R Analysis Example Replication C11

```
# Chapter 11 Longitudinal Analysis HRS data
# Use data sets previously prepared in SAS for this chapter to reduce code burden in R
# Complete Case 1 Wave
# 11.3.1 Example: Descriptive Estimation at a Single Wave, Complete Case Analysis Table 11.2
library(survey)
library(haven)
#library (sas7bdat)

hrs_1wave <- read_sas("P:/ASDA 2/Data sets/HRS 2012/HRS 2006_2012 Longitudinal File/cc_1wave.sas7bdat")
names(hrs_1wave)

svyhrs_cc_1 <- svydesign(strata=~STRATUM, id=~SECU, weights=~KWGTR, data=hrs_1wave,nest=T)
ex11_1 <- svymean(~ln_inc08, design=svyhrs_cc_1, se=T, ci=T, keep.vars=T, na.rm=T)
# Exponent of Mean, se, and CI'S
exp(ex11_1)
exp(confint(ex11_1))

# Adjusted Weight 1 Wave
hrs_1wave_adj <- read_sas("P:/ASDA 2/Data sets/HRS 2012/HRS 2006_2012 Longitudinal File/adj_wgt_1wave.sas7bdat")
names(hrs_1wave_adj)
svyhrs_adj_1 <- svydesign(strata=~STRATUM, id=~SECU, weights=~adj_kwgtr, data=hrs_1wave_adj,nest=T)
ex11_1_adj <- svymean(~ln_inc08, design=svyhrs_adj_1, se=T, ci=T, keep.vars=T, na.rm=T)
# Exponent of Mean, se, and CI'S
exp(ex11_1_adj)
exp(confint(ex11_1_adj))

# Multiple Imputation 1 Wave
# Use SAS data set already prepared for this example

b <- read_sas("P:/ASDA 2/Data sets/HRS 2012/HRS 2006_2012 Longitudinal File/wt_deciles_1wave.sas7bdat")
names(b)

b$selfrhealth_06 <- factor(b$selfrhealth_06)
b$marcat_06 <- factor(b$marcat_06)
b$racecat <- factor(b$racecat)
b$edcat <- factor(b$edcat)
b$STRATUM <- factor(b$STRATUM)
b$kwgtr_dec <- factor(b$kwgtr_dec)

# subset variables by number position in data set
(hrs_mi_1wave_sub  <- b[, c(10,11,13,41,42,44,45,46,65,66,67,68,71)])
summary(hrs_mi_1wave_sub)

# use mice to impute missing data
library(mice)

ini <- mice(hrs_mi_1wave_sub, maxiter=0)
summary(ini)

# add a predictor matrix to control imputation model predictors for each imputed variable
pred <- ini$predictorMatrix
pred[,"KWGTR"] <- 0
pred[,"SECU"] <- 0
pred[,"kwgtr_dec"] <- 1
pred

imphrs1wave <- mice(hrs_mi_1wave_sub, m=5, pred=pred, seed=41279)
(imphrs1wave)
```

```
# convert mids to data useable for work in mitools
library(mitools)
hrs_1w_imp <- imputationList(lapply(1:5, complete, x=imphrs1wave))
hrs_1w_imp
summary(hrs_1w_imp)

# set survey design
library(survey)
deshrs_1wave <- svydesign(id=~SECU, strat=~STRATUM, weight=~kwgtr_dec, data=hrs_1w_imp, nest=TRUE)
(deshrs_1wave)

hrs_1w_mean <- with(deshrs_1wave, svymean(~(ln_inc08), se=T, na.rm=T, ci=T ))
hrs_1w_mean

# Use MIcombine for overall combined and design-adjusted mean/se
summary(hrs_1w_comb <- MIcombine(hrs_1w_mean))
# exponent of results for log income ( using decile weight)
exp(10.41648)
exp(10.36628)
exp(10.46668)
```

 # Multiple Imputation using a Selection Model Not Available in R

```
# Complete Case 2 Waves
hrs_2wave <- read_sas("P:/ASDA 2/Data sets/HRS 2012/HRS 2006_2012 Longitudinal File/cc_2waves.sas7bdat")
names(hrs_2wave)
svyhrs_cc_2 <- svydesign(strata=~STRATUM, id=~SECU, weights=~KWGTR, data=hrs_2wave,nest=T)
ex11_2 <- svymean(~incdiff_06_10, design=svyhrs_cc_2, se=T, ci=T, keep.vars=T, na.rm=T)
show(ex11_2)
confint(ex11_2)

# Adjusted Weight 2 Wave
hrs_2waves_adj <- read_sas("P:/ASDA 2/Data sets/HRS 2012/HRS 2006_2012 Longitudinal File/adj_wgt_2waves.sas7bdat")
names(hrs_2waves_adj)
svyhrs_adj_1 <- svydesign(strata=~STRATUM, id=~SECU, weights=~adj_kwgtr, data=hrs_2waves_adj,nest=T)
show(ex11_2_adj <- svymean(~incdiff_06_10, design=svyhrs_adj_1, se=T, ci=T, keep.vars=T, na.rm=T))
confint(ex11_2_adj)

# Multiple Imputation for 2 Waves of Data

# Multiple Imputation for 2 Waves, SAS data set already prepared for this example
hrs_a  <- read.table(file="P:/ASDA 2/Data sets/HRS 2012/HRS 2006_2012 Longitudinal File/wt_deciles_2waves.csv", sep = ",", header = T,
as.is=T)
names(hrs_a)
summary(hrs_a)

hrs_a$selfrhealth_06 <- factor(hrs_a$selfrhealth_06)
hrs_a$marcat_06 <- factor(hrs_a$marcat_06)
hrs_a$racecat <- factor(hrs_a$racecat)
hrs_a$edcat <- factor(hrs_a$edcat)
hrs_a$STRATUM <- factor(hrs_a$STRATUM)
hrs_a$kwgtr_dec <- factor(hrs_a$kwgtr_dec)

# subset variables by number position in data set
# subset key variables for imputation
(hrs_mi_2waves_sub  <- hrs_a [, c(10,11,13,41,42,44,45,46,65,66,67,69,71)])
names(hrs_mi_2waves_sub)

# use mice to impute missing data
library(mice)
# Dry run to prepare the predictor matrix
```

```
ini <- mice(hrs_mi_2waves_sub, maxiter=0)
summary(ini)

# add a predictor matrix to control imputation model predictors for each imputed variable
pred <- ini$predictorMatrix
pred[,"KWGTR"] <- 0
pred[,"SECU"] <- 0
pred[,"kwgtr_dec"] <- 1
pred[,"ln_inc06"] <- 1
pred

# use same variables as from C11 Stata example, use norm.nob for imputation method
imphrs2waves1 <- mice(hrs_mi_2waves_sub, pred=pred, m=5, seed=41279, method="norm.nob", print=FALSE)
imphrs2waves1

# convert mids (MI data) to data useable for work in mitools
library(mitools)
hrs_2w_imp1 <- imputationList(lapply(1:5, complete, x=imphrs2waves1))
hrs_2w_imp1

# set survey design
# Use 2006 individual weight for survey design setup
library(survey)
library(mitools)

deshrs_2waves <- svydesign(id=~SECU, strat=~STRATUM, weight=~KWGTR, data=(hrs_2w_imp1), nest=TRUE)
deshrs_2waves

deshrs_2waves <- update(deshrs_2waves, ln_inc10a=ifelse(ln_inc10 > 14.92, 14.92, ln_inc10), inc10=exp(ln_inc10a), inc06=exp(ln_inc06))
deshrs_2waves <- update(deshrs_2waves, new_chg0610=(inc10 - inc06))
deshrs_2waves <- update(deshrs_2waves, new_chg0610a=ifelse(new_chg0610 < -12300000, -12300000, new_chg0610),
new_chg0610b=ifelse(new_chg0610a > 2062968, 2062968, new_chg0610a))

hrs_2w_meandiff <- with(deshrs_2waves, svymean(~(new_chg0610b), se=T, ci=T ))
hrs_2w_meandiff

# Use MIcombine for overall combined and design-adjusted mean/se
(hrs_2w_comb <- MIcombine(hrs_2w_meandiff))
summary(hrs_2w_comb)

# Calibration Method
hrs_2waves_cal <- read_sas("P:/ASDA 2/Data sets/HRS 2012/HRS 2006_2012 Longitudinal File/calibration_2waves.sas7bdat")
names(hrs_2waves_cal)
# need subset of cases without missing weight variable
hrs_2waves_calsub <- hrs_2waves_cal[ which(hrs_2waves_cal$kwgtr_cal > 0),]
summary(hrs_2waves_calsub$kwgtr_cal)
svyhrs_cal_2 <- svydesign(strata=~STRATUM, id=~SECU, weights=~kwgtr_cal, data=hrs_2waves_calsub,nest=T, na.rm=T)
show(ex11_2_cal <- svymean(~incdiff_06_10, design=svyhrs_cal_2, se=T, ci=T, keep.vars=T, na.rm=T))
confint(ex11_2_cal)
```

```
