Project 4

```
heat <- read_csv("heating_data.csv")
```

```
## Warning: Missing column names filled in: 'X1' [1]
```

```
## Parsed with column specification:
## cols(
##   X1 = col_double(),
##   consumer_id = col_double(),
##   choice = col_character(),
##   ic_gc = col_double(),
##   ic_gr = col_double(),
##   ic_ec = col_double(),
##   ic_er = col_double(),
##   ic_hp = col_double(),
##   oc_gc = col_double(),
##   oc_gr = col_double(),
##   oc_ec = col_double(),
##   oc_er = col_double(),
##   oc_hp = col_double(),
##   income = col_double(),
##   age = col_double(),
##   rooms = col_double(),
##   region = col_character()
## )
```

```
heat %>% summarise_if(is.numeric, mean)
```

```
## # A tibble: 1 x 15
##      X1 consumer_id ic_gc ic_gr ic_ec ic_er ic_hp oc_gc oc_gr oc_ec oc_er
##   <dbl>       <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl>
## 1  450.        450.  777.  922.  825.  984. 1046.  172.  154.  477.  430.
## # ... with 4 more variables: oc_hp <dbl>, income <dbl>, age <dbl>,
## #   rooms <dbl>
```

```
prop.table(table(heat$choice))
```

```
##
##         ec         er         gc         gr         hp
## 0.07111111 0.09333333 0.63666667 0.14333333 0.05555556
```

7.1 % choose ec, 9.3 % choose er, 63.67 % choose gc, 14.33% choose gr, and 5.55 % choose hp

Multinomial logit model of demand ver.1

IncludE only installation cost and operating cost as product characteristics.

```
beta <- c(.1,.1)
multi_logit <- function(beta){
  beta1 = beta[1]
  beta2 = beta[2]
  heat <- heat %>% mutate(v_igc = beta1*ic_gc+beta2*oc_gc,
                          v_igr = beta1*ic_gr+beta2*oc_gr,
                          v_iec = beta1*ic_ec+beta2*oc_ec,
                          v_ier = beta1*ic_er+beta2*oc_er,
```

```
                          v_ihp = beta1*ic_hp+beta2*oc_hp)


  heat <- heat %>% mutate(p_igc=exp(v_igc)/(exp(v_igc)+exp(v_igr)+exp(v_iec)+exp(v_ier)+exp(v_ihp)),
                          p_igr = exp(v_igr)/(exp(v_igc)+exp(v_igr)+exp(v_iec)+exp(v_ier)+exp(v_ihp)),
                          p_iec = exp(v_iec)/(exp(v_igc)+exp(v_igr)+exp(v_iec)+exp(v_ier)+exp(v_ihp)),
                          p_ier = exp(v_ier)/(exp(v_igc)+exp(v_igr)+exp(v_iec)+exp(v_ier)+exp(v_ihp)),
                          p_ihp = exp(v_ihp)/(exp(v_igc)+exp(v_igr)+exp(v_iec)+exp(v_ier)+exp(v_ihp)))

  heat <- heat %>% mutate(pbeta = ifelse(choice == "gc", p_igc, ifelse(choice=="gr",p_igr,ifelse(choice=

  heat$pbeta <- as.numeric(heat$pbeta)

  lbeta = sum(log(heat$pbeta))
  return(-lbeta)
}
optim(beta,multi_logit)
```

```
## $par
## [1] -0.006231293 -0.004574586
##
## $value
## [1] 1095.237
##
## $counts
## function gradient
##       77       NA
##
## $convergence
## [1] 0
##
## $message
## NULL
```

β_1 is -0.006231 and β_2 is -0.004754.

