**ORIGINAL MODEL RESULTS:**

summary(testmodel)

Call:

svyglm(formula = outcome ~ W2wthlessYP + locus + W2plann16YP3 +

sum\_beh\_any4 + sum\_subs\_any2 + +W2yschat2 + W2sportYP2 + W2bulev +

W2ethgrpYP.x2 + W2jobYP + W2careYP + IMDRSCORE2 + W2SexYP,

design = sdesign, family = quasibinomial(link = "logit"),

data = test)

Survey design:

svydesign(ids = ~test$SampPSU.x, strata = test$SampStratum.x,

data = test, weights = test$W8FINWT)

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) -2.37538 0.61047 -3.89 0.00011 \*\*\*

W2wthlessYPNo more than usual 0.07471 0.23277 0.32 0.74834

W2wthlessYPRather more than usual -0.01418 0.25597 -0.06 0.95583

W2wthlessYPMuch more than usual 0.53030 0.26253 2.02 0.04387 \*

locusexternal 0.60192 0.30339 1.98 0.04776 \*

W2plann16YP3loworuncertain 0.26778 0.26281 1.02 0.30870

sum\_beh\_any4 0.00295 0.21162 0.01 0.98887

sum\_subs\_any2Yes -0.02431 0.20614 -0.12 0.90617

W2yschat2 -0.03567 0.01207 -2.96 0.00326 \*\*

W2sportYP2Moderate 0.05694 0.25800 0.22 0.82542

W2sportYP2Low or never 0.59797 0.27067 2.21 0.02757 \*

W2bulevYes 0.48712 0.22908 2.13 0.03391 \*

W2ethgrpYP.x2Other -0.23543 0.22042 -1.07 0.28594

W2jobYPNo 0.51451 0.22314 2.31 0.02150 \*

W2careYPNo -0.46699 0.31503 -1.48 0.13882

IMDRSCORE2 0.02753 0.00507 5.43 8.4e-08 \*\*\*

W2SexYPFemale 0.41244 0.21913 1.88 0.06034 .

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for quasibinomial family taken to be 1)

Number of Fisher Scoring iterations: 5

**IMPUTED MODEL RESULTS:**

> screenreg(convertModel(model3))

================================================

Model 1

------------------------------------------------

(Intercept) -2.04 \*\*\*

(0.51)

W2wthlessYPNo more than usual 0.02

(0.22)

W2wthlessYPRather more than usual -0.11

(0.24)

W2wthlessYPMuch more than usual 0.60 \*

(0.24)

locusexternal 0.32

(0.30)

W2plann16YP3loworuncertain 0.23

(0.24)

sum\_subs\_any2Yes -0.07

(0.19)

sum\_beh\_any4Yes 0.15

(0.19)

W2yschat2 -0.02 \*

(0.01)

W2sportYP2Moderate 0.04

(0.23)

W2sportYP2Low or never 0.63 \*\*

(0.24)

W2bulevNo -0.48 \*

(0.21)

W2ethgrpYP.x2Other -0.38

(0.20)

W2jobYPNo 0.44 \*

(0.21)

W2careYPNo -0.56 \*

(0.28)

IMDRSCORE2 0.03 \*\*\*

(0.00)

W2SexYPFemale 0.41 \*

(0.20)

------------------------------------------------

R2 -109.05

BIC (null) 10578

N 2224

================================================

\*\*\* p < 0.001; \*\* p < 0.01; \* p < 0.05

**THE SCRIPT:**

test2 <- subset(dataset3, select = c(W2alceverYP, W2cignowYP, W2canntryYP, W2sprayYP,

W2smashYP, W2shopYP, W2fightYP, W2bulev, W2plann16YP3,

W2wthlessYP, locus, IMDRSCORE2, W2ethgrpYP.x2, W2careYP, W2jobYP,

W2sportYP2, W2SexYP, W2yschat2, outcome,

SampPSU.x, SampStratum.x, W8FINWT))

imputationstest2 <- mice(test2, m=10, maxit=30, printFlag = FALSE)

lm\_svy\_mi2 <- function(formula, imputationstest2) {

#setting up null objects allows us to easily add results

#later

b <- se <- R2 <- NULL

#now loop through our imputations and run the model

for(i in 1:imputationstest2$m) {

#grab the complete dataset

imputation2 <- complete(imputationstest2, i)

#include derivation process here

imputation2$W2alcohol <- rep(NA, 2224)

imputation2$W2alcohol[which(imputation2$W2alceverYP == "Yes")] <- 1

imputation2$W2alcohol[which(imputation2$W2alceverYP == "No ")] <- 0

imputation2$W2cig <- rep(NA, 2224)

imputation2$W2cig[which(imputation2$W2cignowYP == "Yes")] <- 1

imputation2$W2cig[which(imputation2$W2cignowYP == "No ")] <- 0

imputation2$W2cann <- rep(NA, 2224)

imputation2$W2cann[which(imputation2$W2canntryYP == "Yes")] <- 1

imputation2$W2cann[which(imputation2$W2canntryYP == "No ")] <- 0

#sum the rows

imputation2$sum <- rowSums(imputation2[, c("W2cann", "W2cig", "W2alcohol")], na.rm = FALSE)

