

## Appendix A: R Code and All Visualizations of the Data

### Generalized Additive Models for Boundary and Prominence

```
1 library(ggplot2)
2 library(mgcv)
3 library(visreg)
4 library(parallel)
5
6
7
8 setwd("/data/Dropbox/current papers/Active/Lab Phonology/Modeling 2016 May/")
9
10 nc <- 2 ## cluster size, set for example portability
11 if (detectCores()>1) { ## no point otherwise
12   cl <- makeCluster(nc)
13   ## could also use makeForkCluster, but read warnings first!
14 } else cl <- NULL
15
16
17 boundGAM = bam(rating ~ s(word.phonerate) + pos + s(log.f0.sv.max.norm) + s
18   (int.sv.norm) +
19     s(log.wordfreq.switchboard)+ s(durpostpause) +
20     s(word,bs="re") +
21     s(subject, word.phonerate,bs="fs",m=1)+
22     s(subject, log.f0.sv.max.norm, bs="fs",m=1)+
23     s(subject, int.sv.norm, bs="fs",m=1)+
24     s(subject,log.wordfreq.switchboard,bs="fs",m=1)+
25     s(subject,durpostpause,bs="fs",m=1)+
26     s(subject,pos,bs="re"),
27   family="binomial", data=labuslong, cluster=cl)
28
29
30
31 promGAM = bam(rating ~ s(word.phonerate) + s(durpostpause)+
32   s(log.f0.sv.max.norm) + s(log.wordfreq.switchboard)+ s(int.
33     sv.norm) + pos + boundarym +
34     s(word,bs="re") +
35     s(subject,durpostpause,bs="fs",m=1)+
36     s(subject, pos, bs="re") +
37     s(subject, boundarym,bs="re")+
38     s(subject, word.phonerate,bs="fs",m=1)+
39     s(subject, log.wordfreq.switchboard, bs="fs",m=1)+
40     s(subject, log.f0.sv.max.norm, bs="fs",m=1)+
41     s(subject, int.sv.norm, bs="fs",m=1),
42   family="binomial", data=labuslong2, cluster=cl)
43
44 setwd("/data/Big Gamms/")
45 save(boundGAM,promGAM,labuslong,labuslong2, file="gammsNEW.rda", compress='xz
46   ')
47
48 plotGAMSSubject(gam1=boundGAM,type="Boundary")
49 plotGAMSSubject(gam1=boundGAM,type="Boundary_all_in_one",ovt=TRUE)
50 plotGAMSSubject(gam1=promGAM,type="Prominence")
51 plotGAMSSubject(gam1=promGAM,type="Prominence_all_in_one",ovt=TRUE)
```

```

50 plotGAMSSubject(gam1=boundGAM,type="Boundary",subj=FALSE)
51 plotGAMSSubject(gam1=promGAM,type="Prominence",subj=FALSE)
52
53 plotGAMSSubject(gam1=boundGAM,type="ToBI_Boundary",subj=FALSE, condL = list(
  subject="33"))
54 plotGAMSSubject(gam1=promGAM,type="ToBI_Prominence",subj=FALSE, condL=list(
  subject="33"))
55
56
57
58 summary(boundGAM)
59 summary(promGAM)

```

## Visualizing GAMs

```

1 plotGAMSSubject <- function(gam1, type, pDir= "/data/Dropbox/current papers/
  Active/Lab Phonology/Modeling 2016 May/pics/", ovt=FALSE,subj=TRUE,
  condL=NULL){
2   require(stringr)
3   require(reshape2)
4   require(visreg)
5   formula = as.character(gam1$formula)[3]
6   terms = strsplit(formula, "\\+")
7   if(subj){
8     t2 = grep("subject", terms[[1]],value=TRUE)
9     outm=colsplit(string=t2, pattern=",", names=c("Subject", "term","fit","m"))
10    outm2 = trimws(outm$term)
11    setwd(pDir)
12
13
14    for(i in 1:length(outm2)){
15      title = paste0(outm2[i],type,"_bySubject.png")
16      png(file = title, bg = "white",units="in",width = 8, height = 11, res =
        72)
17
18      plot = visreg(gam1,type="conditional",xvar = outm2[i], by="subject",line=
        list(col="black"),ps = 12, cex = 1, cex.main = 1,
19        main=type,xlab=outm2[i],ylab=paste0("P(",type,")"),cond=
        condL, scale="response",overlay=ovt)
20
21      dev.off()
22    }else{
23      t1 = terms[[1]]
24      t2 = t1[!grepl("subject", t1)]
25      t2 = t2[!grepl("word", t2)]
26      outm2 = trimws(t2)
27      setwd(pDir)
28      outm2 = sub('s\\(', ' ', outm2)
29      outm2 = sub('\\)', ' ', outm2)
30
31      for(i in 1:length(outm2)){
32        title = paste0(outm2[i],type,"_overall.png")
33        png(file = title, bg = "white",units="in",width = 8, height = 11, res =
          72)
34
35        plot = visreg(gam1,type="conditional",xvar = outm2[i],line=list(col="
          black"),cond = condL, ps = 12, cex = 1, cex.main = 1,
36          main=type,xlab=outm2[i],ylab=paste0("P(",type,")"), scale
          ="response",ylim = c(0,1.00))
37
38        dev.off()
39      }
40    }

```

## Visualizations of Data

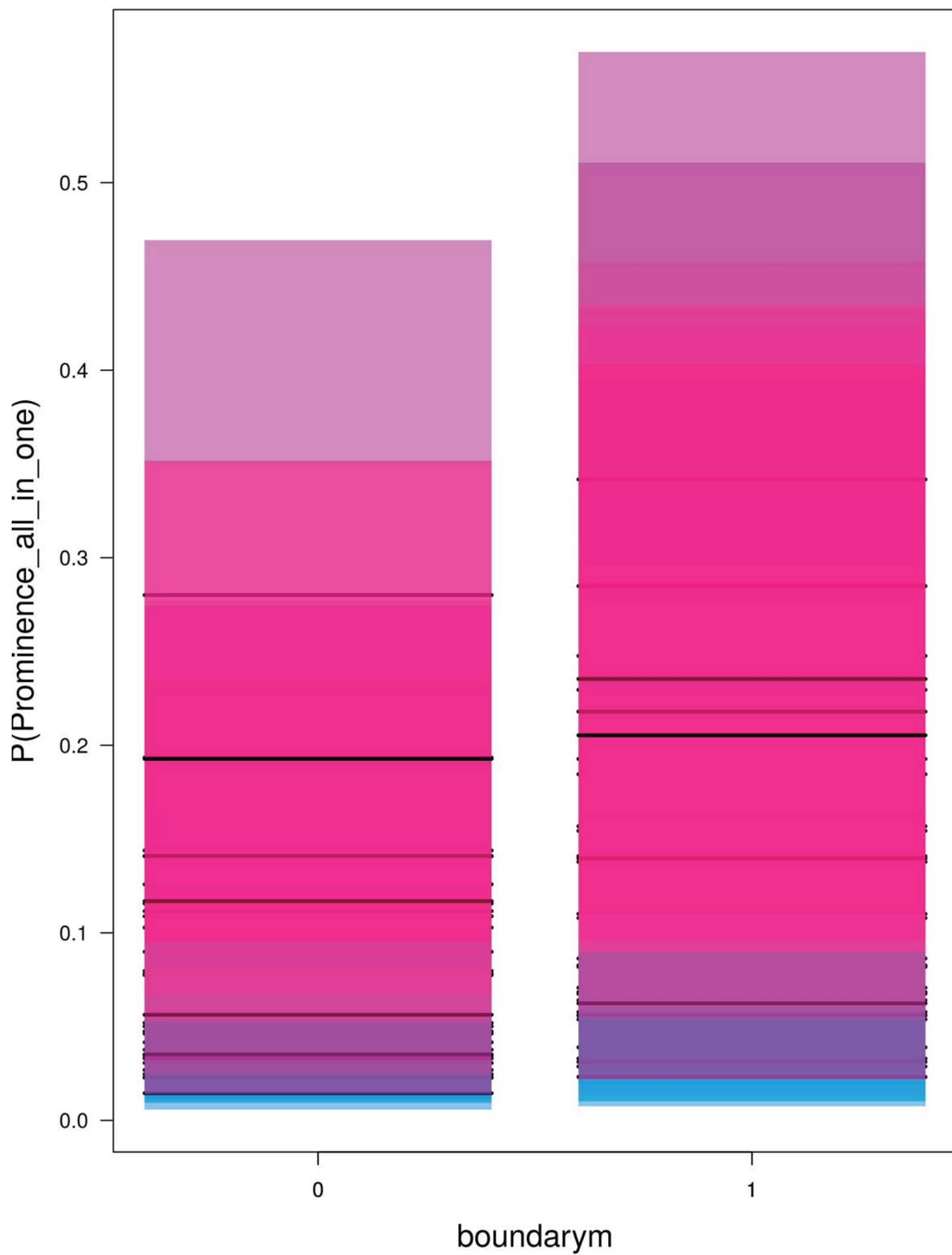
All the visualizations of the data are produced with the `visreg()` package in R: Breheny & Burchett (version 2.4-1,2017).

The full color graphs are all individual effects on one plot. The multi-facted plots are the predicted values with the label representing each annotator number and ToBI labelling marked. The overall effect for every main predictor is plotted on a separate chart.

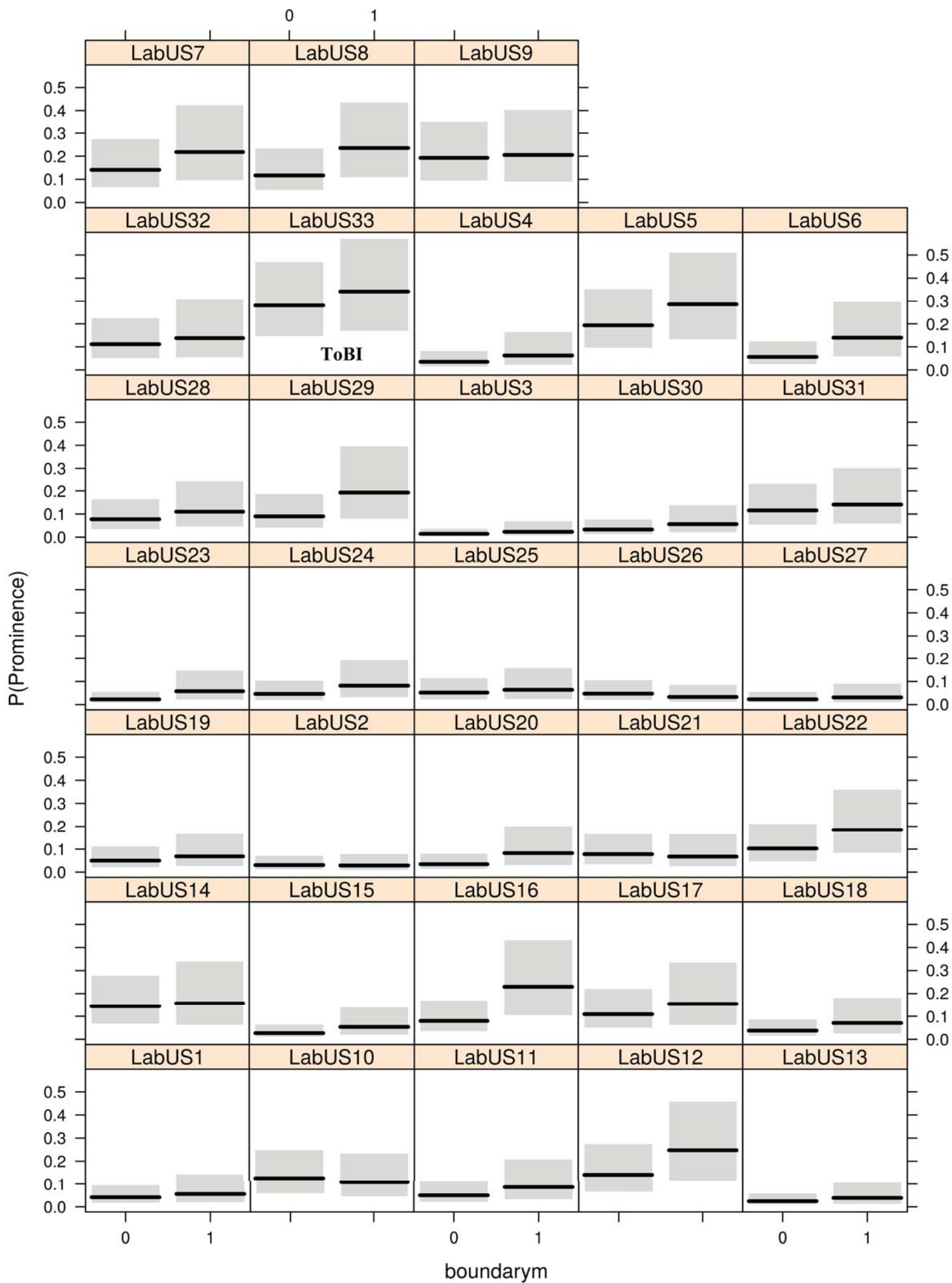
## References

Breheny, P. & Burchett, W. (2017). *visreg: Visualization of Regression Models*. Retrieved from <https://CRAN.R-project.org/package=visreg>