If installation cost increases by 1 dollar, the operation cost decreases by 1.36 dollar, and this seems to be reasonable.

```
# own elasticities

  beta1 = -0.006231293
  beta2 = -0.004574586
  heat <- heat %>% mutate(v_igc = beta1*ic_gc+beta2*oc_gc,
                          v_igr = beta1*ic_gr+beta2*oc_gr,
                          v_iec = beta1*ic_ec+beta2*oc_ec,
                          v_ier = beta1*ic_er+beta2*oc_er,
                          v_ihp = beta1*ic_hp+beta2*oc_hp)


  heat <- heat %>% mutate(p_igc=exp(v_igc)/(exp(v_igc)+exp(v_igr)+exp(v_iec)+exp(v_ier)+exp(v_ihp)),
                          p_igr = exp(v_igr)/(exp(v_igc)+exp(v_igr)+exp(v_iec)+exp(v_ier)+exp(v_ihp)),
                          p_iec = exp(v_iec)/(exp(v_igc)+exp(v_igr)+exp(v_iec)+exp(v_ier)+exp(v_ihp)),
                          p_ier = exp(v_ier)/(exp(v_igc)+exp(v_igr)+exp(v_iec)+exp(v_ier)+exp(v_ihp)),
                          p_ihp = exp(v_ihp)/(exp(v_igc)+exp(v_igr)+exp(v_iec)+exp(v_ier)+exp(v_ihp)))
```

```
  heat <- heat %>% mutate(e_gcic = beta1*(1-p_igc)*ic_gc,
                          e_gcoc = beta2*(1-p_igc)*oc_gc,
                          e_gric = beta1*(1-p_igr)*ic_gr,
                          e_groc = beta2*(1-p_igr)*oc_gr,
                          e_ecic = beta1*(1-p_iec)*ic_ec,
                          e_ecoc = beta2*(1-p_iec)*oc_ec,
                          e_eric = beta1*(1-p_ier)*ic_er,
                          e_eroc = beta2*(1-p_ier)*oc_er,
                          e_hpic = beta1*(1-p_ihp)*ic_hp,
                          e_hpoc = beta2*(1-p_ihp)*oc_hp,
                          )
heat %>% summarise_if(is.numeric, mean)
```

```
## # A tibble: 1 x 35
##      X1 consumer_id ic_gc ic_gr ic_ec ic_er ic_hp oc_gc oc_gr oc_ec oc_er
##   <dbl>       <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl>
## 1  450.        450.  777.  922.  825.  984. 1046.  172.  154.  477.  430.
## # ... with 24 more variables: oc_hp <dbl>, income <dbl>, age <dbl>,
## #   rooms <dbl>, v_igc <dbl>, v_igr <dbl>, v_iec <dbl>, v_ier <dbl>,
## #   v_ihp <dbl>, p_igc <dbl>, p_igr <dbl>, p_iec <dbl>, p_ier <dbl>,
## #   p_ihp <dbl>, e_gcic <dbl>, e_gcoc <dbl>, e_gric <dbl>, e_groc <dbl>,
## #   e_ecic <dbl>, e_ecoc <dbl>, e_eric <dbl>, e_eroc <dbl>, e_hpic <dbl>,
## #   e_hpoc <dbl>
```

own elasticity gc for installation cost: -2.35, operation cost: -0.38 own elasticity gr for installation cost: -4.40, operation cost: -0.54 own elasticity ec for installation cost: -4.63, operation cost: -1.96 own elasticity er for installation cost: -5.84, operation cost: -1.87 own elasticity hp for installation cost: -5.98, operation cost: -0.92

Consumers are relatively price elastic toward installtion costs and are relatively price inelastic toward operation costs.

Multinomial logit model ver.2: with product fixed effects

```
params <- c(.1,.1,.1,.1,.1,.1)
multi_logit2 <- function(params){
  delta1 = params[1]
  delta2 = params[2]
  delta3 = params[3]
  delta4 = params[4]
  beta1 = params[5]
  beta2 = params[6]
  heat <- heat %>% mutate(v_igc = delta1 + beta1*ic_gc+beta2*oc_gc,
                          v_igr = delta2 + beta1*ic_gr+beta2*oc_gr,
                          v_iec = delta3 + beta1*ic_ec+beta2*oc_ec,
                          v_ier = delta4 + beta1*ic_er+beta2*oc_er,
                          v_ihp = beta1*ic_hp+beta2*oc_hp)


  heat <- heat %>% mutate(p_igc=exp(v_igc)/(exp(v_igc)+exp(v_igr)+exp(v_iec)+exp(v_ier)+exp(v_ihp)),
                          p_igr = exp(v_igr)/(exp(v_igc)+exp(v_igr)+exp(v_iec)+exp(v_ier)+exp(v_ihp)),
                          p_iec = exp(v_iec)/(exp(v_igc)+exp(v_igr)+exp(v_iec)+exp(v_ier)+exp(v_ihp)),
                          p_ier = exp(v_ier)/(exp(v_igc)+exp(v_igr)+exp(v_iec)+exp(v_ier)+exp(v_ihp)),
                          p_ihp = exp(v_ihp)/(exp(v_igc)+exp(v_igr)+exp(v_iec)+exp(v_ier)+exp(v_ihp)))
```

```
  heat <- heat %>% mutate(pbeta = ifelse(choice == "gc", p_igc, ifelse(choice=="gr",p_igr,ifelse(choice=

  heat$pbeta <- as.numeric(heat$pbeta)

  lbeta = sum(log(heat$pbeta))
  return(-lbeta)
}
optim(params,multi_logit2)
```

```
## $par
## [1]  0.833463275 -0.418421176  0.913483138  1.477046944 -0.003312859
## [6] -0.007490285
##
## $value
## [1] 1017.651
##
## $counts
## function gradient
##      469       NA
##
## $convergence
## [1] 0
##
## $message
## NULL
```

