chains=4 , cores=4 , log_lik=TRUE )
```

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

$$\text{logit}(p_i) = \bar{\alpha}_{A[i]} + \alpha_{A[i],T[i]} + \bar{\beta}_{B[i]} + \beta_{B[i],T[i]}$$

$$\alpha = \left(\text{diag}(\mathbf{S}_A)\mathbf{L}_A\mathbf{Z}_{T,A}\right)^\intercal$$

$$\beta = \left(\text{diag}(\mathbf{S}_B)\mathbf{L}_B\mathbf{Z}_{T,B}\right)^\intercal$$

$$\mathbf{Z}_{T,A} \sim \text{Normal}(0,1)$$

$$\mathbf{Z}_{T,B} \sim \text{Normal}(0,1)$$

$$z_{\bar{A},j}, z_{\bar{B},k} \sim \text{Normal}(0,1)$$

$$\bar{\alpha} = z_{\bar{A}}\tau_A, \bar{\beta} = z_{\bar{B}}\tau_B$$

$$S_{A,j}, S_{B,j}, \tau_A, \tau_B \sim \text{Exponential}(1)$$

$$\mathbf{R}_A, \mathbf{R}_B \sim \text{LKJcorr}(4)$$

```
m14.3 <- ulam(
alist(
    P ~ bernoulli(p),
    logit(p) <- abar[A] + a[A,T] + bbar[B] + b[B,T],

    # adaptive priors - non-centered
    transpars> matrix[A,4]:a <-
            compose_noncentered( sigma_A , L_Rho_A , zA ),
    transpars> matrix[B,4]:b <-
            compose_noncentered( sigma_B , L_Rho_B , zB ),
    matrix[4,A]:zA ~ normal( 0 , 1 ),
    matrix[4,B]:zB ~ normal( 0 , 1 ),