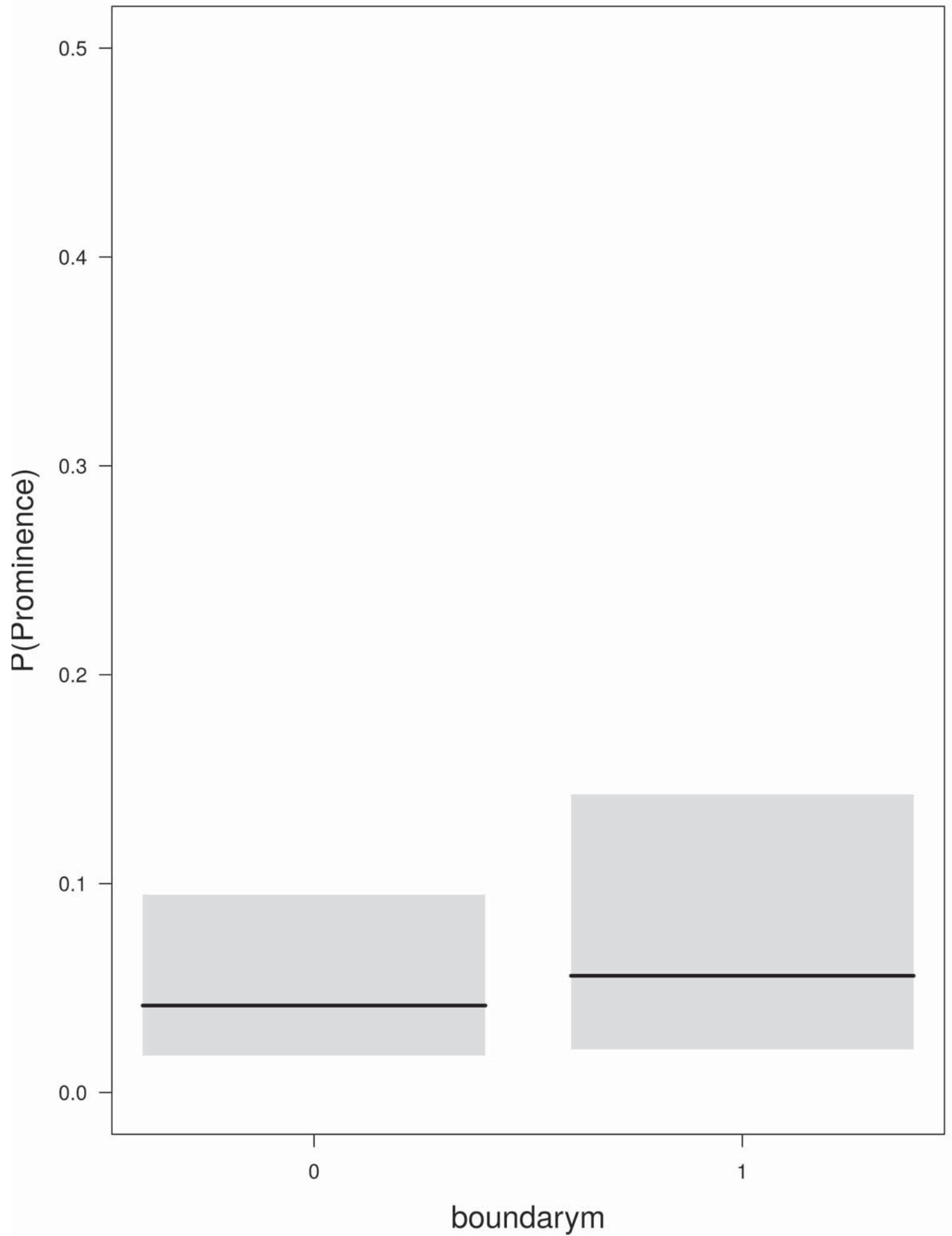
# Prominence\_all\_in\_one



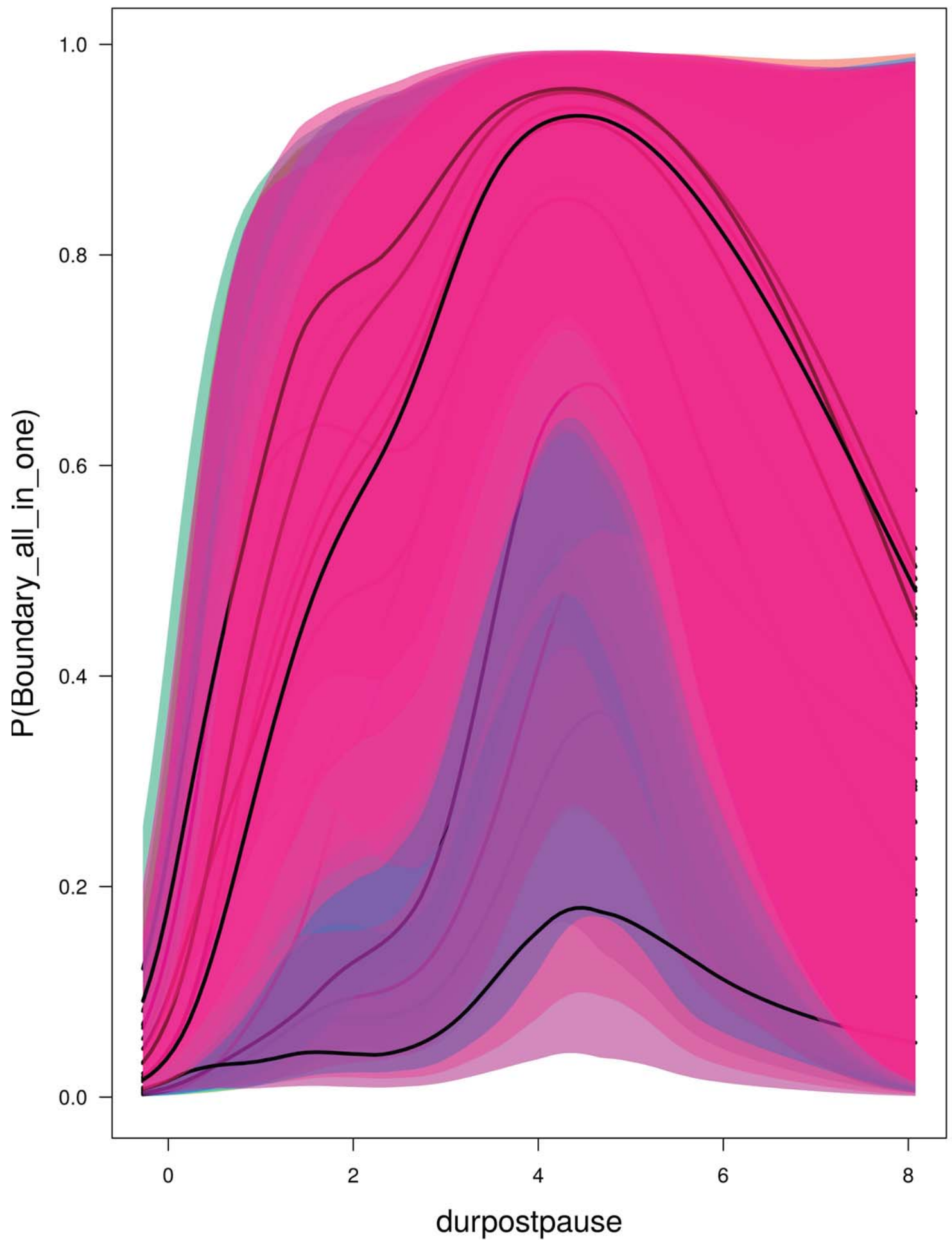
# Prominence



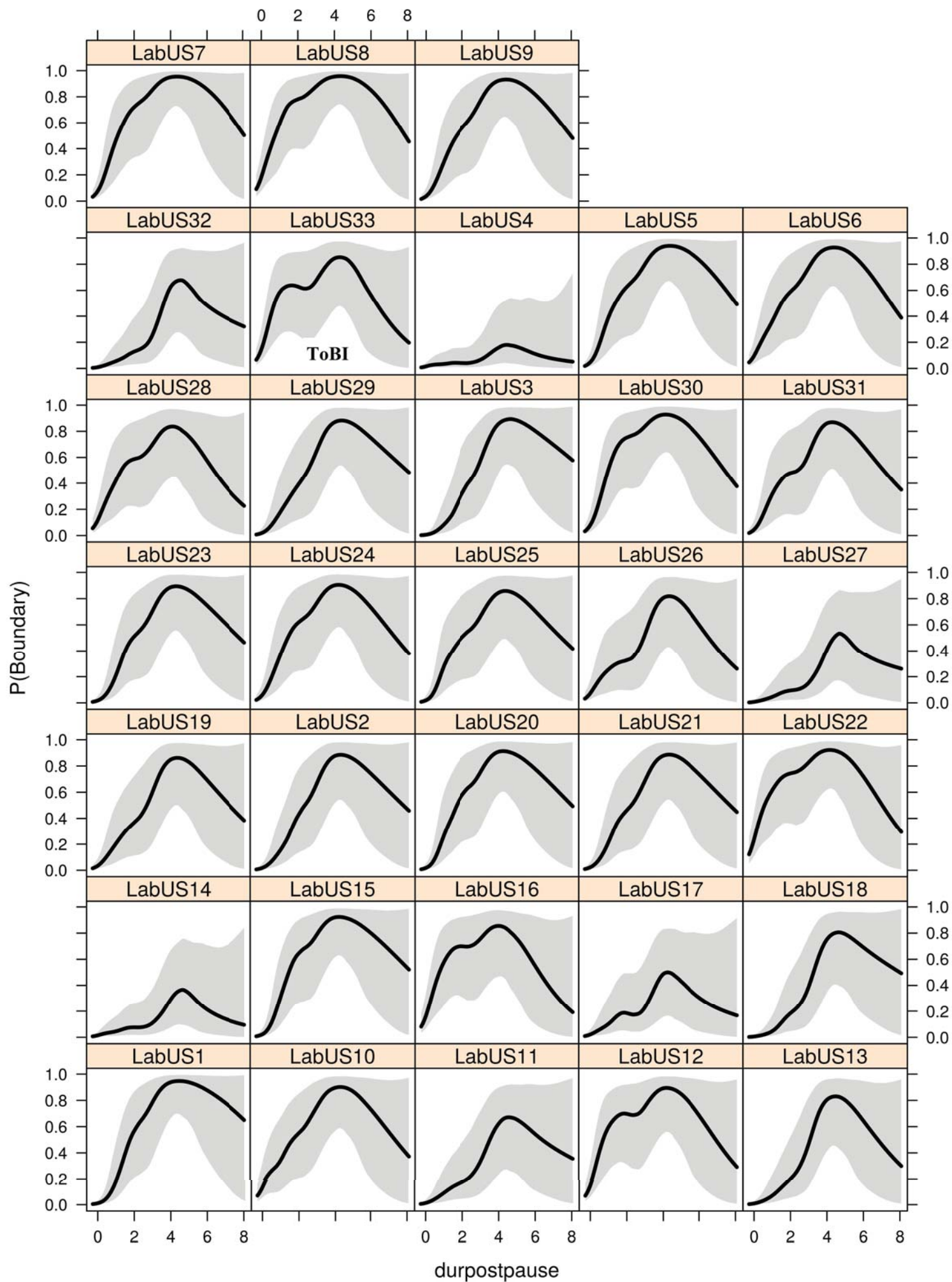
# Prominence



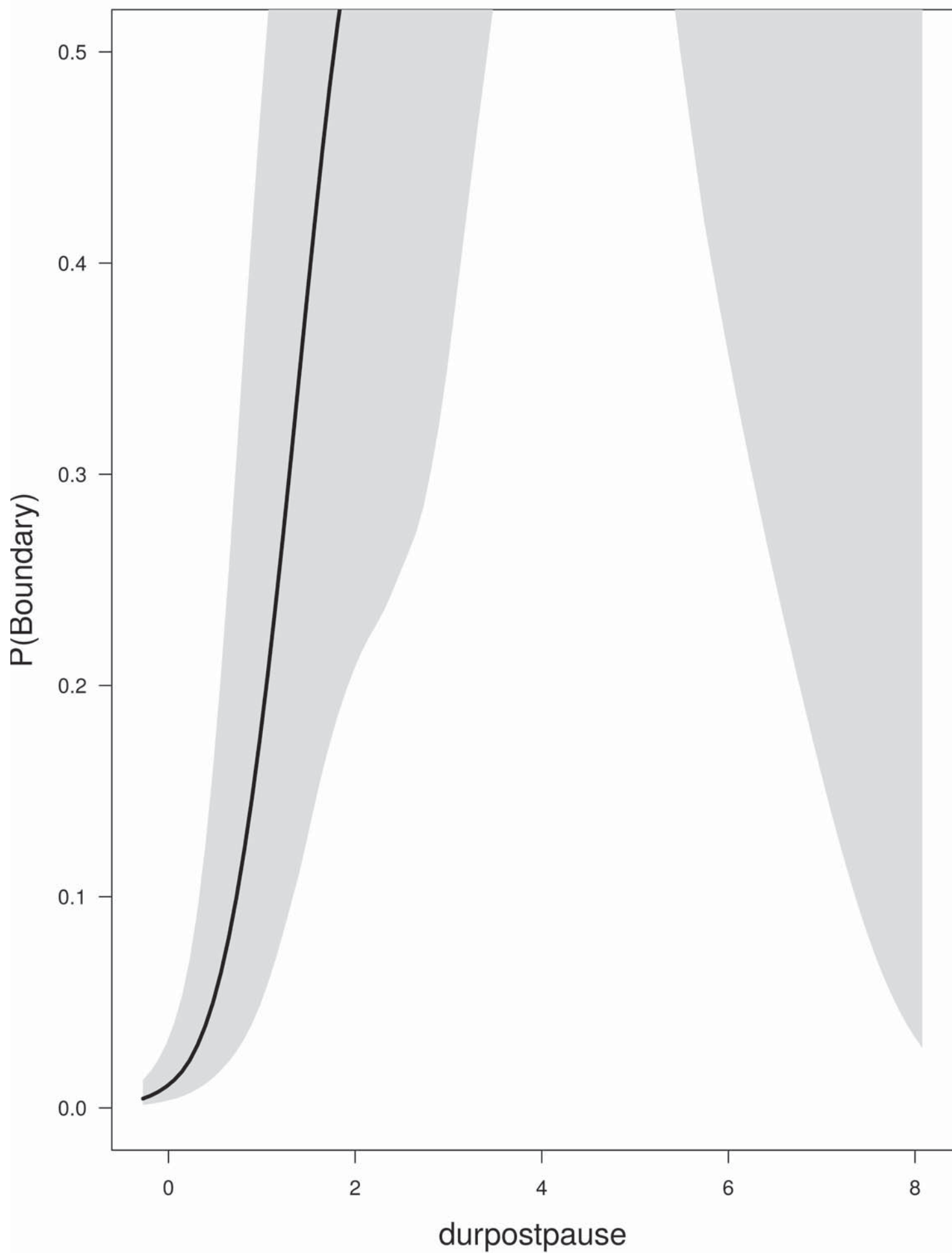
# Boundary\_all\_in\_one



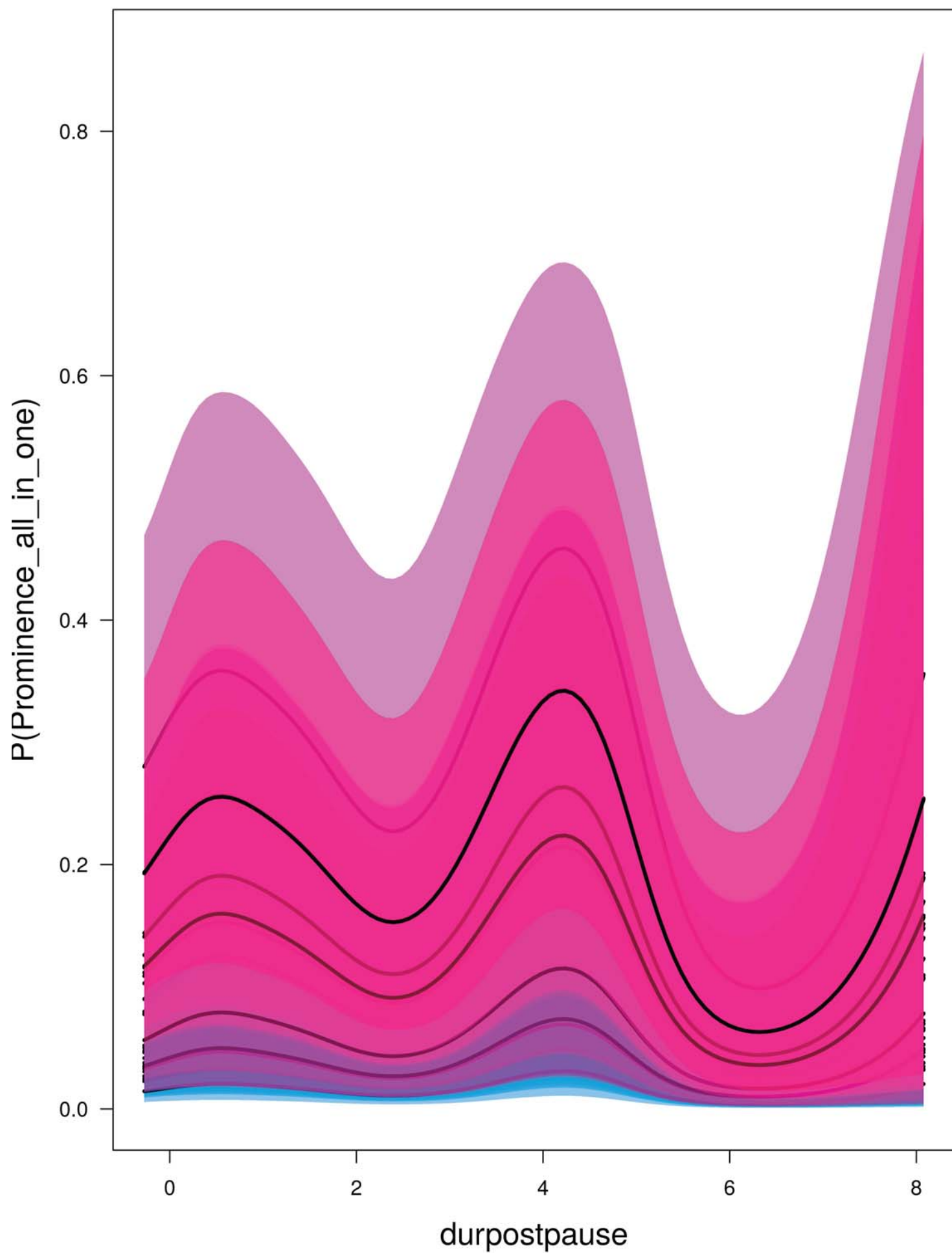
