Analysis Examples Replication Chapter 8

R Code

```
# data production prior to analysis examples
# need ncsr part1 part2 and population weight data
# nhanes subset of >=18 years of age
# hrs hh weight and financial respondents
# hrs res weight for individual level analyses
# can create subsets at the design data set level

setwd("P:/ASDA3/Replication R/Chapter 8")
getwd()

# data production
# load survey package and sas7bdat package
library (sas7bdat)
library (survey)

#nhanes first
library(haven)
# example of how to read in from SAS, not used here
#data <- read_sas("<path to your SAS file>")
#nhanesdata <- read_sas("P:\\ASDA3\\Replication SAS\\nhanes1112_f.sas7bdat")

# load nhanes1112.rdata
load ("P:/ASDA3/Data Sets for Analysis Examples and Stata R Code/nhanes1112.rdata")
nhanesdata <- nhanes1112
summary(nhanesdata)

#create factor variables
nhanesdata$racec <- factor(nhanesdata$ridreth1, levels = 1: 5 , labels =c("Mexican", "Other Hispanic", "White", 
"Black", "Other"))
nhanesdata$marcatc <- factor(nhanesdata$marcat, levels = 1: 3, labels =c("Married", "Previously Married", "Never 
Married"))
nhanesdata$edcatc <- factor(nhanesdata$edcat, levels = 1: 4, labels =c("0-11", "12", "13-15","16+"))
nhanesdata$bp_catc <- factor(nhanesdata$bp_cat, levels = 1: 4, labels =c("Normal", "Pre-HBP", "Stage 1 
HBP","Stage 2 HBP"))
#nhanesdata$agecsq <- (nhanesdata$age * nhanesdata$age)

names(nhanesdata)

nhanessvy2 <- svydesign(strata=~sdmvstra, id=~sdmvpsu, weights=~wtmec2yr, data=nhanesdata, nest=T)
subnhanes <- subset(nhanessvy2 , age >= 18)
names (nhanessvy2)

####
#ncs-r next
# read in NCSR data
# ncsr <- read_sas("P:\\ASDA3\\Replication SAS\\ncsr_f.sas7bdat")


load ("P:/ASDA3/Data Sets for Analysis Examples and Stata R Code/ncsr.rdata")
names(ncsr)
summary(ncsr)

#create factor versions with labels
ncsr$racec <- factor(ncsr$racecat, levels = 1: 4, labels =c("Other", "Hispanic", "Black", "White"))
ncsr$mar3catc <- factor(ncsr$mar3cat, levels = 1: 3, labels =c("Married", "Previously Married", "Never 
Married"))
```

```
ncsr$ed4catc <- factor(ncsr$ed4cat, levels = 1: 4, labels =c("0-11", "12", "13-15","16+"))
ncsr$sexc <- factor(ncsr$sex, levels = 1:2, labels=c("Male","Female"))
ncsr$ag4catc <- factor(ncsr$ag4cat, levels = 1:4, labels=c("18-29", "30-44", "45-59", "60+"))
ncsr$mdec <- factor(ncsr$mde, level = 1:2, labels=c("No","Yes"))

# part 1 data n=9282
ncsrsvyp1 <- svydesign(strata=~sestrat, id=~seclustr, weights=~ncsrwtsh, data=ncsr, nest=T)
names (ncsrsvyp1)

# part 2 data n=5692
ncsrp2 <- subset(ncsr, !is.na(ncsrwtlg))
ncsrsvyp2 <- svydesign(strata=~sestrat, id=~seclustr, weights=~ncsrwtlg, data=ncsrp2, nest=T)
names (ncsrsvyp2)

# rescale weight to sum to population in 2001
ncsr$popweight <- (ncsr$ncsrwtsh*(209128094/9282))

ncsrsvypop <- svydesign(strata=~sestrat, id=~seclustr, weights=~popweight, data=ncsr, nest=T)
summary(ncsrsvypop)

###
# HRS 2012 data
# both hh and resp weights are needed plus financial respondent for hh level analysis

# informational, if need to import from SAS
# library(haven)
# hrs <- read_sas("P:\\ASDA3\\Replication SAS\\hrs12_f.sas7bdat")
# summary(hrs)

# get/load data
load ("P:/ASDA3/Data Sets for Analysis Examples and Stata R Code/hrs12.rdata")
names(hrs12)

hrssvyhh <- svydesign(strata=~stratum, id=~secu, weights=~nwgthh , data=hrs12, nest=T)
summary(hrssvyhh)

hrssvysub <-subset(hrssvyhh, nfinr==1)
summary(hrssvysub)

hrssvyr <- svydesign(strata=~stratum, id=~secu, weights=~nwgtr , data=hrs12, nest=T)
summary(hrssvyr)
```

```
###
# Chapter 8 ASDA2 Analysis Examples Replication
# Bivariate chisq tests for Example 8.1
# NCSR Data

# MDE by Sex, Age, Alcohol Dependence, Education, and Marital Status
ex8_1sex <- svyby(~factor(mde), ~sexc, svymean, design=ncsrsvyp2, se=T, na.rm=T, ci=T)
ex8_1sex

ex8_1age <- svyby(~factor(mde), ~ag4catc, svymean, design=ncsrsvyp2, se=T, na.rm=T, ci=T)
ex8_1age

ex8_1ald <- svyby(~factor(mde), ~ald, svymean, design=ncsrsvyp2, se=T, na.rm=T, ci=T)
ex8_1ald

ex8_1ed <- svyby(~factor(mde), ~ed4catc, svymean, design=ncsrsvyp2, se=T, na.rm=T, ci=T)
ex8_1ed

ex8_1mar <- svyby(~factor(mde), ~mar3catc, svymean, design=ncsrsvyp2, se=T, na.rm=T, ci=T)
ex8_1mar

# ChiSq Tests
svychisq(~mde + ag4cat, ncsrsvyp2)
svychisq(~mde + sexc, ncsrsvyp2)
svychisq(~mde + ald, ncsrsvyp2)
svychisq(~mde + ed4cat, ncsrsvyp2)
svychisq(~mde + mar3cat, ncsrsvyp2)

# Model 8.1
mod81 <- svyglm(mde ~ factor(ag4catc) + sexm + ald + factor(ed4catc) + factor(mar3catc), family=quasibinomial,
design=ncsrsvyp2)
summary(mod81)
#exp of coefficients
exp(mod81$coef)
# Lumley AIC (design-adjusted) test for first model
AIC(mod81)

# tests of parameters
regTermTest(mod81, ~factor(ag4catc))
regTermTest(mod81, ~factor(ed4catc))
regTermTest(mod81, ~factor(mar3catc))

# FORTHCOMING! Check website for example of creating average marginal effects of ALD on MDE by Age Groups in R

# add sex X other predictors interactions to model
mod81_int <- svyglm(mde ~ factor(ag4catc) + sexm + ald + factor(ed4catc) + factor(mar3catc) +
  sexm*factor(ag4catc) + sexm*ald + sexm*factor(ed4catc) + sexm*factor(mar3catc),
  family=quasibinomial, design=ncsrsvyp2)
summary(mod81_int)
# AIC for 2nd model
AIC(mod81_int)

# Test interactions of sex by all other predictors
regTermTest(mod81_int, ~sexm:factor(ag4catc))
regTermTest(mod81_int, ~sexm:ald)
regTermTest(mod81_int, ~sexm:factor(ed4catc))
regTermTest(mod81_int, ~sexm:factor(mar3catc))

# Design-Adjusted LRT test in R, tests model 1 v. model 2 with interactions
```

```
anova(mod81, mod81_int, test="Chisq", method="LRT")

##
# Alternative Bayesian method
# make sure brms and cssampling packages are loaded
library(brms)
library(csSampling)

# create new factor variable of sex
ncsr$sexc <- factor(ncsr$sex, levels = 1:2, labels=c("Male","Female"))
ncsr$sexc <- relevel(ncsr$sexc, ref="Female")


# create complete data set with variables of interest
ncsr.red <- ncsr[, c("mde", "mar3catc", "ed4catc", "sexc", "ag4catc", "ald", "seclustr", "sestrat", "ncsrwtlg")]
ncsr.red <- ncsr.red[complete.cases(ncsr.red),]

# need to normalize NCS-R weights to match what is done for Stan modeling
ncsr.red$wtsc <- ncsr.red$ncsrwtlg / mean(ncsr.red$ncsrwtlg)

# survey design object
ncsr.des <- svydesign(id = ~seclustr, strata = ~sestrat, weights = ~wtsc, nest = T, data = ncsr.red)

# Bayesian approach, flat prior
set.seed(41279)