#rearrange the new derived variable---yes to any of the three variables

imputation2$sum\_subs\_any3 <- rep(NA, 2224)

imputation2$sum\_subs\_any3[which(imputation2$sum == 1)] <- "Yes"

imputation2$sum\_subs\_any3[which(imputation2$sum == 2)] <- "Yes"

imputation2$sum\_subs\_any3[which(imputation2$sum == 3)] <- "Yes"

imputation2$sum\_subs\_any3[which(imputation2$sum == 0)] <- "No"

#yes to any two of the variables

imputation2$sum\_subs\_any2 <- rep(NA, 2224)

imputation2$sum\_subs\_any2[which(imputation2$sum == 1)] <- "No"

imputation2$sum\_subs\_any2[which(imputation2$sum == 2)] <- "Yes"

imputation2$sum\_subs\_any2[which(imputation2$sum == 3)] <- "Yes"

imputation2$sum\_subs\_any2[which(imputation2$sum == 0)] <- "No"

#behavioral problems derivation

imputation2$W2spray <- rep(NA, 2224)

imputation2$W2spray[which(imputation2$W2sprayYP == "Yes")] <- 1

imputation2$W2spray[which(imputation2$W2sprayYP == "No ")] <- 0

imputation2$W2smash <- rep(NA, 2224)

imputation2$W2smash[which(imputation2$W2smashYP == "Yes")] <- 1

imputation2$W2smash[which(imputation2$W2smashYP == "No ")] <- 0

imputation2$W2shop <- rep(NA, 2224)

imputation2$W2shop[which(imputation2$W2shopYP == "Yes")] <- 1

imputation2$W2shop[which(imputation2$W2shopYP == "No ")] <- 0

imputation2$W2fight <- rep(NA, 2224)

imputation2$W2fight[which(imputation2$W2fightYP == "Yes")] <- 1

imputation2$W2fight[which(imputation2$W2fightYP == "No ")] <- 0

imputation2$sum\_beh <- rowSums(imputation2[, c("W2spray", "W2smash", "W2shop", "W2fight")], na.rm = FALSE)

#rearrange the new derived variable-----any 4

imputation2$sum\_beh\_any4 <- rep(NA, 2224)

imputation2$sum\_beh\_any4[which(imputation2$sum\_beh == 1)] <- "Yes"

imputation2$sum\_beh\_any4[which(imputation2$sum\_beh == 2)] <- "Yes"

imputation2$sum\_beh\_any4[which(imputation2$sum\_beh == 3)] <- "Yes"

imputation2$sum\_beh\_any4[which(imputation2$sum\_beh == 4)] <- "Yes"

imputation2$sum\_beh\_any4[which(imputation2$sum\_beh == 0)] <- "No"

#create the design effect object

imputation.svy2 <- svydesign(ids=~SampPSU.x, strata=~SampStratum.x, weight=~W8FINWT,

data=imputation2)

#run the model

model2 <- svyglm(outcome ~ W2wthlessYP+locus+W2plann16YP3+

sum\_subs\_any2+sum\_beh\_any4+W2yschat2+W2sportYP2+W2bulev+

W2ethgrpYP.x2+W2jobYP+W2careYP+IMDRSCORE2+W2SexYP, design=imputation.svy2,

family=quasibinomial(link = "logit"))

#collect the results

b <- cbind(b, coef(model2))

se <- cbind(se, summary(model2)$coef[,2])

#We should get R squared too. Sadly, svyglm won't give

#it to us by default, but we can get it from some of the

#slots in the model output

SSR <- sum((model2$residuals)^2)

SSY <- sum((model2$y-mean(model2$y))^2)

R2 <- c(R2,1-SSR/SSY)

}

#now pool the results

b.pool <- apply(b, 1, mean)

between.var <- apply(b, 1, var)

within.var <- apply(se^2, 1, mean)

se.pool <- sqrt(within.var+between.var+between.var/imputationstest2$m)

t.pool <- b.pool/se.pool

pvalue.pool <- (1-pnorm(abs(t.pool)))\*2

coefficients <- data.frame(b.pool, se.pool, t.pool, pvalue.pool)

#lets take the mean R2 value

r.squared <- mean(R2)

#we can also grap n and p from the last model since

#they should be the same across all iterations

n <- nobs(model2)

p <- length(model2$coefficients)-1

#go ahead and calculate BIC.null

bic.null <- n\*log(1-r.squared)+p\*log(n)

#return everything in a list

return(list(coef=coefficients,

n=n,

r.squared=r.squared,

bic.null=bic.null))

}

lm\_svy\_mi2(outcome ~ W2wthlessYP+locus+W2plann16YP3+

sum\_subs\_any2+sum\_beh2+W2yschat2+W2sportYP2+W2bulev+

W2ethgrpYP.x2+W2jobYP+W2careYP+IMDRSCORE2+W2SexYP, imputationstest2)

install.packages("texreg")

library(texreg)

convertModel <- function(model2) {

tr <- createTexreg(

coef.names = rownames(model2$coef),

coef = model2$coef$b.pool,

se = model2$coef$se.pool,

pvalues = model2$coef$pvalue.pool,

gof.names = c("R2","BIC (null)","N"),

gof = c(model2$r.squared, model2$bic.null, model2$n),

gof.decimal = c(T,F,F)

)

}

model3 <- lm\_svy\_mi2(outcome ~ W2wthlessYP+locus+W2plann16YP3+

sum\_subs\_any2+sum\_beh2+W2yschat2+W2sportYP2+W2bulev+

W2ethgrpYP.x2+W2jobYP+W2careYP+IMDRSCORE2+W2SexYP, imputationstest2)

screenreg(convertModel(model3))