# Boundary



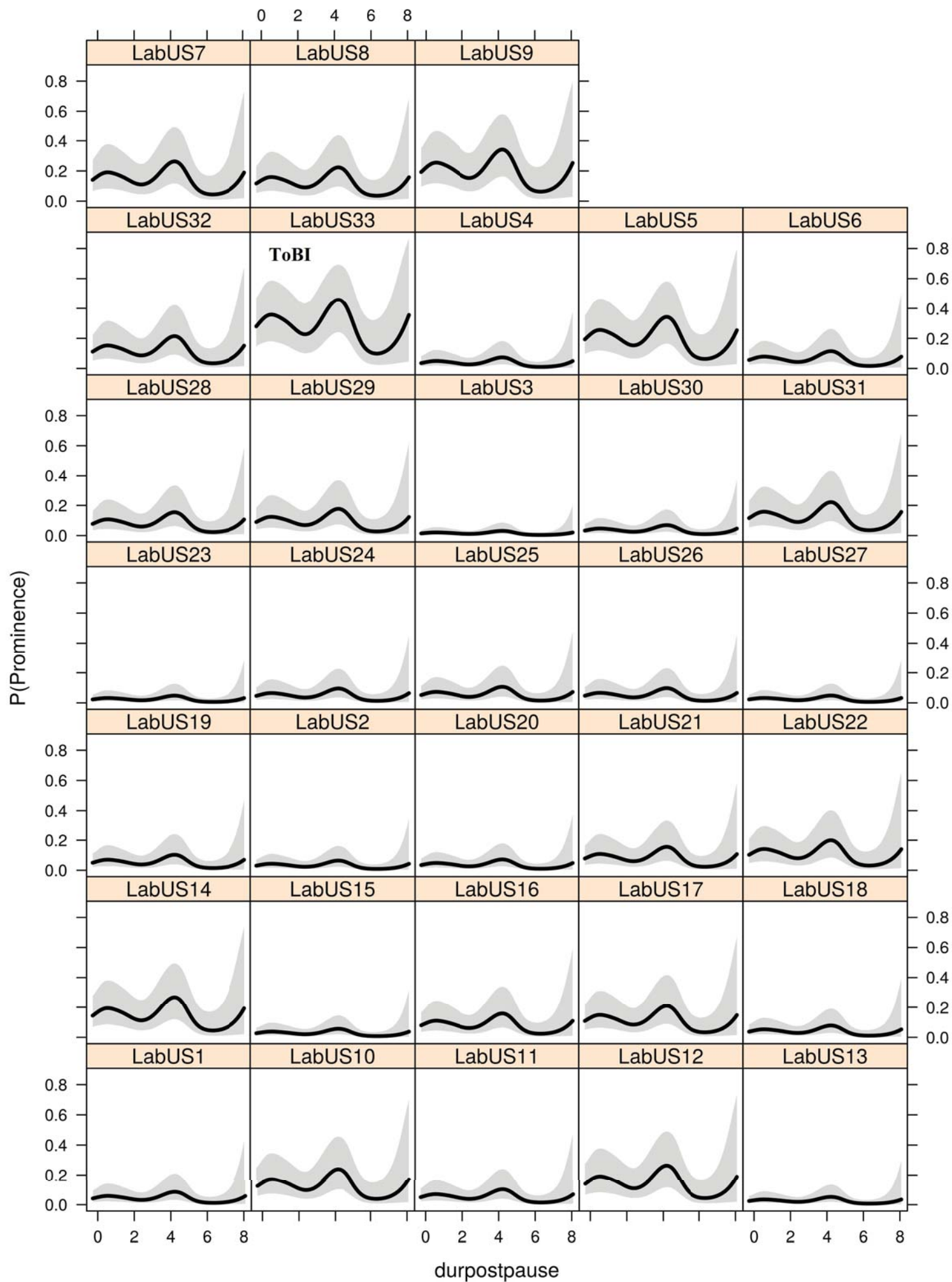
# Boundary



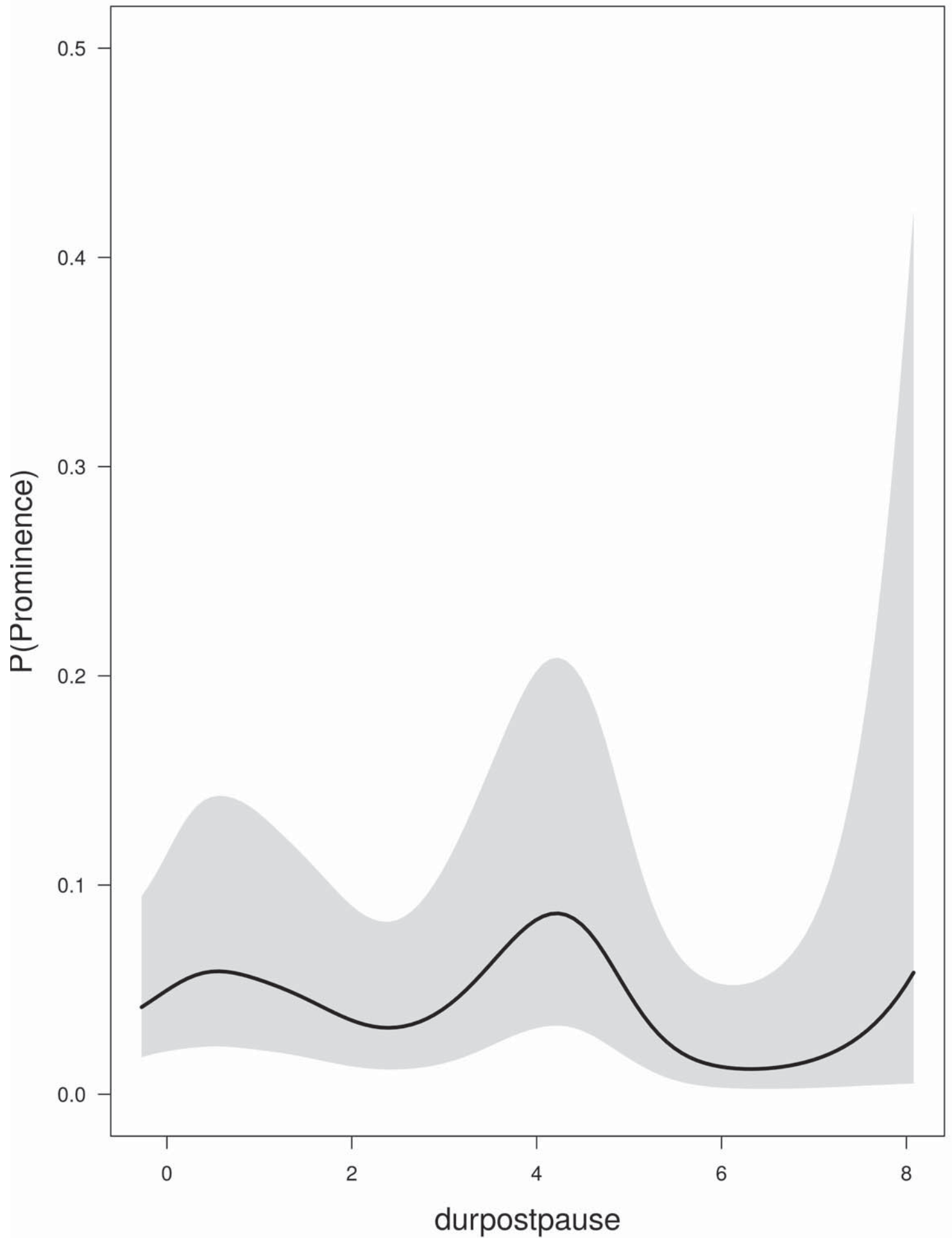
# Prominence\_all\_in\_one



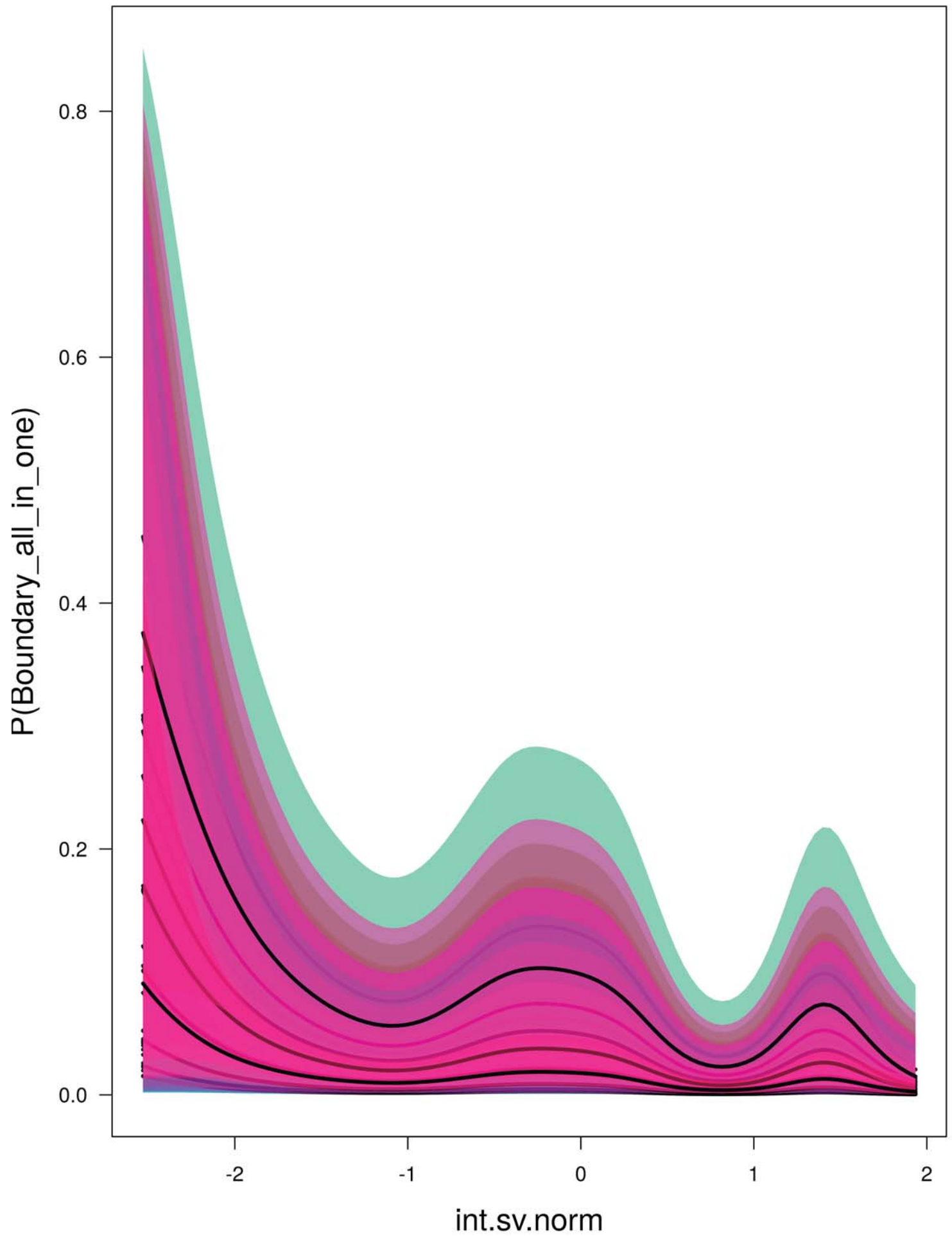
# Prominence



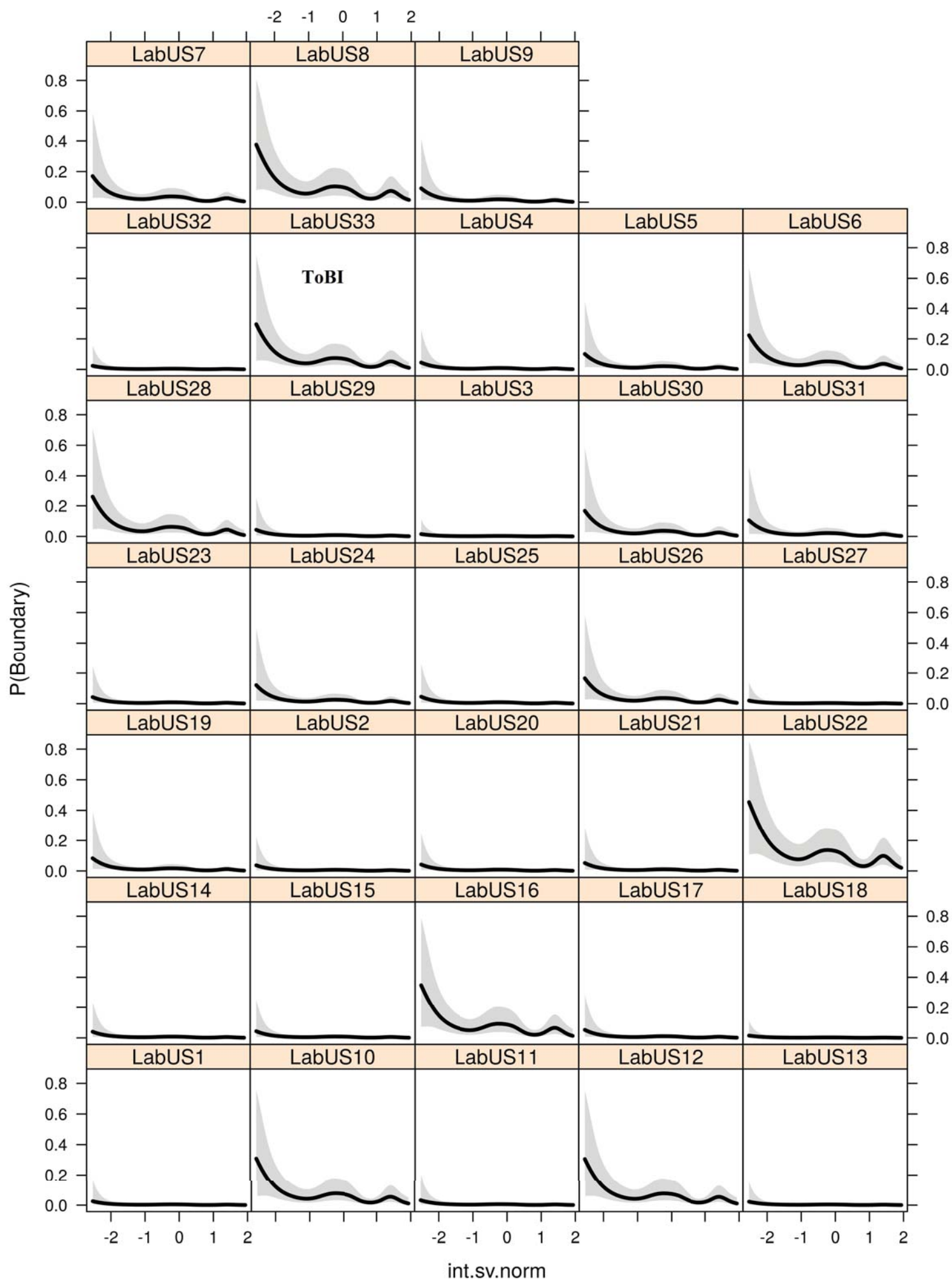
# Prominence



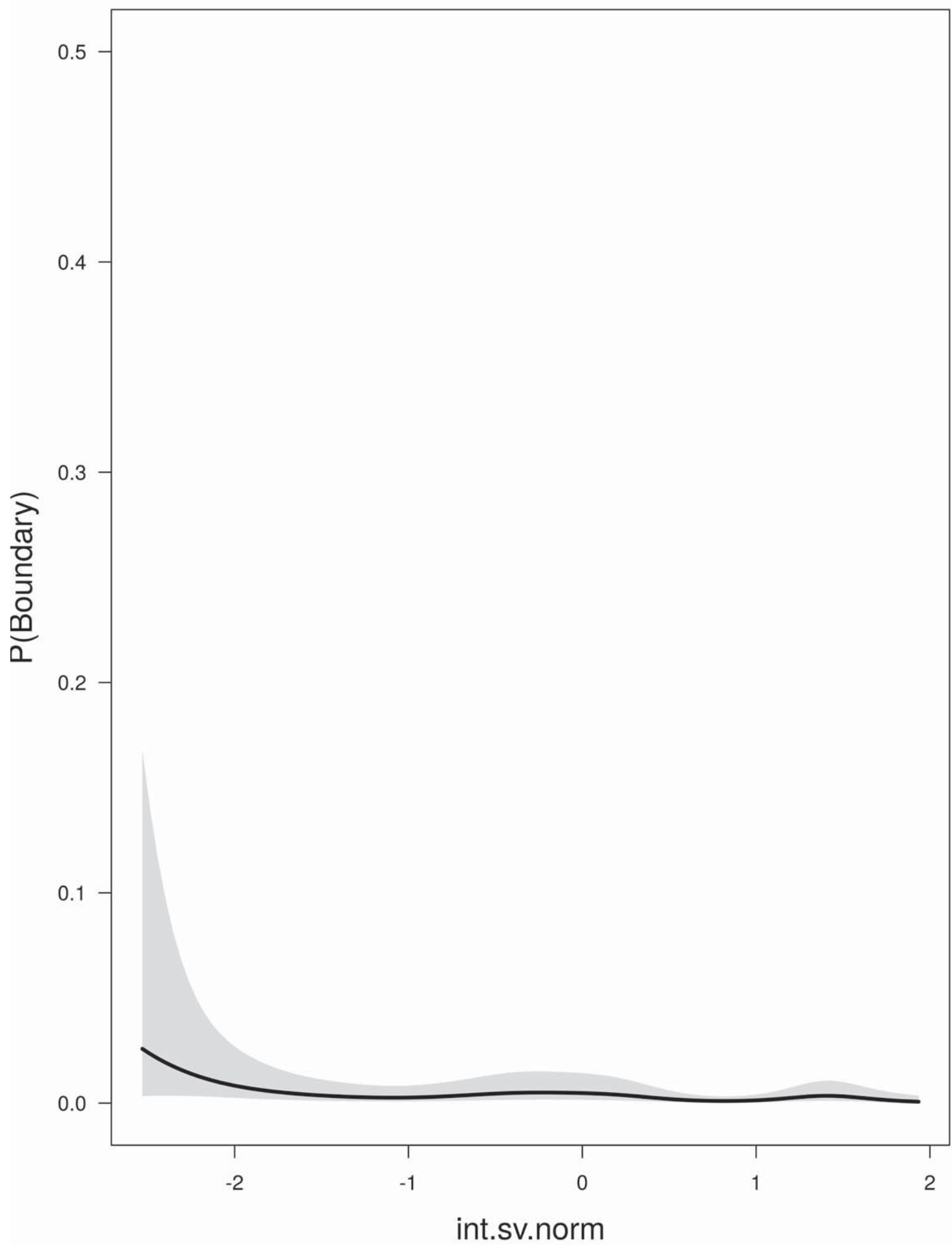
# Boundary\_all\_in\_one



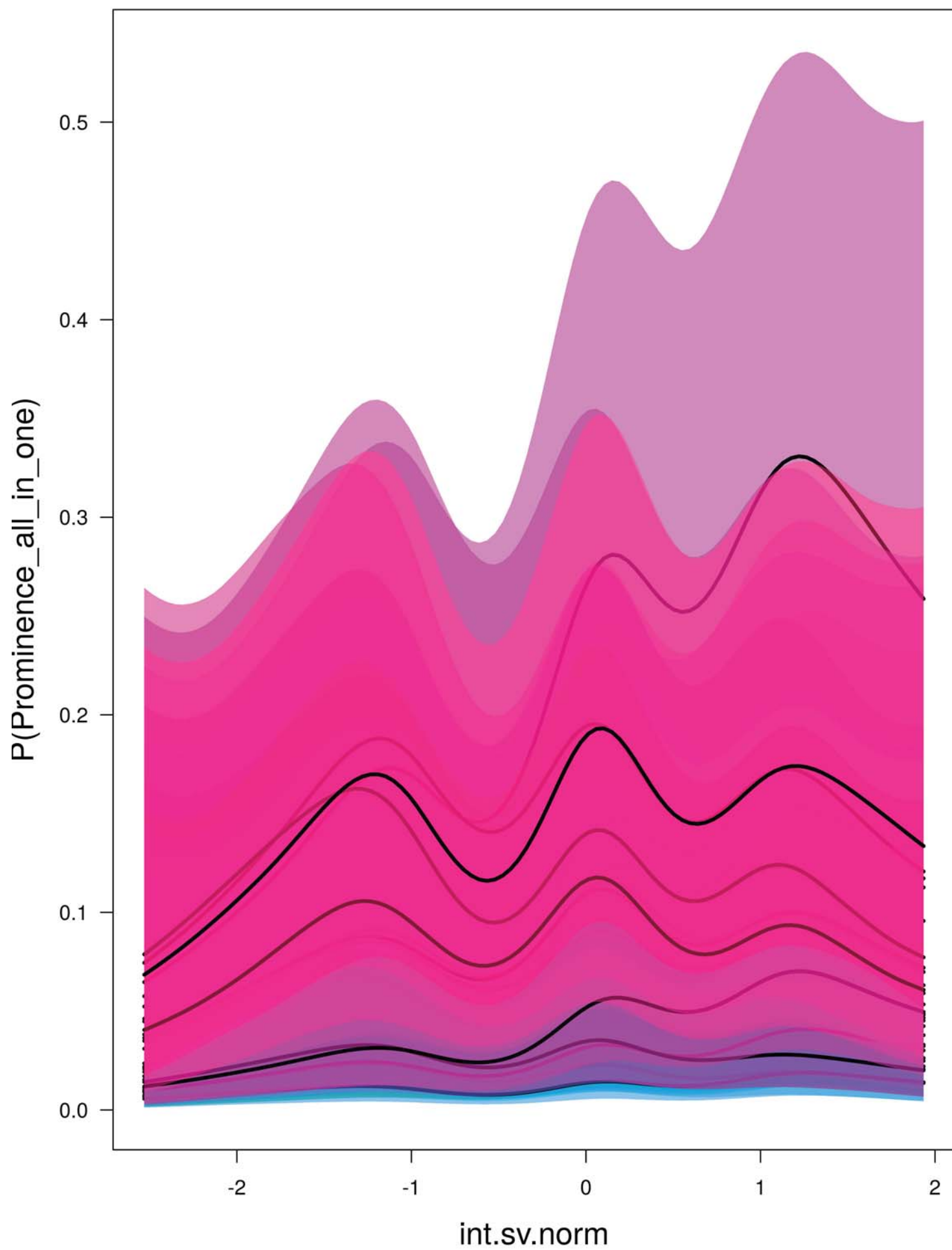
