# Labor Supply Analysis Using 2016 Canadian Census Data

Bahareh Aghababaei

2023-01-28

## Data Cleaning & Preparation

The data set was restricted in terms of province and age. British Columbia and the age between 18 and 65 were selected in the data set. In different stages, the unavailable and inapplicable data of each variable, e.g., code 88 and 99 in CFINC, were removed from the dataset to have a better estimation.

```
load("C:/Users/bahar/Documents/MSc courses/ECON 562/Assignment 2/indiv_pumf06_v2.RData")

names(table)
```

```
##   [1] "ABOID"     "AGEGRP"    "AGEIMM"    "ATTSCH"    "BFNMEMB"   "CFINC"
##   [7] "CFINC_AT"  "CFINEF"    "CFSIZE"    "CFSTAT"    "CHDBN"     "CIP"
##  [13] "CITIZEN"   "CITOTH"    "CMA"       "CONDO"     "COW"       "CQPPB"
##  [19] "DIST"      "EFINC"     "EFINC_AT"  "EFNOTCF"   "EFSIZE"    "EICBN"
##  [25] "EMPIN"     "ETHDER"    "FOL"       "FPTWK"     "GENSTAT"   "GOVTI"
##  [31] "GROSRT"    "GTRFS"     "HDGREE"    "HHCLASS"   "HHINC"     "HHINC_AT"
##  [37] "HHSIZE"    "HHTYPE"    "HLAEN"     "HLAFR"     "HLANO"     "HLBEN"
##  [43] "HLBFR"     "HLBNO"     "HRSWRK"    "IMMSTAT"   "INCTAX"    "INVST"
##  [49] "KOL"       "LFACT"     "LICO"      "LICO_AT"   "LOCSTUD"   "LSTWRK"
##  [55] "LWAEN"     "LWAFR"     "LWANO"     "LWBEN"     "LWBFR"     "LWBNO"
##  [61] "MARST"     "MARSTH"    "MFS"       "MOB1"      "MOB5"      "MODE"
##  [67] "MRKINC"    "MSI"       "MTNEN"     "MTNFR"     "MTNNO"     "NAICS"
##  [73] "NOCHRD"    "NOCS"      "NOL"       "NONCFINHH" "OASGI"     "OMP"
##  [79] "OTINC"     "PKID0_1"   "PKID15_24" "PKID2_5"   "PKID25"    "PKID6_14"
##  [85] "PKIDHH"    "POB"       "POBF"      "POBM"      "POWST"     "PPSORT"
##  [91] "PR"        "PR1"       "PR5"       "PRIHM"     "PWPR"      "REGIND"
##  [97] "REPAIR"    "RETIR"     "ROOM"      "SEMPI"     "SEX"       "SSGRAD"
## [103] "TENUR"     "TOTINC"    "TOTINC_AT" "UPHWRK"    "UPKID"     "UPSR"
## [109] "VALUE"     "VISMIN"    "VISMINH"   "WAGES"     "WEIGHT"    "WKSWRK"
## [115] "WRKACT"    "WT1"       "WT2"       "WT3"       "WT4"       "WT5"
## [121] "WT6"       "WT7"       "WT8"       "YRIMM"
```

```
nrow(table)
```

```
## [1] 844476
```

```
census_2006 <- subset(table, PR==59 & AGEGRP>=7 & AGEGRP<=16 & CFINC>=1 & CFINC<=28)
nrow(census_2006)
```

```
## [1] 58086
```

# Non-Labor Income Calculation

```r
new_dataset <- subset(census_2006,select = c("WAGES", "CFINC","HRSWRK","SEX"))

range_f <- c(1000,3500,6000,8500,11000,13500,16000,18500,22500,27500,32500,37500,42500,47500,52500,5750

for (i in 1:58086){
  new_dataset$Cfamily_income[i]= range_f[new_dataset$CFINC[i]]
  new_dataset$non_labour_INC[i]=new_dataset$Cfamily_income[i]-new_dataset$WAGES[i]

}
```

To find the non-labor income, I used economic family income groups (CFINC) among other family variables. The reason for choosing this variable is that CFINC can better represent family income compared to EFINC and HHINC. By definition, all persons who are members of a census family are also members of an economic family. For example: two co-resident census families who are related to one another are considered one economic family; and, nieces or nephews living with aunts or uncles are considered one economic family. However, Census family is defined as a married couple and the children, or a couple living common law and the children. Thus, this index is a viable representative of a family income.

