Identification of Heterogeneous Causal Effects

Falco J. Bargagli-Stoffi

14/12/2020

Identifying the Causal Effects

In this section we estimate the causal effects of additional funding on students performances. Namely, our outcome variables are `certificate` (which is a dummy variable that assumes value 1 if the student gets an "A" certificate: the highest possible grade in the Flanders) and `progress school` (which is a dummy recording whether or not the student got school retention).

The following analyses were performed using the algorithm proposed in the paper "Heterogeneous causal effects with imperfect compliance: a novel Bayesian machine learning approach" by Falco J. Bargagli Stoffi, Kristof De Witte and Giorgio Gnecco.

First Outcome: Certificate

We now focus on the effects when we sample units in a bandwidth of 0.035 (which is the optimal bandwidth for the outcome *certificate*) around the threshold (10%).

```
students_data_randomized_03_2011 <-
  students_data_2011[which(students_data_2011$GOKpercentage >= .065
                           & students_data_2011$GOKpercentage <= .135),]
```

Moreover, from every school we sample a number students in order to increase the balance in the covariate and to guarantee an equal representation to all the schools, avoiding biases related to the over-representation of biggest schools' students.

In the first case, we sample 50 students from each school.

```
schools <-
students_data_randomized_03_2011$school[which(
!duplicated(students_data_randomized_03_2011$school))]

sample_students <- as.data.frame(matrix(data = NA, nrow = 50*length(schools),
                                        ncol = ncol(students_data_randomized_03_2011)))
colnames(sample_students) <- colnames(students_data_randomized_03_2011)

for (j in (0:(length(schools)-1))){
  set.seed(j  + 2020)
  sample_students[(1+(j*50)):(50+(j*50)),]  <-
  students_data_randomized_03_2011[which(
  students_data_randomized_03_2011$school %in%
  schools[j+1]),][sample(1:nrow(students_data_randomized_03_2011[which(
  students_data_randomized_03_2011$school %in% schools[j+1]),]),
  50, replace = FALSE ),]
}
```

```
sample_student <- round(sample_students[, -(1:4)], 0)
sample_students <- cbind(sample_students[,1:4], sample_student)
```

Then, we run our BCF-IV algorithm on this sample of units.

```
# Attaching the Sample and the Covariates
attach(sample_students)
x <- cbind(primary_retention , man , BULO,
           leerkracht_age , leerkracht_seniority,
           leerkracht_diploma,
           directie_age, directie_seniority
)
z <- as.matrix(eligible_dummy)
y <- as.matrix(certificate)
w <- as.matrix(GOKschool)
detach(sample_students)
```