# Boundary



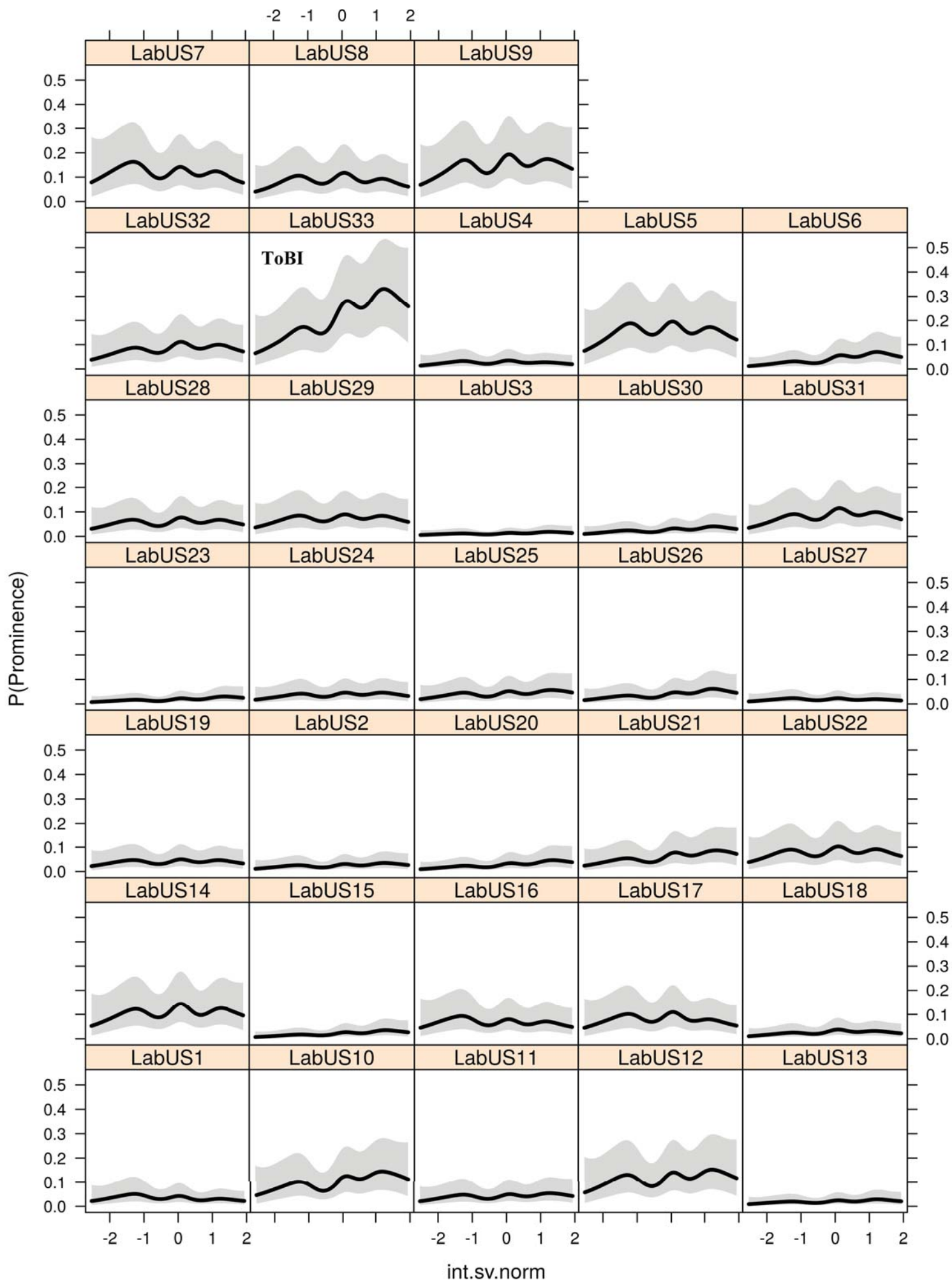
# Boundary



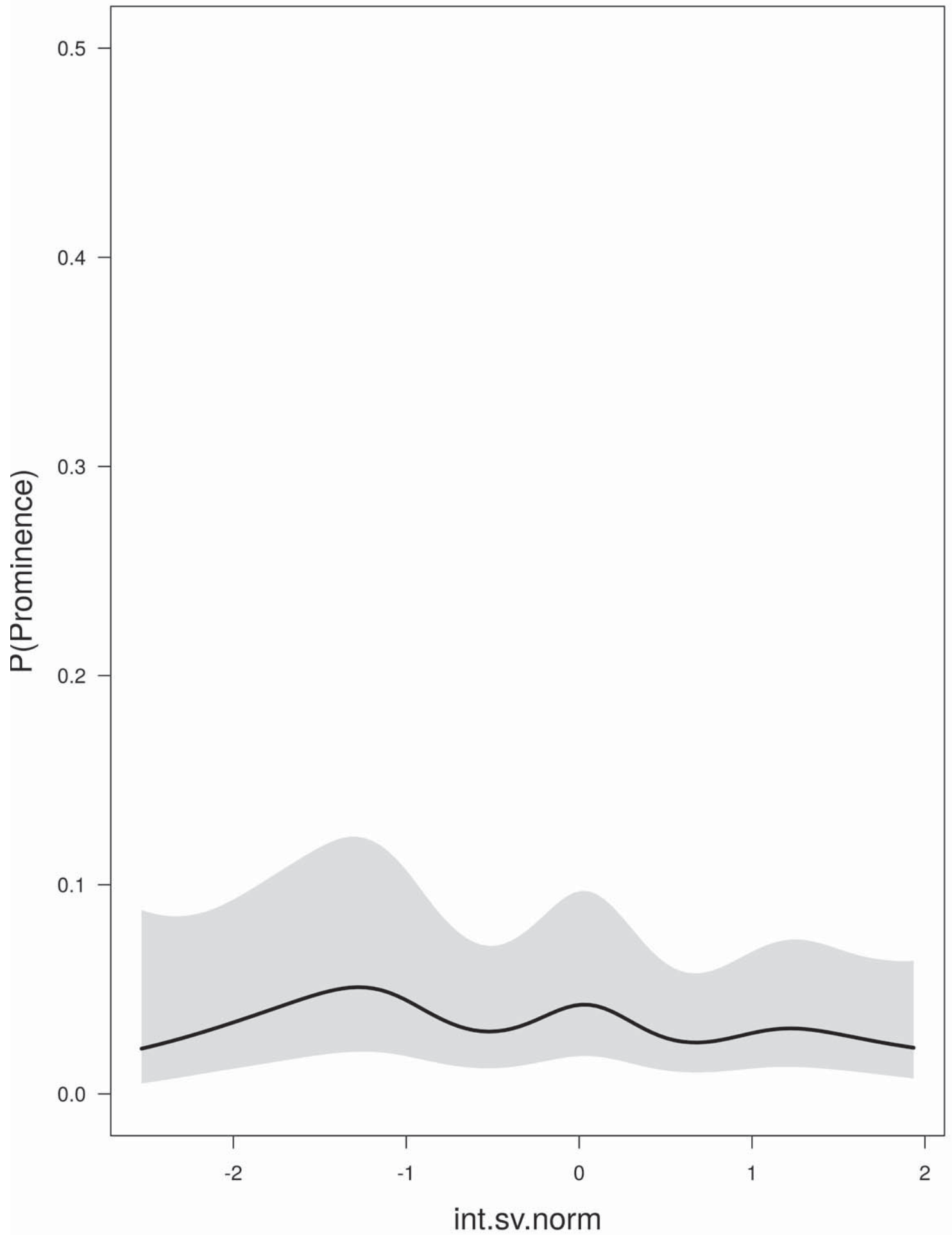
# Prominence\_all\_in\_one



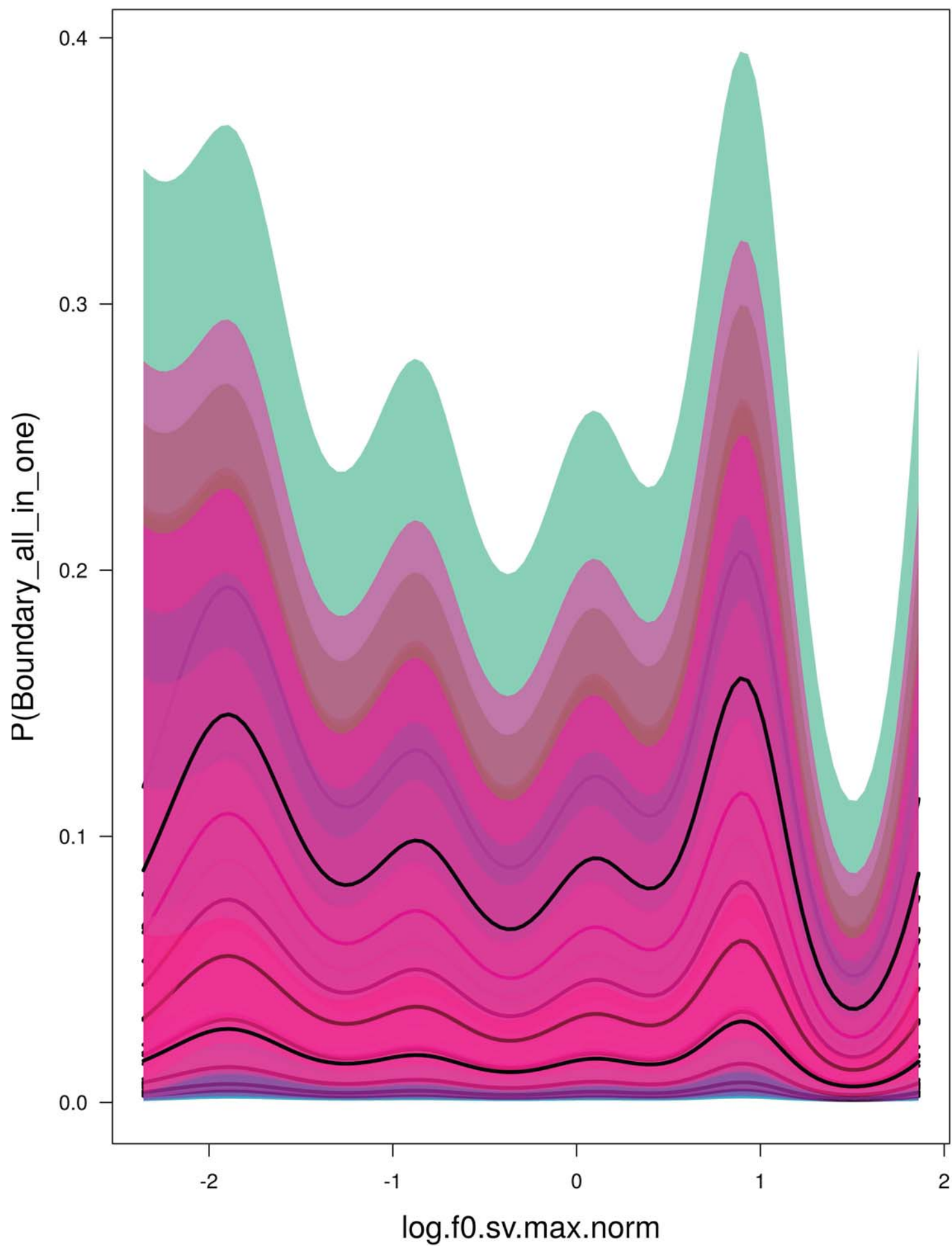
# Prominence



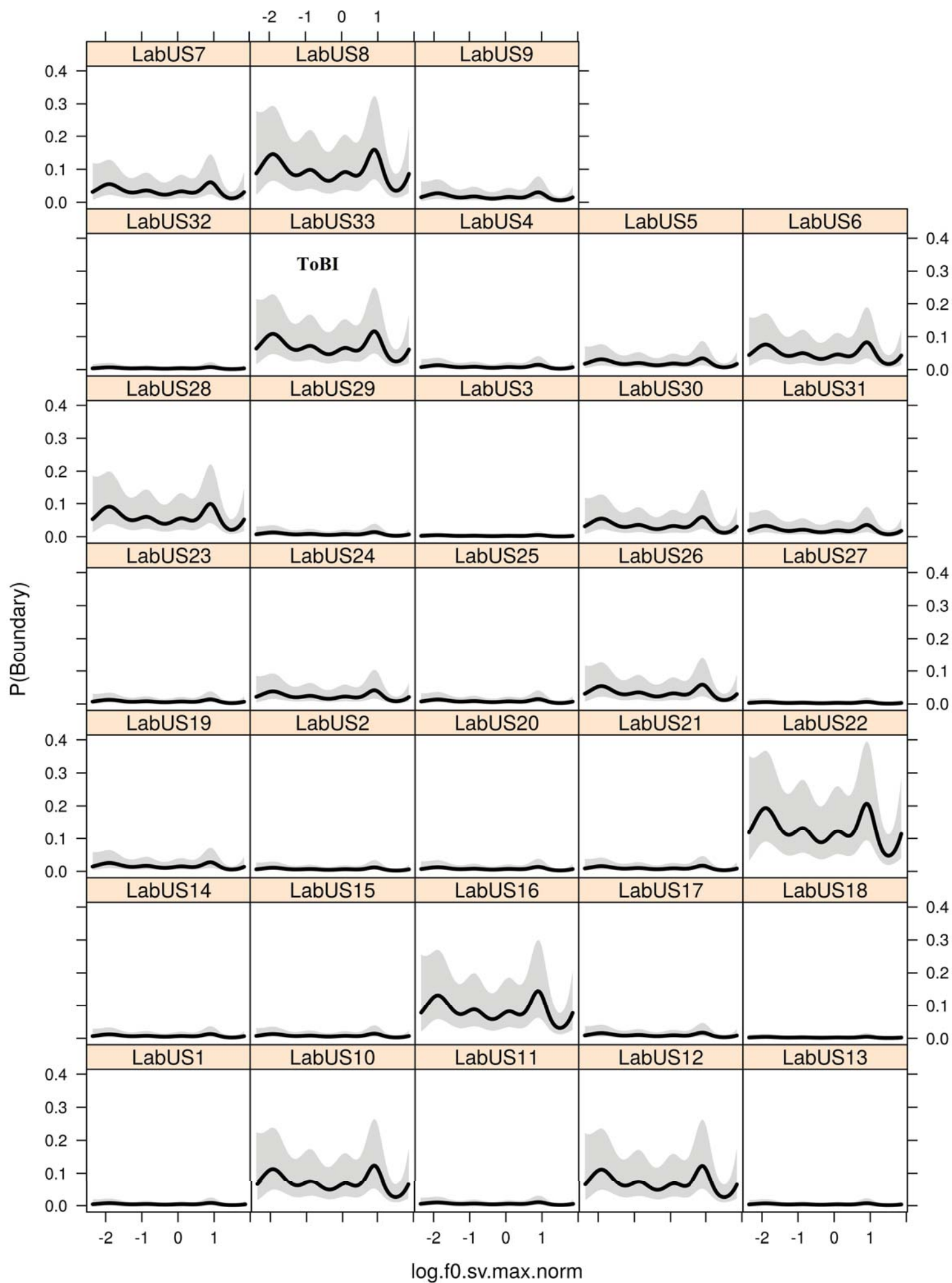
# Prominence



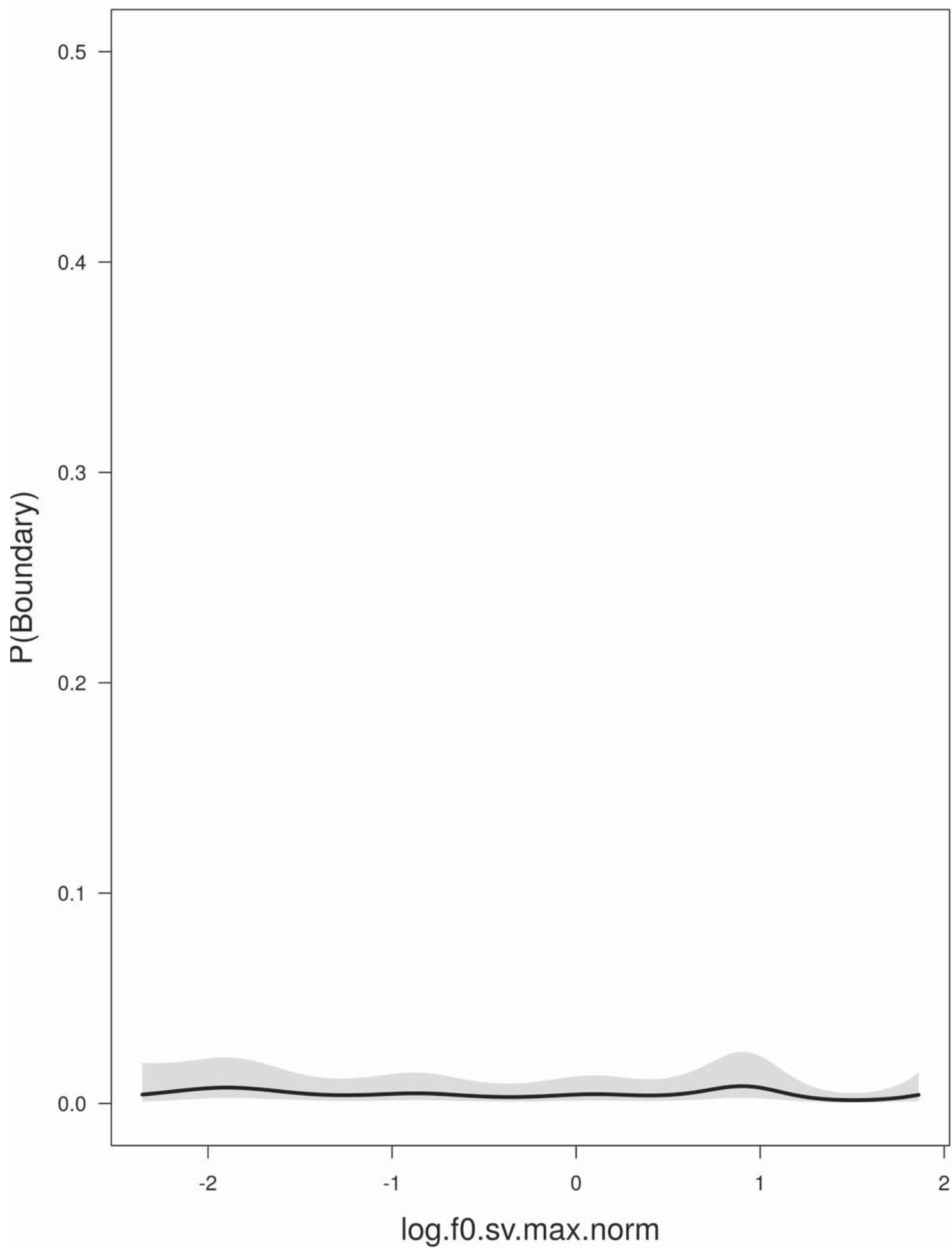