# Labor Supply Elasticity

```r
Reg_Model <- lm(HRSWRK~WAGES+non_labour_INC,data=new_dataset)

summary(Reg_Model)
```

```
##
## Call:
## lm(formula = HRSWRK ~ WAGES + non_labour_INC, data = new_dataset)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -96.329 -23.394   6.351  14.240  76.071
##
## Coefficients:
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)    2.186e+01  1.596e-01  136.97   <2e-16 ***
## WAGES          7.529e-05  1.678e-06   44.87   <2e-16 ***
## non_labour_INC 6.816e-05  1.666e-06   40.92   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 21.16 on 58083 degrees of freedom
## Multiple R-squared:  0.03395,    Adjusted R-squared:  0.03392
## F-statistic:  1021 on 2 and 58083 DF,  p-value: < 2.2e-16
```

```r
income_elasticity <-summary(Reg_Model)$coefficient[3,1]*mean(new_dataset$non_labour_INC)/mean(new_datas
income_elasticity
```

```
## [1] 0.1202086
```

```r
substitution_elasticity <-summary(Reg_Model)$coefficient[2,1]*mean(new_dataset$WAGES)/mean(new_dataset$I
substitution_elasticity
```

```
## [1] 0.08759095
```

```r
compensated_elasticity <- substitution_elasticity-(income_elasticity)*(mean(new_dataset$WAGES)*mean(new_
compensated_elasticity
```

```
## [1] -2.100509
```

when we first run the regression, the static elasticity of substitution, income elasticity , and the compensated elasticity of substitution are 0.088, 0.120, AND -2.1 respectively. Substitution elasticity shows if we increase the wage rate, the workers tend to increase their hours of work (or decrease their leisure time) about 8% . Based on the elasticity of income, if the non-labor income increases, the labors working hours grow by 12%, which in turn decrease their leisure time. The compensated elasticity of substitution is negative which shows the effect of elasticity of non-labor income is greater than elasticity of substitution.

## Gender-Based Regression Analysis (women)

```r
new_dataset_women <-subset(new_dataset,SEX==1)
Reg_Model_women <- lm(HRSWRK~WAGES+non_labour_INC,data=new_dataset_women)
summary(Reg_Model_women)
```

```
##
## Call:
## lm(formula = HRSWRK ~ WAGES + non_labour_INC, data = new_dataset_women)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -74.020 -19.623   1.989  16.396  80.306
##
## Coefficients:
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)    1.763e+01  2.010e-01   87.75   <2e-16 ***
## WAGES          6.596e-05  2.188e-06   30.15   <2e-16 ***
## non_labour_INC 6.038e-05  2.130e-06   28.35   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 19.63 on 30344 degrees of freedom
## Multiple R-squared:  0.02911,    Adjusted R-squared:  0.02905
## F-statistic: 454.9 on 2 and 30344 DF,  p-value: < 2.2e-16
```

```r
income_elasticity <-summary(Reg_Model_women)$coefficient[3,1]*mean(new_dataset_women$non_labour_INC)/mea
income_elasticity
```

```
## [1] 0.1466164
```

```r
substitution_elasticity <-summary(Reg_Model_women)$coefficient[2,1]*mean(new_dataset_women$WAGES)/mean(r

substitution_elasticity
```

```
## [1] 0.0695905
```

```r
compensated_elasticity <- substitution_elasticity-(income_elasticity)*(mean(new_dataset_women$WAGES)*mea

compensated_elasticity
```

```
## [1] -1.363572
```

## Gender-Based Regression Analysis (Men)

```r
new_dataset_men <-subset(new_dataset,SEX==2)

Reg_Model_men <- lm(HRSWRK~WAGES+non_labour_INC,data=new_dataset_men)

summary(Reg_Model_men)
```

```
##
## Call:
## lm(formula = HRSWRK ~ WAGES + non_labour_INC, data = new_dataset_men)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -88.391 -19.797   5.873  11.504  70.702
##
## Coefficients:
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)    2.723e+01  2.398e-01  113.53   <2e-16 ***
## WAGES          7.408e-05  2.440e-06   30.36   <2e-16 ***
## non_labour_INC 6.845e-05  2.469e-06   27.73   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 21.49 on 27736 degrees of freedom
## Multiple R-squared:  0.03259,    Adjusted R-squared:  0.03252
## F-statistic: 467.2 on 2 and 27736 DF,  p-value: < 2.2e-16
```

```r
income_elasticity <-summary(Reg_Model_men)$coefficient[3,1]*mean(new_dataset_men$non_labour_INC)/mean(no
income_elasticity
```

```
## [1] 0.08697926
```

```
substitution_elasticity <-summary(Reg_Model_men)$coefficient[2,1]*mean(new_dataset_men$WAGES)/mean(new_d

substitution_elasticity
```

```
## [1] 0.09214333
```

```
compensated_elasticity <- substitution_elasticity-(income_elasticity)*(mean(new_dataset_men$WAGES)*mean

compensated_elasticity
```

```
## [1] -2.732148
```