Below we run the BCF-IV algorithm.

```
set.seed(2020)
bcf_iv(y, w, z, x, max_depth = 2, n_burn = 1000, n_sim = 1000,
       inference_ratio = 0.50 , inference_ratio = 0.50 )
```

```
## Loading required package: lattice

## Loading required package: rattle

## Rattle: A free graphical interface for data science with R.
## Version 5.2.0 Copyright (c) 2006-2018 Togaware Pty Ltd.
## Type 'rattle()' to shake, rattle, and roll your data.

## Loading required package: bartCause

## The effect on the overall sample is -0.0041
## P-value 0.921232654011692
## P-value Weak-Instrument Test 2.78133165300155e-100
## Proportion of observations in the node:  1.00
## Intention-to-treat effect:  -0.00346470548203343
## Proportion of compliers in the node:  0.837209302325581
##
##  node number: 2
##    root
##    directie_age>=6.5
## The conditional effect on the subpopulation is -0.0288
## P-value 0.816181164228318
## P-value Weak-Instrument Test 6.93555073690141e-23
## Proportion of observations in the node:  0.381085271317829
## Intention-to-treat effect:  -0.0173581265136767
## Proportion of compliers in the node:  0.602929210740439
##
##  node number: 4
##    root
##    directie_age>=6.5
##    man< 0.5
## The conditional effect on the subpopulation is 0.1864
## P-value 0.312385114987977
## P-value Weak-Instrument Test 1.47823903628575e-11
```

```
## Proportion of observations in the node:  0.20062015503876
## Intention-to-treat effect:  0.108319691246813
## Proportion of compliers in the node:  0.581143740340031
##
##  node number: 5
##    root
##    directie_age>=6.5
##    man>=0.5
## The conditional effect on the subpopulation is -0.2373
## P-value 0.163085297178804
## P-value Weak-Instrument Test 5.70156647620314e-13
## Proportion of observations in the node:  0.18046511627907
## Intention-to-treat effect:  -0.148844340690133
## Proportion of compliers in the node:  0.627147766323024
##
##  node number: 3
##    root
##    directie_age< 6.5
## The conditional effect on the subpopulation is -8e-04
## P-value 0.985054288856483
## P-value Weak-Instrument Test 4.43362622577243e-79
## Proportion of observations in the node:  0.618914728682171
## Intention-to-treat effect:  -0.000495232837176779
## Proportion of compliers in the node:  0.644288577154309
##
##  node number: 6
##    root
##    directie_age< 6.5
##    leerkracht_seniority< 3.5
## The conditional effect on the subpopulation is 0.0511
## P-value 0.282584822269171
## P-value Weak-Instrument Test 1.63166760042691e-64
## Proportion of observations in the node:  0.102325581395349
## Intention-to-treat effect:  0.0455345241268989
## Proportion of compliers in the node:  0.890909090909091
##
##  node number: 7
##    root
##    directie_age< 6.5
##    leerkracht_seniority>=3.5
## The conditional effect on the subpopulation is -0.0388
## P-value 0.475164983587273
## P-value Weak-Instrument Test 8.0804994794541e-49
## Proportion of observations in the node:  0.516589147286822
## Intention-to-treat effect:  -0.0231255980821084
## Proportion of compliers in the node:  0.595438175270108
```

As a robustness check, we run the BCF-ITT algorithm.

```
set.seed(2020)
bcf_itt(y, w, z, x, max_depth = 2, n_burn = 1000, n_sim = 1000,
        inference_ratio = 0.50 , inference_ratio = 0.50 )
```

```
## The effect on the overall sample is -0.0041
## P-value 0.921232654011692
```

```
## P-value Weak-Instrument Test 2.78133165300155e-100
## Proportion of observations in the node:  1.00
##
##  node number: 2
##    root
##    leerkracht_seniority< 3.5
## The conditional effect on the subpopulation is -0.0083
## P-value 0.882750498947076
## P-value Weak-Instrument Test 1.13404124528579e-54
## Proportion of observations in the node:  0.138604651162791
##
##  node number: 4
##    root
##    leerkracht_seniority< 3.5
##    leerkracht_age>=3.5
## The conditional effect on the subpopulation is -0.0694
## P-value 0.475914543044949
## P-value Weak-Instrument Test 2.53016192911411e-25
## Proportion of observations in the node:  0.105116279069767
##
##  node number: 5
##    root
##    leerkracht_seniority< 3.5
##    leerkracht_age< 3.5
## The conditional effect on the subpopulation is 0.0575
## P-value 0.253628255961621
## P-value Weak-Instrument Test 0
## Proportion of observations in the node:  0.0334883720930233
##
##  node number: 3
##    root
##    leerkracht_seniority>=3.5
## The conditional effect on the subpopulation is -0.0111
## P-value 0.830289189926922
## P-value Weak-Instrument Test 1.06860054018598e-69
## Proportion of observations in the node:  0.861395348837209
##
##  node number: 6
##    root
##    leerkracht_seniority>=3.5
##    primary_retention>=0.5
## The conditional effect on the subpopulation is -0.6667
## P-value 0.180654276380246
## P-value Weak-Instrument Test 0.000799845861653577
## Proportion of observations in the node:  0.0368992248062016
##
##  node number: 7
##    root
##    leerkracht_seniority>=3.5
##    primary_retention< 0.5
## The conditional effect on the subpopulation is 0.0304
## P-value 0.540980195769663
## P-value Weak-Instrument Test 3.35869725340887e-67
## Proportion of observations in the node:  0.824496124031008
```