# Boundary\_all\_in\_one



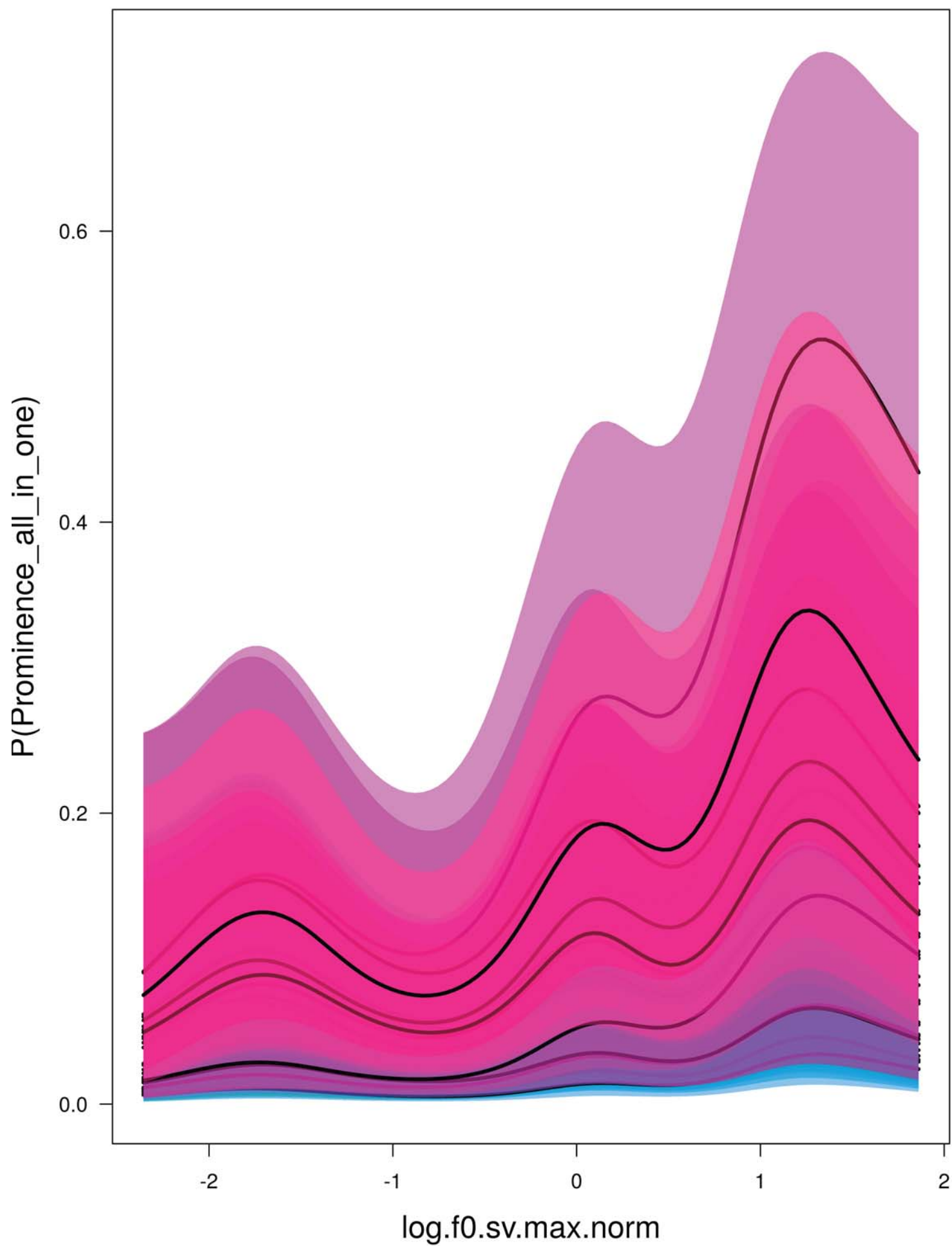
# Boundary



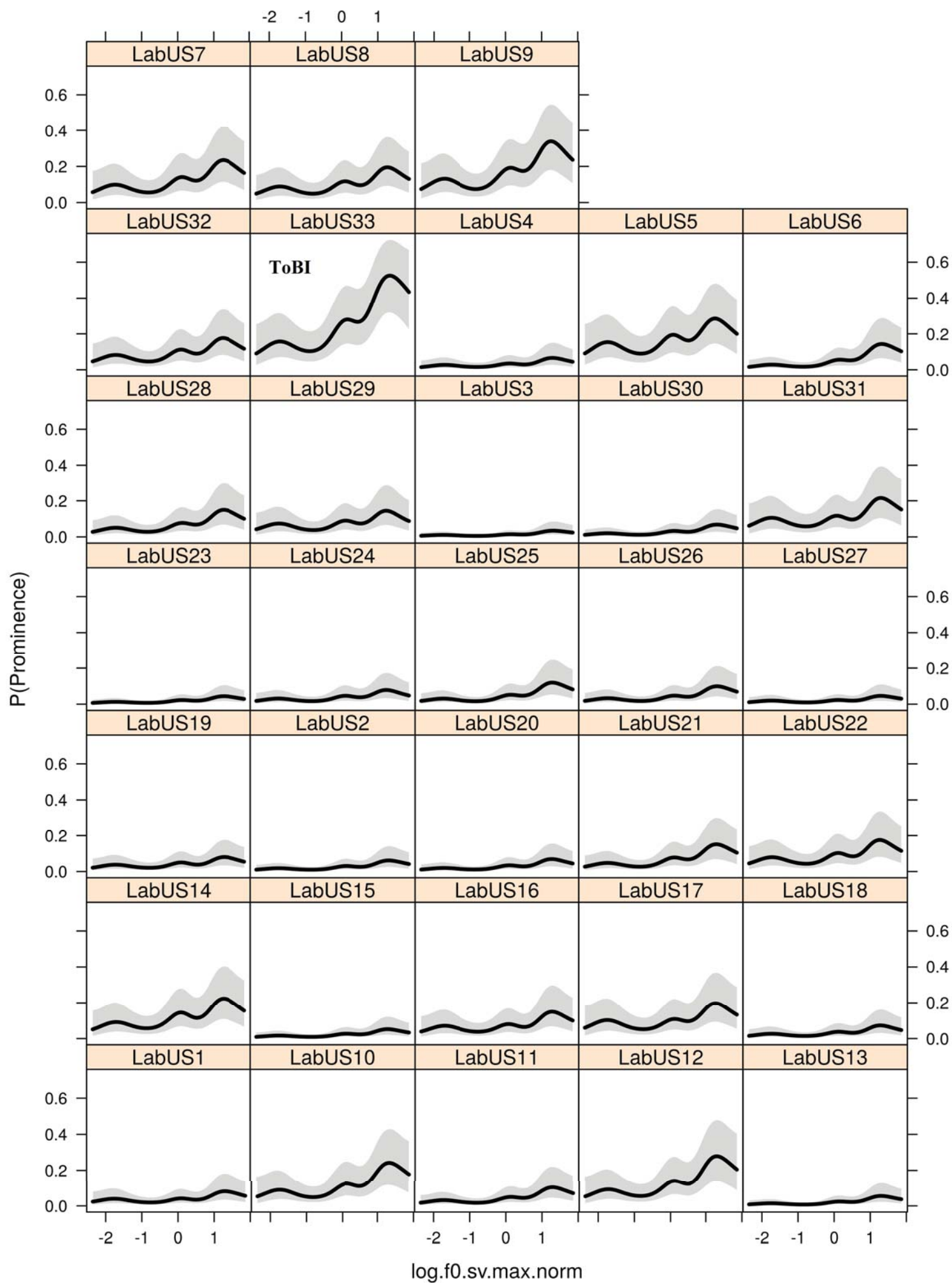
# Boundary



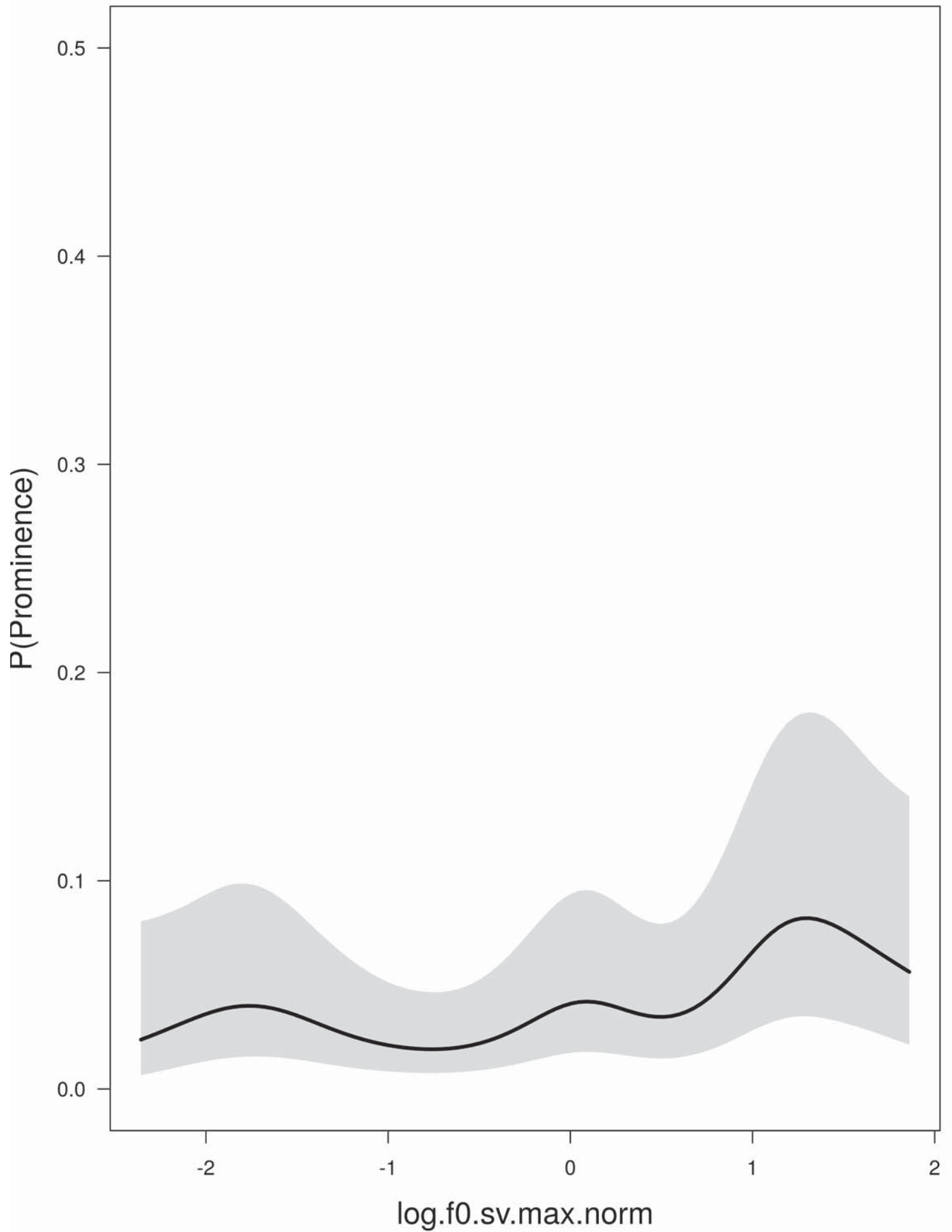
# Prominence\_all\_in\_one



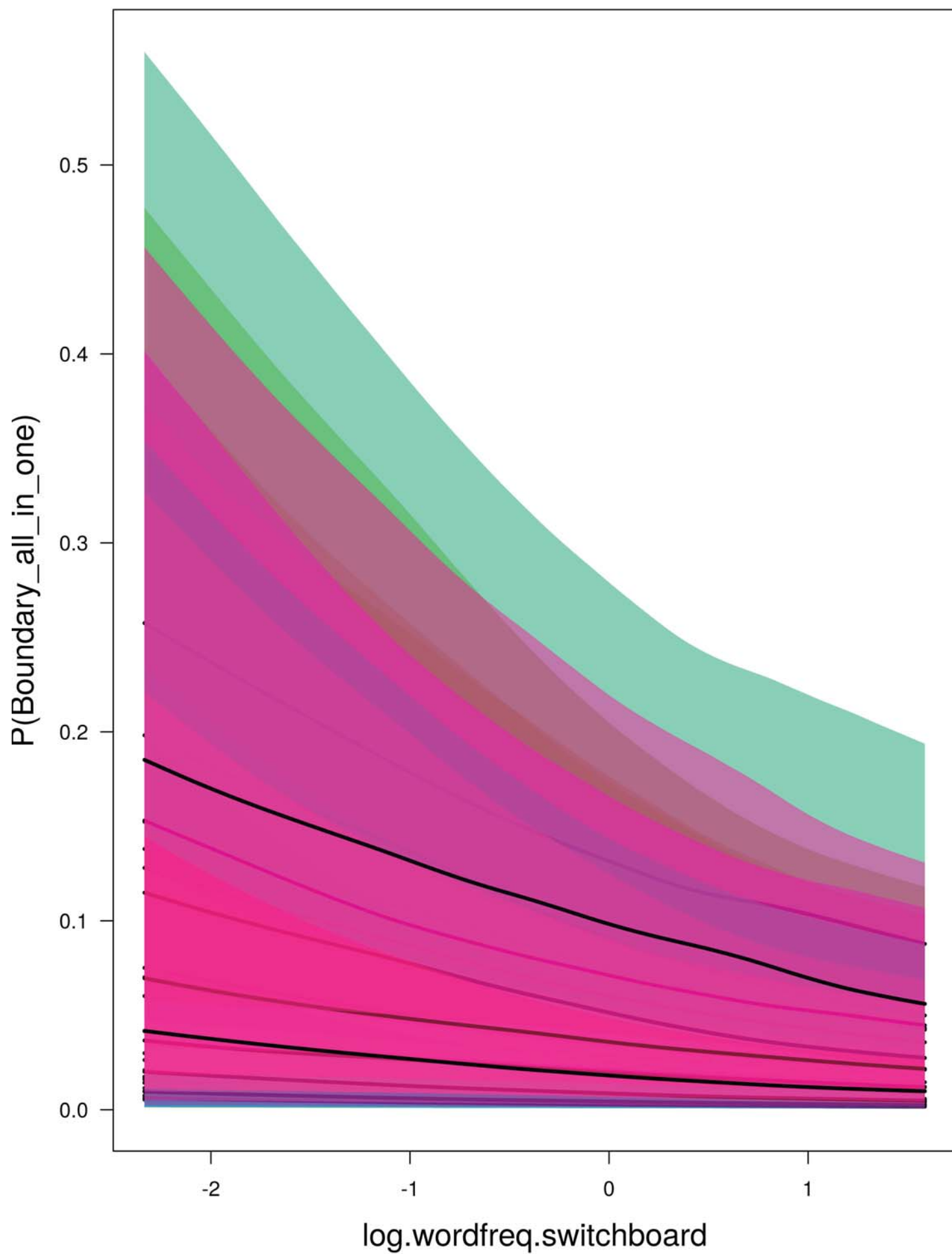
# Prominence



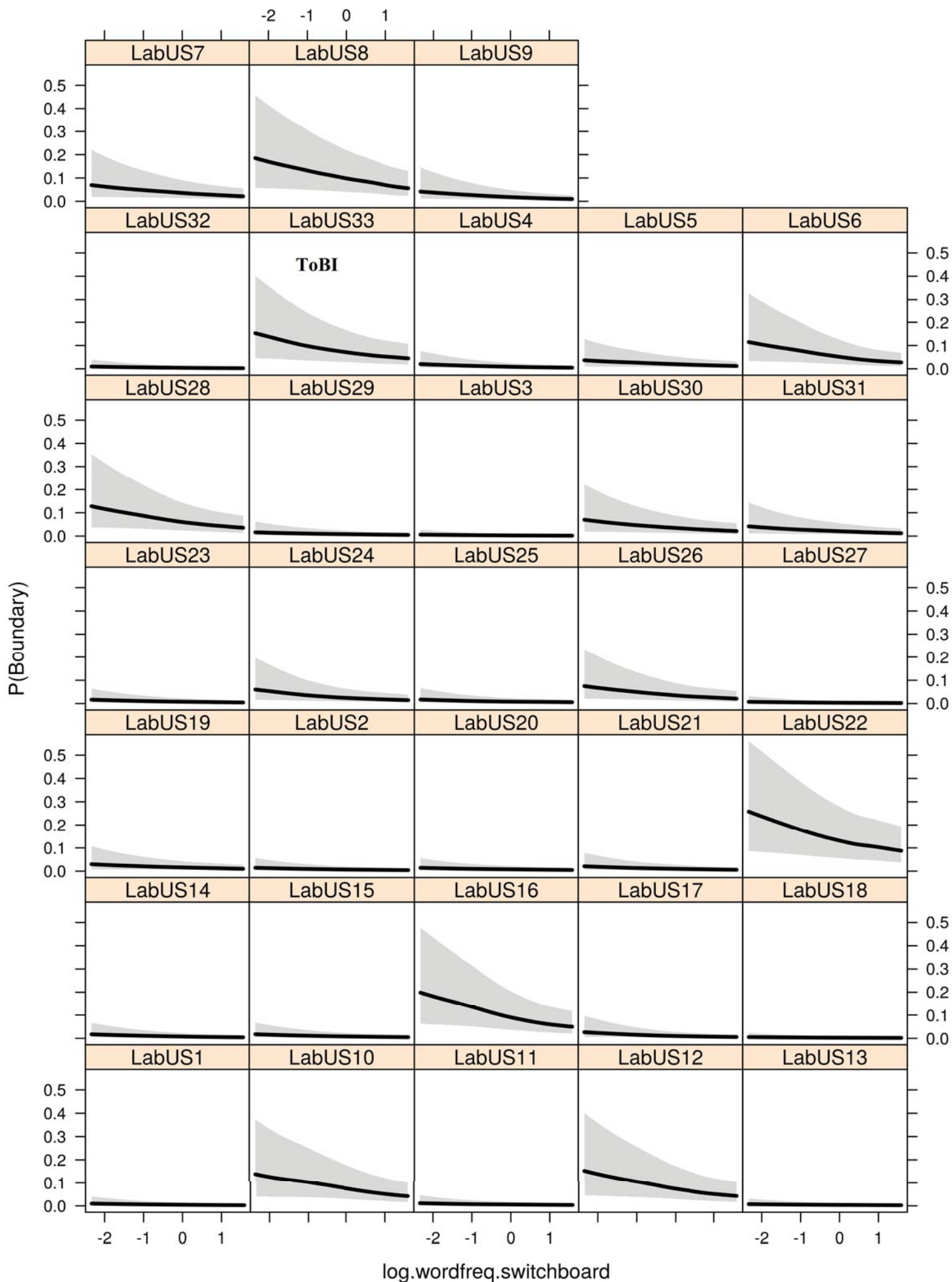