We repeated the same regression model for both men and women separately.Based on the results, women and men are different in terms of income elasticity and compensated elasticity of substitution, while there are almost the same in elasticity of substitution.

## New set of Features

```
census_2006 <- subset(table, PR==59 &  AGEGRP>=7 & AGEGRP<=16 & CFINC>=1 & CFINC<=28)


census_2006$AGE_SQURE <- census_2006$AGEGRP*census_2006$AGEGRP

new_dataset <- subset(census_2006,select = c("WAGES", "CFINC","HRSWRK","SEX","AGEGRP","PKID0_1","PKID15_


range_f <- c(1000,3500,6000,8500,11000,13500,16000,18500,22500,27500,32500,37500,42500,47500,52500,5750

for (i in 1:58086){
  new_dataset$Cfamily_income[i]= range_f[new_dataset$CFINC[i]]
  new_dataset$non_labour_INC[i]=new_dataset$Cfamily_income[i]-new_dataset$WAGES[i]

}
```

Based on the definition of intertemporal elasticity of substitution defined by MaCurdy(1981), the intertemporal substitution effect is interpreted as an elasticity that is associated with a particular kind of parametric wage change. In particular, it determines the response of hours of work at age t to a shift in the age t wage rate holding X or the marginal utility of wealth constant. As a result, we need to introduce the variables showing people background or wage profile. Accordingly, I have considered these variables: Province, Occupation, Census family income groups, education, Investment income, Household income groups, value of the dwelling, TENUR, immigrant status, Generation status, and Aboriginal identity. For part b, we need to define some variables representing people with different wage profile , so age and age-squared can be considered to show the effect of a shift on a wage profile.

## Feature Selection with LASSO

```r
library(glmnet)
```

```
## Loading required package: Matrix
```

```
## Loaded glmnet 4.1-8
```