In the second case we sample 62 students from every school (where 62 units is the size of the smallest school).

```
schools <-
students_data_randomized_03_2011$school[which(
!duplicated(students_data_randomized_03_2011$school))]

sample_students <- as.data.frame(matrix(data = NA, nrow = 62*length(schools),
                                        ncol = ncol(students_data_randomized_03_2011)))
colnames(sample_students) <- colnames(students_data_randomized_03_2011)

for (j in (0:(length(schools)-1))){
  set.seed(j  + 2020)
  sample_students[(1+(j*62)):(62+(j*62)),]  <-
  students_data_randomized_03_2011[which(
  students_data_randomized_03_2011$school %in%
  schools[j+1]),][sample(1:nrow(students_data_randomized_03_2011[which(
  students_data_randomized_03_2011$school %in% schools[j+1]),]),
  62, replace = FALSE ),]
}

sample_student <- round(sample_students[, -(1:4)], 0)
sample_students <- cbind(sample_students[,1:4], sample_student)
```

```
# Attaching the Sample and the Covariates
attach(sample_students)
x <- cbind(primary_retention , man , BULO,
           leerkracht_age , leerkracht_seniority,
           directie_age, directie_seniority
)
z <- as.matrix(eligible_dummy)
w <- as.matrix(GOKschool)
y <- as.matrix(certificate)
detach(sample_students)
```

Below we run the BCF-IV algorithm.

```
set.seed(2020)
bcf_iv(y, w, z, x, max_depth = 2, n_burn = 1000, n_sim = 1000,
       inference_ratio = 0.50 , inference_ratio = 0.50 )
```

```
## The effect on the overall sample is -0.0192
## P-value 0.609503889529807
## P-value Weak-Instrument Test 2.17505867538481e-127
## Proportion of observations in the node:  1.00
## Intention-to-treat effect:  -0.0150562749316396
## Proportion of compliers in the node:  0.784994138335287
##
##  node number: 2
##    root
##    directie_age>=7.5
## The conditional effect on the subpopulation is -0.0623
## P-value 0.449072929776276
## P-value Weak-Instrument Test 4.89771807038385e-23
## Proportion of observations in the node:  0.0914419695193435
## Intention-to-treat effect:  -0.0465082343131129
```

```
## Proportion of compliers in the node:  0.746153846153846
##
##  node number: 4
##    root
##    directie_age>=7.5
##    man< 0.5
## The conditional effect on the subpopulation is -0.1524
## P-value 0.229155874084574
## P-value Weak-Instrument Test 1.89413171493092e-09
## Proportion of observations in the node:  0.047831184056272
## Intention-to-treat effect:  -0.105322128851542
## Proportion of compliers in the node:  0.691176470588235
##
##  node number: 5
##    root
##    directie_age>=7.5
##    man>=0.5
## The conditional effect on the subpopulation is -0.0099
## P-value 0.92636807249041
## P-value Weak-Instrument Test 1.34023685633917e-16
## Proportion of observations in the node:  0.0436107854630715
## Intention-to-treat effect:  -0.00800051203276873
## Proportion of compliers in the node:  0.806451612903226
##
##  node number: 3
##    root
##    directie_age< 7.5
## The conditional effect on the subpopulation is -0.0102
## P-value 0.803392203676569
## P-value Weak-Instrument Test 2.98339062128499e-109
## Proportion of observations in the node:  0.908558030480657
## Intention-to-treat effect:  -0.00626046785179109
## Proportion of compliers in the node:  0.613161290322581
```