model_formula <- formula("mde|weights(wtsc) ~ ag4catc + sexc + ald + ed4catc + mar3catc")
mod.brms <- cs_sampling_brms(svydes = ncsr.des, brmsmod = brmsformula(model_formula), data = ncsr.red, family =
bernoulli())
mod.brms$stan_fit

# Bayesian with informative prior for ALD coefficient
mod.brms <- cs_sampling_brms(svydes = ncsr.des, brmsmod = brmsformula(model_formula), prior =
c(set_prior("normal(1.0,0.2)", class = "b", coef="ald")), data = ncsr.red, family = bernoulli())
mod.brms$stan_fit


# Model 8.2 with logit, probit, and cloglog comparison
summary(ex82_logit <- svyglm(ald ~ factor(ag4catc) + sexm + factor(ed4catc) + factor(mar3catc),
family=quasibinomial, design=ncsrsvyp2))
regTermTest(ex82_logit, ~factor(ag4catc))
regTermTest(ex82_logit, ~factor(ed4catc))
regTermTest(ex82_logit, ~factor(mar3catc))

summary(ex82_probit <- svyglm(ald ~ factor(ag4catc) + sexm + factor(ed4catc) + factor(mar3catc),
family=quasibinomial(link=probit),  design=ncsrsvyp2))
regTermTest(ex82_probit, ~factor(ag4catc))
regTermTest(ex82_probit, ~factor(ed4catc))
regTermTest(ex82_probit, ~factor(mar3catc))

summary(ex82_cloglog <- svyglm(ald ~ factor(ag4catc) + sexm + factor(ed4catc) + factor(mar3catc),
family=quasibinomial(link=cloglog),  design=ncsrsvyp2))
regTermTest(ex82_cloglog, ~factor(ag4catc))
regTermTest(ex82_cloglog, ~factor(ed4catc))
regTermTest(ex82_cloglog, ~factor(mar3catc))> # Chapter Analysis Examples Replication Chapter 8
```

R Results

```
> # data production prior to analysis examples
> # need ncsr part1 part2 and population weight data
> # nhanes subset of >=18 years of age
> # hrs hh weight and financial respondents
> # hrs res weight for individual level analyses
> # can create subsets at the design data set level
> 
> setwd("P:/ASDA3/Replication R/Chapter 8")
> getwd()
[1] "P:/ASDA3/Replication R/Chapter 8"
> 
> # data production
> # load survey package and sas7bdat package
> library (sas7bdat)
> library (survey)
> 
> #nhanes first
> library(haven)
> # example of how to read in from SAS, not used here
> #data <- read_sas("<path to your SAS file>")
> #nhanesdata <- read_sas("P:\\ASDA3\\Replication SAS\\nhanes1112_f.sas7bdat")
> 
> # load nhanes1112.rdata
> load ("P:/ASDA3/Data Sets for Analysis Examples and Stata R Code/nhanes1112.rdata")
> nhanesdata <- nhanes1112
> summary(nhanesdata)
      seqn          ridstatr        riagendr        ridreth1        dmdmartl     
 Min.   :62161   Min.   :1.000   Min.   :1.000   Min.   :1.000   Min.   : 1.000  
 1st Qu.:64600   1st Qu.:2.000   1st Qu.:1.000   1st Qu.:3.000   1st Qu.: 1.000  
 Median :67039   Median :2.000   Median :2.000   Median :3.000   Median : 2.000  
 Mean   :67039   Mean   :1.957   Mean   :1.502   Mean   :3.229   Mean   : 2.749  
 3rd Qu.:69477   3rd Qu.:2.000   3rd Qu.:2.000   3rd Qu.:4.000   3rd Qu.: 5.000  
 Max.   :71916   Max.   :2.000   Max.   :2.000   Max.   :5.000   Max.   :99.000  
                                                                 NA's   :4196    
    wtint2yr         wtmec2yr         sdmvpsu         sdmvstra         indfmpir    
 Min.   :  3321   Min.   :     0   Min.   :1.000   Min.   : 90.00   Min.   :0.000  
 1st Qu.: 11352   1st Qu.: 11174   1st Qu.:1.000   1st Qu.: 92.00   1st Qu.:0.860  
 Median : 18098   Median : 18090   Median :2.000   Median : 96.00   Median :1.630  
 Mean   : 31426   Mean   : 31426   Mean   :1.643   Mean   : 95.87   Mean   :2.205  
 3rd Qu.: 34887   3rd Qu.: 34792   3rd Qu.:2.000   3rd Qu.: 99.00   3rd Qu.:3.580  
 Max.   :220233   Max.   :222580   Max.   :3.000   Max.   :103.00   Max.   :5.000  
                                                                    NA's   :840    
     bpxsy1          bpxdi1          bpxsy2          bpxdi2           bpxsy3     
 Min.   : 74.0   Min.   :  0.0   Min.   : 74.0   Min.   :  0.00   Min.   : 74.0  
 1st Qu.:106.0   1st Qu.: 60.0   1st Qu.:106.0   1st Qu.: 58.00   1st Qu.:106.0  
 Median :116.0   Median : 68.0   Median :116.0   Median : 68.00   Median :116.0  
 Mean   :119.2   Mean   : 66.9   Mean   :118.7   Mean   : 66.28   Mean   :118.2  
 3rd Qu.:128.0   3rd Qu.: 76.0   3rd Qu.:128.0   3rd Qu.: 76.00   3rd Qu.:128.0  
 Max.   :238.0   Max.   :120.0   Max.   :234.0   Max.   :134.00   Max.   :232.0  
 NA's   :3000    NA's   :3000    NA's   :2848    NA's   :2848     NA's   :2839   
     bpxdi3           bpxsy4         bpxdi4           bmxbmi           lbxtc      
 Min.   :  0.00   Min.   : 78    Min.   :  0.00   Min.   :12.40   Min.   : 59.0  
 1st Qu.: 58.00   1st Qu.:104    1st Qu.: 64.00   1st Qu.:19.30   1st Qu.:154.0  
 Median : 68.00   Median :116    Median : 72.00   Median :24.50   Median :179.0  
 Mean   : 65.91   Mean   :119    Mean   : 71.78   Mean   :25.34   Mean   :183.2  
 3rd Qu.: 76.00   3rd Qu.:130    3rd Qu.: 82.00   3rd Qu.:29.80   3rd Qu.:209.0  
 Max.   :128.00   Max.   :226    Max.   :130.00   Max.   :82.10   Max.   :523.0  
```

```
 NA's   :2839     NA's   :9309   NA's   :9309     NA's   :1154    NA's   :2768
     age18p         irregular          edcat             age           marcat
 Min.   :0.0000   Min.   :0.0000   Min.   :1.000   Min.   : 0.0   Min.   :1.000
 1st Qu.:0.0000   1st Qu.:0.0000   1st Qu.:1.000   1st Qu.: 9.0   1st Qu.:1.000
 Median :1.0000   Median :0.0000   Median :2.000   Median :26.0   Median :1.000
 Mean   :0.6011   Mean   :0.0134   Mean   :2.102   Mean   :31.4   Mean   :1.652
 3rd Qu.:1.0000   3rd Qu.:0.0000   3rd Qu.:3.000   3rd Qu.:52.0   3rd Qu.:2.000
 Max.   :1.0000   Max.   :1.0000   Max.   :4.000   Max.   :80.0   Max.   :3.000
                  NA's   :729      NA's   :1602                   NA's   :4203
    pre_hibp          bp_cat          ag1829          ag3044           ag4559
 Min.   :0.0000   Min.   :1.000   Min.   :0.000   Min.   :0.0000   Min.   :0.0000
 1st Qu.:0.0000   1st Qu.:1.000   1st Qu.:0.000   1st Qu.:0.0000   1st Qu.:0.0000
 Median :0.0000   Median :1.000   Median :1.000   Median :0.0000   Median :0.0000
 Mean   :0.4357   Mean   :1.546   Mean   :0.532   Mean   :0.1461   Mean   :0.1384
 3rd Qu.:1.0000   3rd Qu.:2.000   3rd Qu.:1.000   3rd Qu.:0.0000   3rd Qu.:0.0000
 Max.   :1.0000   Max.   :4.000   Max.   :1.000   Max.   :1.0000   Max.   :1.0000
 NA's   :2701     NA's   :2701
      ag60             mex              othhis           white            black
 Min.   :0.0000   Min.   :0.0000   Min.   :0.0000   Min.   :0.0000   Min.   :0.000