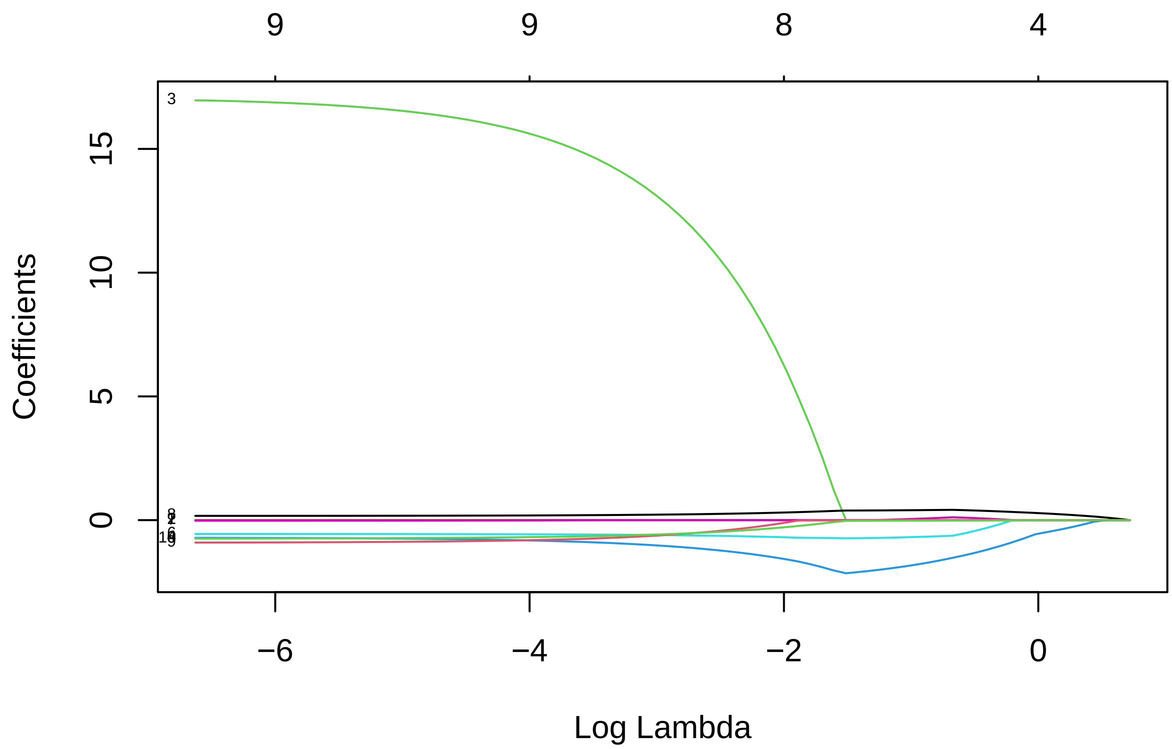
```r
x1 <- subset(new_dataset, ROOM<12 & HDGREE<14   & PKID0_1<2 & PKID15_24<2 & PKID2_5<2 & PKID25<2 & PKID
x1$CHILD_NUM <- rowSums(x1[,c("PKID0_1","PKID2_5","PKID6_14","PKID15_24","PKID2_5")],na.rm=TRUE )

y <- as.matrix(subset(x1,select = "HRSWRK"))
x <- as.matrix(subset(x1,select =c("WAGES","non_labour_INC","AGEGRP","MARST","PR","IMMSTAT","ROOM","HDGR


lassofit <- glmnet(x,y, alpha=1, nlambda=100 )
summary(lassofit)
```

```
##           Length Class     Mode
## a0         80    -none-    numeric
## beta      800    dgCMatrix S4
## df         80    -none-    numeric
## dim         2    -none-    numeric
## lambda     80    -none-    numeric
## dev.ratio  80    -none-    numeric
## nulldev     1    -none-    numeric
## npasses     1    -none-    numeric
## jerr        1    -none-    numeric
## offset      1    -none-    logical
## call        5    -none-    call
## nobs        1    -none-    numeric
```

```r
plot(lassofit, xvar ='lambda',label=T )
```




```r
coef(lassofit, s=0.03)
```

```
## 11 x 1 sparse Matrix of class "dgCMatrix"
##                           s1
## (Intercept)    -5.332902e+01
## WAGES           6.859678e-05
## non_labour_INC  6.285298e-05
## AGEGRP          1.468617e+01
## MARST          -8.943035e-01
## PR              .
## IMMSTAT        -5.798172e-01
## ROOM           -3.293553e-04
## HDGREE          2.019713e-01
## CHILD_NUM      -7.433054e-01
## AGE_SQURE      -6.464876e-01
```

Based on the plot,as the lambda increases, more number of variables goes to zero and be excluded from the model. The plot shows age is a key variable and shouldn't be excluded from the model.