# 3+ Waves of Data ##############################################################################################
# Example 11.3.3 Weighted Multilevel Modeling not available in R Survey Package
# Example 11.3.3.1 Veiga Method for multi-level modeling not available in R Survey Package


# Example 11.3.4 Weighted GEE Analysis using Geepack from R (See geepack.pdf for details)

# read data from SAS
# install and load package
library(geepack)

hrs_3w_gee <- read_sas("P:/ASDA 2/Data sets/HRS 2012/HRS 2006_2012 Longitudinal File/wgt_gee_3pwaves.sas7bdat")
names(hrs_3w_gee)

# set factor variables
hrs_3w_gee$GENDER <- as.factor(hrs_3w_gee$GENDER)
hrs_3w_gee$STRATUM <- as.factor(hrs_3w_gee$STRATUM)
hrs_3w_gee$year <- as.factor(hrs_3w_gee$year)

# model for Example 11.3.4
# model formula
mf <- formula(ln_inc~yrssince06 + GENDER + yrs06sq + (yrssince06*GENDER) + (yrs06sq*GENDER) + STRATUM)
mf

# Run model using geeglm with weight
ex11_3_4 <- geeglm(mf, data=hrs_3w_gee, id=newid_num, weight=casewt, family=gaussian("identity"), corstr="exchangeable")
summary(ex11_3_4)
```

Output R Analysis Example Replication C11

```
> # Complete Case 1 Wave
> # 11.3.1 Example: Descriptive Estimation at a Single Wave, Complete Case Analysis Table 11.2
> library(survey)
> library(haven)
> #library (sas7bdat)
>
> hrs_1wave <- read_sas("P:/ASDA 2/Data sets/HRS 2012/HRS 2006_2012 Longitudinal File/cc_1wave.sas7bdat")
> names(hrs_1wave)
 [1] "HHID"           "PN"             "KFINR"          "KCO01"          "KCO10"          "KCO70"          "GENDER"         "HISPANIC"
"SCHLYRS"        "SECU"           "STRATUM"        "KMARST"         "KWGTR"
[14] "KWHYORWT"       "LFINR"          "LMARST"         "LWGTR"          "MFINR"          "MMARST"         "MWGTR"          "NFINR"
"NMARST"         "NWGTR"          "H8ATOTA"        "H9ATOTA"        "H10ATOTA"
[27] "H11ATOTA"       "H8ITOT"         "H9ITOT"         "H10ITOT"        "H11ITOT"        "LCO01"          "LCO10"          "LCO70"
"MCO01"          "MCO10"          "MCO70"          "NCO01"          "NCO10"
[40] "NCO70"          "marcat_06"      "diabetes_06"    "numfalls24_06"  "arthritis_06"   "selfrhealth_06" "age_06"
"marcat_08"      "diabetes_08"    "numfalls24_08"  "arthritis_08"   "selfrhealth_08" "age_08"
[53] "marcat_10"      "diabetes_10"    "numfalls24_10"  "arthritis_10"   "selfrhealth_10" "age_10"         "marcat_12"
"diabetes_12"    "numfalls24_12"  "arthritis_12"   "selfrhealth_12" "age_12"         "edcat"
[66] "racecat"        "ln_inc06"       "ln_inc08"       "ln_inc10"       "ln_inc12"
>
> svyhrs_cc_1 <- svydesign(strata=~STRATUM, id=~SECU, weights=~KWGTR, data=hrs_1wave,nest=T)
> ex11_1 <- svymean(~ln_inc08, design=svyhrs_cc_1, se=T, ci=T, keep.vars=T, na.rm=T)
> # Exponent of Mean, se, and CI'S
> exp(ex11_1)
          mean     SE
ln_inc08 34224 0.0263
> exp(confint(ex11_1))
           2.5 %   97.5 %
ln_inc08 32504.8 36034.07
```

```
> # Adjusted Weight 1 Wave

> hrs_1wave_adj <- read_sas("P:/ASDA 2/Data sets/HRS 2012/HRS 2006_2012 Longitudinal File/adj_wgt_1wave.sas7bdat")

> names(hrs_1wave_adj)
 [1] "HHID"           "PN"             "KFINR"          "KC001"          "KC010"          "KC070"          "GENDER"         "HISPANIC"
"SCHLYRS"        "SECU"           "STRATUM"        "KMARST"         "KWGTR"
[14] "KWHYORWT"       "LFINR"          "LMARST"         "LWGTR"          "MFINR"          "MMARST"         "MWGTR"          "NFINR"
"NMARST"         "NWGTR"          "H8ATOTA"        "H9ATOTA"        "H10ATOTA"
[27] "H11ATOTA"       "H8ITOT"         "H9ITOT"         "H10ITOT"        "H11ITOT"        "LC001"          "LC010"          "LC070"
"MC001"          "MC010"          "MC070"          "NC001"          "NC010"
[40] "NC070"          "marcat_06"      "diabetes_06"    "numfalls24_06"  "arthritis_06"   "selfrhealth_06" "age_06"
"marcat_08"      "diabetes_08"    "numfalls24_08"  "arthritis_08"   "selfrhealth_08" "age_08"
[53] "marcat_10"      "diabetes_10"    "numfalls24_10"  "arthritis_10"   "selfrhealth_10" "age_10"         "marcat_12"
"diabetes_12"    "numfalls24_12"  "arthritis_12"   "selfrhealth_12" "age_12"         "edcat"
[66] "racecat"        "ln_inc06"       "ln_inc08"       "ln_inc10"       "ln_inc12"       "resp08"         "_LEVEL_"        "phat"
"dec"            "mean_phat"      "adj_kwgtr"

> svyhrs_adj_1 <- svydesign(strata=~STRATUM, id=~SECU, weights=~adj_kwgtr, data=hrs_1wave_adj,nest=T)

> ex11_1_adj <- svymean(~ln_inc08, design=svyhrs_adj_1, se=T, ci=T, keep.vars=T, na.rm=T)

> # Exponent of Mean, se, and CI'S
> exp(ex11_1_adj)
          mean     SE
ln_inc08 33309 0.0266

> exp(confint(ex11_1_adj))
            2.5 %   97.5 %
ln_inc08 31616.23 35092.42
```

```
> # Multiple Imputation 1 Wave
> # Use SAS data set already prepared for this example

> b <- read_sas("P:/ASDA 2/Data sets/HRS 2012/HRS 2006_2012 Longitudinal File/wt_deciles_1wave.sas7bdat")

> names(b)
 [1] "HHID"          "PN"            "KFINR"         "KCO01"         "KCO10"         "KCO70"         "GENDER"        "HISPANIC"
"SCHLYRS"       "SECU"          "STRATUM"       "KMARST"        "KWGTR"
[14] "KWHYORWT"      "LFINR"         "LMARST"        "LWGTR"         "MFINR"         "MMARST"        "MWGTR"         "NFINR"
"NMARST"        "NWGTR"         "H8ATOTA"       "H9ATOTA"       "H10ATOTA"
[27] "H11ATOTA"      "H8ITOT"        "H9ITOT"        "H10ITOT"       "H11ITOT"       "LC001"         "LC010"         "LC070"
"MC001"         "MC010"         "MC070"         "NC001"         "NC010"
[40] "NC070"         "marcat_06"     "diabetes_06"   "numfalls24_06" "arthritis_06"  "selfrhealth_06" "age_06"
"marcat_08"     "diabetes_08"   "numfalls24_08" "arthritis_08"  "selfrhealth_08" "age_08"
[53] "marcat_10"     "diabetes_10"   "numfalls24_10" "arthritis_10"  "selfrhealth_10" "age_10"        "marcat_12"
"diabetes_12"   "numfalls24_12" "arthritis_12"  "selfrhealth_12" "age_12"        "edcat"
[66] "racecat"       "ln_inc06"      "ln_inc08"      "ln_inc10"      "ln_inc12"      "kwgtr_dec"

> b$selfrhealth_06 <- factor(b$selfrhealth_06)
> b$marcat_06 <- factor(b$marcat_06)
> b$racecat <- factor(b$racecat)
> b$edcat <- factor(b$edcat)
> b$STRATUM <- factor(b$STRATUM)
> b$kwgtr_dec <- factor(b$kwgtr_dec)

> # subset variables by number position in data set

> (hrs_mi_1wave_sub  <- b[, c(10,11,13,41,42,44,45,46,65,66,67,68,71)])