 1st Qu.:0.0000   1st Qu.:0.0000   1st Qu.:0.0000   1st Qu.:0.0000   1st Qu.:0.000
 Median :0.0000   Median :0.0000   Median :0.0000   Median :0.0000   Median :0.000
 Mean   :0.1836   Mean   :0.1389   Mean   :0.1103   Mean   :0.3047   Mean   :0.275
 3rd Qu.:0.0000   3rd Qu.:0.0000   3rd Qu.:0.0000   3rd Qu.:1.0000   3rd Qu.:1.000
 Max.   :1.0000   Max.   :1.0000   Max.   :1.0000   Max.   :1.0000   Max.   :1.000

     other            lbdhdd          lbdhddsi        lbdtcsi
 Min.   :0.0000   Min.   : 14.00   Min.   :0.360   Min.   : 1.530
 1st Qu.:0.0000   1st Qu.: 43.00   1st Qu.:1.110   1st Qu.: 3.980
 Median :0.0000   Median : 51.00   Median :1.320   Median : 4.630
 Mean   :0.1711   Mean   : 52.63   Mean   :1.361   Mean   : 4.738
 3rd Qu.:0.0000   3rd Qu.: 60.00   3rd Qu.:1.550   3rd Qu.: 5.400
 Max.   :1.0000   Max.   :175.00   Max.   :4.530   Max.   :13.520
                  NA's   :2767     NA's   :2767    NA's   :2768

> #create factor variables
> nhanesdata$racec <- factor(nhanesdata$ridreth1, levels = 1: 5 , labels =c("Mexican", "Other Hispanic",
"White", "Black", "Other"))
> nhanesdata$marcatc <- factor(nhanesdata$marcat, levels = 1: 3, labels =c("Married", "Previously Married",
"Never Married"))
> nhanesdata$edcatc <- factor(nhanesdata$edcat, levels = 1: 4, labels =c("0-11", "12", "13-15","16+"))
> nhanesdata$bp_catc <- factor(nhanesdata$bp_cat, levels = 1: 4, labels =c("Normal", "Pre-HBP", "Stage 1
HBP","Stage 2 HBP"))
> #nhanesdata$agecsq <- (nhanesdata$age * nhanesdata$age)
>
> names(nhanesdata)
 [1] "seqn"      "ridstatr"  "riagendr"  "ridreth1"  "dmdmartl"  "wtint2yr"  "wtmec2yr"
 [8] "sdmvpsu"   "sdmvstra"  "indfmpir"  "bpxsy1"    "bpxdi1"    "bpxsy2"    "bpxdi2"
[15] "bpxsy3"    "bpxdi3"    "bpxsy4"    "bpxdi4"    "bmxbmi"    "lbxtc"     "age18p"
[22] "irregular" "edcat"     "age"       "marcat"    "pre_hibp"  "bp_cat"    "ag1829"
[29] "ag3044"    "ag4559"    "ag60"      "mex"       "othhis"    "white"     "black"
[36] "other"     "lbdhdd"    "lbdhddsi"  "lbdtcsi"   "racec"     "marcatc"   "edcatc"
[43] "bp_catc"

> nhanessvy2 <- svydesign(strata=~sdmvstra, id=~sdmvpsu, weights=~wtmec2yr, data=nhanesdata, nest=T)
> subnhanes <- subset(nhanessvy2 , age >= 18)
> names (nhanessvy2)
[1] "cluster"    "strata"     "has.strata" "prob"       "allprob"    "call"       "variables"
[8] "fpc"        "pps"
```

```
> #ncs-r next
> # read in NCSR data
> # ncsr <- read_sas("P:\\ASDA3\\Replication SAS\\ncsr_f.sas7bdat")


> load ("P:/ASDA3/Data Sets for Analysis Examples and Stata R Code/ncsr.rdata")
> names(ncsr)
 [1] "caseid"   "dsm_so"   "age"      "region"   "mar3cat"  "ed4cat"   "obese6ca" "ncsrwtsh"
 [9] "ncsrwtlg" "sex"      "wkstat3c" "sestrat"  "seclustr" "ag4cat"   "racecat"  "mde"
[17] "ald"      "sexf"     "sexm"
> summary(ncsr)
     caseid         dsm_so           age            region         mar3cat
 Min.   :   1   Min.   :1.000   Min.   :18.00   Min.   :1.00   Min.   :1.000
 1st Qu.:2321   1st Qu.:5.000   1st Qu.:30.00   1st Qu.:2.00   1st Qu.:1.000
 Median :4642   Median :5.000   Median :43.00   Median :3.00   Median :1.000
 Mean   :4642   Mean   :4.507   Mean   :44.73   Mean   :2.57   Mean   :1.636
 3rd Qu.:6962   3rd Qu.:5.000   3rd Qu.:57.00   3rd Qu.:3.00   3rd Qu.:2.000
 Max.   :9282   Max.   :5.000   Max.   :99.00   Max.   :4.00   Max.   :3.000

     ed4cat         obese6ca       ncsrwtsh         ncsrwtlg           sex
 Min.   :1.000   Min.   :1.00   Min.   :0.1694   Min.   : 0.114   Min.   :1.000
 1st Qu.:2.000   1st Qu.:2.00   1st Qu.:0.6899   1st Qu.: 0.463   1st Qu.:1.000
 Median :3.000   Median :3.00   Median :0.8722   Median : 0.644   Median :2.000
 Mean   :2.661   Mean   :2.93   Mean   :1.0000   Mean   : 1.000   Mean   :1.554
 3rd Qu.:4.000   3rd Qu.:4.00   3rd Qu.:1.1616   3rd Qu.: 1.083   3rd Qu.:2.000
 Max.   :4.000   Max.   :6.00   Max.   :7.1361   Max.   :10.102   Max.   :2.000
                 NA's   :176                     NA's   :3590
    wkstat3c        sestrat         seclustr         ag4cat        racecat
 Min.   :1.000   Min.   : 1.00   Min.   :1.000   Min.   :1.00   Min.   :1.000
 1st Qu.:1.000   1st Qu.:20.00   1st Qu.:1.000   1st Qu.:2.00   1st Qu.:3.000
 Median :1.000   Median :28.00   Median :2.000   Median :2.00   Median :4.000
 Mean   :1.592   Mean   :26.31   Mean   :1.509   Mean   :2.44   Mean   :3.524
 3rd Qu.:3.000   3rd Qu.:36.00   3rd Qu.:2.000   3rd Qu.:3.00   3rd Qu.:4.000
 Max.   :3.000   Max.   :42.00   Max.   :2.000   Max.   :4.00   Max.   :4.000
 NA's   :2649
      mde             ald               sexf             sexm
 Min.   :0.000   Min.   :0.00000   Min.   :0.0000   Min.   :0.0000
 1st Qu.:0.000   1st Qu.:0.00000   1st Qu.:0.0000   1st Qu.:0.0000
 Median :0.000   Median :0.00000   Median :1.0000   Median :0.0000
 Mean   :0.197   Mean   :0.04783   Mean   :0.5541   Mean   :0.4459
 3rd Qu.:0.000   3rd Qu.:0.00000   3rd Qu.:1.0000   3rd Qu.:1.0000
 Max.   :1.000   Max.   :1.00000   Max.   :1.0000   Max.   :1.0000


> #create factor versions with labels
> ncsr$racec <- factor(ncsr$racecat, levels = 1: 4, labels =c("Other", "Hispanic", "Black", "White"))
> ncsr$mar3catc <- factor(ncsr$mar3cat, levels = 1: 3, labels =c("Married", "Previously Married", "Never
Married"))
> ncsr$ed4catc <- factor(ncsr$ed4cat, levels = 1: 4, labels =c("0-11", "12", "13-15","16+"))
> ncsr$sexc <- factor(ncsr$sex, levels = 1:2, labels=c("Male","Female"))
> ncsr$ag4catc <- factor(ncsr$ag4cat, levels = 1:4, labels=c("18-29", "30-44", "45-59", "60+"))
> ncsr$mdec <- factor(ncsr$mde, level = 1:2, labels=c("No","Yes"))

> # part 1 data n=9282
> ncsrsvyp1 <- svydesign(strata=~sestrat, id=~seclustr, weights=~ncsrwtsh, data=ncsr, nest=T)
> names (ncsrsvyp1)
[1] "cluster"    "strata"     "has.strata" "prob"       "allprob"    "call"       "variables"
[8] "fpc"        "pps"
```

```
> # part 2 data n=5692
> ncsrp2 <- subset(ncsr, !is.na(ncsrwtlg))
> ncsrsvyp2 <- svydesign(strata=~sestrat, id=~seclustr, weights=~ncsrwtlg, data=ncsrp2, nest=T)
> names (ncsrsvyp2)
[1] "cluster"    "strata"     "has.strata" "prob"       "allprob"    "call"       "variables"
[8] "fpc"        "pps"

> # rescale weight to sum to population in 2001
> ncsr$popweight <- (ncsr$ncsrwtsh*(209128094/9282))

> ncsrsvypop <- svydesign(strata=~sestrat, id=~seclustr, weights=~popweight, data=ncsr, nest=T)
> summary(ncsrsvypop)
Stratified 1 - level Cluster Sampling design (with replacement)
With (84) clusters.