    zAbar[A] ~ normal(0,1),
    zBbar[B] ~ normal(0,1),
    transpars> vector[A]:abar <<- zAbar*tau_A,
    transpars> vector[B]:bbar <<- zBbar*tau_B,

    # fixed priors
    c(tau_A,tau_B) ~ exponential(1),
    vector[4]:sigma_A ~ exponential(1),
    cholesky_factor_corr[4]:L_Rho_A ~ lkj_corr_cholesky(4),
    vector[4]:sigma_B ~ exponential(1),
    cholesky_factor_corr[4]:L_Rho_B ~ lkj_corr_cholesky(4),

    # compute ordinary correlation matrixes
    gq> matrix[4,4]:Rho_A <<- Chol_to_Corr(L_Rho_A),
    gq> matrix[4,4]:Rho_B <<- Chol_to_Corr(L_Rho_B)
) , data=dat , chains=4 , cores=4 , log_lik=TRUE )
```

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

$$\text{logit}(p_i) = \bar{\alpha}_{A[i]} + \alpha_{A[i],T[i]} + \bar{\beta}_{B[i]} + \beta_{B[i],T[}$$

$$\alpha = \left(\text{diag}(\mathbf{S}_A)\mathbf{L}_A\mathbf{Z}_{T,A}\right)^\intercal$$

$$\beta = \left(\text{diag}(\mathbf{S}_B)\mathbf{L}_B\mathbf{Z}_{T,B}\right)^\intercal$$

$$\mathbf{Z}_{T,A} \sim \text{Normal}(0,1)$$

$$\mathbf{Z}_{T,B} \sim \text{Normal}(0,1)$$

$$z_{\bar{A},j}, z_{\bar{B},k} \sim \text{Normal}(0,1)$$