As a robustness check, we run the BCF-ITT algorithm.

```
set.seed(2020)
bcf_itt(y, w, z, x, max_depth = 2, n_burn = 1000, n_sim = 1000,
        inference_ratio = 0.50 , inference_ratio = 0.50 )
```

```
## The effect on the overall sample is -0.0299
## P-value 0.442123397494213
## P-value Weak-Instrument Test 4.46272651712928e-119
## Proportion of observations in the node:  1.00
##
##  node number: 2
##    root
##    directie_age>=5.5
## The conditional effect on the subpopulation is -0.0476
## P-value 0.332890644491161
## P-value Weak-Instrument Test 1.86595964869224e-80
## Proportion of observations in the node:  0.699674918729682
##
##  node number: 4
##    root
```

```
##    directie_age>=5.5
##    directie_age>=7.5
## The conditional effect on the subpopulation is -0.0415
## P-value 0.632061338528029
## P-value Weak-Instrument Test 2.59584030099857e-21
## Proportion of observations in the node:  0.0930232558139535
##
##  node number: 5
##    root
##    directie_age>=5.5
##    directie_age< 7.5
## The conditional effect on the subpopulation is -0.044
## P-value 0.436223792610796
## P-value Weak-Instrument Test 7.71172785179552e-64
## Proportion of observations in the node:  0.606651662915729
##
##  node number: 3
##    root
##    directie_age< 5.5
## The conditional effect on the subpopulation is 0.0044
## P-value 0.944766807950413
## P-value Weak-Instrument Test 4.41867552642894e-40
## Proportion of observations in the node:  0.300325081270318
##
##  node number: 6
##    root
##    directie_age< 5.5
##    leerkracht_seniority< 3.5
## The conditional effect on the subpopulation is 0.0194
## P-value 0.652734040427151
## P-value Weak-Instrument Test 1.53263248896913e-54
## Proportion of observations in the node:  0.0825206301575394
##
##  node number: 7
##    root
##    directie_age< 5.5
##    leerkracht_seniority>=3.5
## The conditional effect on the subpopulation is -0.026
## P-value 0.87843992747154
## P-value Weak-Instrument Test 4.67622409638456e-12
## Proportion of observations in the node:  0.217804451112778
```

Second Outcome: School Progress

We now focus on the effects when we sample units in a bandwidth of 0.035 (which is the optimal bandwidth for the outcome *progress_school*) around the threshold (10%).

```
students_data_randomized_03_2011 <-
  students_data_2011[which(students_data_2011$GOKpercentage >= .063
                           & students_data_2011$GOKpercentage <= .137),]
```

Moreover, from every school we sample a number students in order to increase the balance in the covariate and to guarantee an equal representation to all the schools, avoiding biases related to the over-representation of biggest schools' students. In the first case, we sample 50 students from each school.

```
schools <-
students_data_randomized_03_2011$school[which(
!duplicated(students_data_randomized_03_2011$school))]

sample_students <- as.data.frame(matrix(data = NA, nrow = 50*length(schools),
                                        ncol = ncol(students_data_randomized_03_2011)))
colnames(sample_students) <- colnames(students_data_randomized_03_2011)

for (j in (0:(length(schools)-1))){
  set.seed(j  + 2020)
  sample_students[(1+(j*50)):(50+(j*50)),]  <-
  students_data_randomized_03_2011[which(
  students_data_randomized_03_2011$school %in%
  schools[j+1]),][sample(1:nrow(students_data_randomized_03_2011[which(
  students_data_randomized_03_2011$school %in% schools[j+1]),]),
  50, replace = FALSE ),]
}

sample_student <- round(sample_students[, -(1:4)], 0)
sample_students <- cbind(sample_students[,1:4], sample_student)
```

Then, we run our BCF-IV algorithm on this sample of units.

```
# Attaching the Sample and the Covariates
attach(sample_students)
x <- cbind(primary_retention , man , BULO,
           leerkracht_age , leerkracht_seniority,
           directie_age, directie_seniority
)
z <- as.matrix(eligible_dummy)
w <- as.matrix(GOKschool)
y <- as.matrix(progress_school)
detach(sample_students)
```