# A tibble: 11,789 × 13
    SECU STRATUM KWGTR marcat_06 diabetes_06 arthritis_06 selfrhealth_06 age_06  edcat racecat  ln_inc06  ln_inc08 kwgtr_dec
   <dbl>  <fctr> <dbl>    <fctr>       <dbl>        <dbl>         <fctr>  <dbl> <fctr>  <fctr>     <dbl>     <dbl>    <fctr>
1      1      40  4093         1           0            0              3     70      2       2 10.596360 10.691968         5
2      2       1  7434         3           0            0              4     66      2       2  9.229064  9.236106         8
3      2       1  5217         1           0            1              4     66      4       2 11.348652 10.981914         7
4      2       1  5373         2           1            0              3     68      2       2  9.569203 17.910095         7
5      2       1  5440         2           0            0              2     58      3       2 10.918736  9.785154         7
6      2       2  5217         1           0            1              2     70      4       2 12.213053 12.057045         7
7      2       2  5778         2           0            0              2     64      4       4 10.865917 11.596430         7
8      2       2  5400         1           0            0              2     78      4       2 13.040722 13.368287         7
9      2       1  1799         2           1            0              4     68      1       1  9.244259        NA         1
10     2       1  3282         1           1            1              4     69      1       1 10.141283  9.479833         4
# ... with 11,779 more rows

> summary(hrs_mi_1wave_sub)
      SECU          STRATUM         KWGTR       marcat_06  diabetes_06      arthritis_06    selfrhealth_06     age_06       edcat
racecat     ln_inc06         ln_inc08        kwgtr_dec
 Min.   :1.000   46     : 487   Min.   :  924   1:5502    Min.   :0.0000   Min.   :0.0000   1:1211         Min.   : 52.00   1:2889
1:1039   Min.   : 0.000   Min.   : 0.000   2      :1323
 1st Qu.:1.000   33     : 423   1st Qu.: 2433   2:5802    1st Qu.:0.0000   1st Qu.:0.0000   2:3285         1st Qu.: 62.00   2:3903
2:8627   1st Qu.: 9.647   1st Qu.: 9.686   0      :1185
 Median :2.000   45     : 416   Median : 3723   3: 485    Median :0.0000   Median :1.0000   3:3610         Median : 69.00   3:2438
3:1863   Median :10.348   Median :10.388   8      :1184
 Mean   :1.503   40     : 380   Mean   : 4458             Mean   :0.2129   Mean   :0.6252   4:2626         Mean   : 69.52   4:2559
4: 260   Mean   :10.291   Mean   :10.327   4      :1181
 3rd Qu.:2.000   47     : 379   3rd Qu.: 5296             3rd Qu.:0.0000   3rd Qu.:1.0000   5:1057         3rd Qu.: 77.00
3rd Qu.:11.012   3rd Qu.:11.062   5      :1179
 Max.   :2.000   29     : 357   Max.   :17035             Max.   :1.0000   Max.   :1.0000                  Max.   :104.00
Max.   :17.049   Max.   :17.910   6      :1179
                 (Other):9347
NA's   :1215     (Other):4558
```

```
> # use mice to impute missing data, dry run first
> library(mice)
> ini <- mice(hrs_mi_1wave_sub, maxiter=0)
 iter imp variable
  1   1  ln_inc08
  1   2  ln_inc08
  1   3  ln_inc08
  1   4  ln_inc08
  1   5  ln_inc08
  2   1  ln_inc08
  2   2  ln_inc08
  2   3  ln_inc08
  2   4  ln_inc08
  2   5  ln_inc08
  3   1  ln_inc08
  3   2  ln_inc08
  3   3  ln_inc08
  3   4  ln_inc08
  3   5  ln_inc08
  4   1  ln_inc08
  4   2  ln_inc08
  4   3  ln_inc08
  4   4  ln_inc08
  4   5  ln_inc08
  5   1  ln_inc08
  5   2  ln_inc08
  5   3  ln_inc08
  5   4  ln_inc08
  5   5  ln_inc08


> summary(ini)
Multiply imputed data set
Call:
mice(data = hrs_mi_1wave_sub, maxiter = 0)
Number of multiple imputations:  5
Missing cells per column:
          SECU        STRATUM          KWGTR      marcat_06    diabetes_06   arthritis_06 selfrhealth_06         age_06          edcat
racecat       ln_inc06       ln_inc08      kwgtr_dec
             0              0              0              0              0              0              0              0              0
0              0           1215              0
Imputation methods:
          SECU        STRATUM          KWGTR      marcat_06    diabetes_06   arthritis_06 selfrhealth_06         age_06          edcat
racecat       ln_inc06       ln_inc08      kwgtr_dec
            ""             ""             ""             ""             ""             ""             ""             ""             ""
""             ""          "pmm"             ""
VisitSequence:
ln_inc08
      12
PredictorMatrix:
               SECU STRATUM KWGTR marcat_06 diabetes_06 arthritis_06 selfrhealth_06 age_06 edcat racecat ln_inc06 ln_inc08 kwgtr_dec
SECU              0       0     0         0           0            0              0      0     0       0        0        0         0
STRATUM           0       0     0         0           0            0              0      0     0       0        0        0         0
KWGTR             0       0     0         0           0            0              0      0     0       0        0        0         0
marcat_06         0       0     0         0           0            0              0      0     0       0        0        0         0
diabetes_06       0       0     0         0           0            0              0      0     0       0        0        0         0
arthritis_06      0       0     0         0           0            0              0      0     0       0        0        0         0
selfrhealth_06    0       0     0         0           0            0              0      0     0       0        0        0         0
age_06            0       0     0         0           0            0              0      0     0       0        0        0         0
edcat             0       0     0         0           0            0              0      0     0       0        0        0         0
```

```
racecat           0       0     0         0           0           0              0      0     0       0        0        0         0
ln_inc06          0       0     0         0           0           0              0      0     0       0        0        0         0
ln_inc08          1       1     1         1           1           1              1      1     1       1        1        0         1
kwgtr_dec         0       0     0         0           0           0              0      0     0       0        0        0         0
Random generator seed value:  NA