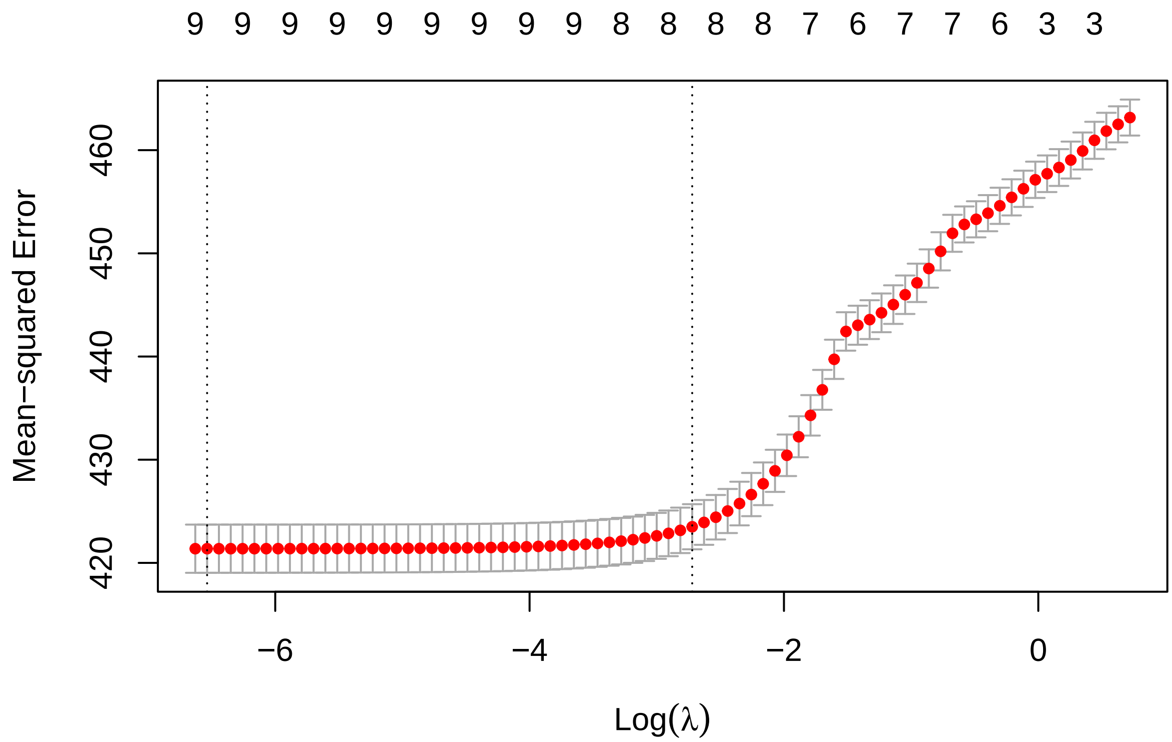
## Cross Validation to Find the Optimal Lambda

After the cross-validation, we found the optimal value for lambda, the result of lasso model based on this optimal value is as follows:

```
x1 <- subset(new_dataset, ROOM<12 & HDGREE<14   & PKID0_1<2 & PKID15_24<2 & PKID2_5<2 & PKID25<2 & PKID
x1$CHILD_NUM <- rowSums(x1[,c("PKID0_1","PKID2_5","PKID6_14","PKID15_24","PKID2_5")],na.rm=TRUE )

y <- as.matrix(subset(x1,select = "HRSWRK"))
x <- as.matrix(subset(x1,select =c("WAGES","non_labour_INC","AGEGRP","MARST","PR","IMMSTAT","ROOM","HDG


cvfit <- cv.glmnet(x, y,type.measure="deviance", alpha=1, nlambda=100)
plot(cvfit)
```




```
cvfit$lambda.min
```

```
## [1] 0.001450549
```

```
#coefficients
coef(cvfit, s = "lambda.min")
```

```
## 11 x 1 sparse Matrix of class "dgCMatrix"
##                            s1
## (Intercept)    -6.651958e+01
## WAGES           7.268674e-05
## non_labour_INC  6.707122e-05
## AGEGRP          1.695935e+01
## MARST          -7.128583e-01
```

```
## PR                .
## IMMSTAT        -5.593850e-01
## ROOM           -2.292286e-02
## HDGREE          1.724998e-01
## CHILD_NUM      -9.048989e-01
## AGE_SQURE      -7.425986e-01
```

Based on the result, age, number of children, marital status, immigrant status and education affect the labor supply and should be included in the model while the values of wage,province, non-labor income and number of rooms are almost zero, thus they should be excluded from the model.The reason for province is that we already limited our dataset to consider only one province, so the model cannot compare different provinces with each other.