svydesign(strata = ~sestrat, id = ~seclustr, weights = ~popweight,
    data = ncsr, nest = T)
Probabilities:
     Min.   1st Qu.    Median      Mean   3rd Qu.      Max.
6.220e-06 3.821e-05 5.089e-05 5.313e-05 6.433e-05 2.621e-04
Stratum Sizes:
            1   2   3   4   5   6   7   8   9  10  11  12 13 14  15  16 17 18  19  20  21  22
obs        91 108 123 134 124 113 188 157 140 126 107 108 98 79 109 107 74 86 113 260 309 297
design.PSU  2   2   2   2   2   2   2   2   2   2   2   2  2  2   2   2  2  2   2   2   2   2
actual.PSU  2   2   2   2   2   2   2   2   2   2   2   2  2  2   2   2  2  2   2   2   2   2
             23  24  25  26  27  28  29  30  31  32  33  34  35  36  37  38  39  40  41  42
obs         269 286 378 272 279 214 336 239 393 289 149 361 343 305 333 366 401 313 346 359
design.PSU    2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2
actual.PSU    2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2
Data variables:
 [1] "caseid"    "dsm_so"    "age"       "region"    "mar3cat"   "ed4cat"    "obese6ca"
 [8] "ncsrwtsh"  "ncsrwtlg"  "sex"       "wkstat3c"  "sestrat"   "seclustr"  "ag4cat"
[15] "racecat"   "mde"       "ald"       "sexf"      "sexm"      "racec"     "mar3catc"
[22] "ed4catc"   "sexc"      "ag4catc"   "mdec"      "popweight"


> # HRS 2012 data
> # both hh and resp weights are needed plus financial respondent for hh level analysis

> # informational, if need to import from SAS
> # library(haven)
> # hrs <- read_sas("P:\\ASDA3\\Replication SAS\\hrs12_f.sas7bdat")
> # summary(hrs)

> # get/load data
> load ("P:/ASDA3/Data Sets for Analysis Examples and Stata R Code/hrs12.rdata")
> names(hrs12)
 [1] "hhid"       "pn"         "r11bmi"     "nage"       "nfinr"      "gender"
 [7] "secu"       "stratum"    "nwgthh"     "nwgtr"      "h11atota"   "h11itot"
[13] "marcat"     "edcat"      "racecat"    "diabetes"   "numfalls24" "age65p"
[19] "arthritis"

> hrssvyhh <- svydesign(strata=~stratum, id=~secu, weights=~nwgthh , data=hrs12, nest=T)
> summary(hrssvyhh)
Stratified 1 - level Cluster Sampling design (with replacement)
With (112) clusters.
svydesign(strata = ~stratum, id = ~secu, weights = ~nwgthh, data = hrs12,
    nest = T)
Probabilities:
     Min.   1st Qu.    Median      Mean   3rd Qu.      Max.
```

```
5.594e-05 1.806e-04 2.978e-04       Inf 5.727e-04       Inf 
Stratum Sizes: 
             1   2   3   4   5   6   7   8   9  10  11  12  13  14  15  16  17  18  19  20
obs        352 627 366 333 276 334 384 403 367 334 261 370 209 159 177 136 276 185 295 371
design.PSU   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2
actual.PSU   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2
            21  22  23 24 25  26  27  28  29  30  31  32  33  34  35  36  37  38  39  40  41
obs        332 182 107 68 83 302 588 432 499 533 514 457 560 483 306 542 196 370 366 669 510
design.PSU   2   2   2  2  2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2
actual.PSU   2   2   2  2  2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2
            42  43  44  45  46  47  48  49  50  51  52  53 54  55  56
obs        489 777 691 601 687 447 297 333 685 548 153 145 56 179 152
design.PSU   2   2   2   2   2   2   2   2   2   2   2   2  2   2   2
actual.PSU   2   2   2   2   2   2   2   2   2   2   2   2  2   2   2
Data variables:
 [1] "hhid"       "pn"         "r11bmi"     "nage"       "nfinr"      "gender"    
 [7] "secu"       "stratum"    "nwgthh"     "nwgtr"      "h11atota"   "h11itot"   
[13] "marcat"     "edcat"      "racecat"    "diabetes"   "numfalls24" "age65p"    
[19] "arthritis" 

> hrssvysub <-subset(hrssvyhh, nfinr==1)
> summary(hrssvysub)
Stratified 1 - level Cluster Sampling design (with replacement)
With (112) clusters.
subset(hrssvyhh, nfinr == 1)
Probabilities:
     Min.   1st Qu.    Median      Mean   3rd Qu.      Max. 
5.594e-05 1.959e-04 3.110e-04       Inf 5.961e-04       Inf 
Stratum Sizes: 
             1   2   3   4   5   6   7   8   9  10  11  12  13  14  15 16  17  18  19  20  21
obs        257 409 247 228 194 245 281 275 253 228 185 253 134 105 125 96 192 123 204 273 233
design.PSU   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2  2   2   2   2   2   2
actual.PSU   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2  2   2   2   2   2   2
            22 23 24 25  26  27  28  29  30  31  32  33  34  35  36  37  38  39  40  41  42
obs        129 77 52 61 217 424 308 350 364 364 314 405 339 217 358 136 258 242 454 345 349
design.PSU   2  2  2  2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2
actual.PSU   2  2  2  2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2
            43  44  45  46  47  48  49  50  51  52 53 54  55 56
obs        498 445 406 467 308 203 244 463 387 111 96 40 124 96
design.PSU   2   2   2   2   2   2   2   2   2   2  2  2   2  2
actual.PSU   2   2   2   2   2   2   2   2   2   2  2  2   2  2
Data variables:
 [1] "hhid"       "pn"         "r11bmi"     "nage"       "nfinr"      "gender"    
 [7] "secu"       "stratum"    "nwgthh"     "nwgtr"      "h11atota"   "h11itot"   
[13] "marcat"     "edcat"      "racecat"    "diabetes"   "numfalls24" "age65p"    
[19] "arthritis" 

> hrssvyr <- svydesign(strata=~stratum, id=~secu, weights=~nwgtr , data=hrs12, nest=T)
> summary(hrssvyr)
Stratified 1 - level Cluster Sampling design (with replacement)
With (112) clusters.
svydesign(strata = ~stratum, id = ~secu, weights = ~nwgtr, data = hrs12, 
    nest = T)
Probabilities:
     Min.   1st Qu.    Median      Mean   3rd Qu.      Max. 