> # add a predictor matrix to control imputation model predictors for each imputed variable
> pred <- ini$predictorMatrix
> pred[,"KWGTR"] <- 0
> pred[,"SECU"] <- 0
> pred[,"kwgtr_dec"] <- 1
> pred
               SECU STRATUM KWGTR marcat_06 diabetes_06 arthritis_06 selfrhealth_06 age_06 edcat racecat ln_inc06 ln_inc08 kwgtr_dec
SECU              0       0     0         0           0           0              0      0     0       0        0        0         1
STRATUM           0       0     0         0           0           0              0      0     0       0        0        0         1
KWGTR             0       0     0         0           0           0              0      0     0       0        0        0         1
marcat_06         0       0     0         0           0           0              0      0     0       0        0        0         1
diabetes_06       0       0     0         0           0           0              0      0     0       0        0        0         1
arthritis_06      0       0     0         0           0           0              0      0     0       0        0        0         1
selfrhealth_06    0       0     0         0           0           0              0      0     0       0        0        0         1
age_06            0       0     0         0           0           0              0      0     0       0        0        0         1
edcat             0       0     0         0           0           0              0      0     0       0        0        0         1
racecat           0       0     0         0           0           0              0      0     0       0        0        0         1
ln_inc06          0       0     0         0           0           0              0      0     0       0        0        0         1
ln_inc08          0       1     0         1           1           1              1      1     1       1        1        0         1
kwgtr_dec         0       0     0         0           0           0              0      0     0       0        0        0         1
>
> imphrs1wave <- mice(hrs_mi_1wave_sub, m=5, pred=pred, seed=41279)

 iter imp variable
  1   1  ln_inc08
  1   2  ln_inc08
  1   3  ln_inc08
  1   4  ln_inc08
  1   5  ln_inc08
  2   1  ln_inc08
  2   2  ln_inc08
  2   3  ln_inc08
  2   4  ln_inc08
  2   5  ln_inc08
  3   1  ln_inc08
  3   2  ln_inc08
  3   3  ln_inc08
  3   4  ln_inc08
  3   5  ln_inc08
  4   1  ln_inc08
  4   2  ln_inc08
  4   3  ln_inc08
  4   4  ln_inc08
  4   5  ln_inc08
  5   1  ln_inc08
  5   2  ln_inc08
  5   3  ln_inc08
  5   4  ln_inc08
  5   5  ln_inc08
```

```
> (imphrs1wave)
Multiply imputed data set
Call:
mice(data = hrs_mi_1wave_sub, m = 5, predictorMatrix = pred,
    seed = 41279)
Number of multiple imputations:  5
Missing cells per column:
          SECU        STRATUM          KWGTR      marcat_06    diabetes_06   arthritis_06 selfrhealth_06         age_06          edcat
racecat       ln_inc06       ln_inc08      kwgtr_dec
             0              0              0              0              0              0              0              0              0
0              0           1215              0
Imputation methods:
          SECU        STRATUM          KWGTR      marcat_06    diabetes_06   arthritis_06 selfrhealth_06         age_06          edcat
racecat       ln_inc06       ln_inc08      kwgtr_dec
            ""             ""             ""             ""             ""             ""             ""             ""             ""
""             ""          "pmm"             ""
VisitSequence:
ln_inc08
      12
PredictorMatrix:
              SECU STRATUM KWGTR marcat_06 diabetes_06 arthritis_06 selfrhealth_06 age_06 edcat racecat ln_inc06 ln_inc08 kwgtr_dec
SECU             0       0     0         0           0            0              0      0     0       0        0        0         0
STRATUM          0       0     0         0           0            0              0      0     0       0        0        0         0
KWGTR            0       0     0         0           0            0              0      0     0       0        0        0         0
marcat_06        0       0     0         0           0            0              0      0     0       0        0        0         0
diabetes_06      0       0     0         0           0            0              0      0     0       0        0        0         0
arthritis_06     0       0     0         0           0            0              0      0     0       0        0        0         0
selfrhealth_06   0       0     0         0           0            0              0      0     0       0        0        0         0
age_06           0       0     0         0           0            0              0      0     0       0        0        0         0
edcat            0       0     0         0           0            0              0      0     0       0        0        0         0
racecat          0       0     0         0           0            0              0      0     0       0        0        0         0
ln_inc06         0       0     0         0           0            0              0      0     0       0        0        0         0
ln_inc08         0       1     0         1           1            1              1      1     1       1        1        0         1
kwgtr_dec        0       0     0         0           0            0              0      0     0       0        0        0         0
Random generator seed value:  41279

> # convert mids to data useable for work in mitools
> library(mitools)
> hrs_1w_imp <- imputationList(lapply(1:5, complete, x=imphrs1wave))
> hrs_1w_imp
MI data with 5 datasets
Call: imputationList(lapply(1:5, complete, x = imphrs1wave))
> summary(hrs_1w_imp)
            Length Class  Mode
imputations 5      -none- list
call        2      -none- call

> # set survey design
> library(survey)
> deshrs_1wave <- svydesign(id=~SECU, strat=~STRATUM, weight=~kwgtr_dec, data=hrs_1w_imp, nest=TRUE)
Error in 1/as.matrix(weights) : non-numeric argument to binary operator
> (deshrs_1wave)
Multiple (5) imputations: svydesign(id = ~SECU, strat = ~STRATUM, weight = ~kwgtr_dec,
    data = hrs_1w_imp, nest = TRUE)

> hrs_1w_mean <- with(deshrs_1wave, svymean(~(ln_inc08), se=T, na.rm=T, ci=T ))
> hrs_1w_mean
[[1]]
           mean     SE
ln_inc08 10.418 0.0253
```

```
[[2]]
           mean     SE
