Hypothesis test

Felix Schönbrodt

Thu Nov 24 07:23:43 2016

```r
# plotting of results
library(ggplot2)
library(dplyr)
library(tidyr)
library(reshape2)

load("res.wide.red.RData")
res.hyp <- res.wide.red %>% select(1:8, b0_estimate, b0_p.value, skewtest_p.value, 41:45) %>% filter(!method %in% c("pcurve", "pcurve.lack

# define critical p-value for each method
res.hyp$p.crit <- .05

# merge two p-value columns into one
res.hyp$p.value <- ifelse(!is.na(res.hyp$b0_p.value), res.hyp$b0_p.value, res.hyp$skewtest_p.value)
res.hyp <- res.hyp %>% select(-b0_p.value, -skewtest_p.value)

# compute rejection:
# Reject H0 if test is significant AND estimate in correct direction.
# In case of p-curve skewness tests, there is no estimate; estimate is set to NA there.
res.hyp$H0.reject <- (res.hyp$p.value < res.hyp$p.crit) & (is.na(res.hyp$b0_estimate) | res.hyp$b0_estimate > 0)

# Add combined hypothesis test: PETPEESE + 3PSM
PP3 <- res.hyp %>% filter(method %in% c("PETPEESE.lm", "3PSM")) %>% group_by(id, condition, k, delta, qrpEnv, selProp, tau, delta.label, k
    H0.reject = H0.reject[1] & H0.reject[2] & !is.na(H0.reject[1]) & !is.na(H0.reject[2]),
    method="PP3",
    p.value = NA
)

res.hyp <- bind_rows(res.hyp, PP3)