# Prominence



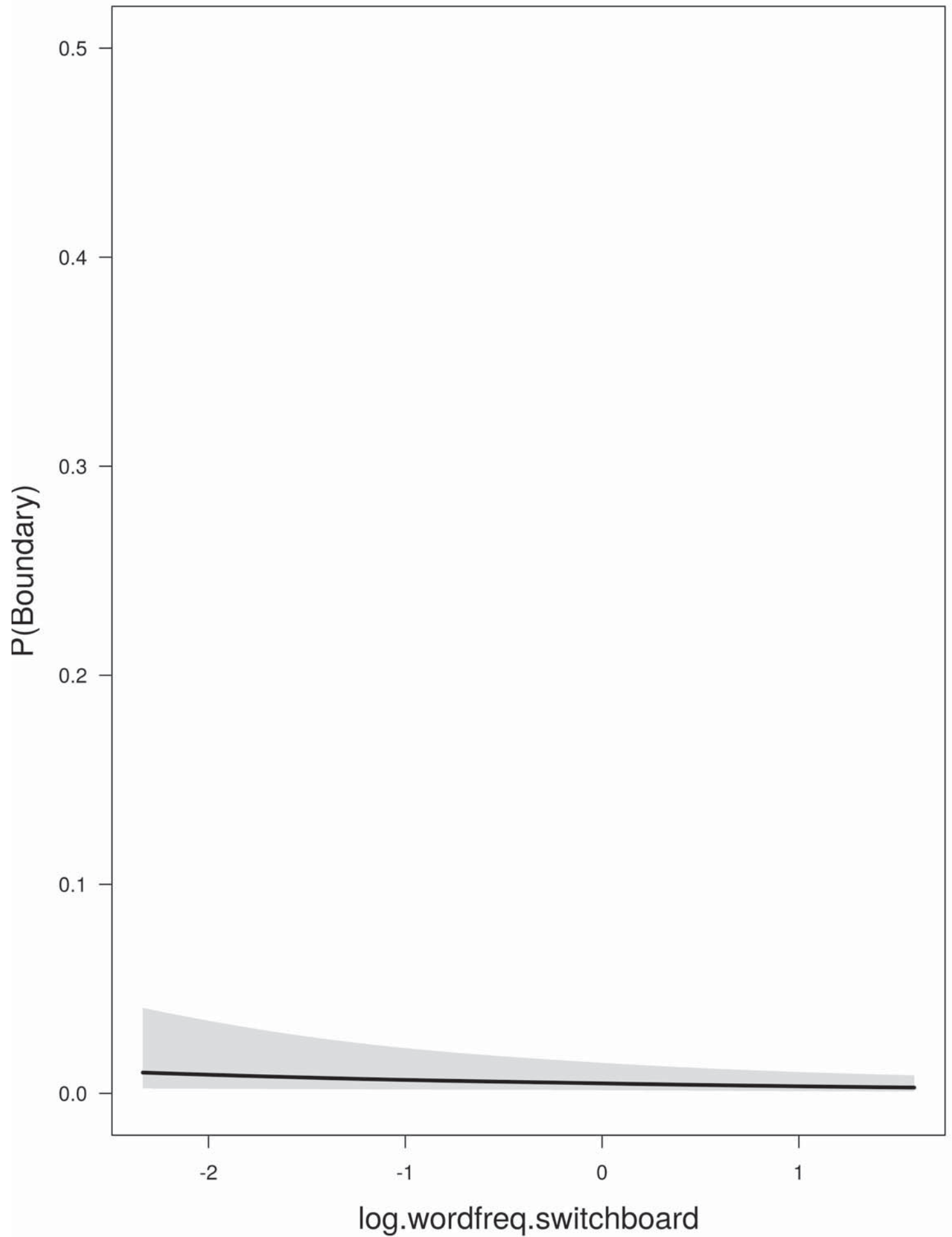
# Boundary\_all\_in\_one



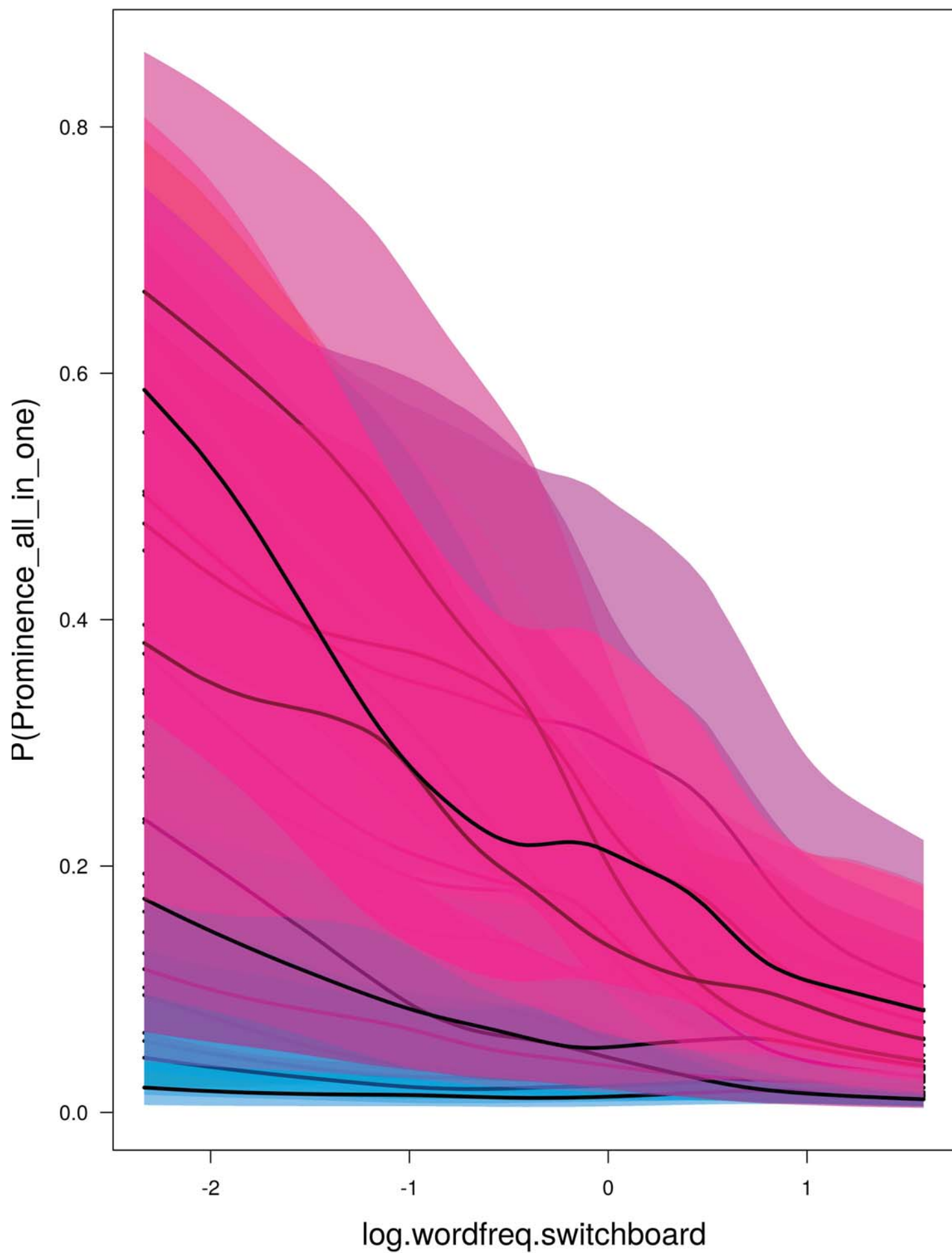
# Boundary



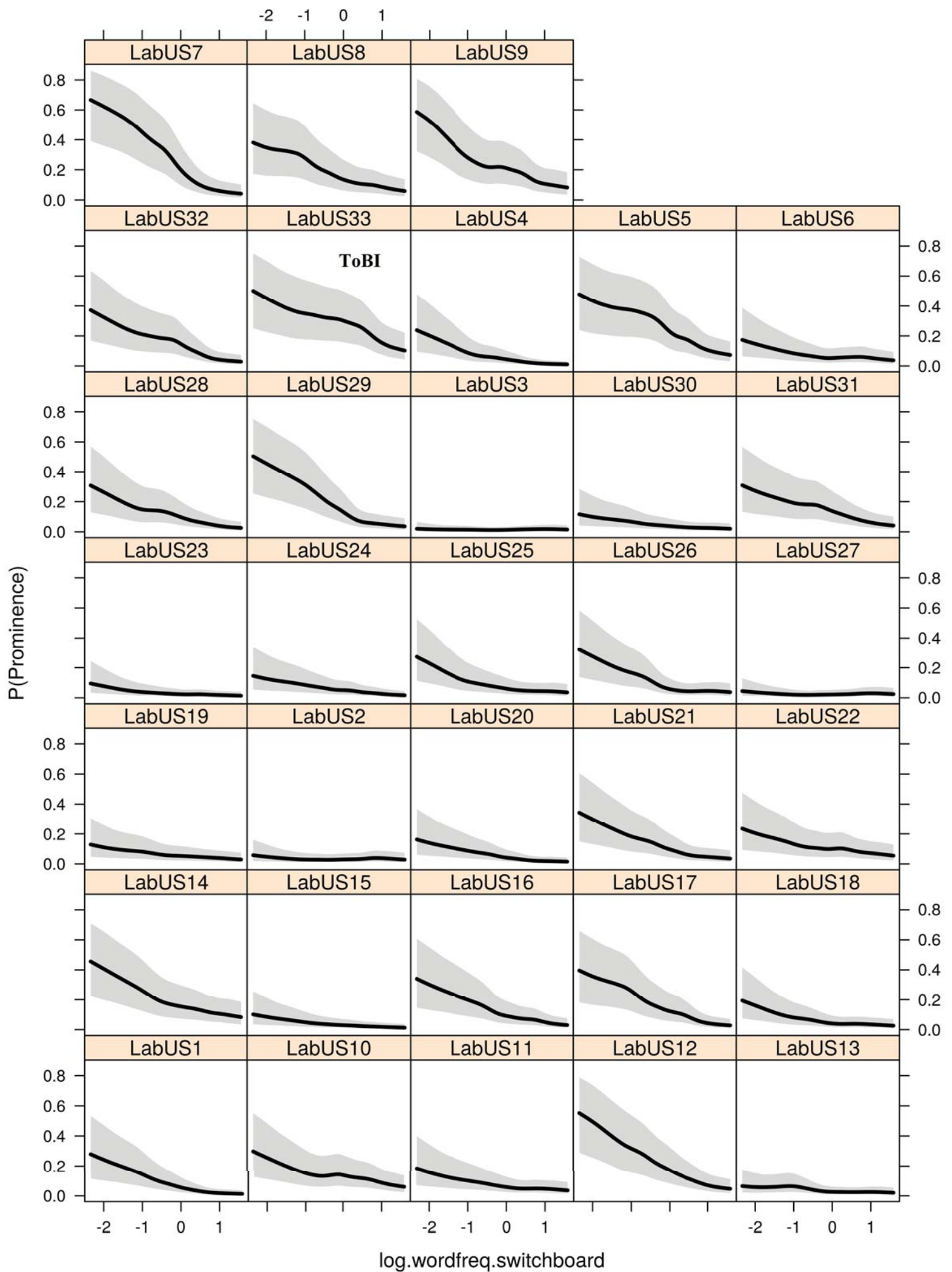
# Boundary



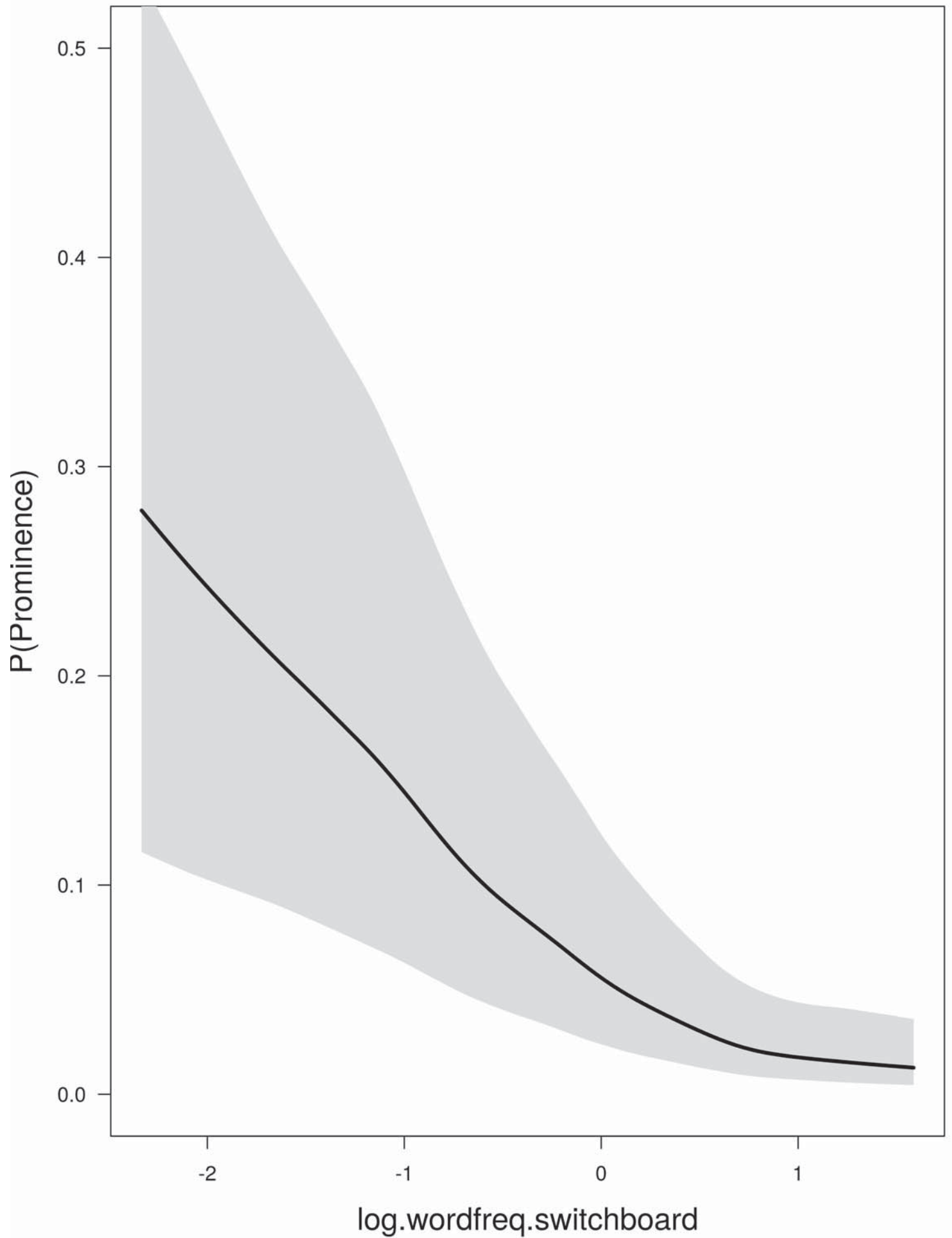
# Prominence\_all\_in\_one



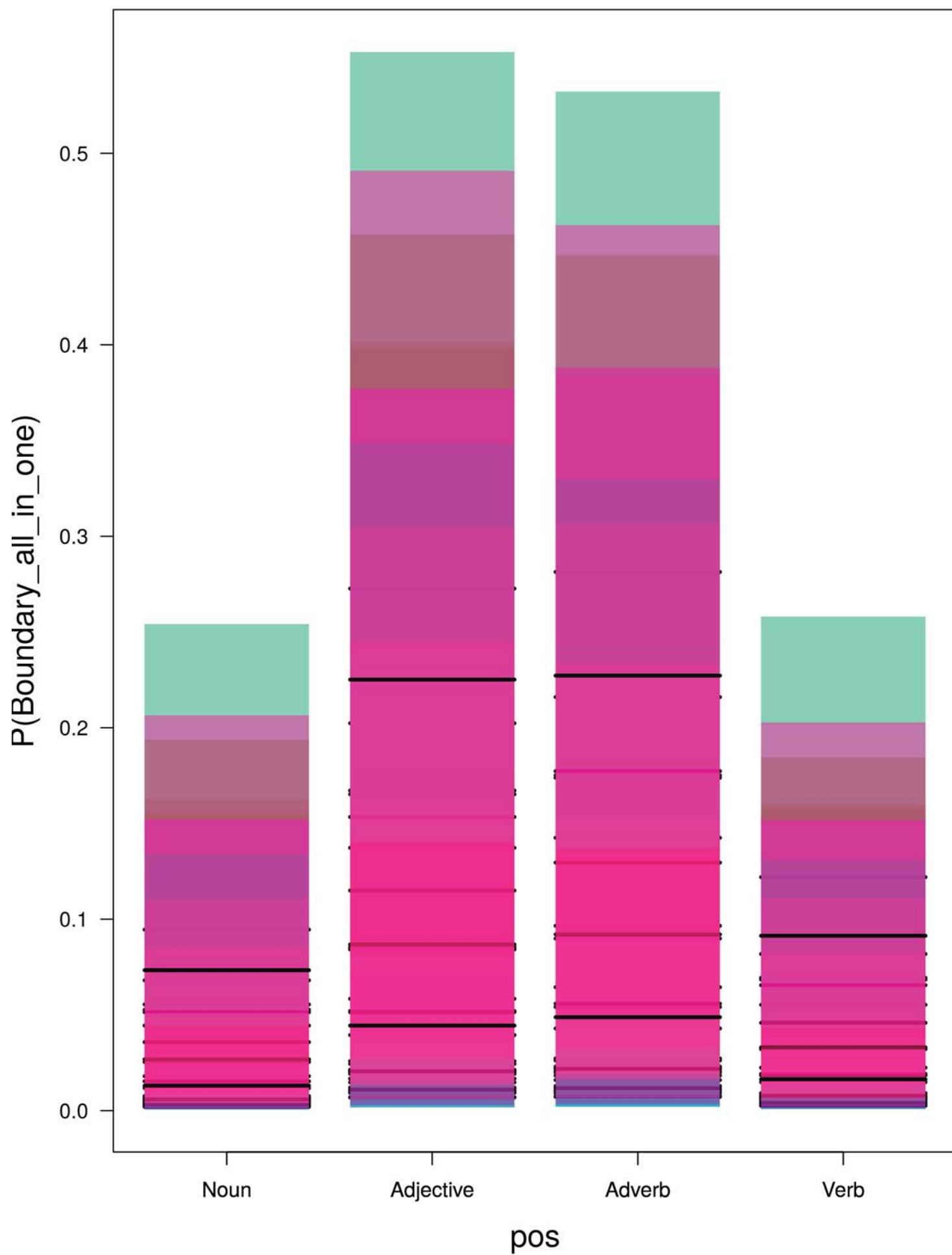
# Prominence



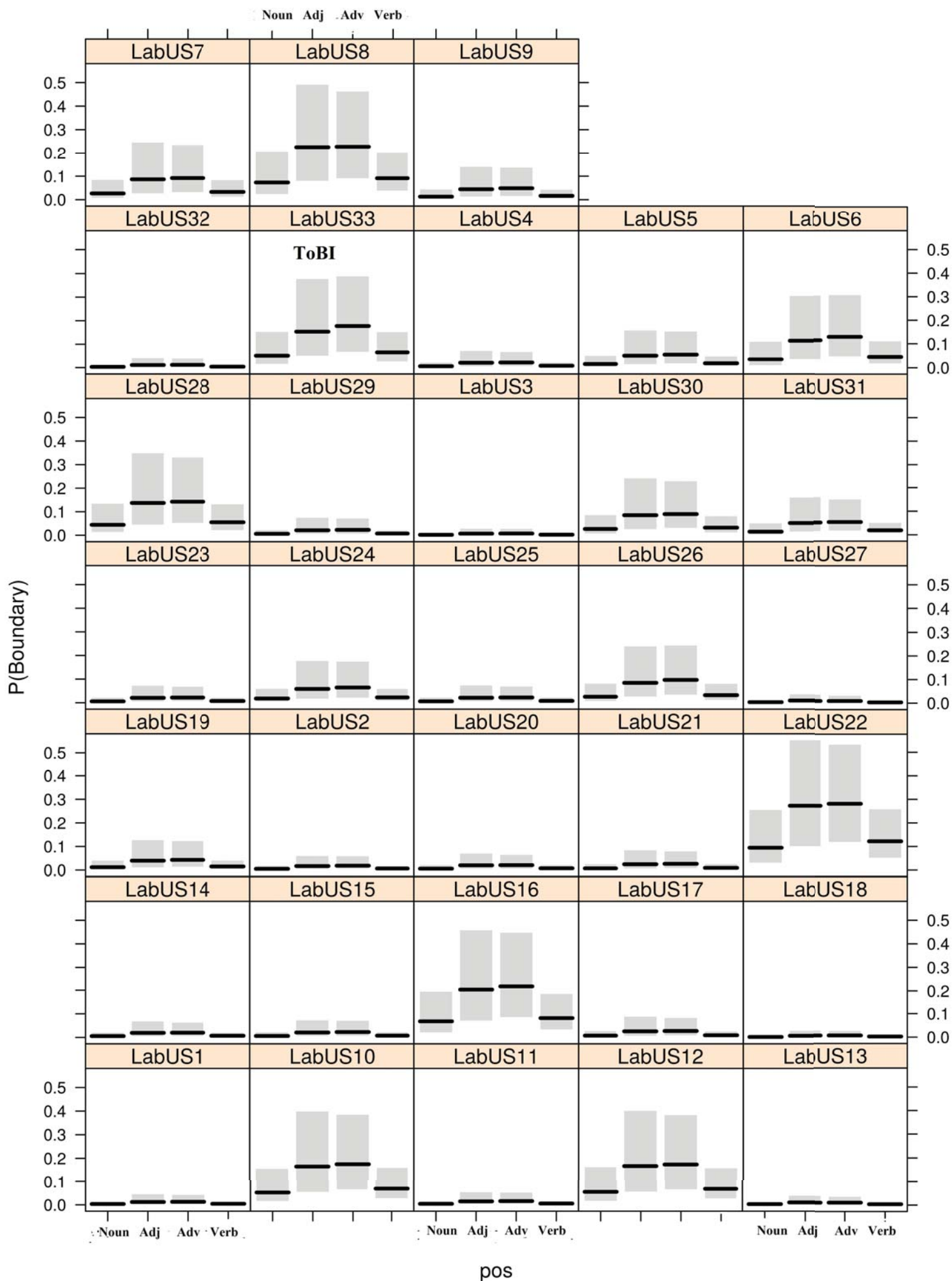