ln_inc08 10.413 0.0255

[[3]]
           mean     SE
ln_inc08 10.419 0.0252

[[4]]
           mean     SE
ln_inc08 10.419 0.0249

[[5]]
           mean    SE
ln_inc08 10.413 0.026

attr(,"call")
with(deshrs_1wave, svymean(~(ln_inc08), se = T, na.rm = T, ci = T))

> # Use MIcombine for overall combined and design-adjusted mean/se
> summary(hrs_1w_comb <- MIcombine(hrs_1w_mean))
Multiple imputation results:
      with(deshrs_1wave, svymean(~(ln_inc08), se = T, na.rm = T, ci = T))
      MIcombine.default(hrs_1w_mean)
          results         se   (lower   upper) missInfo
ln_inc08 10.41648 0.02561051 10.36628 10.46668      2 %
> # exponent of results for log income ( using decile weight)
> exp(10.41648)
[1] 33405.64
> exp(10.36628)
[1] 31770.07
> exp(10.46668)
[1] 35125.41
```

> # Multiple Imputation using a Selection Model Not Available in R

```
> ######################### 2 Waves of Data ###################################################################
> # Complete Case 2 Waves
> hrs_2wave <- read_sas("P:/ASDA 2/Data sets/HRS 2012/HRS 2006_2012 Longitudinal File/cc_2waves.sas7bdat")
> names(hrs_2wave)

 [1] "HHID"           "PN"             "KFINR"          "KCO01"          "KCO10"          "KCO70"          "GENDER"         "HISPANIC"
"SCHLYRS"        "SECU"           "STRATUM"        "KMARST"         "KWGTR"

[14] "KWHYORWT"       "LFINR"          "LMARST"         "LWGTR"          "MFINR"          "MMARST"         "MWGTR"          "NFINR"
"NMARST"         "NWGTR"          "H8ATOTA"        "H9ATOTA"        "H10ATOTA"

[27] "H11ATOTA"       "H8ITOT"         "H9ITOT"         "H10ITOT"        "H11ITOT"        "LCO01"          "LCO10"          "LCO70"
"MCO01"          "MCO10"          "MCO70"          "NCO01"          "NCO10"

[40] "NCO70"          "marcat_06"      "diabetes_06"    "numfalls24_06"  "arthritis_06"   "selfrhealth_06" "age_06"
"marcat_08"      "diabetes_08"    "numfalls24_08"  "arthritis_08"   "selfrhealth_08" "age_08"

[53] "marcat_10"      "diabetes_10"    "numfalls24_10"  "arthritis_10"   "selfrhealth_10" "age_10"         "marcat_12"
"diabetes_12"    "numfalls24_12"  "arthritis_12"   "selfrhealth_12" "age_12"         "edcat"
[66] "racecat"        "ln_inc06"       "ln_inc08"       "ln_inc10"       "ln_inc12"       "incdiff_06_10"  "resp10"
> svyhrs_cc_2 <- svydesign(strata=~STRATUM, id=~SECU, weights=~KWGTR, data=hrs_2wave,nest=T)
> ex11_2 <- svymean(~incdiff_06_10, design=svyhrs_cc_2, se=T, ci=T, keep.vars=T, na.rm=T)
> show(ex11_2)
                 mean     SE
incdiff_06_10 -6551.4 1866.1
> confint(ex11_2)
                  2.5 %    97.5 %
incdiff_06_10 -10208.96 -2893.845


> # Adjusted Weight 2 Wave
> hrs_2waves_adj <- read_sas("P:/ASDA 2/Data sets/HRS 2012/HRS 2006_2012 Longitudinal File/adj_wgt_2waves.sas7bdat")
> names(hrs_2waves_adj)

 [1] "HHID"           "PN"             "KFINR"          "KCO01"          "KCO10"          "KCO70"          "GENDER"         "HISPANIC"
"SCHLYRS"        "SECU"           "STRATUM"        "KMARST"         "KWGTR"

[14] "KWHYORWT"       "LFINR"          "LMARST"         "LWGTR"          "MFINR"          "MMARST"         "MWGTR"          "NFINR"
"NMARST"         "NWGTR"          "H8ATOTA"        "H9ATOTA"        "H10ATOTA"

[27] "H11ATOTA"       "H8ITOT"         "H9ITOT"         "H10ITOT"        "H11ITOT"        "LCO01"          "LCO10"          "LCO70"
"MCO01"          "MCO10"          "MCO70"          "NCO01"          "NCO10"

[40] "NCO70"          "marcat_06"      "diabetes_06"    "numfalls24_06"  "arthritis_06"   "selfrhealth_06" "age_06"
"marcat_08"      "diabetes_08"    "numfalls24_08"  "arthritis_08"   "selfrhealth_08" "age_08"

[53] "marcat_10"      "diabetes_10"    "numfalls24_10"  "arthritis_10"   "selfrhealth_10" "age_10"         "marcat_12"
"diabetes_12"    "numfalls24_12"  "arthritis_12"   "selfrhealth_12" "age_12"         "edcat"

[66] "racecat"        "ln_inc06"       "ln_inc08"       "ln_inc10"       "ln_inc12"       "incdiff_06_10"  "resp10"         "_LEVEL_"
"phat1"          "dec"            "mean_phat"      "adj_kwgtr"
> svyhrs_adj_1 <- svydesign(strata=~STRATUM, id=~SECU, weights=~adj_kwgtr, data=hrs_2waves_adj,nest=T)
> show(ex11_2_adj <- svymean(~incdiff_06_10, design=svyhrs_adj_1, se=T, ci=T, keep.vars=T, na.rm=T))
               mean   SE
incdiff_06_10 -6120 1703
> confint(ex11_2_adj)
                 2.5 %   97.5 %
incdiff_06_10 -9457.72 -2782.22
```

```
# Multiple Imputation for 2 Waves of Data

> # Multiple Imputation for 2 Waves, SAS data set already prepared for this example
> hrs_a  <- read.table(file="P:/ASDA 2/Data sets/HRS 2012/HRS 2006_2012 Longitudinal File/wt_deciles_2waves.csv", sep = ",", header =
T, as.is=T)

> names(hrs_a)
 [1] "HHID"           "PN"             "KFINR"          "KCO01"          "KCO10"
 [6] "KCO70"          "GENDER"         "HISPANIC"       "SCHLYRS"        "SECU"
[11] "STRATUM"        "KMARST"         "KWGTR"          "KWHYORWT"       "LFINR"
[16] "LMARST"         "LWGTR"          "MFINR"          "MMARST"         "MWGTR"
[21] "NFINR"          "NMARST"         "NWGTR"          "H8ATOTA"        "H9ATOTA"
[26] "H10ATOTA"       "H11ATOTA"       "H8ITOT"         "H9ITOT"         "H10ITOT"
[31] "H11ITOT"        "LCO01"          "LCO10"          "LCO70"          "MCO01"
[36] "MCO10"          "MCO70"          "NCO01"          "NCO10"          "NCO70"
[41] "marcat_06"      "diabetes_06"    "numfalls24_06"  "arthritis_06"   "selfrhealth_06"
[46] "age_06"         "marcat_08"      "diabetes_08"    "numfalls24_08"  "arthritis_08"
[51] "selfrhealth_08" "age_08"         "marcat_10"      "diabetes_10"    "numfalls24_10"
[56] "arthritis_10"   "selfrhealth_10" "age_10"         "marcat_12"      "diabetes_12"
[61] "numfalls24_12"  "arthritis_12"   "selfrhealth_12" "age_12"         "edcat"
[66] "racecat"        "ln_inc06"       "ln_inc08"       "ln_inc10"       "ln_inc12"
[71] "kwgtr_dec"      "incdiff_06_10"

> hrs_a$selfrhealth_06 <- factor(hrs_a$selfrhealth_06)
> hrs_a$marcat_06 <- factor(hrs_a$marcat_06)
> hrs_a$racecat <- factor(hrs_a$racecat)
> hrs_a$edcat <- factor(hrs_a$edcat)
> hrs_a$STRATUM <- factor(hrs_a$STRATUM)
> hrs_a$kwgtr_dec <- factor(hrs_a$kwgtr_dec)

> # subset variables by number position in data set
> # subset key variables for imputation
> hrs_mi_2waves_sub  <- hrs_a [, c(10,11,13,41,42,44,45,46,65,66,67,69,71)]