Althuogh both cost coefficients (the last two numbers in the ouput) are negative, which are consistent with the previous estimates, their difference in magnitude still indicates that they are meaningfully different. Fixed effects are the first four numbers in the set of outputs. In the ranking of preferences from high to low are: electric room, electric central, gas central, heat pump, and gas room.

Multinomial logit model ver.3: include interactions between income and the two cost variables.

```
params2 <- c(.1,.1,.1,.1,.1,.1,.01,.01)
multi_logit3 <- function(params){
  delta1 = params[1]
  delta2 = params[2]
  delta3 = params[3]
  delta4 = params[4]
  beta1 = params[5]
  beta2 = params[6]
  beta3 = params[7]
  beta4 = params[8]

  heat <- heat %>% mutate(v_igc = delta1 + beta1*ic_gc + beta2*oc_gc + beta3*ic_gc*income + beta4*oc_gc
                          v_igr = delta2 + beta1*ic_gr+beta2*oc_gr + beta3*ic_gr*income + beta4*oc_gr*i
                          v_iec = delta3 + beta1*ic_ec+beta2*oc_ec + beta3*ic_ec*income + beta4*oc_ec*i
                          v_ier = delta4 + beta1*ic_er+beta2*oc_er + beta3*ic_er*income + beta4*oc_er*i
                          v_ihp = beta1*ic_hp + beta2*oc_hp + beta3*ic_hp*income + beta4*oc_hp*income)


  heat <- heat %>% mutate(p_igc=exp(v_igc)/(exp(v_igc)+exp(v_igr)+exp(v_iec)+exp(v_ier)+exp(v_ihp)),
```

```
                          p_igr = exp(v_igr)/(exp(v_igc)+exp(v_igr)+exp(v_iec)+exp(v_ier)+exp(v_ihp)),
                          p_iec = exp(v_iec)/(exp(v_igc)+exp(v_igr)+exp(v_iec)+exp(v_ier)+exp(v_ihp)),
                          p_ier = exp(v_ier)/(exp(v_igc)+exp(v_igr)+exp(v_iec)+exp(v_ier)+exp(v_ihp)),
                          p_ihp = exp(v_ihp)/(exp(v_igc)+exp(v_igr)+exp(v_iec)+exp(v_ier)+exp(v_ihp)))

  heat <- heat %>% mutate(pbeta = ifelse(choice == "gc", p_igc, ifelse(choice=="gr",p_igr,ifelse(choice=

  heat$pbeta <- as.numeric(heat$pbeta)

  lbeta = sum(log(heat$pbeta))
  return(-lbeta)
}
optim(params2,multi_logit3)
```

```
## $par
## [1] -1.254721e-01  2.435476e-01  4.160495e-01  1.661204e-02 -5.962660e-03
## [6] -5.086794e-03 -2.434778e-04 -6.843212e-05
##
## $value
## [1] 1152.604
##
## $counts
## function gradient
##      501       NA
##
## $convergence
## [1] 1
##
## $message
## NULL
```

Using the estimates with the highest income (7) and the lowest income (2), the difference is 1.12, so I would say that the price is not sensitive relative to income. This doesn't seem to be very reasonable.

```
-1.254721e-01 + (-5.962660e-03*heat$ic_gc[1]) + (-5.086794e-03*heat$oc_gc[1]) + (-0.001298310*heat$ic_g
```

```
## [1] -9.406195
```

```
-1.254721e-01 + (-5.962660e-03*heat$ic_gc[1]) + (-5.086794e-03*heat$oc_gc[1]) + (-0.001298310*heat$ic_g
```

```
## [1] -17.15939
```

Multinomial logit model ver.4: include interactions between income and the product fixed effects.