Below we run the BCF-IV algorithm.

```
set.seed(2020)
bcf_iv(y, w, z, x, max_depth = 2, n_burn = 1000, n_sim = 1000,
       inference_ratio = 0.50 , inference_ratio = 0.50 )
```

```
## The effect on the overall sample is -0.005
## P-value 0.74555800512467
## P-value Weak-Instrument Test 2.83497505110029e-112
## Proportion of observations in the node:  1.00
## Intention-to-treat effect:  -0.00461716028538263
## Proportion of compliers in the node:  0.914927768860353
##
##  node number: 2
##    root
##    man>=0.5
## The conditional effect on the subpopulation is 0.0117
## P-value 0.688593506973482
## P-value Weak-Instrument Test 6.60233901830994e-51
## Proportion of observations in the node:  0.489245585874799
## Intention-to-treat effect:  0.00750870483713419
```

```
## Proportion of compliers in the node:  0.643700787401575
##
##  node number: 4
##    root
##    man>=0.5
##    directie_seniority>=6.5
## The conditional effect on the subpopulation is 0.0142
## P-value 0.85497081143672
## P-value Weak-Instrument Test 1.92015047852439e-11
## Proportion of observations in the node:  0.152166934189406
## Intention-to-treat effect:  0.0096859432990476
## Proportion of compliers in the node:  0.681434599156118
##
##  node number: 5
##    root
##    man>=0.5
##    directie_seniority< 6.5
## The conditional effect on the subpopulation is 0.0158
## P-value 0.627246212107553
## P-value Weak-Instrument Test 2.58296695396726e-38
## Proportion of observations in the node:  0.337078651685393
## Intention-to-treat effect:  0.00992102965007537
## Proportion of compliers in the node:  0.626666666666667
##
##  node number: 3
##    root
##    man< 0.5
## The conditional effect on the subpopulation is -0.0211
## P-value 0.189556113829234
## P-value Weak-Instrument Test 3.20717376764302e-62
## Proportion of observations in the node:  0.510754414125201
## Intention-to-treat effect:  -0.0133673286203198
## Proportion of compliers in the node:  0.632306725329981
##
##  node number: 6
##    root
##    man< 0.5
##    directie_seniority< 6.5
## The conditional effect on the subpopulation is -0.0157
## P-value 0.217247180178665
## P-value Weak-Instrument Test 3.18578871079386e-50
## Proportion of observations in the node:  0.345746388443018
## Intention-to-treat effect:  -0.0101729498455837
## Proportion of compliers in the node:  0.646239554317549
##
##  node number: 7
##    root
##    man< 0.5
##    directie_seniority>=6.5
## The conditional effect on the subpopulation is -0.0501
## P-value 0.416588320152841
## P-value Weak-Instrument Test 1.87989513225369e-13
## Proportion of observations in the node:  0.165008025682183
## Intention-to-treat effect:  -0.0301968664143559
```

```
## Proportion of compliers in the node:  0.603112840466926
```