5.034e-05 1.757e-04 2.874e-04       Inf 5.754e-04       Inf 
Stratum Sizes: 
             1   2   3   4   5   6   7   8   9  10  11  12  13  14  15  16  17  18  19  20
obs        352 627 366 333 276 334 384 403 367 334 261 370 209 159 177 136 276 185 295 371
```

```
design.PSU   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2
actual.PSU   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2
            21  22  23 24 25  26  27  28  29  30  31  32  33  34  35  36  37  38  39  40  41
obs        332 182 107 68 83 302 588 432 499 533 514 457 560 483 306 542 196 370 366 669 510
design.PSU   2   2   2  2  2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2
actual.PSU   2   2   2  2  2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2
            42  43  44  45  46  47  48  49  50  51  52  53 54  55  56
obs        489 777 691 601 687 447 297 333 685 548 153 145 56 179 152
design.PSU   2   2   2   2   2   2   2   2   2   2   2   2  2   2   2
actual.PSU   2   2   2   2   2   2   2   2   2   2   2   2  2   2   2
Data variables:
 [1] "hhid"       "pn"         "r11bmi"     "nage"       "nfinr"      "gender"    
 [7] "secu"       "stratum"    "nwgthh"     "nwgtr"      "h11atota"   "h11itot"   
[13] "marcat"     "edcat"      "racecat"    "diabetes"   "numfalls24" "age65p"    
[19] "arthritis" 

> # Chapter 8 ASDA2 Analysis Examples Replication
> # Bivariate chisq tests for Example 8.1
> # NCSR Data

> # MDE by Sex, Age, Alcohol Dependence, Education, and Marital Status

> ex8_1sex <- svyby(~factor(mde), ~sexc, svymean, design=ncsrsvyp2, se=T, na.rm=T, ci=T)
> ex8_1sex
         sexc factor(mde)0 factor(mde)1 se.factor(mde)0 se.factor(mde)1
Male     Male    0.8471074    0.1528926     0.009137590     0.009137590
Female Female    0.7738295    0.2261705     0.006727609     0.006727609
> 

> ex8_1age <- svyby(~factor(mde), ~ag4catc, svymean, design=ncsrsvyp2, se=T, na.rm=T, ci=T)
> ex8_1age
      ag4catc factor(mde)0 factor(mde)1 se.factor(mde)0 se.factor(mde)1
18-29   18-29    0.8160119    0.1839881     0.008853615     0.008853615
30-44   30-44    0.7712342    0.2287658     0.011044532     0.011044532
45-59   45-59    0.7766733    0.2233267     0.012606600     0.012606600
60+       60+    0.8893912    0.1106088     0.009562657     0.009562657


> ex8_1ald <- svyby(~factor(mde), ~ald, svymean, design=ncsrsvyp2, se=T, na.rm=T, ci=T)
> ex8_1ald
  ald factor(mde)0 factor(mde)1 se.factor(mde)0 se.factor(mde)1
0   0    0.8230860    0.1769140     0.006506907     0.006506907
1   1    0.5484091    0.4515909     0.029025055     0.029025055


> ex8_1ed <- svyby(~factor(mde), ~ed4catc, svymean, design=ncsrsvyp2, se=T, na.rm=T, ci=T)
> ex8_1ed
      ed4catc factor(mde)0 factor(mde)1 se.factor(mde)0 se.factor(mde)1
0-11     0-11    0.8369163    0.1630837      0.01212127      0.01212127
12         12    0.8145033    0.1854967      0.00827191      0.00827191
13-15   13-15    0.7875147    0.2124853      0.01043948      0.01043948
16+       16+    0.8033254    0.1966746      0.01087591      0.01087591


> ex8_1mar <- svyby(~factor(mde), ~mar3catc, svymean, design=ncsrsvyp2, se=T, na.rm=T, ci=T)
> ex8_1mar
                             mar3catc factor(mde)0 factor(mde)1 se.factor(mde)0
Married                       Married    0.8267385    0.1732615     0.007419619
Previously Married Previously Married    0.7609794    0.2390206     0.014493869
```

```
Never Married           Never Married    0.8060058    0.1939942     0.011549336
                   se.factor(mde)1
Married                0.007419619
Previously Married     0.014493869
Never Married          0.011549336


> # ChiSq Tests
> svychisq(~mde + ag4cat, ncsrsvyp2)

        Pearson's X^2: Rao & Scott adjustment

data:  svychisq(~mde + ag4cat, ncsrsvyp2)
F = 26.39, ndf = 2.7612, ddf = 115.9704, p-value = 1.965e-12

> svychisq(~mde + sexc, ncsrsvyp2)

        Pearson's X^2: Rao & Scott adjustment

data:  svychisq(~mde + sexc, ncsrsvyp2)
F = 44.834, ndf = 1, ddf = 42, p-value = 3.965e-08

> svychisq(~mde + ald, ncsrsvyp2)

        Pearson's X^2: Rao & Scott adjustment

data:  svychisq(~mde + ald, ncsrsvyp2)
F = 120.03, ndf = 1, ddf = 42, p-value = 6.86e-14

> svychisq(~mde + ed4cat, ncsrsvyp2)

        Pearson's X^2: Rao & Scott adjustment

data:  svychisq(~mde + ed4cat, ncsrsvyp2)
F = 4.3043, ndf = 2.903, ddf = 121.925, p-value = 0.006916

> svychisq(~mde + mar3cat, ncsrsvyp2)

        Pearson's X^2: Rao & Scott adjustment

data:  svychisq(~mde + mar3cat, ncsrsvyp2)
F = 11.085, ndf = 1.8987, ddf = 79.7446, p-value = 7.616e-05
```

```
> # Model 8.1
> mod81 <- svyglm(mde ~ factor(ag4catc) + sexm + ald + factor(ed4catc) + factor(mar3catc), family=quasibinomial,
design=ncsrsvyp2)
> summary(mod81)

Call:
svyglm(formula = mde ~ factor(ag4catc) + sexm + ald + factor(ed4catc) +
    factor(mar3catc), design = ncsrsvyp2, family = quasibinomial)

Survey design:
svydesign(strata = ~sestrat, id = ~seclustr, weights = ~ncsrwtlg,
    data = ncsrp2, nest = T)

Coefficients:
                                   Estimate Std. Error t value Pr(>|t|)
(Intercept)                        -1.58308    0.12066 -13.120 2.01e-14 ***
factor(ag4catc)30-44                0.25562    0.09438   2.708   0.0108 *
factor(ag4catc)45-59                0.20645    0.09153   2.256   0.0311 *
factor(ag4catc)60+                 -0.67579    0.14130  -4.783 3.74e-05 ***
sexm                               -0.57735    0.07722  -7.477 1.64e-08 ***
ald                                 1.42368    0.15416   9.235 1.53e-10 ***
factor(ed4catc)12                   0.07925    0.09690   0.818   0.4194
factor(ed4catc)13-15                0.23051    0.09307   2.477   0.0187 *
factor(ed4catc)16+                  0.16293    0.11061   1.473   0.1505
factor(mar3catc)Previously Married  0.48642    0.08542   5.694 2.63e-06 ***
factor(mar3catc)Never Married       0.11558    0.10787   1.071   0.2920
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for quasibinomial family taken to be 1.001564)

Number of Fisher Scoring iterations: 4

> #exp of coefficients
> exp(mod81$coef)
                       (Intercept)               factor(ag4catc)30-44
                         0.2053424                          1.2912600
              factor(ag4catc)45-59                 factor(ag4catc)60+
                         1.2293019                          0.5087563
                              sexm                                ald
                         0.5613867                          4.1523575
                 factor(ed4catc)12               factor(ed4catc)13-15
                         1.0824803                          1.2592434
                factor(ed4catc)16+ factor(mar3catc)Previously Married
                         1.1769489                          1.6264870
     factor(mar3catc)Never Married
                         1.1225236

> # Lumley AIC (design-adjusted) test for first model
> AIC(mod81)
       eff.p          AIC     deltabar
   9.7757533 5288.0774889    0.9775753

> # tests of parameters
> regTermTest(mod81, ~factor(ag4catc))
Wald test for factor(ag4catc)
 in svyglm(formula = mde ~ factor(ag4catc) + sexm + ald + factor(ed4catc) +
    factor(mar3catc), design = ncsrsvyp2, family = quasibinomial)
F =  19.98292  on  3  and  32  df: p= 1.7536e-07
```

```
> regTermTest(mod81, ~factor(ed4catc))
Wald test for factor(ed4catc)
 in svyglm(formula = mde ~ factor(ag4catc) + sexm + ald + factor(ed4catc) + 
    factor(mar3catc), design = ncsrsvyp2, family = quasibinomial)
F =  2.236337  on  3  and  32  df: p= 0.10303 
> regTermTest(mod81, ~factor(mar3catc))
Wald test for factor(mar3catc)
 in svyglm(formula = mde ~ factor(ag4catc) + sexm + ald + factor(ed4catc) + 
    factor(mar3catc), design = ncsrsvyp2, family = quasibinomial)
F =  17.00766  on  2  and  32  df: p= 9.2915e-06 


> # FORTHCOMING! Check website for example of creating average marginal effects of ALD on MDE by Age Groups in R


> # add sex X other predictors interactions to model
> mod81_int <- svyglm(mde ~ factor(ag4catc) + sexm + ald + factor(ed4catc) + factor(mar3catc) +
+   sexm*factor(ag4catc) + sexm*ald + sexm*factor(ed4catc) + sexm*factor(mar3catc),
+   family=quasibinomial, design=ncsrsvyp2)
> summary(mod81_int)

Call:
svyglm(formula = mde ~ factor(ag4catc) + sexm + ald + factor(ed4catc) + 
    factor(mar3catc) + sexm * factor(ag4catc) + sexm * ald + 
    sexm * factor(ed4catc) + sexm * factor(mar3catc), design = ncsrsvyp2, 
    family = quasibinomial)

Survey design:
svydesign(strata = ~sestrat, id = ~seclustr, weights = ~ncsrwtlg, 
    data = ncsrp2, nest = T)

Coefficients:
                                         Estimate Std. Error t value Pr(>|t|)    
(Intercept)                             -1.599889   0.134010 -11.939 2.45e-11 ***
factor(ag4catc)30-44                     0.220404   0.113761   1.937  0.06507 .  