```

```
## Warning in bind_rows_(x, .id): binding factor and character vector,
## coercing into character vector
```

```r
save(res.hyp, file="res.hyp.RData", compress="gzip")
#load(file="res.hyp.RData")

# Compute summary measures across replications
hyp.summ <- res.hyp %>% group_by(condition, k, k.label, delta, delta.label, qrpEnv, qrp.label, selProp, selProp.label, tau, tau.label, met
    H0.reject = sum(H0.reject, na.rm=TRUE)/sum(!is.na(H0.reject)),
    n.simulations = n()
)

#  loop plot for H0

#  show partly loop style
library(gtools)

# new labels for meaningful order
#hyp.summ$qrp.label <- factor(hyp.summ$qrpEnv, labels=c("QRP = 0 (none)", "QRP = 1 (med)", "QRP = 2 (high)"))

# order two variables into one loop
hyp.summ <- hyp.summ %>%  mutate(loop = paste0(tau, "_", k))

# order loop factor alphabetically
hyp.summ$loop <- factor(hyp.summ$loop, levels = mixedsort(unique(hyp.summ$loop)))

# H0
```
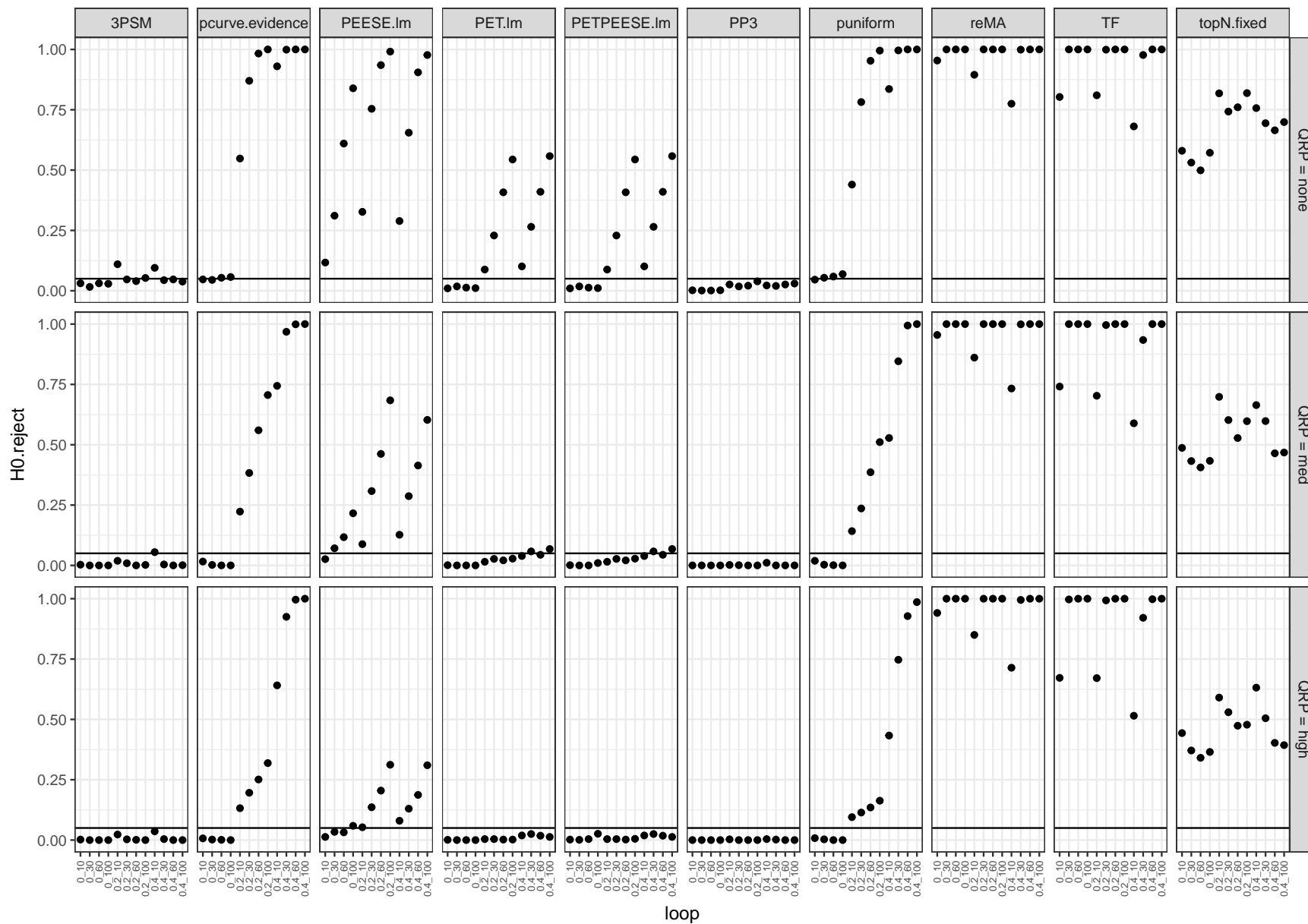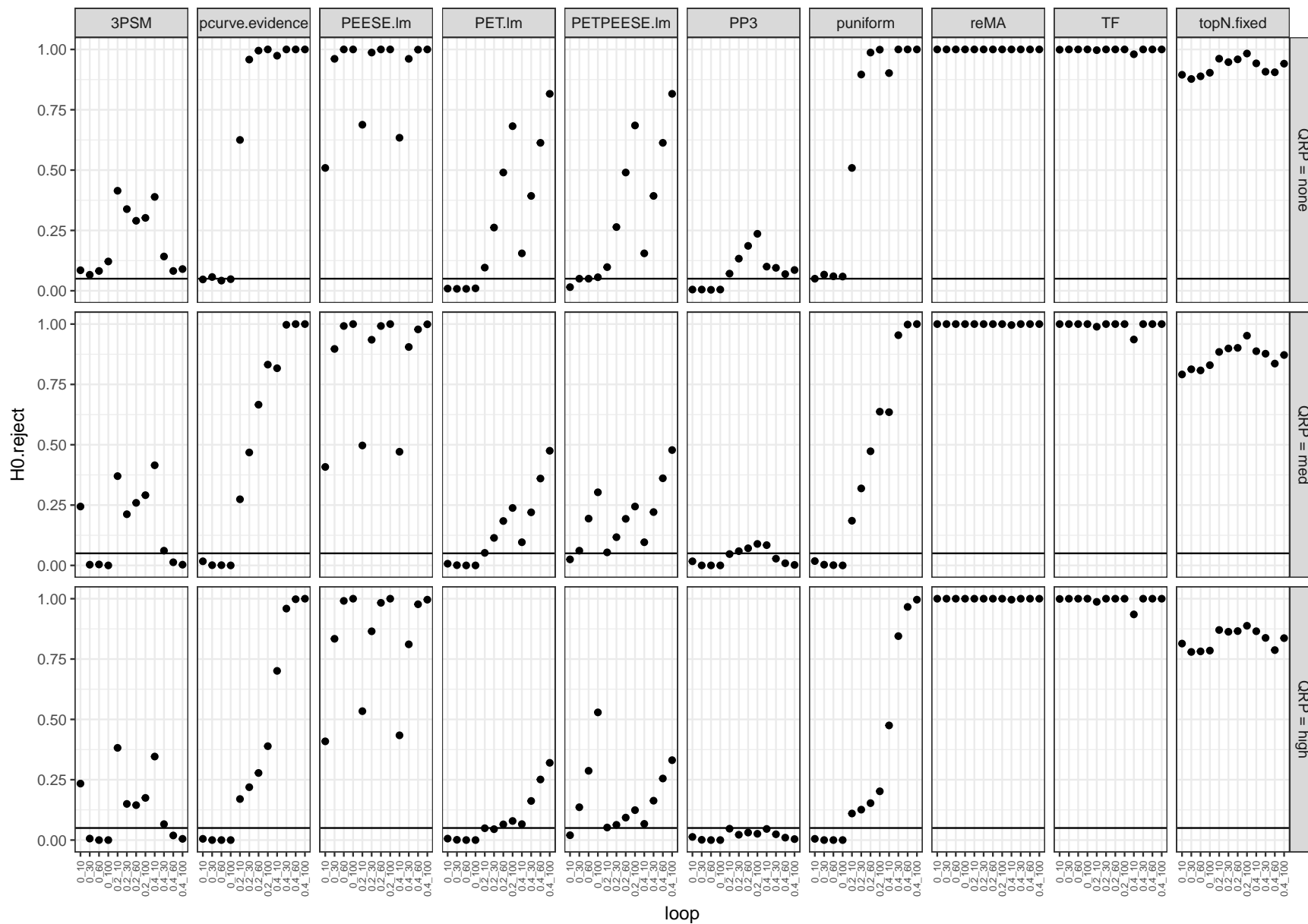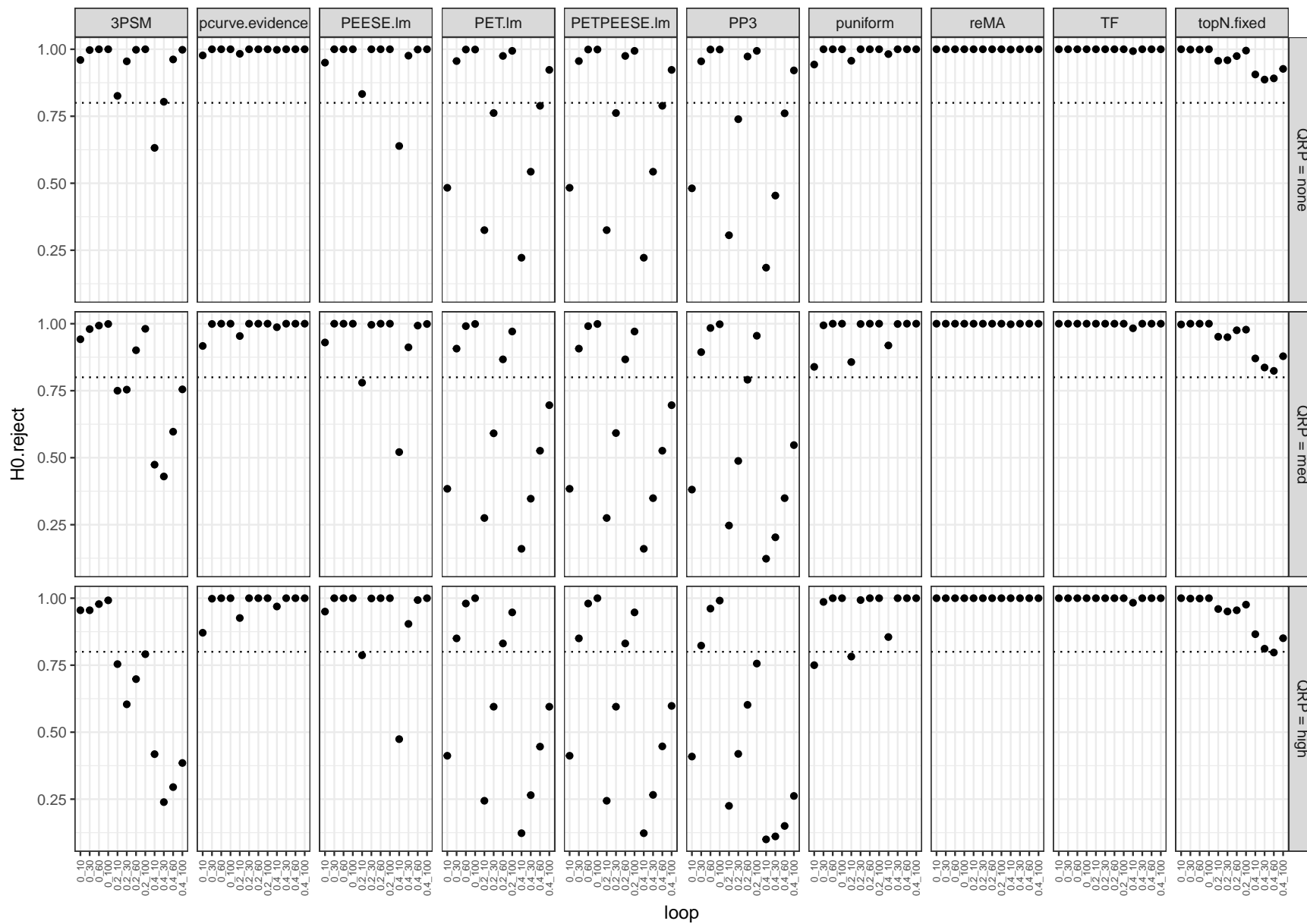
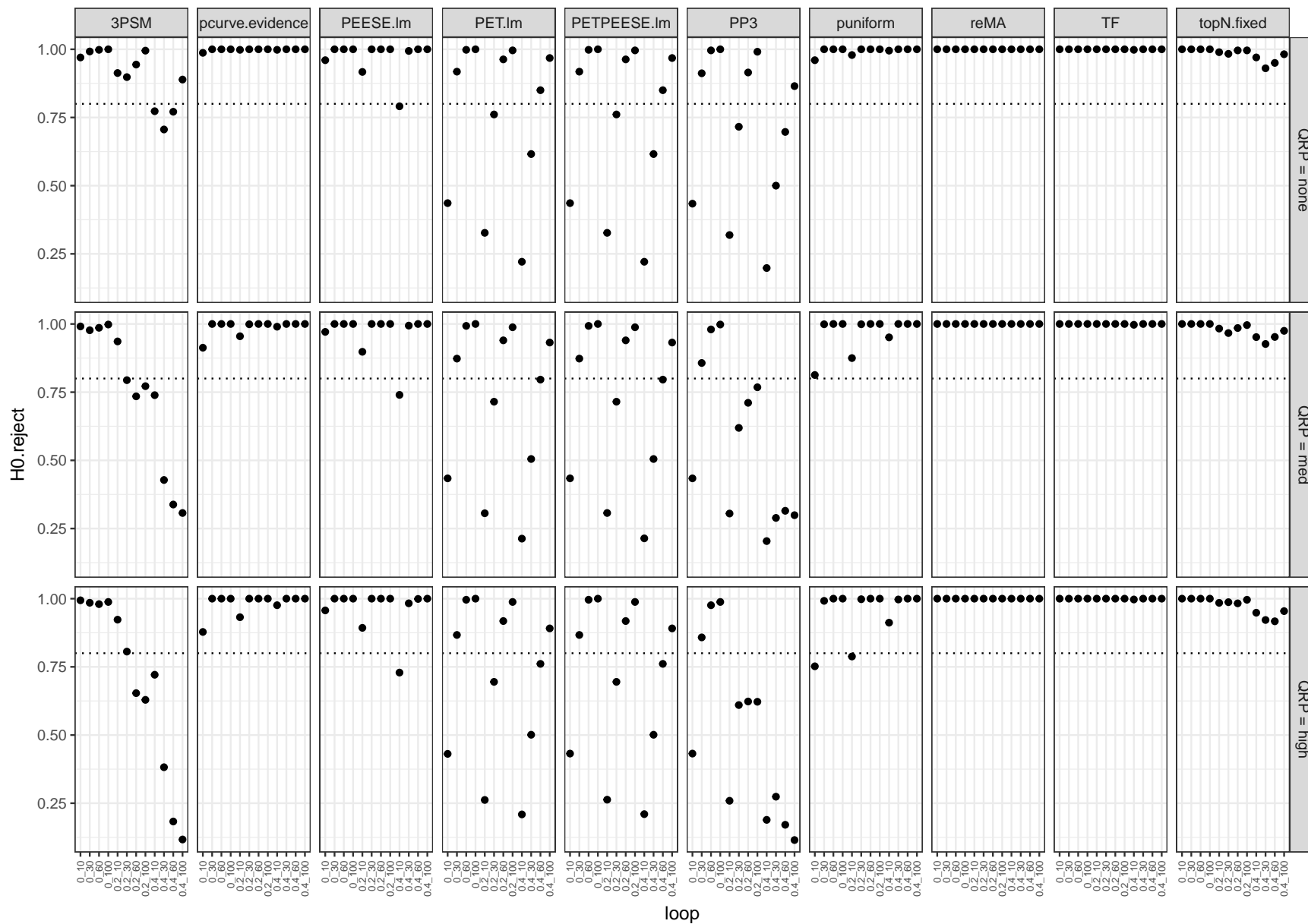Under H0 – nominal Type I error? (for selProp = 60%)

Under H0 – nominal Type I error? (for selProp = 90%)

Under H1 – Power? (for selProp = 60%, delta=0.5)

Under H1 – Power? (for selProp = 90%, delta=0.5)

```
# Compute rejection ratios

RR <- hyp.summ %>% ungroup() %>%
    select(condition, k, delta, qrp.label, selProp, tau, loop, method, H0.reject)

RR$TypeI <- RR$H0.reject
RR$TypeI[RR$delta!=0] <- NA

RR$TypeII <- 1-RR$H0.reject
RR$TypeII[RR$delta==0] <- NA
RR$Power <- 1-RR$TypeII

# compare delta==0 against delta==0.5
H1 <- 0.5
RR.H0 <- RR %>% filter(delta == 0) %>% select(condition, k, qrp.label, selProp, tau, loop, method, TypeI)
RR.H1 <- RR %>% filter(delta == H1) %>% select(k, qrp.label, selProp, tau, loop, method, Power)

glimpse(RR.H0)
```

