Default Tips Part 1: Optimal RDD estimations

1. Loading

We use the `optrdd` package to estimate the the fare's discontinuity policy effect using convex optimization methods. Below are the packages used:

We load pre-cleaned data with fares between 12 and 18 US dollars; the threshold is 15 dollars.

```
# Clear the data environment
rm(list = ls())
# Read in data
user = Sys.info()[["user"]]
if (user == "ellamao"){ ## Ella's path
  data <- read.csv("/Users/ellamao/Dropbox/Default Tips Project/Data/Intermediate/fare_1218_recoded.csv
}
if (user == "51989"){ ## Bruno's PC path
  data = read.csv(paste0("C:/Users/51989/OneDrive/Escritorio/Dropbox/Default Tips Project/Data/Intermedi
}

n = nrow(data)
threshold = 15
X = data$fare
W = as.numeric(X >= threshold)
Y = data$tip_zero
```

2. Estimation

Below we estimate the optimal RDD models for values of B between 0.025 and 0.3

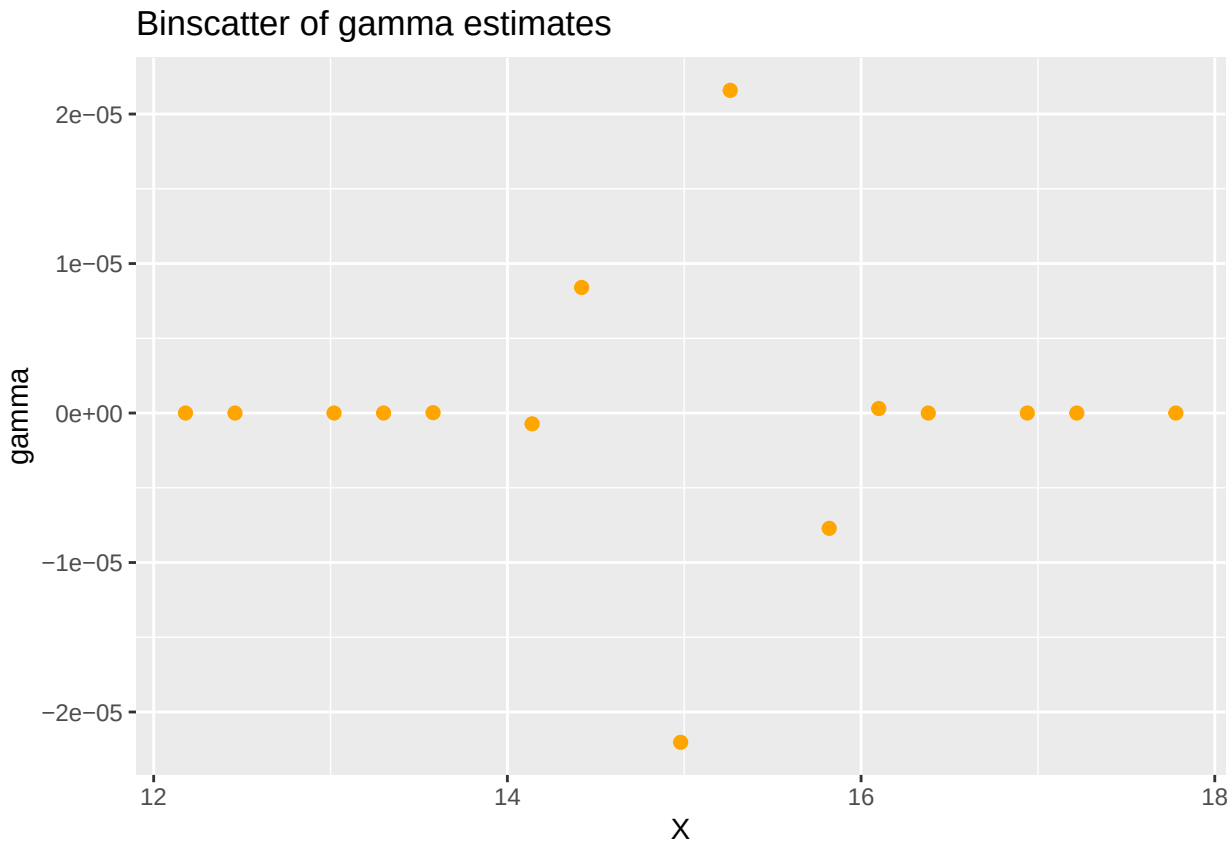
```
set.seed(1234)
ind = (X<18 & X>12) # Keeping those with fare between 14 and 16

for (i in seq(0.025, 0.3, by = 0.025)){
    # optrdd estimation
    out = optrdd(X=X[ind], Y=Y[ind], W=W[ind],
                 max.second.derivative = i,
                 estimation.point = NULL,
                 optimizer = "quadprog",
                 try.elnet.for.sigma.sq = TRUE )
    #plot.gamma = plot(out,main = paste0("max.second.derivative = ", i)) # Keeping this as comment sinc

    # store results
    tau.hat = out$tau.hat
    tau.ul = out$tau.hat + out$tau.plusminus
    tau.ll = out$tau.hat - out$tau.plusminus
    b = i
    est <- cbind(tau.hat, tau.ul, tau.ll, b)
    if (i == 0.025) {
        est.matrix <- est
```