# Prominence



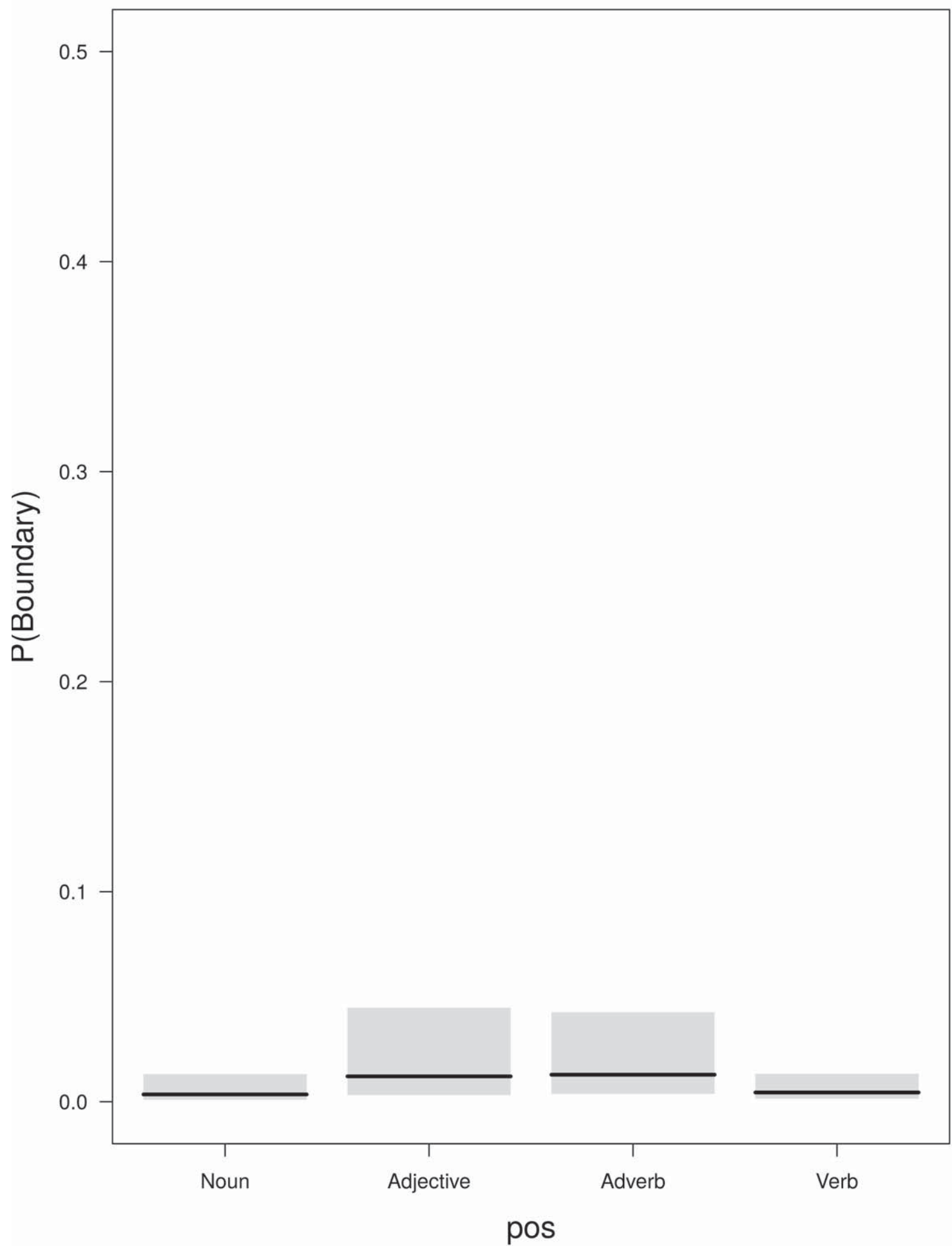
# Boundary\_all\_in\_one



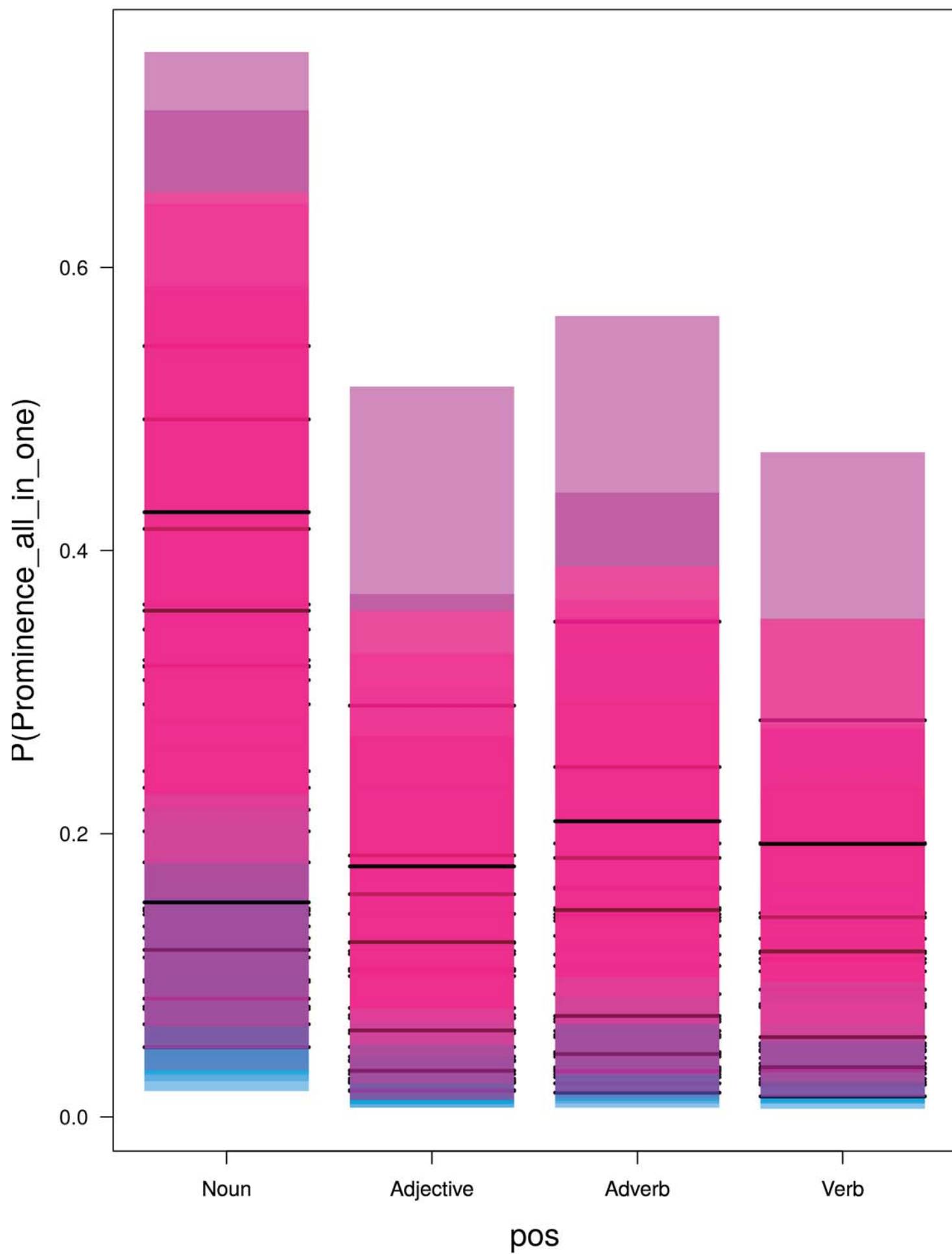
# Boundary



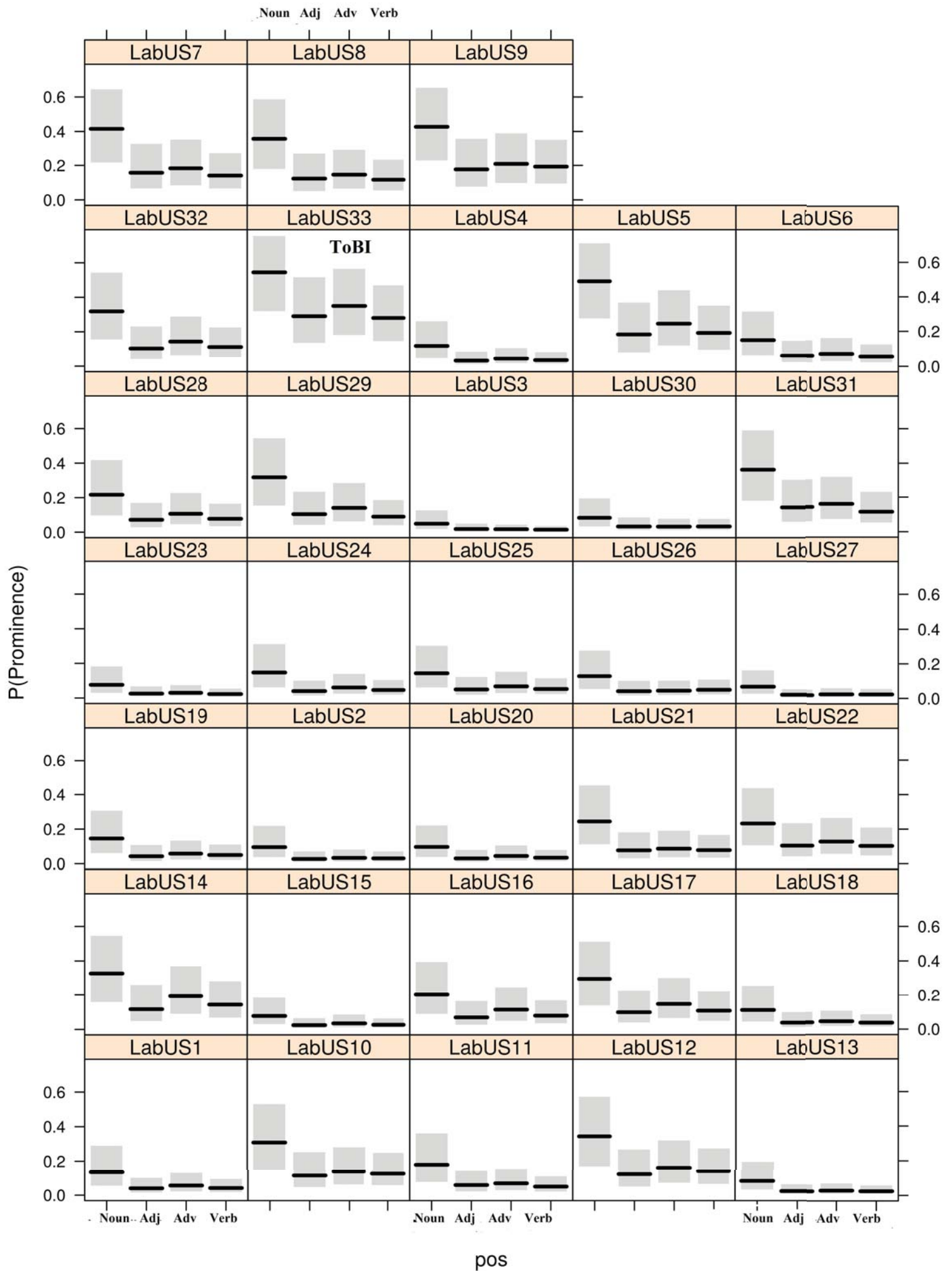
# Boundary



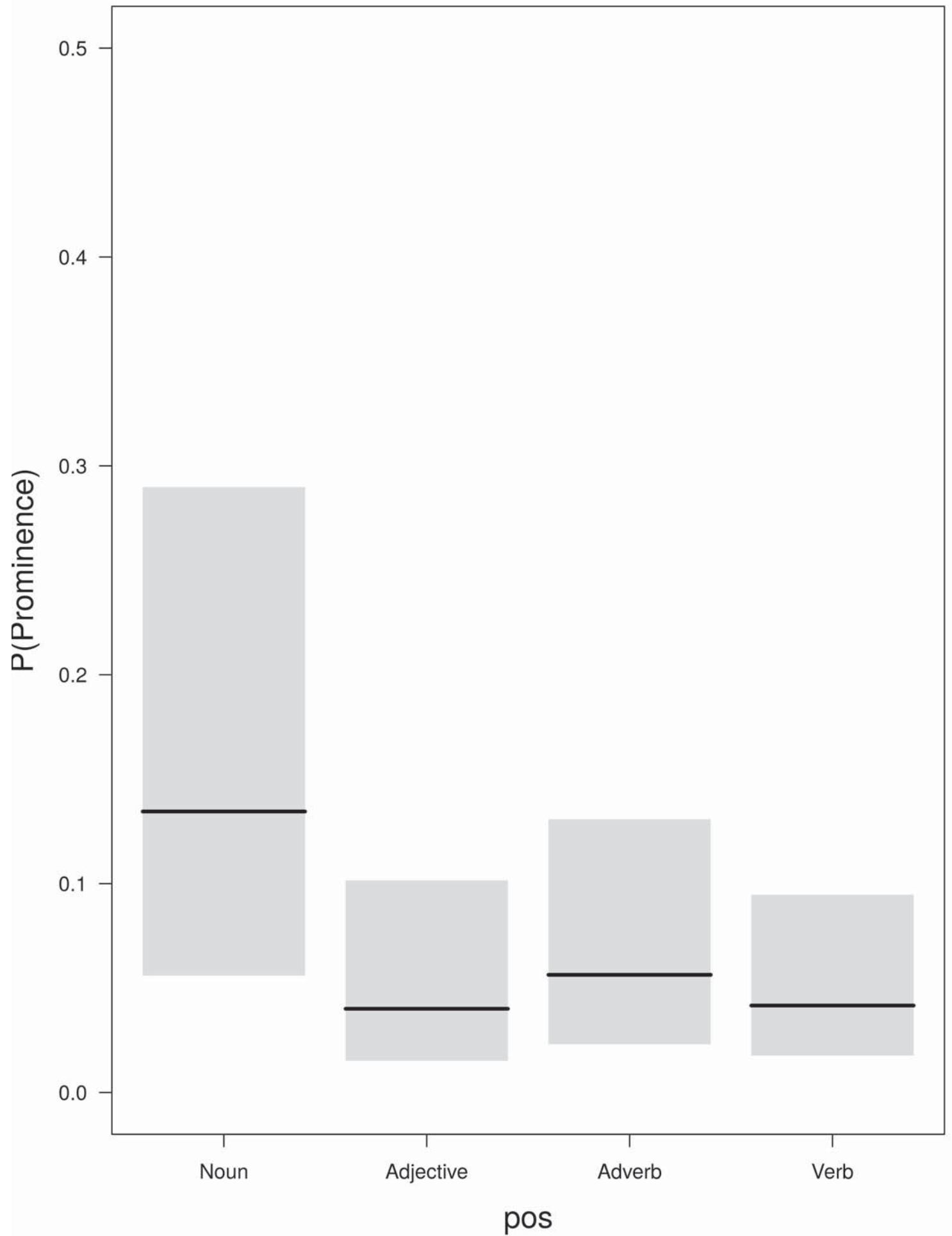
# Prominence\_all\_in\_one



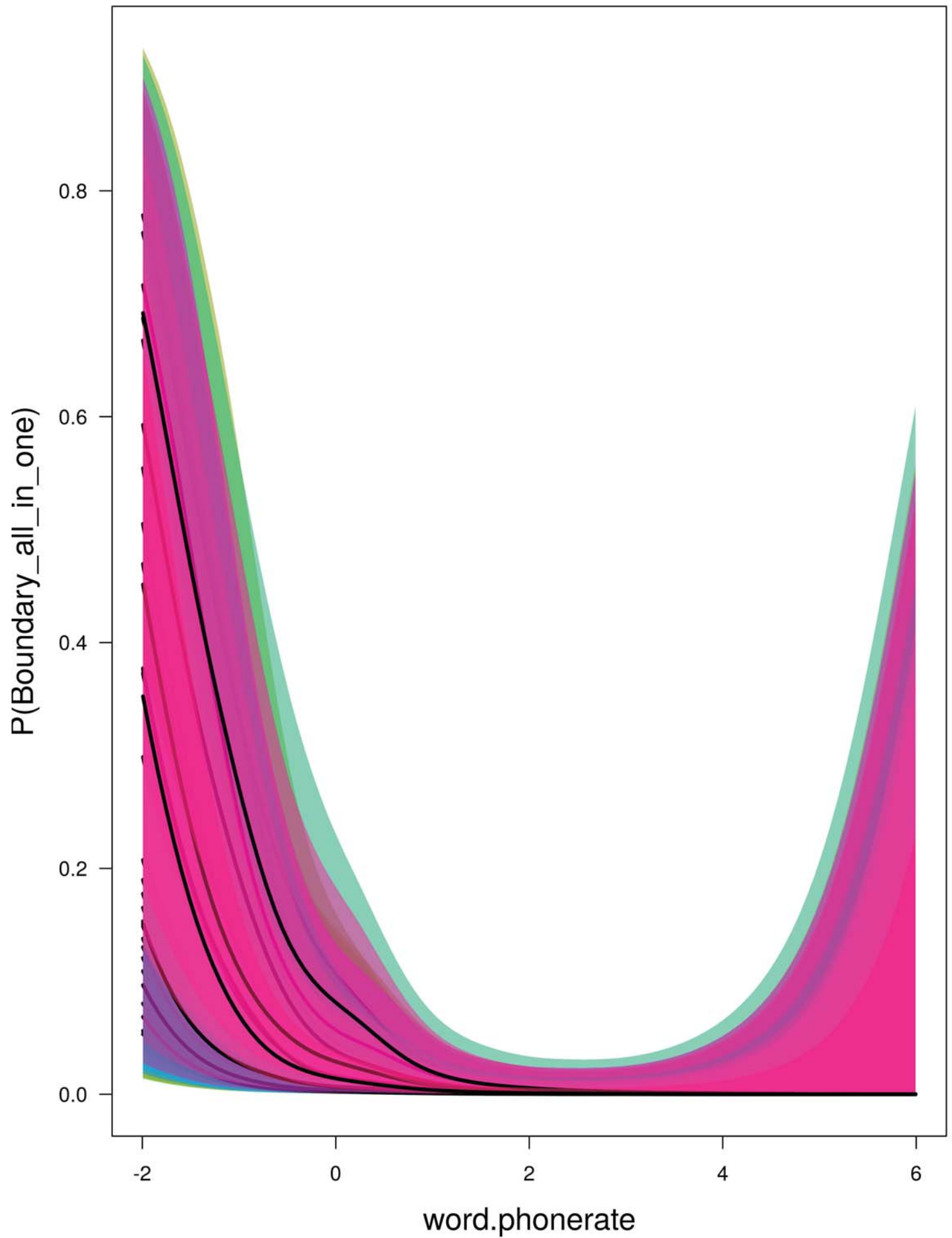