```
## Observations: 1,078
## Variables: 8
## $ condition <int> 1, 1, 1, 1, 1, 1, 1, 1, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2...
## $ k         <dbl> 10, 10, 10, 10, 10, 10, 10, 10, 30, 30, 30, 30, 30, ...
## $ qrp.label <fctr> QRP = none, QRP = none, QRP = none, QRP = none, QRP...
## $ selProp   <dbl> 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0...
## $ tau       <dbl> 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0...
## $ loop      <fctr> 0_10, 0_10, 0_10, 0_10, 0_10, 0_10, 0_10, 0_10, 0_3...
## $ method    <chr> "3PSM", "PEESE.lm", "PET.lm", "PETPEESE.lm", "PP3", ...
## $ TypeI     <dbl> 0.00900000, 0.03000000, 0.03000000, 0.01800000, 0.00...
```

```
glimpse(RR.H1)
```

```
## Observations: 1,080
## Variables: 7
## $ k         <dbl> 10, 10, 10, 10, 10, 10, 10, 10, 10, 10, 30, 30, 30, ...
## $ qrp.label <fctr> QRP = none, QRP = none, QRP = none, QRP = none, QRP...
## $ selProp   <dbl> 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0...
## $ tau       <dbl> 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0...
## $ loop      <fctr> 0_10, 0_10, 0_10, 0_10, 0_10, 0_10, 0_10, 0_10, 0_1...
```

```
## $ method     <chr> "3PSM", "pcurve.evidence", "PEESE.lm", "PET.lm", "PE...
## $ Power      <dbl> 0.9739740, 0.9107884, 0.8950000, 0.5010000, 0.501000...
```

```
RR.wide <- inner_join(RR.H0, RR.H1)
```

```
## Joining, by = c("k", "qrp.label", "selProp", "tau", "loop", "method")
```

```
glimpse(RR.wide)
```