$$\bar{\alpha} = z_{\bar{A}}\tau_A, \bar{\beta} = z_{\bar{B}}\tau_B$$

$$\mathrm{S}_{A,j}, \mathrm{S}_{B,j}, \tau_A, \tau_B \sim \text{Exponential}(1)$$

$$\mathbf{R}_A, \mathbf{R}_B \sim \text{LKJcorr}(4)$$

```
m14.3 <- ulam(
alist(
    P ~ bernoulli(p),
    logit(p) <- abar[A] + a[A,T] + bbar[B] + b[B,T],

    # adaptive priors - non-centered
    transpars> matrix[A,4]:a <-
            compose_noncentered( sigma_A , L_Rho_A , zA ),
    transpars> matrix[B,4]:b <-
            compose_noncentered( sigma_B , L_Rho_B , zB ),
    matrix[4,A]:zA ~ normal( 0 , 1 ),
    matrix[4,B]:zB ~ normal( 0 , 1 ),
    zAbar[A] ~ normal(0,1),
    zBbar[B] ~ normal(0,1),
    transpars> vector[A]:abar <<- zAbar*tau_A,
    transpars> vector[B]:bbar <<- zBbar*tau_B,

    # fixed priors
    c(tau_A,tau_B) ~ exponential(1),
    vector[4]:sigma_A ~ exponential(1),
    cholesky_factor_corr[4]:L_Rho_A ~ lkj_corr_cholesky(4),
    vector[4]:sigma_B ~ exponential(1),
    cholesky_factor_corr[4]:L_Rho_B ~ lkj_corr_cholesky(4),

    # compute ordinary correlation matrixes
    gq> matrix[4,4]:Rho_A <<- Chol_to_Corr(L_Rho_A),
    gq> matrix[4,4]:Rho_B <<- Chol_to_Corr(L_Rho_B)