## Regression Analysis on Labor Force Participation Using Linear Probability Model

```r
new_dataset <- subset(census_2006,select = c("WAGES", "CFINC","HRSWRK","SEX","AGEGRP","PKID0_1","PKID15_

range_f <- c(1000,3500,6000,8500,11000,13500,16000,18500,22500,27500,32500,37500,42500,47500,52500,5750

for (i in 1:58086){
  new_dataset$Cfamily_income[i]= range_f[new_dataset$CFINC[i]]
  new_dataset$non_labour_INC[i]=new_dataset$Cfamily_income[i]-new_dataset$WAGES[i]

}

new_dataset$dummy_labour <- ifelse(new_dataset$LFACT==1&2,1,0)
new_dataset$CHILD_NUM <- rowSums(new_dataset[,c("PKID0_1","PKID2_5","PKID6_14","PKID15_24","PKID2_5")],

linear_prb_model <- lm(dummy_labour~WAGES+non_labour_INC+AGEGRP+AGE_SQURE+IMMSTAT+MARST+HDGREE+INVST+RO

options("scipen" = 100, "digits"=4)
summary(linear_prb_model )
```

```
##
## Call:
## lm(formula = dummy_labour ~ WAGES + non_labour_INC + AGEGRP +
##     AGE_SQURE + IMMSTAT + MARST + HDGREE + INVST + ROOM + +ABOID +
##     KOL + CHILD_NUM, data = new_dataset)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -1.947 -0.479  0.202  0.302  1.028
##
## Coefficients:
##                      Estimate    Std. Error t value             Pr(>|t|)
## (Intercept)    -0.7544445684  0.0466656547  -16.17 < 0.0000000000000002 ***
## WAGES           0.0000026407  0.0000000457   57.84 < 0.0000000000000002 ***
## non_labour_INC  0.0000008140  0.0000000396   20.54 < 0.0000000000000002 ***
## AGEGRP          0.2494307506  0.0070540324   35.36 < 0.0000000000000002 ***
```

```
## AGE_SQURE      -0.0113647943  0.0002939344  -38.66 < 0.0000000000000002 ***
## IMMSTAT        -0.0166837996  0.0038151134   -4.37    0.000012271813993 ***
## MARST          -0.0020950902  0.0023283056   -0.90                 0.37
## HDGREE          0.0021869694  0.0002895857    7.55    0.000000000000043 ***
## INVST          -0.0000018541  0.0000000458  -40.51 < 0.0000000000000002 ***
## ROOM            0.0012337154  0.0007834509    1.57                 0.12
## ABOID           0.0219644429  0.0019390367   11.33 < 0.0000000000000002 ***
## KOL            -0.0293400228  0.0026253844  -11.18 < 0.0000000000000002 ***
## CHILD_NUM      -0.0028525470  0.0006852195   -4.16    0.000031458974100 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.428 on 58073 degrees of freedom
## Multiple R-squared:  0.109,  Adjusted R-squared:  0.109
## F-statistic:  591 on 12 and 58073 DF,  p-value: <0.0000000000000002
```

In the linear probability model, all variables expect marital status, the number of rooms and porvince are statistically significantly different from zero

## Regression Analysis on Labor Force Participation Using Probit Model

```r
myprobit <- glm(dummy_labour~WAGES+non_labour_INC+AGEGRP+AGE_SQURE+IMMSTAT+MARST+HDGREE+INVST+ROOM++ABOI

summary(myprobit)
```

```
##
## Call:
## glm(formula = dummy_labour ~ WAGES + non_labour_INC + AGEGRP +
##     AGE_SQURE + IMMSTAT + MARST + HDGREE + INVST + ROOM + +ABOID +
##     KOL + CHILD_NUM, family = binomial(link = "probit"), data = new_dataset)
##
## Coefficients:
##                    Estimate   Std. Error z value             Pr(>|z|)
## (Intercept)    -3.483143014  0.147540371  -23.61 < 0.0000000000000002 ***
## WAGES           0.000013202  0.000000242   54.54 < 0.0000000000000002 ***
## non_labour_INC  0.000002377  0.000000134   17.71 < 0.0000000000000002 ***
## AGEGRP          0.663398288  0.022389530   29.63 < 0.0000000000000002 ***
## AGE_SQURE      -0.030382147  0.000933936  -32.53 < 0.0000000000000002 ***
## IMMSTAT        -0.031344400  0.012319776   -2.54              0.01095 *
## MARST          -0.001207119  0.007548991   -0.16              0.87296
## HDGREE          0.007682414  0.001082514    7.10      0.0000000000013 ***
## INVST          -0.000010860  0.000000247  -44.01 < 0.0000000000000002 ***
## ROOM            0.005895683  0.002536993    2.32              0.02013 *
## ABOID           0.059014527  0.005919525    9.97 < 0.0000000000000002 ***
## KOL            -0.076210898  0.008208802   -9.28 < 0.0000000000000002 ***
## CHILD_NUM      -0.007841309  0.002135369   -3.67              0.00024 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
```

```
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 69883  on 58085  degrees of freedom
## Residual deviance: 61971  on 58073  degrees of freedom
## AIC: 61997
##
## Number of Fisher Scoring iterations: 11
```

In the probit model, all variables except immigrant status, marital status, province, and the number of rooms are statically significant.