factor(ag4catc)45-59                     0.214641   0.102494   2.094  0.04746 *  
factor(ag4catc)60+                      -0.645556   0.175192  -3.685  0.00123 ** 
sexm                                    -0.546442   0.357168  -1.530  0.13967    
ald                                      1.553140   0.211023   7.360 1.74e-07 ***
factor(ed4catc)12                        0.130518   0.083710   1.559  0.13261    
factor(ed4catc)13-15                     0.297324   0.117042   2.540  0.01829 *  
factor(ed4catc)16+                       0.242218   0.151870   1.595  0.12438    
factor(mar3catc)Previously Married       0.417786   0.110525   3.780  0.00097 ***
factor(mar3catc)Never Married            0.017337   0.129782   0.134  0.89489    
factor(ag4catc)30-44:sexm                0.096743   0.200840   0.482  0.63458    
factor(ag4catc)45-59:sexm                0.002637   0.212817   0.012  0.99022    
factor(ag4catc)60+:sexm                 -0.037809   0.302028  -0.125  0.90146    
sexm:ald                                -0.200417   0.242241  -0.827  0.41654    
sexm:factor(ed4catc)12                  -0.137780   0.271016  -0.508  0.61603    
sexm:factor(ed4catc)13-15               -0.168790   0.269350  -0.627  0.53705    
sexm:factor(ed4catc)16+                 -0.194018   0.344159  -0.564  0.57838    
sexm:factor(mar3catc)Previously Married  0.182504   0.207867   0.878  0.38903    
sexm:factor(mar3catc)Never Married       0.231898   0.211963   1.094  0.28526    
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for quasibinomial family taken to be 1.001859)

Number of Fisher Scoring iterations: 4
```

```
> # AIC for 2nd model
> AIC(mod81_int)
     eff.p        AIC   deltabar
  19.51813 5303.89335    1.02727


> # Test interactions of sex by all other predictors
> regTermTest(mod81_int, ~sexm:factor(ag4catc))
Wald test for sexm:factor(ag4catc)
 in svyglm(formula = mde ~ factor(ag4catc) + sexm + ald + factor(ed4catc) +
    factor(mar3catc) + sexm * factor(ag4catc) + sexm * ald +
    sexm * factor(ed4catc) + sexm * factor(mar3catc), design = ncsrsvyp2,
    family = quasibinomial)
F =  0.259966  on  3  and  23  df: p= 0.85346
> regTermTest(mod81_int, ~sexm:ald)
Wald test for sexm:ald
 in svyglm(formula = mde ~ factor(ag4catc) + sexm + ald + factor(ed4catc) +
    factor(mar3catc) + sexm * factor(ag4catc) + sexm * ald +
    sexm * factor(ed4catc) + sexm * factor(mar3catc), design = ncsrsvyp2,
    family = quasibinomial)
F =  0.6845  on  1  and  23  df: p= 0.41654
> regTermTest(mod81_int, ~sexm:factor(ed4catc))
Wald test for sexm:factor(ed4catc)
 in svyglm(formula = mde ~ factor(ag4catc) + sexm + ald + factor(ed4catc) +
    factor(mar3catc) + sexm * factor(ag4catc) + sexm * ald +
    sexm * factor(ed4catc) + sexm * factor(mar3catc), design = ncsrsvyp2,
    family = quasibinomial)
F =  0.1318233  on  3  and  23  df: p= 0.94015
> regTermTest(mod81_int, ~sexm:factor(mar3catc))
Wald test for sexm:factor(mar3catc)
 in svyglm(formula = mde ~ factor(ag4catc) + sexm + ald + factor(ed4catc) +
    factor(mar3catc) + sexm * factor(ag4catc) + sexm * ald +
    sexm * factor(ed4catc) + sexm * factor(mar3catc), design = ncsrsvyp2,
    family = quasibinomial)
F =  0.783767  on  2  and  23  df: p= 0.4685
>
>
> # Design-Adjusted LRT test in R, tests model 1 v. model 2 with interactions
> anova(mod81, mod81_int, test="Chisq", method="LRT")
Working (Rao-Scott) LRT for factor(ag4catc):sexm sexm:ald sexm:factor(ed4catc) sexm:factor(mar3catc)
 in svyglm(formula = mde ~ factor(ag4catc) + sexm + ald + factor(ed4catc) +
    factor(mar3catc) + sexm * factor(ag4catc) + sexm * ald +
    sexm * factor(ed4catc) + sexm * factor(mar3catc), design = ncsrsvyp2,
    family = quasibinomial)
Working 2logLR =  3.412191 p= 0.90812
(scale factors:  3 1.5 1.1 0.84 0.78 0.65 0.57 0.34 0.23 )
```

```
> # Alternative Bayesian method
> # make sure brms and cssampling packages are loaded
> library(brms)
> library(csSampling)
>
> # create new factor variable of sex
> ncsr$sexc <- factor(ncsr$sex, levels = 1:2, labels=c("Male","Female"))
> ncsr$sexc <- relevel(ncsr$sexc, ref="Female")


> # create complete data set with variables of interest
> ncsr.red <- ncsr[, c("mde", "mar3catc", "ed4catc", "sexc", "ag4catc", "ald", "seclustr", "sestrat",
"ncsrwtlg")]
> ncsr.red <- ncsr.red[complete.cases(ncsr.red),]

> # need to normalize NCS-R weights to match what is done for Stan modeling
> ncsr.red$wtsc <- ncsr.red$ncsrwtlg / mean(ncsr.red$ncsrwtlg)

> # survey design object
> ncsr.des <- svydesign(id = ~seclustr, strata = ~sestrat, weights = ~wtsc, nest = T, data = ncsr.red)

> # Bayesian approach, flat prior
> set.seed(41279)
> model_formula <- formula("mde|weights(wtsc) ~ ag4catc + sexc + ald + ed4catc + mar3catc")
> mod.brms <- cs_sampling_brms(svydes = ncsr.des, brmsmod = brmsformula(model_formula), data = ncsr.red, family
= bernoulli())
[1] "compiling stan model"
[1] "stan fitting"

SAMPLING FOR MODEL 'anon_model' NOW (CHAIN 1).