```
## Observations: 1,078
## Variables: 9
## $ condition <int> 1, 1, 1, 1, 1, 1, 1, 1, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2...
## $ k         <dbl> 10, 10, 10, 10, 10, 10, 10, 10, 30, 30, 30, 30, 30, ...
## $ qrp.label <fctr> QRP = none, QRP = none, QRP = none, QRP = none, QRP...
## $ selProp   <dbl> 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0...
## $ tau       <dbl> 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0...
## $ loop      <fctr> 0_10, 0_10, 0_10, 0_10, 0_10, 0_10, 0_10, 0_10, 0_3...
## $ method    <chr> "3PSM", "PEESE.lm", "PET.lm", "PETPEESE.lm", "PP3", ...
## $ TypeI     <dbl> 0.00900000, 0.03000000, 0.03000000, 0.01800000, 0.00...
## $ Power     <dbl> 0.9739740, 0.8950000, 0.5010000, 0.5010000, 0.495000...
```
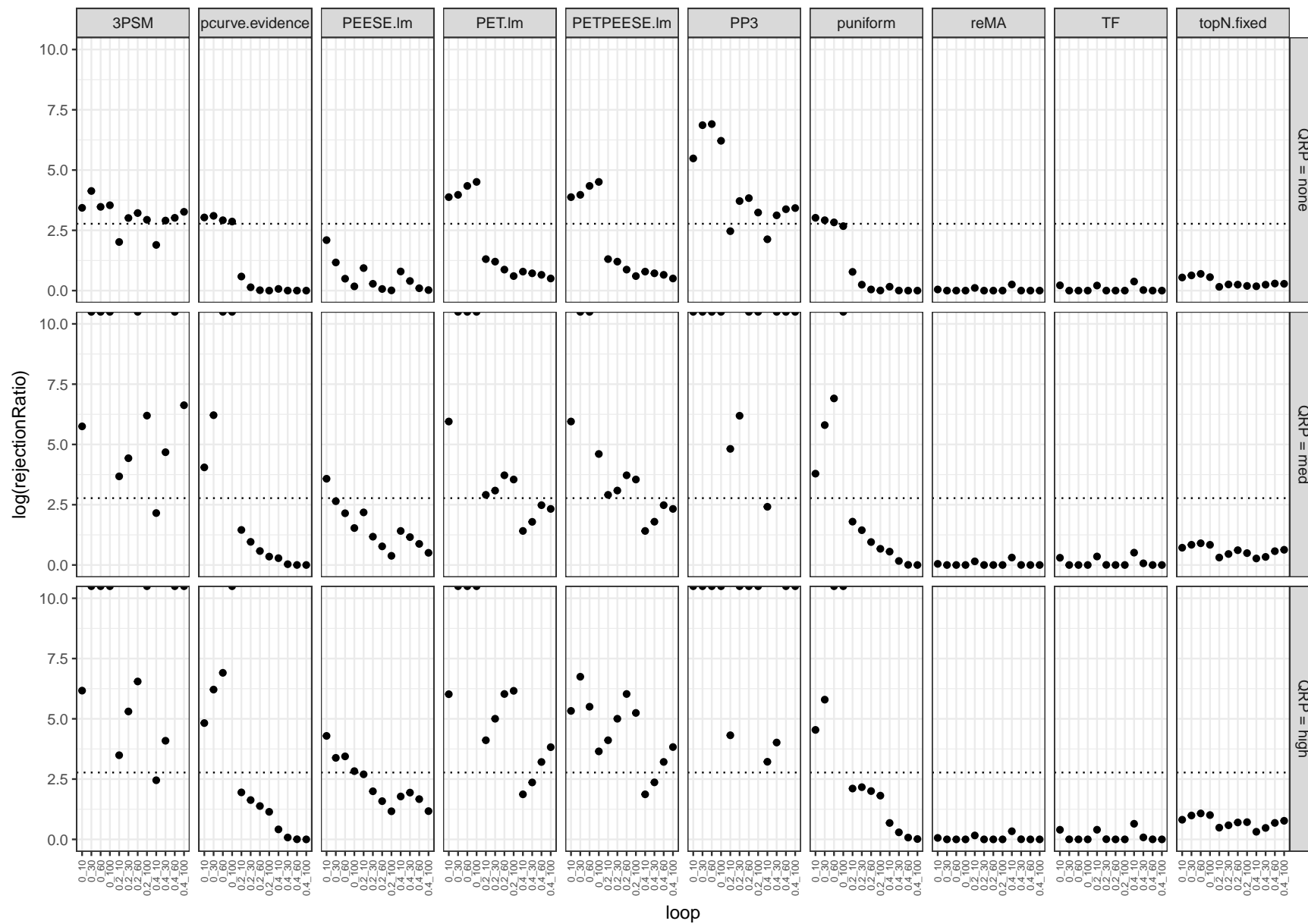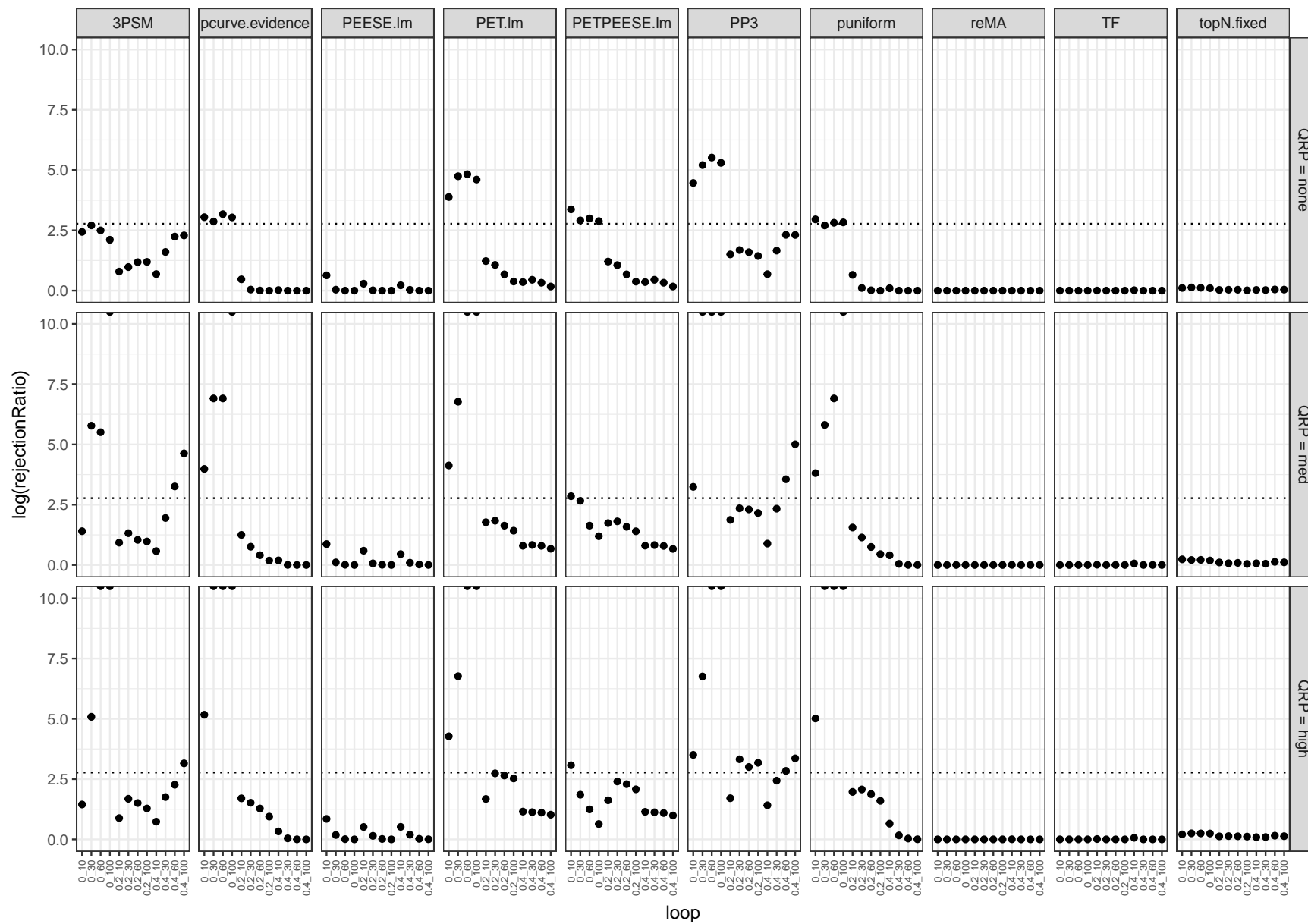
```
RR.wide$rejectionRatio <- RR.wide$Power/RR.wide$TypeI
```

Rejection Ratio (for selProp = 60%, H0 against delta=0.5)

Rejection Ratio (for selProp = 90%, H0 against delta=0.5)

```r
# Rate of significant results in the wrong direction

wrongSig <- res.wide.red %>% select(1:8, b0_estimate, b0_p.value) %>%
    filter(!method %in% c("pcurve.evidence", "pcurve.lack"), delta < 0.5, !is.na(b0_estimate), !is.na(b0_p.value)) %>%
    group_by(condition, k, delta, qrpEnv, selProp, tau, method) %>%
    summarise(
        wrongSig = sum((b0_estimate < 0) & (b0_p.value < .05)) / n()
    ) %>%
    mutate(loop = paste0(k, "_", tau))

# order loop factor alphabetically
wrongSig$loop <- factor(wrongSig$loop, levels = mixedsort(unique(wrongSig$loop)))
```

Percentage of significant estimates in the *wrong* direction