> # use mice to impute missing data
> library(mice)
> # Dry run to prepare the predictor matrix
> ini <- mice(hrs_mi_2waves_sub, maxiter=0)

 iter imp variable
  1   1  ln_inc10
  1   2  ln_inc10
  1   3  ln_inc10
  1   4  ln_inc10
  1   5  ln_inc10
  2   1  ln_inc10
  2   2  ln_inc10
  2   3  ln_inc10
  2   4  ln_inc10
  2   5  ln_inc10
  3   1  ln_inc10
  3   2  ln_inc10
  3   3  ln_inc10
  3   4  ln_inc10
  3   5  ln_inc10
  4   1  ln_inc10
  4   2  ln_inc10
  4   3  ln_inc10
  4   4  ln_inc10
  4   5  ln_inc10
  5   1  ln_inc10
```

```
  5   2  ln_inc10
  5   3  ln_inc10
  5   4  ln_inc10
  5   5  ln_inc10
> summary(ini)
Multiply imputed data set
Call:
mice(data = hrs_mi_2waves_sub, maxiter = 0)
Number of multiple imputations:  5
Missing cells per column:
          SECU        STRATUM          KWGTR      marcat_06    diabetes_06   arthritis_06
             0              0              0              0              0              0
selfrhealth_06         age_06          edcat        racecat       ln_inc06       ln_inc10
             0              0              0              0              0           2387
     kwgtr_dec
             0
Imputation methods:
          SECU        STRATUM          KWGTR      marcat_06    diabetes_06   arthritis_06
            ""             ""             ""             ""             ""             ""
selfrhealth_06         age_06          edcat        racecat       ln_inc06       ln_inc10
            ""             ""             ""             ""             ""          "pmm"
     kwgtr_dec
            ""
VisitSequence:
ln_inc10
      12
PredictorMatrix:
               SECU STRATUM KWGTR marcat_06 diabetes_06 arthritis_06 selfrhealth_06 age_06 edcat
SECU              0       0     0         0           0            0              0      0     0
STRATUM           0       0     0         0           0            0              0      0     0
KWGTR             0       0     0         0           0            0              0      0     0
marcat_06         0       0     0         0           0            0              0      0     0
diabetes_06       0       0     0         0           0            0              0      0     0
arthritis_06      0       0     0         0           0            0              0      0     0
selfrhealth_06    0       0     0         0           0            0              0      0     0
age_06            0       0     0         0           0            0              0      0     0
edcat             0       0     0         0           0            0              0      0     0
racecat           0       0     0         0           0            0              0      0     0
ln_inc06          0       0     0         0           0            0              0      0     0
ln_inc10          1       1     1         1           1            1              1      1     1
kwgtr_dec         0       0     0         0           0            0              0      0     0
               racecat ln_inc06 ln_inc10 kwgtr_dec
SECU                 0        0         0         0
STRATUM              0        0         0         0
KWGTR                0        0         0         0
marcat_06            0        0         0         0
diabetes_06          0        0         0         0
arthritis_06         0        0         0         0
selfrhealth_06       0        0         0         0
age_06               0        0         0         0
edcat                0        0         0         0
racecat              0        0         0         0
ln_inc06             0        0         0         0
ln_inc10             1        1         0         1
kwgtr_dec            0        0         0         0
Random generator seed value:  NA
>
> # add a predictor matrix to control imputation model predictors for each imputed variable
> pred <- ini$predictorMatrix
> pred[,"KWGTR"] <- 0
> pred[,"SECU"] <- 0
```

```
> pred[,"kwgtr_dec"] <- 1
> pred[,"ln_inc06"] <- 1
> pred
               SECU STRATUM KWGTR marcat_06 diabetes_06 arthritis_06 selfrhealth_06 age_06 edcat
SECU              0       0     0         0           0            0              0      0     0
STRATUM           0       0     0         0           0            0              0      0     0
KWGTR             0       0     0         0           0            0              0      0     0
marcat_06         0       0     0         0           0            0              0      0     0
diabetes_06       0       0     0         0           0            0              0      0     0
arthritis_06      0       0     0         0           0            0              0      0     0
selfrhealth_06    0       0     0         0           0            0              0      0     0
age_06            0       0     0         0           0            0              0      0     0
edcat             0       0     0         0           0            0              0      0     0
racecat           0       0     0         0           0            0              0      0     0
ln_inc06          0       0     0         0           0            0              0      0     0
ln_inc10          0       1     0         1           1            1              1      1     1
kwgtr_dec         0       0     0         0           0            0              0      0     0
               racecat ln_inc06 ln_inc10 kwgtr_dec
SECU                 0        1         0         1
STRATUM              0        1         0         1
KWGTR                0        1         0         1
marcat_06            0        1         0         1
diabetes_06          0        1         0         1
arthritis_06         0        1         0         1
selfrhealth_06       0        1         0         1
age_06               0        1         0         1
edcat                0        1         0         1
racecat              0        1         0         1
ln_inc06             0        1         0         1
ln_inc10             1        1         0         1
kwgtr_dec            0        1         0         1

> # use same variables as from C11 Stata example, use norm.nob (linear regression without Bayesian Method)

> imphrs2waves1 <- mice(hrs_mi_2waves_sub, pred=pred, m=5, seed=41279, method="norm.nob", print=FALSE)
> imphrs2waves1
Multiply imputed data set
Call:
mice(data = hrs_mi_2waves_sub, m = 5, method = "norm.nob", predictorMatrix = pred,