As a robustness check, we run the BCF-ITT algorithm.

```
set.seed(2020)
bcf_itt(y, w, z, x, max_depth = 2, n_burn = 1000, n_sim = 1000,
        inference_ratio = 0.50 , inference_ratio = 0.50 )
```

```
## The effect on the overall sample is -0.004
## P-value 0.779791729837223
## P-value Weak-Instrument Test 1.48434670994665e-121
## Proportion of observations in the node:  1.00
##
##  node number: 2
##    root
##    man< 0.5
## The conditional effect on the subpopulation is -0.0196
## P-value 0.192105110526657
## P-value Weak-Instrument Test 1.8014830931585e-66
## Proportion of observations in the node:  0.511537308960144
##
##  node number: 4
##    root
##    man< 0.5
##    leerkracht_seniority>=3.5
## The conditional effect on the subpopulation is -0.0222
## P-value 0.210162443681802
## P-value Weak-Instrument Test 1.70063155490927e-48
## Proportion of observations in the node:  0.44201378483668
##
##  node number: 5
##    root
##    man< 0.5
##    leerkracht_seniority< 3.5
## The conditional effect on the subpopulation is -0.0115
## P-value 0.625506907068704
## P-value Weak-Instrument Test 1.27961147937025e-31
## Proportion of observations in the node:  0.0695235241234642
##
##  node number: 3
##    root
##    man>=0.5
## The conditional effect on the subpopulation is 0.0122
## P-value 0.642928146906334
## P-value Weak-Instrument Test 3.99739975847916e-56
## Proportion of observations in the node:  0.488462691039856
##
##  node number: 6
##    root
##    man>=0.5
##    directie_seniority>=5.5
## The conditional effect on the subpopulation is 0.0117
## P-value 0.724447650061136
## P-value Weak-Instrument Test 1.85020973808536e-33
## Proportion of observations in the node:  0.30536409949056
```

```
##
##  node number: 7
##    root
##    man>=0.5
##    directie_seniority< 5.5
## The conditional effect on the subpopulation is 0.0207
## P-value 0.633415489032333
## P-value Weak-Instrument Test 3.80270112551981e-23
## Proportion of observations in the node:  0.183098591549296
```

In the second case we sample 62 students from every school (where 62 units is the size of the smallest school).

```
schools <-
students_data_randomized_03_2011$school[which(
!duplicated(students_data_randomized_03_2011$school))]

sample_students <- as.data.frame(matrix(data = NA, nrow = 62*length(schools),
                                        ncol = ncol(students_data_randomized_03_2011)))
colnames(sample_students) <- colnames(students_data_randomized_03_2011)

for (j in (0:(length(schools)-1))){
  set.seed(j  + 2020)
  sample_students[(1+(j*62)):(62+(j*62)),]  <-
  students_data_randomized_03_2011[which(
  students_data_randomized_03_2011$school %in%
  schools[j+1]),][sample(1:nrow(students_data_randomized_03_2011[which(
  students_data_randomized_03_2011$school %in% schools[j+1]),]),
  62, replace = FALSE ),]
}

sample_student <- round(sample_students[, -(1:4)], 0)
sample_students <- cbind(sample_students[,1:4], sample_student)
```

```
# Attaching the Sample and the Covariates
attach(sample_students)
x <- cbind(primary_retention , man , BULO,
           leerkracht_age , leerkracht_seniority,
           directie_age, directie_seniority
)
z <- as.matrix(eligible_dummy)
w <- as.matrix(GOKschool)
y <- as.matrix(progress_school)
detach(sample_students)
```