) , data=dat , chains=4 , cores=4 , log_lik=TRUE )
```

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

$$\text{logit}(p_i) = \bar{\alpha}_{A[i]} + \alpha_{A[i],T[i]} + \bar{\beta}_{B[i]} + \beta_{B[i],T[i]}$$

$$\alpha = \left(\text{diag}(\mathbf{S}_A)\mathbf{L}_A\mathbf{Z}_{T,A}\right)^\intercal$$

$$\beta = \left(\text{diag}(\mathbf{S}_B)\mathbf{L}_B\mathbf{Z}_{T,B}\right)^\intercal$$

$$\mathbf{Z}_{T,A} \sim \text{Normal}(0,1)$$

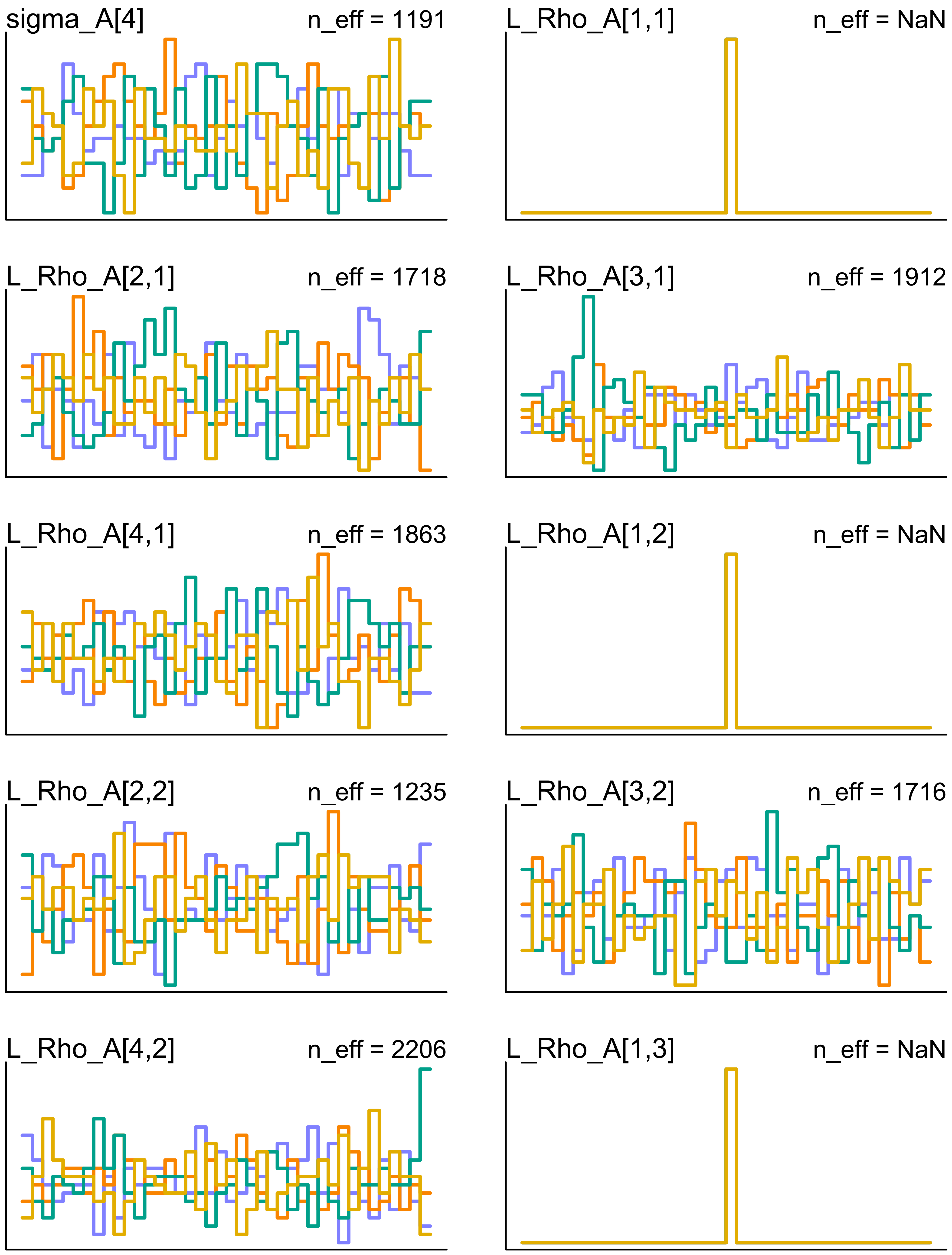
$$\mathbf{Z}_{T,B} \sim \text{Normal}(0,1)$$

$$z_{\bar{A},j}, z_{\bar{B},k} \sim \text{Normal}(0,1)$$

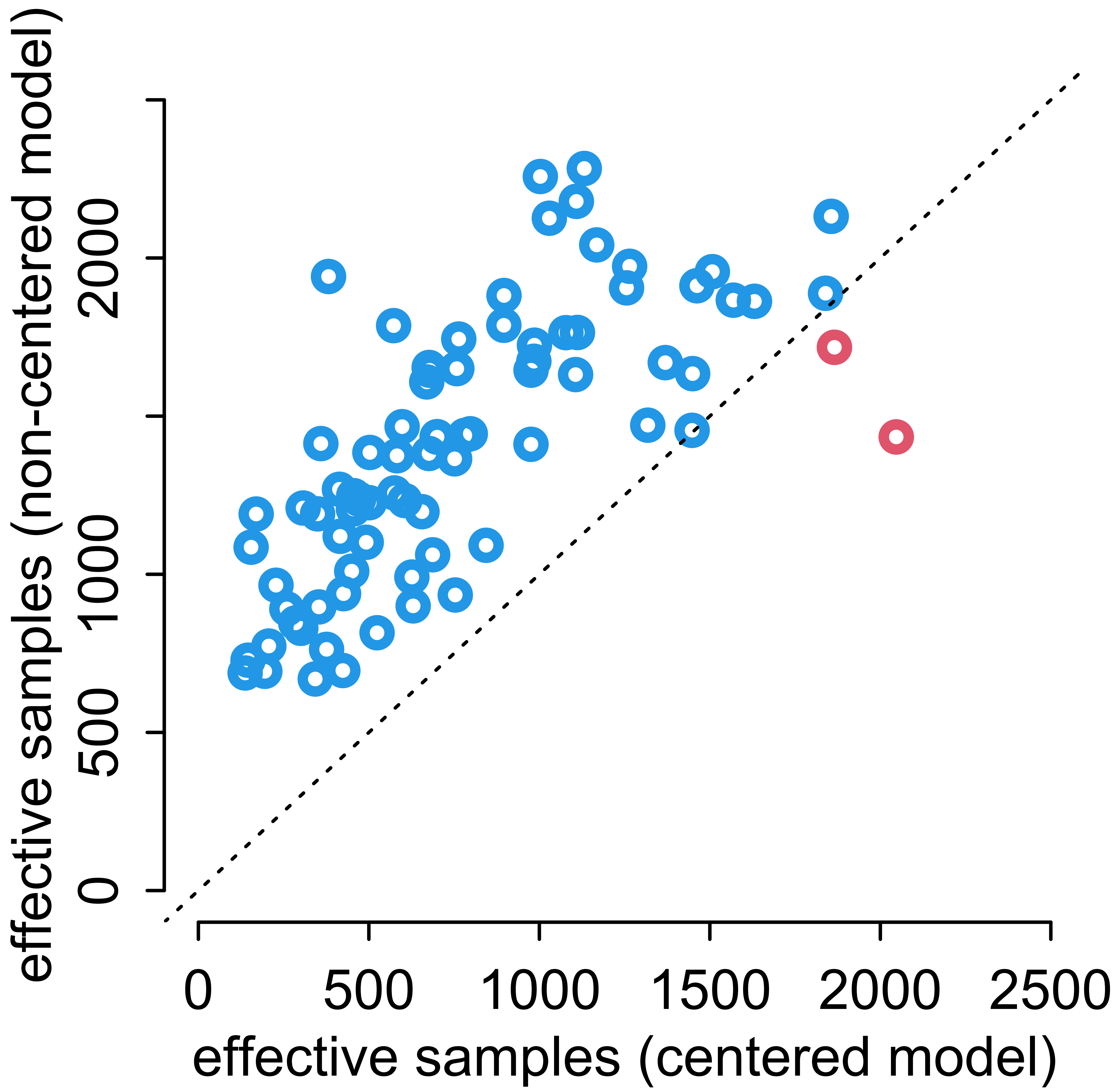
$$\bar{\alpha} = z_{\bar{A}}\tau_A, \bar{\beta} = z_{\bar{B}}\tau_B$$

$$S_{A,j}, S_{B,j}, \tau_A, \tau_B \sim \text{Exponential}(1)$$

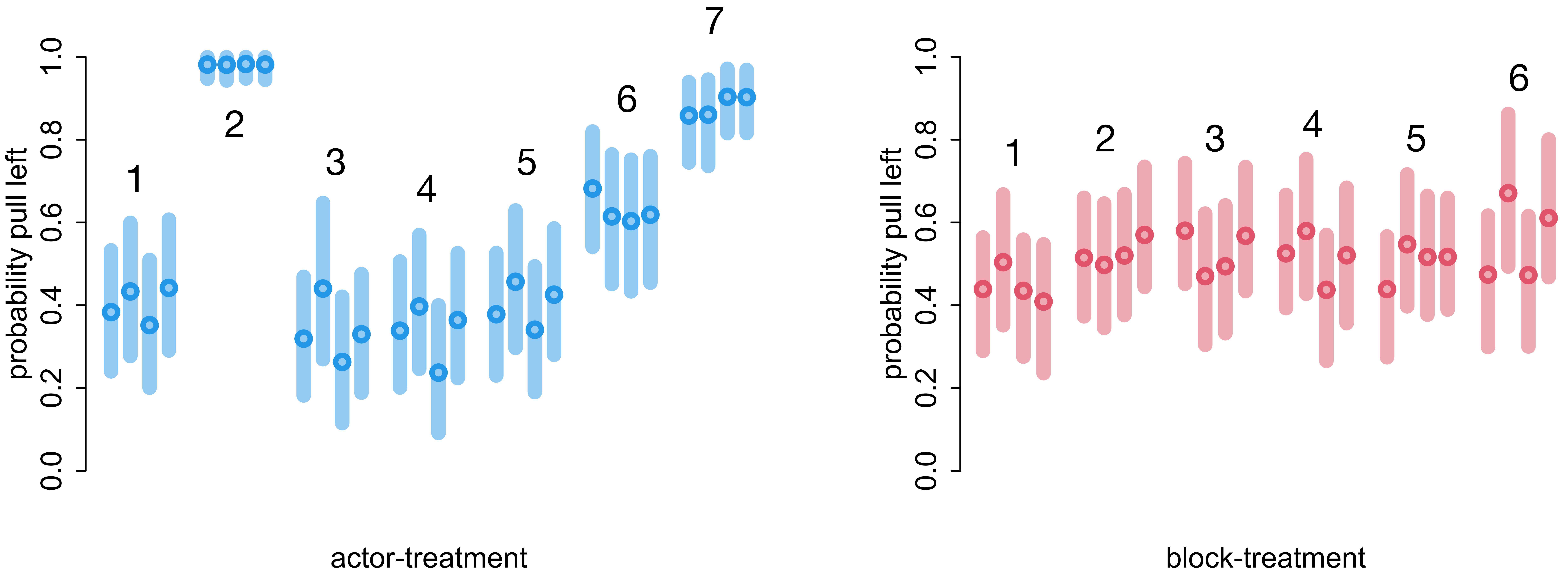
$$\mathbf{R}_A, \mathbf{R}_B \sim \text{LKJcorr}(4)$$



sigma_A[4]
n_eff = 1191
L_Rho_A[1,1]
n_eff = NaN
L_Rho_A[2,1]
n_eff = 1718
L_Rho_A[3,1]
n_eff = 1912
L_Rho_A[4,1]
n_eff = 1863
L_Rho_A[1,2]
n_eff = NaN
L_Rho_A[2,2]
n_eff = 1235
L_Rho_A[3,2]
n_eff = 1716
L_Rho_A[4,2]
n_eff = 2206
L_Rho_A[1,3]
n_eff = NaN