    printFlag = FALSE, seed = 41279)
Number of multiple imputations:  5
Missing cells per column:
          SECU        STRATUM          KWGTR      marcat_06    diabetes_06   arthritis_06
             0              0              0              0              0              0
selfrhealth_06         age_06          edcat        racecat       ln_inc06       ln_inc10
             0              0              0              0              0           2387
     kwgtr_dec
             0
Imputation methods:
          SECU        STRATUM          KWGTR      marcat_06    diabetes_06   arthritis_06
    "norm.nob"     "norm.nob"     "norm.nob"     "norm.nob"     "norm.nob"     "norm.nob"
selfrhealth_06         age_06          edcat        racecat       ln_inc06       ln_inc10
    "norm.nob"     "norm.nob"     "norm.nob"     "norm.nob"     "norm.nob"     "norm.nob"
     kwgtr_dec
    "norm.nob"
```

```
VisitSequence:
ln_inc10
      12
PredictorMatrix:
               SECU STRATUM KWGTR marcat_06 diabetes_06 arthritis_06 selfrhealth_06 age_06 edcat
SECU              0       0     0         0           0            0              0      0     0
STRATUM           0       0     0         0           0            0              0      0     0
KWGTR             0       0     0         0           0            0              0      0     0
marcat_06         0       0     0         0           0            0              0      0     0
diabetes_06       0       0     0         0           0            0              0      0     0
arthritis_06      0       0     0         0           0            0              0      0     0
selfrhealth_06    0       0     0         0           0            0              0      0     0
age_06            0       0     0         0           0            0              0      0     0
edcat             0       0     0         0           0            0              0      0     0
racecat           0       0     0         0           0            0              0      0     0
ln_inc06          0       0     0         0           0            0              0      0     0
ln_inc10          0       1     0         1           1            1              1      1     1
kwgtr_dec         0       0     0         0           0            0              0      0     0
               racecat ln_inc06 ln_inc10 kwgtr_dec
SECU                 0        0        0         0
STRATUM              0        0        0         0
KWGTR                0        0        0         0
marcat_06            0        0        0         0
diabetes_06          0        0        0         0
arthritis_06         0        0        0         0
selfrhealth_06       0        0        0         0
age_06               0        0        0         0
edcat                0        0        0         0
racecat              0        0        0         0
ln_inc06             0        0        0         0
ln_inc10             1        1        0         1
kwgtr_dec            0        0        0         0
Random generator seed value:  41279


> # convert mids (MI data) to data useable for work in mitools
> library(mitools)
> hrs_2w_imp1 <- imputationList(lapply(1:5, complete, x=imphrs2waves1))
> hrs_2w_imp1
MI data with 5 datasets
Call: imputationList(lapply(1:5, complete, x = imphrs2waves1))

> # set survey design
> # Use 2006 individual weight for survey design setup
> library(survey)
> deshrs_2waves <- svydesign(id=~SECU, strat=~STRATUM, weight=~KWGTR, data=(hrs_2w_imp1), nest=TRUE)
> deshrs_2waves
Multiple (5) imputations: svydesign(id = ~SECU, strat = ~STRATUM, weight = ~KWGTR, data = (hrs_2w_imp1),
    nest = TRUE)

> deshrs_2waves <- update(deshrs_2waves, ln_inc10a=ifelse(ln_inc10 > 14.92, 14.92, ln_inc10), inc10=exp(ln_inc10a),
inc06=exp(ln_inc06))
> deshrs_2waves <- update(deshrs_2waves, new_chg0610=(inc10 - inc06))
> deshrs_2waves <- update(deshrs_2waves, new_chg0610a=ifelse(new_chg0610 < -1230000, -1230000, new_chg0610),
new_chg0610b=ifelse(new_chg0610a > 2062968, 2062968, new_chg0610a))

> hrs_2w_meandiff <- with(deshrs_2waves, svymean(~(new_chg0610b), se=T, ci=T ))
> hrs_2w_meandiff
[[1]]
                 mean     SE
new_chg0610b -3381.8 2959.9
```

```
[[2]]
                mean     SE
new_chg0610b -2881.9 2975.4

[[3]]
                mean   SE
new_chg0610b -3312.9 2924

[[4]]
                mean     SE
new_chg0610b -3095.2 2998.6

[[5]]
                mean     SE
new_chg0610b -3570.7 2828.9

attr(,"call")
with(deshrs_2waves, svymean(~(new_chg0610b), se = T, ci = T))
>
> # Use MIcombine for overall combined and design-adjusted mean/se
> (hrs_2w_comb <- MIcombine(hrs_2w_meandiff))
Multiple imputation results:
      with(deshrs_2waves, svymean(~(new_chg0610b), se = T, ci = T))
      MIcombine.default(hrs_2w_meandiff)
               results       se
new_chg0610b -3248.509 2952.429
> summary(hrs_2w_comb)
Multiple imputation results:
      with(deshrs_2waves, svymean(~(new_chg0610b), se = T, ci = T))
      MIcombine.default(hrs_2w_meandiff)
               results       se   (lower   upper) missInfo
new_chg0610b -3248.509 2952.429 -9035.33 2538.312      1 %
```

```
> # Calibration Method
> hrs_2waves_cal <- read_sas("P:/ASDA 2/Data sets/HRS 2012/HRS 2006_2012 Longitudinal File/calibration_2waves.sas7bdat")
> names(hrs_2waves_cal)
 [1] "HHID"           "PN"             "KFINR"          "KCOO1"          "KCO1O"          "KCO7O"
 [7] "GENDER"         "HISPANIC"       "SCHLYRS"        "SECU"           "STRATUM"        "KMARST"
[13] "KWGTR"          "KWHYORWT"       "LFINR"          "LMARST"         "LWGTR"          "MFINR"
[19] "MMARST"         "MWGTR"          "NFINR"          "NMARST"         "NWGTR"          "H8ATOTA"
[25] "H9ATOTA"        "H1OATOTA"       "H11ATOTA"       "H8ITOT"         "H9ITOT"         "H1OITOT"
[31] "H11ITOT"        "LCOO1"          "LCO1O"          "LCO7O"          "MCOO1"          "MCO1O"
[37] "MCO7O"          "NCOO1"          "NCO1O"          "NCO7O"          "marcat_06"      "diabetes_06"
[43] "numfalls24_06"  "arthritis_06"   "selfrhealth_06" "age_06"         "marcat_08"      "diabetes_08"
[49] "numfalls24_08"  "arthritis_08"   "selfrhealth_08" "age_08"         "marcat_10"      "diabetes_10"
[55] "numfalls24_10"  "arthritis_10"   "selfrhealth_10" "age_10"         "marcat_12"      "diabetes_12"