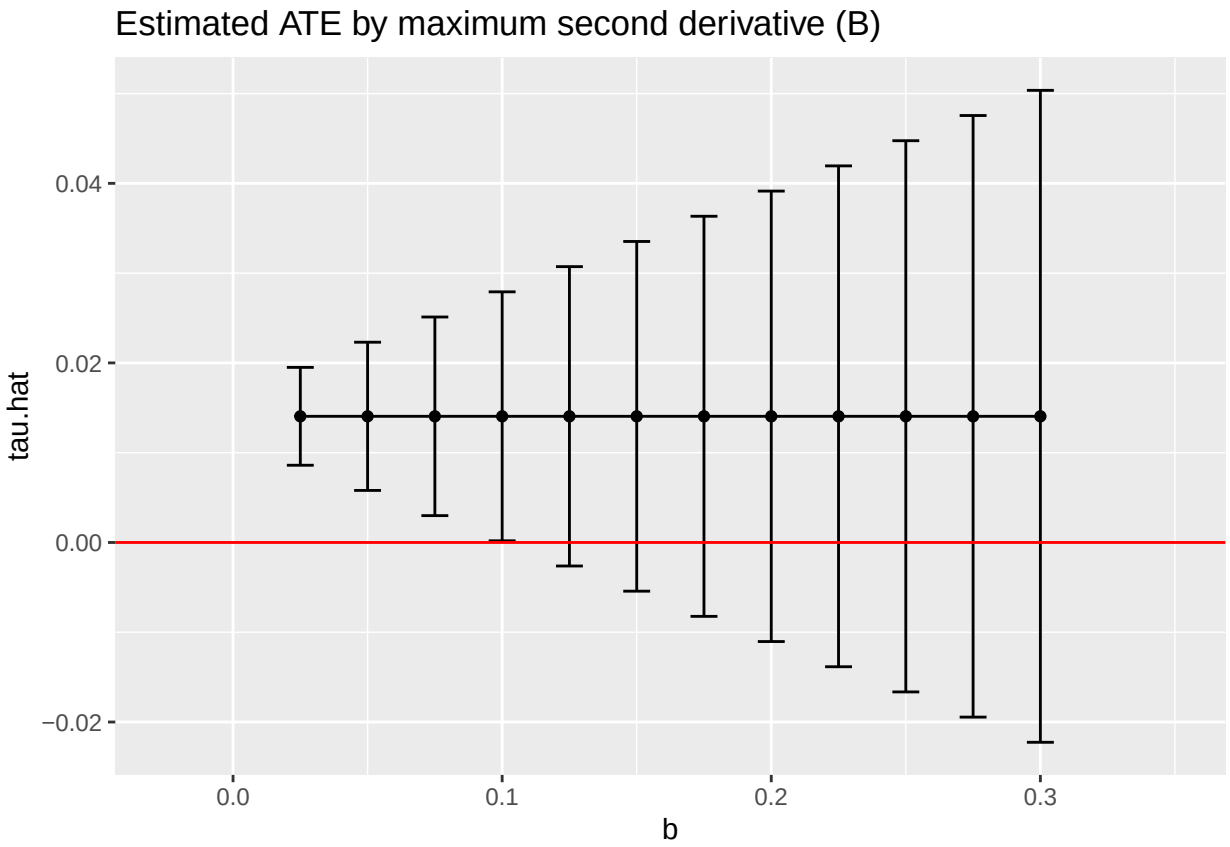
```
        gamma.df = data.frame(gamma=out$gamma, X=X[ind], Y=Y[ind], b=i)
    }
    else{
        if (i == 0.1 | i == 0.2 | i == 0.3){
          df = data.frame(gamma=out$gamma, X=X[ind], Y=Y[ind], b=i)
          gamma.df = rbind(gamma.df,df)
        }
        est.matrix <- rbind(est.matrix,est)
    }
    #plot_list[[i]] = p
}
```

```
## [1] "Running quadrprog with problem of size: 4015 x 59..."
## [1] "Running quadrprog with problem of size: 4015 x 59..."
## [1] "Running quadrprog with problem of size: 4015 x 59..."
## [1] "Running quadrprog with problem of size: 4015 x 59..."
## [1] "Running quadrprog with problem of size: 4015 x 59..."
## [1] "Running quadrprog with problem of size: 4015 x 59..."
## [1] "Running quadrprog with problem of size: 4015 x 59..."
## [1] "Running quadrprog with problem of size: 4015 x 59..."
## [1] "Running quadrprog with problem of size: 4015 x 59..."
## [1] "Running quadrprog with problem of size: 4015 x 59..."
## [1] "Running quadrprog with problem of size: 4015 x 59..."
## [1] "Running quadrprog with problem of size: 4015 x 59..."
```

```
# if we want to produce one figure with various panels
#par(mfrow=c(1,2))
#plot(out.0.1,main = paste0("B = ", i))
```

We plot the estimates of gamma for last estimation (B = 0.3), together with the estimates of all the tau estimates and their 95% confidence intervals.


Binscatter of gamma estimates
gamma
2e−05
1e−05
0e+00
−1e−05
−2e−05
12
14
16
18
X



3. Questions for Stefan

1) Is it OK if we only use the `quadprog` optimizer?

2) Why do we need `try.elnet.for.sigma.sq = TRUE`?

3) The paper's main specification uses multiple fixed effects. Does `optrdd` need this FE for correct identification?

4.1) How should we sensibly set `max.second.derivative`?

4.2) Why are we getting same tau.hat across B? We have noticed this does not happen when we use smaller random samples.