fixed effect + interactions between income and the two cost variables + interactions between income and fixed effect

```
params2 <- c(.1,.1,.1,.1,.1,.01,.01)
multi_logit4 <- function(params){
  delta1 = params[1]
  delta2 = params[2]
  delta3 = params[3]
  delta4 = params[4]
  beta1 = params[5]
  beta2 = params[6]
```

```
  beta3 = params[7]
  beta4 = params[8]

  heat <- heat %>% mutate(v_igc = delta1 + beta1*ic_gc + beta2*oc_gc + beta3*ic_gc*income + beta4*oc_gc
                          v_igr = delta2 + beta1*ic_gr+beta2*oc_gr + beta3*ic_gr*income + beta4*oc_gr*i
                          v_iec = delta3 + beta1*ic_ec+beta2*oc_ec + beta3*ic_ec*income + beta4*oc_ec*i
                          v_ier = delta4 + beta1*ic_er+beta2*oc_er + beta3*ic_er*income + beta4*oc_er*i
                          v_ihp = beta1*ic_hp + beta2*oc_hp + beta3*ic_hp*income + beta4*oc_hp*income)


  heat <- heat %>% mutate(p_igc=exp(v_igc)/(exp(v_igc)+exp(v_igr)+exp(v_iec)+exp(v_ier)+exp(v_ihp)),
                          p_igr = exp(v_igr)/(exp(v_igc)+exp(v_igr)+exp(v_iec)+exp(v_ier)+exp(v_ihp)),
                          p_iec = exp(v_iec)/(exp(v_igc)+exp(v_igr)+exp(v_iec)+exp(v_ier)+exp(v_ihp)),
                          p_ier = exp(v_ier)/(exp(v_igc)+exp(v_igr)+exp(v_iec)+exp(v_ier)+exp(v_ihp)),
                          p_ihp = exp(v_ihp)/(exp(v_igc)+exp(v_igr)+exp(v_iec)+exp(v_ier)+exp(v_ihp)))

  heat <- heat %>% mutate(pbeta = ifelse(choice == "gc", p_igc, ifelse(choice=="gr",p_igr,ifelse(choice

  heat$pbeta <- as.numeric(heat$pbeta)

  lbeta = sum(log(heat$pbeta))
  return(-lbeta)
}
optim(params2,multi_logit4)
```

```
## $par
## [1] -0.118131045  0.094423428  0.441345882  0.128357934 -0.003543583
## [6] -0.002073862 -0.001298310 -0.002134820
##
## $value
## [1] 1421.562
##
## $counts
## function gradient
##      501       NA
##
## $convergence
## [1] 1
##
## $message
## NULL
```

Using the estimates with the highest income (7) and the lowest income (2), the difference is 7.7532, so I would say that the price is sensitive relative to income, which seems to be reasonable. In the ranking of preferences from high to low are: electric central, electric room, gas room, heat pump, and gas central.