Chain 1:
Chain 1: Gradient evaluation took 0.003172 seconds
Chain 1: 1000 transitions using 10 leapfrog steps per transition would take 31.72 seconds.
Chain 1: Adjust your expectations accordingly!
Chain 1:
Chain 1:
Chain 1: Iteration:    1 / 2000 [  0%]  (Warmup)
Chain 1: Iteration:  200 / 2000 [ 10%]  (Warmup)
Chain 1: Iteration:  400 / 2000 [ 20%]  (Warmup)
Chain 1: Iteration:  600 / 2000 [ 30%]  (Warmup)
Chain 1: Iteration:  800 / 2000 [ 40%]  (Warmup)
Chain 1: Iteration: 1000 / 2000 [ 50%]  (Warmup)
Chain 1: Iteration: 1001 / 2000 [ 50%]  (Sampling)
Chain 1: Iteration: 1200 / 2000 [ 60%]  (Sampling)
Chain 1: Iteration: 1400 / 2000 [ 70%]  (Sampling)
Chain 1: Iteration: 1600 / 2000 [ 80%]  (Sampling)
Chain 1: Iteration: 1800 / 2000 [ 90%]  (Sampling)
Chain 1: Iteration: 2000 / 2000 [100%]  (Sampling)
Chain 1:
Chain 1:  Elapsed Time: 29.456 seconds (Warm-up)
Chain 1:                31.245 seconds (Sampling)
Chain 1:                60.701 seconds (Total)
Chain 1:
[1] "gradient evaluation"
Warning message:
In mrbweights(design$cluster, design$strata, design$fpc, ...) :
  Design is sampled with replacement: only first stage used
> mod.brms$stan_fit
Inference for Stan model: anon_model.
```

```
1 chains, each with iter=2000; warmup=1000; thin=1;
post-warmup draws per chain=1000, total post-warmup draws=1000.

                mean se_mean   sd     2.5%      25%      50%      75%    97.5% n_eff Rhat
b[1]            0.25    0.00 0.11     0.04     0.17     0.25     0.33     0.46   822    1
b[2]            0.20    0.00 0.12    -0.02     0.12     0.20     0.29     0.43   727    1
b[3]           -0.68    0.01 0.14    -0.95    -0.77    -0.67    -0.58    -0.41   723    1
b[4]           -0.58    0.00 0.07    -0.71    -0.63    -0.58    -0.53    -0.43  1130    1
b[5]            1.43    0.00 0.12     1.20     1.34     1.43     1.52     1.67  1488    1
b[6]            0.09    0.00 0.11    -0.13     0.01     0.09     0.16     0.29   728    1
b[7]            0.24    0.00 0.11     0.02     0.16     0.23     0.31     0.44   751    1
b[8]            0.17    0.00 0.12    -0.04     0.09     0.17     0.25     0.39   575    1
b[9]            0.49    0.00 0.09     0.32     0.43     0.48     0.54     0.66  1510    1
b[10]           0.12    0.00 0.10    -0.09     0.04     0.12     0.18     0.32   914    1
Intercept      -1.43    0.00 0.03    -1.50    -1.45    -1.43    -1.41    -1.37  1129    1
lprior         -2.12    0.00 0.01    -2.14    -2.13    -2.12    -2.12    -2.11  1124    1
b_Intercept    -1.59    0.01 0.14    -1.85    -1.67    -1.59    -1.50    -1.31   667    1
lp__        -2641.75    0.09 2.23 -2646.68 -2643.04 -2641.54 -2640.07 -2638.31   565    1

Samples were drawn using NUTS(diag_e) at Sun Apr 13 13:18:07 2025.
For each parameter, n_eff is a crude measure of effective sample size,
and Rhat is the potential scale reduction factor on split chains (at
convergence, Rhat=1).


> # Bayesian with informative prior for ALD coefficient
> mod.brms <- cs_sampling_brms(svydes = ncsr.des, brmsmod = brmsformula(model_formula), prior =
c(set_prior("normal(1.0,0.2)", class = "b", coef="ald")), data = ncsr.red, family = bernoulli())
[1] "compiling stan model"
[1] "stan fitting"

SAMPLING FOR MODEL 'anon_model' NOW (CHAIN 1).
Chain 1:
Chain 1: Gradient evaluation took 0.003292 seconds
Chain 1: 1000 transitions using 10 leapfrog steps per transition would take 32.92 seconds.
Chain 1: Adjust your expectations accordingly!
Chain 1:
Chain 1:
Chain 1: Iteration:    1 / 2000 [  0%]  (Warmup)
Chain 1: Iteration:  200 / 2000 [ 10%]  (Warmup)
Chain 1: Iteration:  400 / 2000 [ 20%]  (Warmup)
Chain 1: Iteration:  600 / 2000 [ 30%]  (Warmup)
Chain 1: Iteration:  800 / 2000 [ 40%]  (Warmup)
Chain 1: Iteration: 1000 / 2000 [ 50%]  (Warmup)
Chain 1: Iteration: 1001 / 2000 [ 50%]  (Sampling)
Chain 1: Iteration: 1200 / 2000 [ 60%]  (Sampling)
Chain 1: Iteration: 1400 / 2000 [ 70%]  (Sampling)
Chain 1: Iteration: 1600 / 2000 [ 80%]  (Sampling)
Chain 1: Iteration: 1800 / 2000 [ 90%]  (Sampling)
Chain 1: Iteration: 2000 / 2000 [100%]  (Sampling)
Chain 1:
Chain 1:  Elapsed Time: 30.37 seconds (Warm-up)
Chain 1:                29.98 seconds (Sampling)
Chain 1:                60.35 seconds (Total)
Chain 1:
[1] "gradient evaluation"
Warning message:
In mrbweights(design$cluster, design$strata, design$fpc, ...) :
  Design is sampled with replacement: only first stage used
```

```
> mod.brms$stan_fit
Inference for Stan model: anon_model.
1 chains, each with iter=2000; warmup=1000; thin=1;
post-warmup draws per chain=1000, total post-warmup draws=1000.

                mean se_mean   sd     2.5%      25%      50%      75%    97.5% n_eff Rhat
b[1]            0.26    0.00 0.10     0.06     0.19     0.26     0.33     0.46   720    1
b[2]            0.21    0.00 0.11     0.00     0.14     0.21     0.28     0.41   691    1
b[3]           -0.68    0.01 0.13    -0.93    -0.77    -0.68    -0.59    -0.43   671    1
b[4]           -0.56    0.00 0.07    -0.71    -0.61    -0.56    -0.51    -0.42  1652    1
b[5]            1.30    0.00 0.11     1.09     1.23     1.30     1.37     1.51  1573    1
b[6]            0.07    0.00 0.11    -0.12     0.00     0.07     0.15     0.28   786    1
b[7]            0.22    0.00 0.11     0.01     0.14     0.23     0.30     0.45   817    1
b[8]            0.15    0.00 0.12    -0.08     0.08     0.15     0.23     0.40   759    1
b[9]            0.49    0.00 0.09     0.33     0.43     0.49     0.54     0.66  1404    1
b[10]           0.12    0.00 0.10    -0.07     0.05     0.12     0.18     0.30   961    1
Intercept      -1.43    0.00 0.03    -1.50    -1.45    -1.43    -1.41    -1.37  1407    1
lprior         -2.72    0.02 0.83    -4.70    -3.18    -2.56    -2.09    -1.53  1473    1
b_Intercept    -1.57    0.00 0.13    -1.85    -1.66    -1.57    -1.49    -1.32   742    1
lp__        -2642.64    0.10 2.26 -2647.81 -2644.07 -2642.34 -2640.93 -2639.07   478    1

Samples were drawn using NUTS(diag_e) at Sun Apr 13 13:23:42 2025.
For each parameter, n_eff is a crude measure of effective sample size,
and Rhat is the potential scale reduction factor on split chains (at
convergence, Rhat=1).