effective samples (non-centered model)
0
500
1000
2000
0
500
1000
1500
2000
2500
effective samples (centered model)



probability pull left
0.0
0.2
0.4
0.6
0.8
1.0
1
2
3
4
5
6
7
actor-treatment
probability pull left
0.0
0.2
0.4
0.6
0.8
1.0
1
2
3
4
5
6
block-treatment

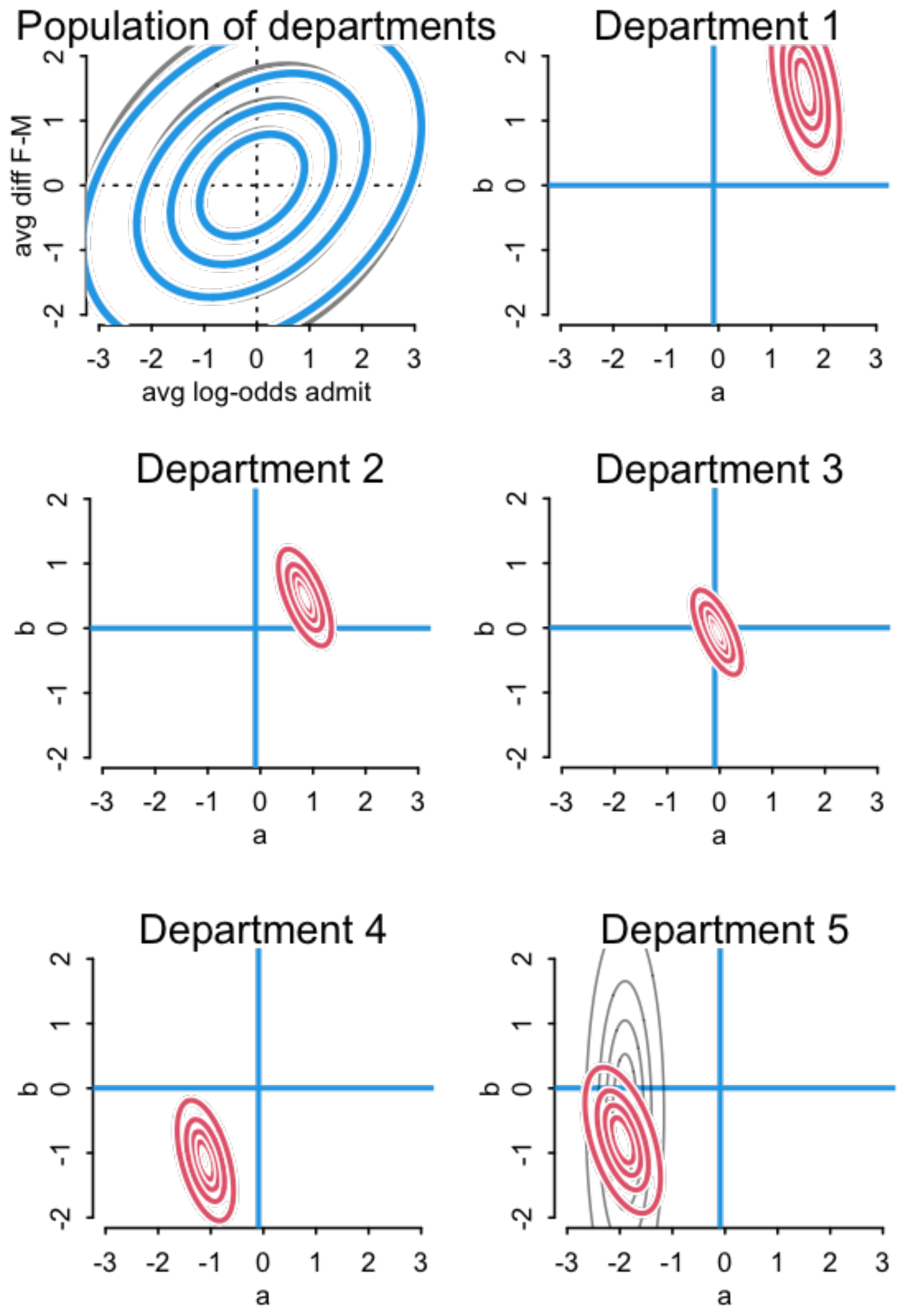

1
2
3
4
5
6
7
probability pull left
0.0
0.2
0.4
0.6
0.8
1.0
1
2
3
4
5
6
probability pull left
0.0
0.2
0.4
0.6
0.8
1.0
by treatment
actor means
Density
0.0
1.0
2.0
0
1
2
3
4
5
standard dev among actors
block means
by treatment
Density
0.0
1.0
2.0
3.0
0
1
2
3
4
5
standard dev among blocks

Correlated Varying Effects

Priors that learn correlation structure:
(1) partial pooling across features
(2) exploit correlations for prediction & causal inference

Varying effects can be correlated even if the prior doesn't learn the correlations!

Ethical obligation to do our best



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 Tactics & Gaussian Processes	Chapter 14
Week 9	Measurement & Missingness	Chapter 15
Week 10	Generalized Linear Madness	Chapter 16

https://github.com/rmcelreath/stat_rethinking_2022