## Gender-Based Probit Model Analysis (women)

```r
new_dataset$SEX <- census_2006$SEX
new_dataset_women <-subset(new_dataset,SEX==1)

myprobit_WOMEN <- glm(dummy_labour~WAGES+non_labour_INC+AGEGRP+AGE_SQURE+IMMSTAT+MARST+HDGREE+INVST+ROO

summary(myprobit_WOMEN )
```

```
##
## Call:
## glm(formula = dummy_labour ~ WAGES + non_labour_INC + AGEGRP +
##     AGE_SQURE + IMMSTAT + MARST + HDGREE + INVST + ROOM + +ABOID +
##     KOL + CHILD_NUM, family = binomial(link = "probit"), data = new_dataset_women)
##
## Coefficients:
##                    Estimate   Std. Error z value             Pr(>|z|)
## (Intercept)    -2.810679752  0.201501822  -13.95 < 0.0000000000000002 ***
## WAGES           0.000021172  0.000000431   49.11 < 0.0000000000000002 ***
## non_labour_INC  0.000001922  0.000000178   10.77 < 0.0000000000000002 ***
## AGEGRP          0.501814080  0.031024502   16.17 < 0.0000000000000002 ***
## AGE_SQURE      -0.023658340  0.001300344  -18.19 < 0.0000000000000002 ***
## IMMSTAT        -0.055058392  0.016562719   -3.32              0.00089 ***
## MARST           0.053042952  0.009853369    5.38      0.0000000731609 ***
## HDGREE          0.005465221  0.001507560    3.63              0.00029 ***
## INVST          -0.000019254  0.000000444  -43.39 < 0.0000000000000002 ***
## ROOM            0.007512061  0.003451149    2.18              0.02950 *
## ABOID           0.053039380  0.007868837    6.74      0.0000000000158 ***
## KOL            -0.076806928  0.010806959   -7.11      0.0000000000012 ***
## CHILD_NUM      -0.011441185  0.002749338   -4.16      0.0000316256577 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 39177  on 30346  degrees of freedom
## Residual deviance: 33994  on 30334  degrees of freedom
## AIC: 34020
##
## Number of Fisher Scoring iterations: 7
```

# Gender-Based Probit Model Analysis (Men)

```r
new_dataset$SEX <- census_2006$SEX
new_dataset_men <-subset(new_dataset,SEX==2)

myprobit_MEN <- glm(dummy_labour~WAGES+non_labour_INC+AGEGRP+AGE_SQURE+IMMSTAT+MARST+HDGREE+INVST+ROOM+

summary(myprobit_MEN )
```

```
##
## Call:
## glm(formula = dummy_labour ~ WAGES + non_labour_INC + AGEGRP +
##     AGE_SQURE + IMMSTAT + MARST + HDGREE + INVST + ROOM + +ABOID +
##     KOL + CHILD_NUM, family = binomial(link = "probit"), data = new_dataset_men)
##
## Coefficients:
##                   Estimate   Std. Error z value             Pr(>|z|)
## (Intercept)    -4.263069060  0.222614887  -19.15 < 0.0000000000000002 ***
## WAGES           0.000008815  0.000000303   29.05 < 0.0000000000000002 ***
## non_labour_INC  0.000003222  0.000000212   15.23 < 0.0000000000000002 ***
## AGEGRP          0.889018748  0.033245240   26.74 < 0.0000000000000002 ***
## AGE_SQURE      -0.040335278  0.001382723  -29.17 < 0.0000000000000002 ***
## IMMSTAT        -0.008421851  0.019018992   -0.44                 0.66
## MARST          -0.117576722  0.012291283   -9.57 < 0.0000000000000002 ***
## HDGREE          0.007458302  0.001592539    4.68      0.0000028232297 ***
## INVST          -0.000005681  0.000000304  -18.70 < 0.0000000000000002 ***
## ROOM            0.005616767  0.003863710    1.45                 0.15
## ABOID           0.063512452  0.009171620    6.92      0.0000000000044 ***
## KOL            -0.063023126  0.013132943   -4.80      0.0000015957311 ***
## CHILD_NUM       0.006341670  0.004144913    1.53                 0.13
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 29674  on 27738  degrees of freedom
## Residual deviance: 26289  on 27726  degrees of freedom
## AIC: 26315
##
## Number of Fisher Scoring iterations: 6
```

The results for women and men are different. For both men and women the number of rooms is insignificant in terms of statistics. While the number of children and immigrant status are a key variable for women, they should not be included for men.