A nested logit version of the previous model

Groups systems into nests based upon whether they are powered by gas or electricity (heat pumps are electric)/a central air system or a room system .

```
# nested logit model grouped by electricity or gas
# gas or electricity. gas : gc,gr; electric: ec,er,hp
params3 <- c(.001,.001,.001,.001,.001,.001,.001,.001,.1,.1)
multi_logit5 <- function(params){
```

```
  delta1 = params[1]
  delta2 = params[2]
  delta3 = params[3]
  delta4 = params[4]
  beta1 = params[5]
  beta2 = params[6]
  beta3 = params[7]
  beta4 = params[8]
  lambda_e = params[9]
  lambda_g = params[10]

  heat <- heat %>% mutate(v_igc = delta1 + beta1*ic_gc + beta2*oc_gc + beta3*ic_gc*income + beta4*oc_gc
                          v_igr = delta2 + beta1*ic_gr+beta2*oc_gr + beta3*ic_gr*income + beta4*oc_gr*i
                          v_iec = delta3 + beta1*ic_ec+beta2*oc_ec + beta3*ic_ec*income + beta4*oc_ec*i
                          v_ier = delta4 + beta1*ic_er+beta2*oc_er + beta3*ic_er*income + beta4*oc_er*i
                          v_ihp = beta1*ic_hp + beta2*oc_hp + beta3*ic_hp*income + beta4*oc_hp*income)




  heat <- heat %>% mutate(p_igc=exp(v_igc/lambda_g)*(exp(v_igc/lambda_g)+exp(v_igr/lambda_g))^(lambda_g
                          p_igr = exp(v_igr/lambda_g)*(exp(v_igc/lambda_g)+exp(v_igr/lambda_g))^(lambda
                          p_iec = exp(v_iec/lambda_e)*(exp(v_iec/lambda_e)+exp(v_ier/lambda_e)+exp(v_ih
                          p_ier = exp(v_ier/lambda_e)*(exp(v_iec/lambda_e)+exp(v_ier/lambda_e)+exp(v_ih
                          p_ihp = exp(v_ihp/lambda_e)*(exp(v_iec/lambda_e)+exp(v_ier/lambda_e)+exp(v_ih

  heat <- heat %>% mutate(pbeta = ifelse(choice == "gc", p_igc, ifelse(choice=="gr",p_igr,ifelse(choice=

  heat$pbeta <- as.numeric(heat$pbeta)

  lbeta = sum(log(heat$pbeta))
  return(-lbeta)
}
optim(params3,multi_logit5)
```

```
## $par
##  [1] -0.0285340528  0.0085508450 -0.0411987586 -0.0317232995  0.0470225962
##  [6]  0.0186383528 -0.0001085649  0.0057204683  0.1179768378  0.1016437857
##
## $value
## [1] -399874.6
##
## $counts
## function gradient
##      501       NA
##
## $convergence
## [1] 1
##
## $message
## NULL
```

```
# nested logit model grouped by central or room
# central : gc,ec,hp; room: gr,er
```

```r
params4 <- c(.001,.001,.001,.001,.001,.001,.001,.001,.1,.1)
multi_logit6 <- function(params){
  delta1 = params[1]
  delta2 = params[2]
  delta3 = params[3]
  delta4 = params[4]
  beta1 = params[5]
  beta2 = params[6]
  beta3 = params[7]
  beta4 = params[8]
  lambda_c = params[9]
  lambda_r = params[10]

  heat <- heat %>% mutate(v_igc = delta1 + beta1*ic_gc + beta2*oc_gc + beta3*ic_gc*income + beta4*oc_gc
                          v_igr = delta2 + beta1*ic_gr+beta2*oc_gr + beta3*ic_gr*income + beta4*oc_gr*i
                          v_iec = delta3 + beta1*ic_ec+beta2*oc_ec + beta3*ic_ec*income + beta4*oc_ec*i
                          v_ier = delta4 + beta1*ic_er+beta2*oc_er + beta3*ic_er*income + beta4*oc_er*i
                          v_ihp = beta1*ic_hp + beta2*oc_hp + beta3*ic_hp*income + beta4*oc_hp*income)




  heat <- heat %>% mutate(p_igc=exp(v_igc/lambda_c)*(exp(v_igc/lambda_c)+exp(v_iec/lambda_c)+exp(v_ihp/
                          p_igr = exp(v_igr/lambda_r)*(exp(v_igr/lambda_r)+exp(v_ier/lambda_r))^(lambda
                          p_iec = exp(v_iec/lambda_c)*(exp(v_igc/lambda_c)+exp(v_iec/lambda_c)+exp(v_ih
                          p_ier = exp(v_ier/lambda_r)*(exp(v_igr/lambda_r)+exp(v_ier/lambda_r))^(lambda
                          p_ihp = exp(v_ihp/lambda_c)*(exp(v_igc/lambda_c)+exp(v_iec/lambda_c)+exp(v_ih

  heat <- heat %>% mutate(pbeta = ifelse(choice == "gc", p_igc, ifelse(choice=="gr",p_igr,ifelse(choice

  heat$pbeta <- as.numeric(heat$pbeta)

  lbeta = sum(log(heat$pbeta))
  return(-lbeta)
}
optim(params4,multi_logit6)
```

```
## $par
##  [1]  0.008076744 -0.004625451 -0.002649738  0.005381268  0.030705520
##  [6]  0.007948802  0.001096232  0.002256536  0.090671353  0.087881891
##
## $value
## [1] -326551.2
##
## $counts
## function gradient
##      501       NA
##
## $convergence
## [1] 1
##
## $message
## NULL
```

```
optim(params3,multi_logit5)$par
```

```
##  [1] -0.0285340528  0.0085508450 -0.0411987586 -0.0317232995  0.0470225962
##  [6]  0.0186383528 -0.0001085649  0.0057204683  0.1179768378  0.1016437857
```

```
optim(params4,multi_logit6)$par
```

```
##  [1]  0.008076744 -0.004625451 -0.002649738  0.005381268  0.030705520
##  [6]  0.007948802  0.001096232  0.002256536  0.090671353  0.087881891
```

Comparing the estimates for the 10 parameters given by the two nested logit model, I would say that nesting decisions do impact my results. While some are relatively close in magnitude, some estimates change their signs and some others have significantly different values.