# Prominence



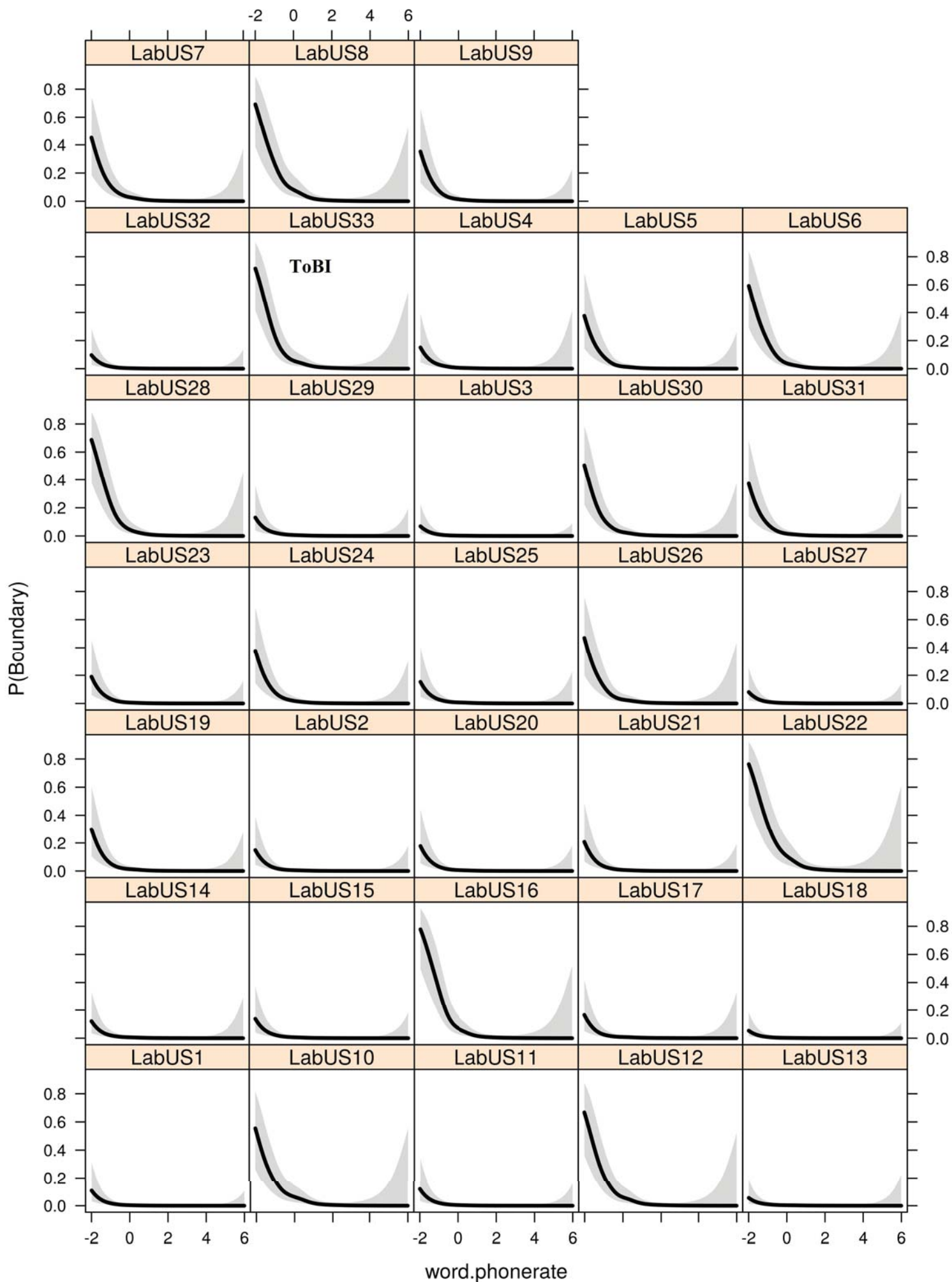
# Prominence



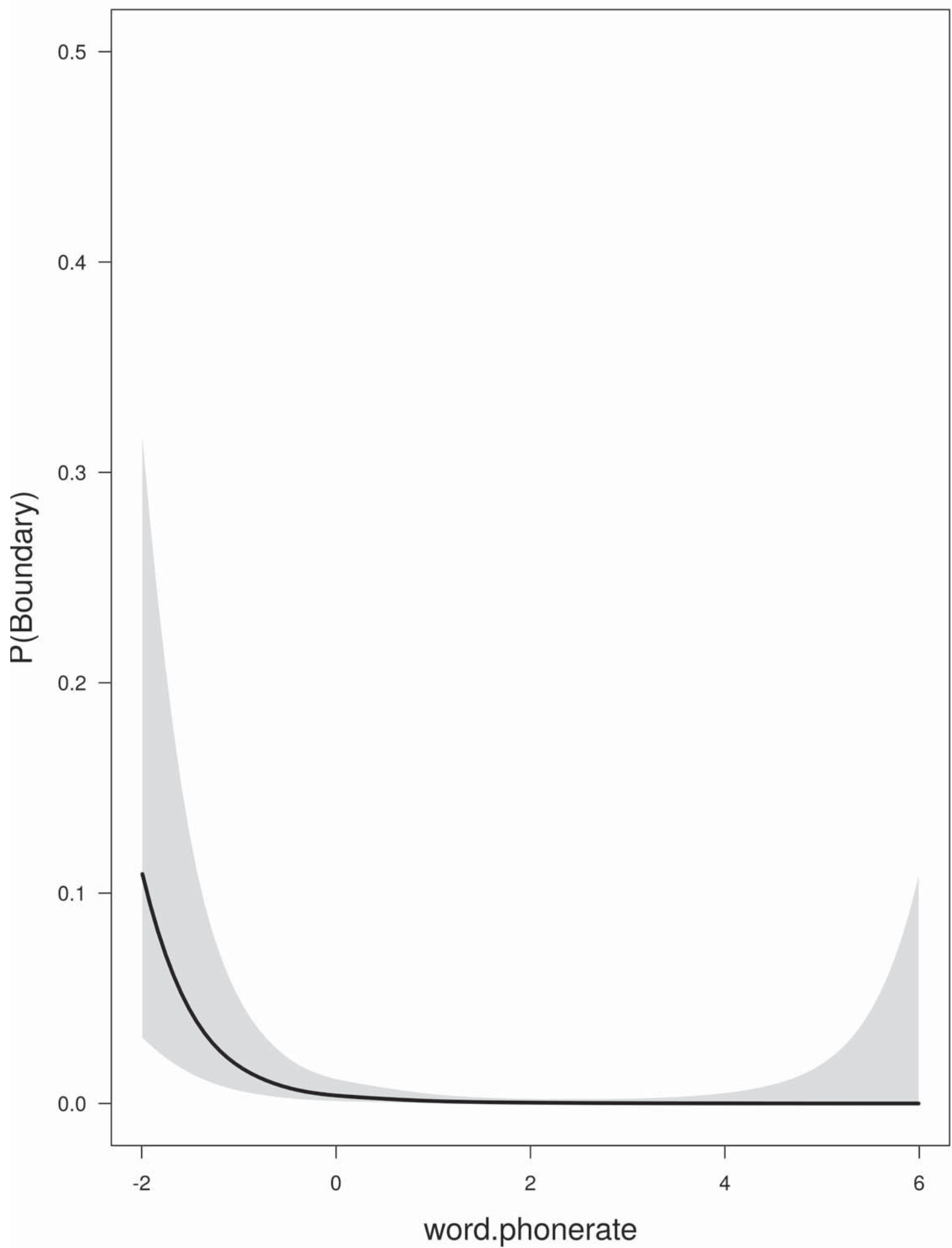
# Boundary\_all\_in\_one



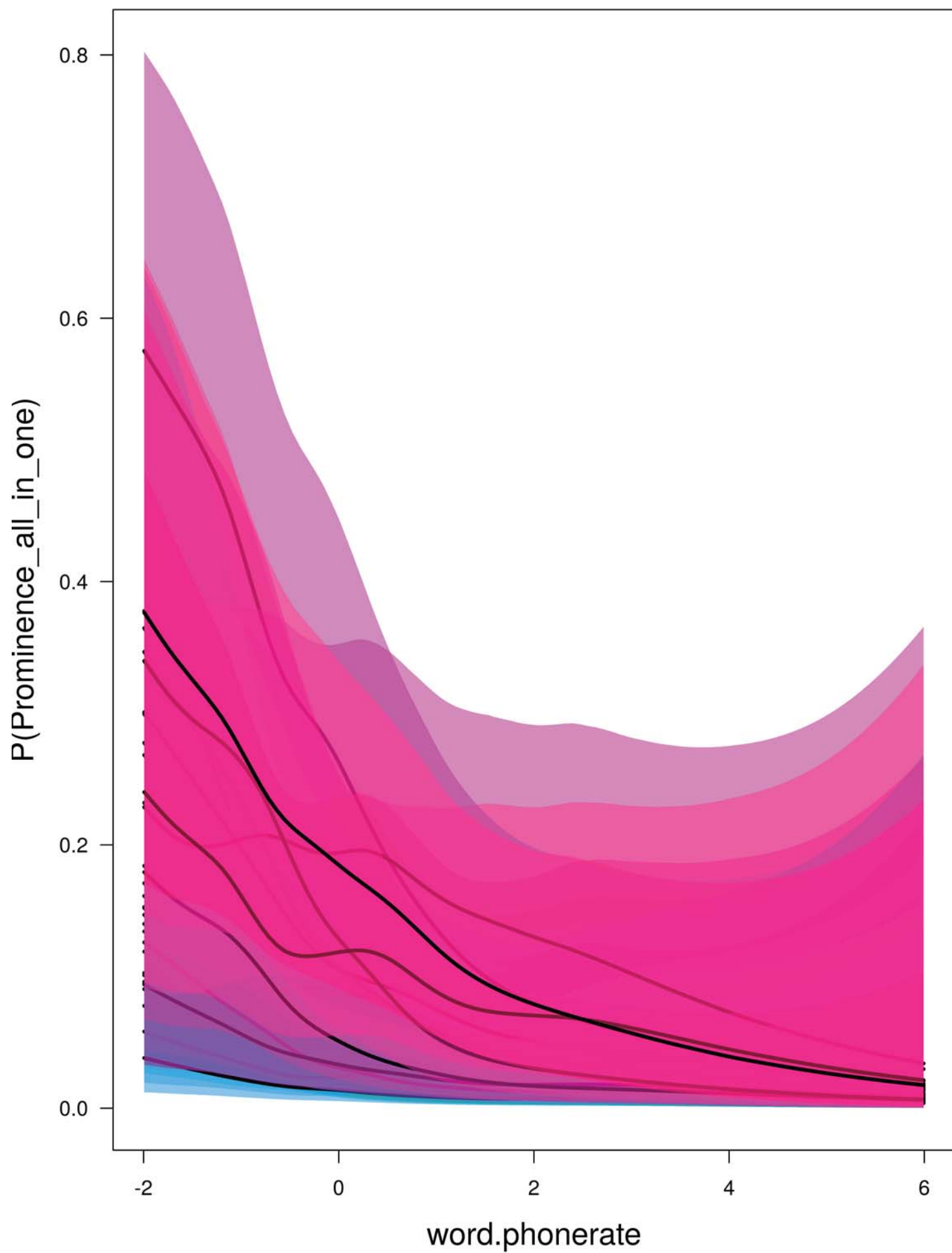
# Boundary



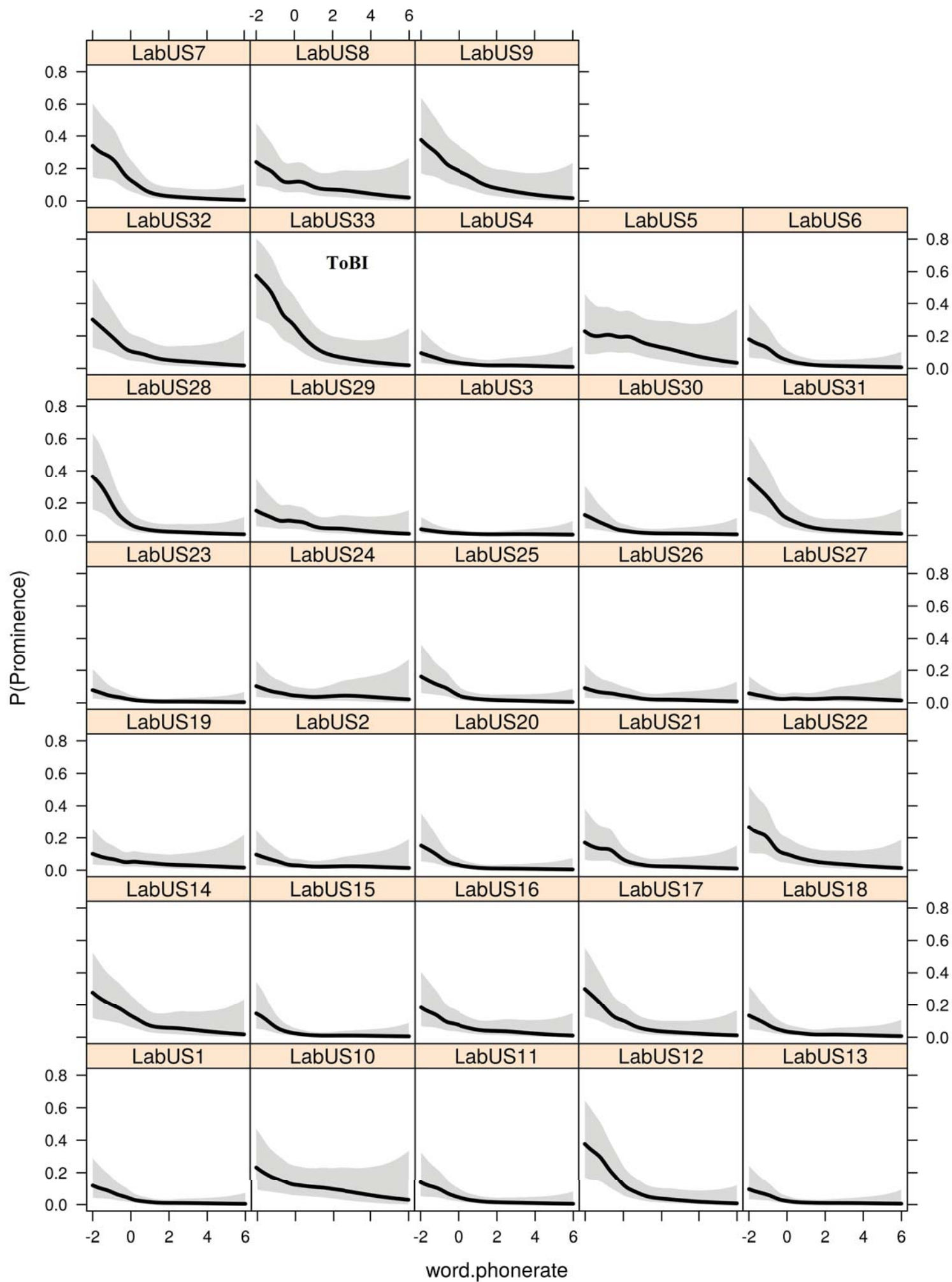
# Boundary



# Prominence\_all\_in\_one



# Prominence



# Prominence