Below we run the BCF-IV algorithm.

```
set.seed(2020)
bcf_iv(y, w, z, x, max_depth = 2, n_burn = 1000, n_sim = 1000,
       inference_ratio = 0.50 , inference_ratio = 0.50 )
```

```
## The effect on the overall sample is -0.003
## P-value 0.807317229834871
## P-value Weak-Instrument Test 2.625529720261e-168
## Proportion of observations in the node:  1.00
## Intention-to-treat effect:  -0.0024002409581728
## Proportion of compliers in the node:  0.800634345265066
```

```
##
##  node number: 2
##    root
##    directie_seniority>=6.5
## The conditional effect on the subpopulation is -0.0807
## P-value 0.0407647185119405
## P-value Weak-Instrument Test 3.72600699022716e-34
## Proportion of observations in the node:  0.314454009968283
## Intention-to-treat effect:  -0.052014189786836
## Proportion of compliers in the node:  0.644812680115274
##
##  node number: 4
##    root
##    directie_seniority>=6.5
##    man< 0.5
## The conditional effect on the subpopulation is -0.0735
## P-value 0.100190006895259
## P-value Weak-Instrument Test 2.46378694051391e-18
## Proportion of observations in the node:  0.163797009515179
## Intention-to-treat effect:  -0.0452445944032444
## Proportion of compliers in the node:  0.615491009681881
##
##  node number: 5
##    root
##    directie_seniority>=6.5
##    man>=0.5
## The conditional effect on the subpopulation is -0.1003
## P-value 0.153736607131768
## P-value Weak-Instrument Test 4.84824280753097e-17
## Proportion of observations in the node:  0.150657000453104
## Intention-to-treat effect:  -0.0678422614649963
## Proportion of compliers in the node:  0.676691729323308
##
##  node number: 3
##    root
##    directie_seniority< 6.5
## The conditional effect on the subpopulation is 0.0142
## P-value 0.257438663978885
## P-value Weak-Instrument Test 2.08914818622444e-130
## Proportion of observations in the node:  0.685545990031717
## Intention-to-treat effect:  0.00905934958097151
## Proportion of compliers in the node:  0.639458030403173
##
##  node number: 6
##    root
##    directie_seniority< 6.5
##    primary_retention< 0.5
## The conditional effect on the subpopulation is 0.0261
## P-value 0.0422194114609425
## P-value Weak-Instrument Test 3.89633444970472e-117
## Proportion of observations in the node:  0.653602174898052
## Intention-to-treat effect:  0.0165656743518995
## Proportion of compliers in the node:  0.634662045060659
##
```

```
##  node number: 7
##    root
##    directie_seniority< 6.5
##    primary_retention>=0.5
## The conditional effect on the subpopulation is -0.0943
## P-value 0.0875344682546614
## P-value Weak-Instrument Test 2.8620838466104e-14
## Proportion of observations in the node:  0.0319438151336656
## Intention-to-treat effect:  -0.0695838351398363
## Proportion of compliers in the node:  0.737588652482269
```

As a robustness check, we run the BCF-ITT algorithm.

```
set.seed(2020)
bcf_itt(y, w, z, x, max_depth = 2, n_burn = 1000, n_sim = 1000,
        inference_ratio = 0.50 , inference_ratio = 0.50 )
```

```
## The effect on the overall sample is -0.0079
## P-value 0.53435754063643
## P-value Weak-Instrument Test 2.1945585850208e-158
## Proportion of observations in the node:  1.00
##
##  node number: 2
##    root
##    leerkracht_seniority>=3.5
## The conditional effect on the subpopulation is -0.014
## P-value 0.33094878726286
## P-value Weak-Instrument Test 2.00445882803938e-117
## Proportion of observations in the node:  0.860802319961334
##
##  node number: 4
##    root
##    leerkracht_seniority>=3.5
##    directie_seniority>=5.5
## The conditional effect on the subpopulation is -0.0149
## P-value 0.354622697206015
## P-value Weak-Instrument Test 1.47251906000849e-97
## Proportion of observations in the node:  0.602706621556307
##
##  node number: 5
##    root
##    leerkracht_seniority>=3.5
##    directie_seniority< 5.5
## The conditional effect on the subpopulation is -0.0126
## P-value 0.683520681054491
## P-value Weak-Instrument Test 3.73036907903811e-25
## Proportion of observations in the node:  0.258095698405027
##
##  node number: 3
##    root
##    leerkracht_seniority< 3.5
## The conditional effect on the subpopulation is 0.0064
## P-value 0.786515794050415
## P-value Weak-Instrument Test 3.6132069982805e-62
## Proportion of observations in the node:  0.139197680038666
```

```
##
##  node number: 6
##    root
##    leerkracht_seniority< 3.5
##    man< 0.5
## The conditional effect on the subpopulation is -0.0088
## P-value 0.645474638479603
## P-value Weak-Instrument Test 1.49256838456971e-37
## Proportion of observations in the node:  0.070082165297245
##
##  node number: 7
##    root
##    leerkracht_seniority< 3.5
##    man>=0.5
## The conditional effect on the subpopulation is 0.0252
## P-value 0.603614476826785
## P-value Weak-Instrument Test 1.26827582266622e-26
## Proportion of observations in the node:  0.069115514741421
```