> # Model 8.2 with logit, probit, and cloglog comparison
> summary(ex82_logit <- svyglm(ald ~ factor(ag4catc) + sexm + factor(ed4catc) + factor(mar3catc),
family=quasibinomial, design=ncsrsvyp2))

Call:
svyglm(formula = ald ~ factor(ag4catc) + sexm + factor(ed4catc) +
    factor(mar3catc), design = ncsrsvyp2, family = quasibinomial)

Survey design:
svydesign(strata = ~sestrat, id = ~seclustr, weights = ~ncsrwtlg,
    data = ncsrp2, nest = T)

Coefficients:
                                   Estimate Std. Error t value Pr(>|t|)    
(Intercept)                        -3.12432    0.22527 -13.869 2.54e-15 ***
factor(ag4catc)30-44                0.14628    0.17813   0.821 0.417443    
factor(ag4catc)45-59               -0.05071    0.14392  -0.352 0.726833    
factor(ag4catc)60+                 -1.12034    0.21248  -5.273 8.25e-06 ***
sexm                                0.99799    0.11910   8.379 1.11e-09 ***
factor(ed4catc)12                  -0.26844    0.19373  -1.386 0.175150    
factor(ed4catc)13-15               -0.26448    0.17614  -1.502 0.142716    
factor(ed4catc)16+                 -0.73623    0.19719  -3.734 0.000712 ***
factor(mar3catc)Previously Married  0.51783    0.14206   3.645 0.000910 ***
factor(mar3catc)Never Married       0.06532    0.16871   0.387 0.701130    
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for quasibinomial family taken to be 0.9751118)

Number of Fisher Scoring iterations: 6

> regTermTest(ex82_logit, ~factor(ag4catc))
```

```
Wald test for factor(ag4catc)
 in svyglm(formula = ald ~ factor(ag4catc) + sexm + factor(ed4catc) +
    factor(mar3catc), design = ncsrsvyp2, family = quasibinomial)
F =  12.66092  on  3  and  33  df: p= 1.1402e-05
> regTermTest(ex82_logit, ~factor(ed4catc))
Wald test for factor(ed4catc)
 in svyglm(formula = ald ~ factor(ag4catc) + sexm + factor(ed4catc) +
    factor(mar3catc), design = ncsrsvyp2, family = quasibinomial)
F =  5.037044  on  3  and  33  df: p= 0.005531
> regTermTest(ex82_logit, ~factor(mar3catc))
Wald test for factor(mar3catc)
 in svyglm(formula = ald ~ factor(ag4catc) + sexm + factor(ed4catc) +
    factor(mar3catc), design = ncsrsvyp2, family = quasibinomial)
F =  6.696711  on  2  and  33  df: p= 0.003622


> summary(ex82_probit <- svyglm(ald ~ factor(ag4catc) + sexm + factor(ed4catc) + factor(mar3catc),
family=quasibinomial(link=probit),  design=ncsrsvyp2))

Call:
svyglm(formula = ald ~ factor(ag4catc) + sexm + factor(ed4catc) +
    factor(mar3catc), design = ncsrsvyp2, family = quasibinomial(link = probit))

Survey design:
svydesign(strata = ~sestrat, id = ~seclustr, weights = ~ncsrwtlg,
    data = ncsrp2, nest = T)

Coefficients:
                                  Estimate Std. Error t value Pr(>|t|)
(Intercept)                       -1.71944    0.10648 -16.149  < 2e-16 ***
factor(ag4catc)30-44               0.06532    0.08541   0.765 0.449852
factor(ag4catc)45-59              -0.03447    0.06734  -0.512 0.612141
factor(ag4catc)60+                -0.53128    0.09372  -5.669 2.56e-06 ***
sexm                               0.47084    0.05695   8.268 1.51e-09 ***
factor(ed4catc)12                 -0.12378    0.09497  -1.303 0.201457
factor(ed4catc)13-15              -0.12438    0.08504  -1.463 0.153028
factor(ed4catc)16+                -0.33956    0.09237  -3.676 0.000836 ***
factor(mar3catc)Previously Married  0.25478    0.07014   3.633 0.000942 ***
factor(mar3catc)Never Married      0.03895    0.07795   0.500 0.620593
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for quasibinomial family taken to be 0.9825414)

Number of Fisher Scoring iterations: 6

> regTermTest(ex82_probit, ~factor(ag4catc))
Wald test for factor(ag4catc)
 in svyglm(formula = ald ~ factor(ag4catc) + sexm + factor(ed4catc) +
    factor(mar3catc), design = ncsrsvyp2, family = quasibinomial(link = probit))
F =  16.00526  on  3  and  33  df: p= 1.3479e-06
> regTermTest(ex82_probit, ~factor(ed4catc))
Wald test for factor(ed4catc)
 in svyglm(formula = ald ~ factor(ag4catc) + sexm + factor(ed4catc) +
    factor(mar3catc), design = ncsrsvyp2, family = quasibinomial(link = probit))
F =  5.005033  on  3  and  33  df: p= 0.0057059
> regTermTest(ex82_probit, ~factor(mar3catc))
Wald test for factor(mar3catc)
 in svyglm(formula = ald ~ factor(ag4catc) + sexm + factor(ed4catc) +
```

```
    factor(mar3catc), design = ncsrsvyp2, family = quasibinomial(link = probit))
F =  6.701662  on  2  and  33  df: p= 0.0036093


> summary(ex82_cloglog <- svyglm(ald ~ factor(ag4catc) + sexm + factor(ed4catc) + factor(mar3catc),
family=quasibinomial(link=cloglog),  design=ncsrsvyp2))

Call:
svyglm(formula = ald ~ factor(ag4catc) + sexm + factor(ed4catc) +
    factor(mar3catc), design = ncsrsvyp2, family = quasibinomial(link = cloglog))

Survey design:
svydesign(strata = ~sestrat, id = ~seclustr, weights = ~ncsrwtlg,
    data = ncsrp2, nest = T)

Coefficients:
                                   Estimate Std. Error t value Pr(>|t|)    
(Intercept)                        -3.14840    0.21716 -14.498 7.19e-16 ***
factor(ag4catc)30-44                0.14304    0.17125   0.835 0.409556    
factor(ag4catc)45-59               -0.04524    0.13967  -0.324 0.748043    
factor(ag4catc)60+                 -1.08287    0.20816  -5.202 1.02e-05 ***
sexm                                0.96522    0.11494   8.398 1.06e-09 ***
factor(ed4catc)12                  -0.26010    0.18467  -1.408 0.168358    
factor(ed4catc)13-15               -0.25556    0.16843  -1.517 0.138700    
factor(ed4catc)16+                 -0.71265    0.19042  -3.743 0.000694 ***
factor(mar3catc)Previously Married  0.49354    0.13522   3.650 0.000898 ***
factor(mar3catc)Never Married       0.06049    0.16369   0.370 0.714073    
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for quasibinomial family taken to be 0.9745877)

Number of Fisher Scoring iterations: 6

> regTermTest(ex82_cloglog, ~factor(ag4catc))
Wald test for factor(ag4catc)
 in svyglm(formula = ald ~ factor(ag4catc) + sexm + factor(ed4catc) +
    factor(mar3catc), design = ncsrsvyp2, family = quasibinomial(link = cloglog))
F =  12.18828  on  3  and  33  df: p= 1.5779e-05
> regTermTest(ex82_cloglog, ~factor(ed4catc))
Wald test for factor(ed4catc)
 in svyglm(formula = ald ~ factor(ag4catc) + sexm + factor(ed4catc) +
    factor(mar3catc), design = ncsrsvyp2, family = quasibinomial(link = cloglog))
F =  5.0184  on  3  and  33  df: p= 0.0056322
> regTermTest(ex82_cloglog, ~factor(mar3catc))
Wald test for factor(mar3catc)
 in svyglm(formula = ald ~ factor(ag4catc) + sexm + factor(ed4catc) +
    factor(mar3catc), design = ncsrsvyp2, family = quasibinomial(link = cloglog))
F =  6.710508  on  2  and  33  df: p= 0.0035866
>
```