[61] "numfalls24_12"  "arthritis_12"   "selfrhealth_12" "age_12"         "edcat"          "racecat"
[67] "ln_inc06"       "ln_inc10"       "_TYPE_"         "_FREQ_"         "popsize"        "sumrespwgts"
[73] "cal_adj"        "resp10"         "kwgtr_cal"      "incdiff_06_10"
> # need subset of cases without missing weight variable
> hrs_2waves_calsub <- hrs_2waves_cal[ which(hrs_2waves_cal$kwgtr_cal > 0),]
> summary(hrs_2waves_calsub$kwgtr_cal)
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
  970.3  3095.0  4611.0  5590.0  6814.0 23510.0
> svyhrs_cal_2 <- svydesign(strata=~STRATUM, id=~SECU, weights=~kwgtr_cal, data=hrs_2waves_calsub,nest=T, na.rm=T)
> show(ex11_2_cal <- svymean(~incdiff_06_10, design=svyhrs_cal_2, se=T, ci=T, keep.vars=T, na.rm=T))
                 mean     SE
incdiff_06_10 -6341.7 1780.6
> confint(ex11_2_cal)
                   2.5 %    97.5 %
incdiff_06_10 -9831.568 -2851.746
```

```
> # 3+ Waves of Data ###############################################################################################
> # Example 11.3.3 Weighted Multilevel Modeling not available in R Survey Package
> # Example 11.3.3.1 Veiga Method for Multi-level Modeling not available in R Survey Package

> # Example 11.3.4 Weighted GEE Analysis using Geepack from R (See geepack.pdf for details)
> library(geepack)
> hrs_3w_gee <- read_sas("P:/ASDA 2/Data sets/HRS 2012/HRS 2006_2012 Longitudinal File/wgt_gee_3pwaves.sas7bdat")
> names(hrs_3w_gee)
 [1] "HHID"         "PN"           "GENDER"       "SECU"         "STRATUM"      "marcat_06"    "diabetes_06"
 [8] "arthritis_06" "edcat"        "racecat"      "cumprob_case" "ln_inc"       "year"         "basewgt"
[15] "casewt"       "yrssince06"   "yrs06sq"      "newid"        "newid_num"
>
> # set factor variables
> hrs_3w_gee$GENDER <- as.factor(hrs_3w_gee$GENDER)
> hrs_3w_gee$STRATUM <- as.factor(hrs_3w_gee$STRATUM)
> hrs_3w_gee$year <- as.factor(hrs_3w_gee$year)
> # model for Example 11.3.4
> # model formula
> mf <- formula(ln_inc~yrssince06 + GENDER + yrs06sq + (yrssince06*GENDER) + (yrs06sq*GENDER) + STRATUM)
> mf
ln_inc ~ yrssince06 + GENDER + yrs06sq + (yrssince06 * GENDER) +
    (yrs06sq * GENDER) + STRATUM
> # Run model using geeglm with weight
> ex11_3_4 <- geeglm(mf, data=hrs_3w_gee, id=newid_num, weight=casewt, family=gaussian("identity"), corstr="exchangeable")
> summary(ex11_3_4)

Call:
geeglm(formula = mf, family = gaussian("identity"), data = hrs_3w_gee,
    weights = casewt, id = newid_num, corstr = "exchangeable")

 Coefficients:
                  Estimate   Std.err    Wald Pr(>|W|)
(Intercept)       9.811754  0.439507 498.382  < 2e-16 ***
yrssince06       -0.083205  0.046595   3.189  0.07415 .
GENDER2          -0.633484  0.102824  37.956 7.24e-10 ***
yrs06sq           0.004962  0.008031   0.382  0.53666
STRATUM2          0.252472  0.495100   0.260  0.61009
STRATUM3          0.756543  0.490653   2.377  0.12310
STRATUM4         -0.294685  1.189041   0.061  0.80426
STRATUM5          0.674029  1.666011   0.164  0.68579
STRATUM6          0.902298  0.615887   2.146  0.14291
STRATUM7          1.477010  0.512809   8.296  0.00397 **
STRATUM8          0.932406  0.480551   3.765  0.05234 .
STRATUM9          0.716107  0.520425   1.893  0.16882
STRATUM10         1.484322  0.459878  10.418  0.00125 **
STRATUM11         0.238604  0.786072   0.092  0.76148
STRATUM12         1.194361  0.545861   4.787  0.02867 *
STRATUM13         1.043130  0.524263   3.959  0.04662 *
STRATUM14         0.981955  0.471417   4.339  0.03725 *
STRATUM15         0.805363  0.511629   2.478  0.11546
STRATUM16         0.837765  0.455406   3.384  0.06583 .
STRATUM17         1.287055  0.483077   7.098  0.00772 **
STRATUM18         1.397770  0.492931   8.041  0.00457 **
STRATUM19         1.072590  0.660809   2.635  0.10456
STRATUM20         1.088616  0.738265   2.174  0.14033
STRATUM21         1.171977  0.453173   6.688  0.00971 **
STRATUM22         1.120087  0.491161   5.201  0.02258 *
STRATUM23         0.557142  0.513011   1.179  0.27747
STRATUM24         0.959569  0.527894   3.304  0.06911 .
STRATUM25         0.924567  0.454931   4.130  0.04212 *
STRATUM26         1.244131  0.535822   5.391  0.02024 *
```

```
STRATUM27           1.227339  0.465830   6.942  0.00842 **
STRATUM28           0.999721  0.464745   4.627  0.03147 *
STRATUM29           1.256069  0.473722   7.030  0.00801 **
STRATUM30           0.972002  0.476475   4.162  0.04135 *
STRATUM31           1.139282  0.464244   6.022  0.01413 *
STRATUM32           1.007486  0.620234   2.639  0.10430
STRATUM33           0.934112  0.571731   2.669  0.10229
STRATUM34           0.391956  0.456093   0.739  0.39013
STRATUM35           1.140609  0.513953   4.925  0.02647 *
STRATUM36           0.769334  0.617221   1.554  0.21260
STRATUM37           0.730089  0.453703   2.589  0.10758
STRATUM38           1.357772  0.535490   6.429  0.01123 *
STRATUM39           0.961475  0.459745   4.374  0.03650 *
STRATUM40           1.418981  0.457088   9.637  0.00191 **
STRATUM41           1.322078  0.508997   6.747  0.00939 **
STRATUM42           0.673105  0.467798   2.070  0.15018
STRATUM43           1.048939  0.478674   4.802  0.02843 *
STRATUM44           1.162143  0.464574   6.258  0.01237 *
STRATUM45           1.375144  0.458912   8.979  0.00273 **
STRATUM46           1.080036  0.475169   5.166  0.02303 *
STRATUM47           0.680106  0.453804   2.246  0.13396
STRATUM48           0.938204  0.458681   4.184  0.04081 *
STRATUM49           0.352605  0.616467   0.327  0.56734
STRATUM50           0.804207  0.457144   3.095  0.07854 .
STRATUM51           0.968136  0.485094   3.983  0.04596 *
STRATUM52           0.076949  0.502231   0.023  0.87823
STRATUM53           0.983399  0.481331   4.174  0.04104 *
STRATUM54           0.292583  0.618121   0.224  0.63597
STRATUM55           0.318202  0.518565   0.377  0.53947
STRATUM56           1.342582  0.475870   7.960  0.00478 **
yrssince06:GENDER2  0.092424  0.077052   1.439  0.23033
GENDER2:yrs06sq    -0.009592  0.012510   0.588  0.44323
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Estimated Scale Parameters:
            Estimate Std.err
(Intercept)    2.286  0.2756

Correlation: Structure = exchangeable  Link = identity

Estimated Correlation Parameters:
      Estimate Std.err
alpha   0.3487 0.09045
Number of clusters:   11789   Maximum cluster size: 4
```