**S1 File. The complete R Code**

*#THIS IS THE R CODE FOR THE ANALYSIS ON GRAINGER ET AL.*

*#Addressing complexity in the assessment of the drivers of consumer food waste*

*#matthew.grainger@ncl.ac.uk*

*#Set working directory*

setwd("C:/Users/Matt/Documents/REFRESH")

*#load packages*

**library**(ggplot2)

**library**(randomForest)

**library**(glmulti)

**library**(MuMIn)

options(na.action = "na.fail")

**library**(car)

*#Get data*

data<-read.csv("Data\_Grainger\_et\_al\_Type1error.csv",header=**TRUE**)

head(data)

names(data)

dim(data)

data<-data[which(data$Study==1),]

data<-data[which(data$IsNA.==0),]

dim(data)

lapply(data, **function**(x) {

**if** (is.numeric(x)) **return**(summary(x))

**if** (is.factor(x)) **return**(table(x))

})

*#remove serial number, Study & is.na columns*

data<-data[,2:55]

dim(data)

*#remove avoidable\_FW & HH\_size 2 columns*

data$Avoidable\_FW<-**NULL**

data$HH\_size\_2<-**NULL**

dim(data)

str(data)

data$HHSize<-as.factor(data$HHSize)

summary(data)

*#######*

*#Full linear model*

*#How many models?*

glmulti(A\_per\_HH~.,data=data,level=1, method='d',confsetsize=10,plotty=T)

*#Error too many predictors*

glmulti(A\_per\_HH~.,data=data,level=2, method='d',confsetsize=10,plotty=T)*#with interaction*

*#Error Oversized candidate set*

mod1<-glm(A\_per\_HH~.,data=data)

sink(file="C:/Users/Matt/Documents/REFRESH/sink.txt")

summary(mod1)

sink()

*#######variable reduction using random forest*

*#discard cooked meat - more than 50 potential categories*

drops <- c("CookMeat")

data<-data[ , !(names(data) %in% drops)]

*#install.packages("party")*

*# library(party)*

*# cforestImpPlot <- function(x) {*

*# cforest\_importance <<- v <- varimp(x)*

*# dotchart(v[order(v)],xlab="Permutation importance")*

*# }*

*# cf1 <- cforest(A\_per\_HH ~ ., data=data,control=cforest\_unbiased())*

*# cforestImpPlot(cf1)*

*# install.packages("Boruta")*

**library**(Boruta)

Boruta1 <- Boruta(A\_per\_HH ~ ., data=data, doTrace = 2, ntree = 500)

print(Boruta1)

plot(Boruta1)

plot(Boruta1, xlab = "", xaxt = "n")

lz<-lapply(1:ncol(Boruta1$ImpHistory),**function**(i)

Boruta1$ImpHistory[is.finite(Boruta1$ImpHistory[,i]),i])

names(lz) <- colnames(Boruta1$ImpHistory)

Labels <- sort(sapply(lz,median))

axis(side = 1,las=2,labels = names(Labels),

at = 1:ncol(Boruta1$ImpHistory), cex.axis = 0.55)

*#selected variables*

myvars<-c("A\_per\_HH","Local.authority", "HHSize", "Q33\_Recoded",

"Discard\_veg", "job\_new", "Q31\_Recoded","Fussy","Discard\_cheese","age\_brackets",

"Discard\_Sellby", "social\_grouping", "children\_3\_to\_11","Q36\_Recoded",

"CupboardTinned")

data1<-data[myvars]

names(data1)

*###look at model set potential*

glmulti(A\_per\_HH~.,data=data1,level=1, method='d',confsetsize=100,plotty=T)

*# 16384 models*

glmulti(A\_per\_HH~.,data=data1,level=2, method='d',confsetsize=100,plotty=T)

*#1e9 models*

fit<-glm(A\_per\_HH~.,data=data1)

*#model selection*

dd<-dredge(fit)

ds<-subset(dd,delta<2)

final<-model.avg(ds)

sink(file="C:/Users/Matt/Documents/REFRESH/sink.txt")

summary(final)

sink()

*###*

*###Get full distribution of p*

modAll<-glmulti(A\_per\_HH~.,data=data1,level=1, method='h',confsetsize=100,plotty=F)

tbl<-weightable(modAll)

nrows <- nrow(tbl)

sink("ModelPvalues.txt",type="output")

**for** (i **in** 1:nrows) {

formulaTest <- as.character(tbl[i,]$model)

model = glm(as.formula(formulaTest), data=data1)

print(paste("Formula: ", formulaTest))

print(summary(model)$coefficients[,4])

}

sink()

*#need to adjust table txt in excel*

pdat<-read.csv("ModelPValues.csv",header=T)

head(pdat)

str(pdat)

graph1 <- **function**(x) {

ggplot(x, aes(x=p)) + geom\_density(fill="red") + xlim(c(0,1))

}

graph1(pdat[pdat$Variable=='social\_groupingDE',])

p<-ggplot(pdat, aes(x=p)) + geom\_density(aes(fill="red"),alpha=.1)+

xlim(0,1)

p <- p + facet\_wrap( ~Variable)

p

*# Generate ggplot2 plot for each level*

plot.list <- by(data = pdat,

INDICES = pdat$Variable,

simplify = **TRUE**,

FUN = **function**(x) {

p %+% x

}

)

**library**(ggplus)

**library**(gridExtra)

*# Build the grobs*

multi.plot <- marrangeGrob(grobs = plot.list,

nrow = 2, ncol = 2,

top = quote(paste('\nPage', g, 'of', pages)))

*# Plot on multiple pages (output plot to R/Rstudio)*

print(multi.plot)

pdf('multi.pdf', width = 28, height = 18)

print(multi.plot)

dev.off()

*####Interactions in the full model*

glmulti(A\_per\_HH~age\_brackets+ Discard\_Sellby+Discard\_veg+Fussy+

HHSize+job\_new+Local.authority+Q31\_Recoded+

Discard\_veg:Fussy+Discard\_veg:HHSize+Discard\_veg:job\_new +Discard\_veg:Local.authority+

Fussy:HHSize+Fussy:job\_new+Fussy:Local.authority+

HHSize:job\_new+Fussy:Local.authority+

job\_new:Local.authority,data=data1,level=1, method='d',confsetsize=100,plotty=T)

*#303227658*

*#256*

fitI<-glm(A\_per\_HH~age\_brackets+ Discard\_Sellby+Discard\_veg+Fussy+

HHSize+job\_new+Local.authority+Q31\_Recoded+

Discard\_veg:Fussy+Discard\_veg:HHSize+Discard\_veg:job\_new +Discard\_veg:Local.authority+

Fussy:HHSize+Fussy:job\_new+Fussy:Local.authority+

HHSize:job\_new+Fussy:Local.authority+

job\_new:Local.authority,data=data1)

dd<-dredge(fitI)

ds<-subset(dd,delta<2)

finalI<-model.avg(ds)

summary(finalI)

*##############remove discard*

names(data)

drops <- c("Discard\_fruit", "Discard\_cooked","Discard\_meal", "Discard\_unused", "Discard\_packs","Discard\_mould", "Discard\_Sellby", "Discard\_veg", "Discard\_Bread", "Discard\_meat",

"Discard\_cheese", "Discard\_milk", "Discard\_ready", "Discard\_home","Discard\_parts" )

data2<-data[ , !(names(data) %in% drops)]

names(data2)

glmulti(A\_per\_HH~.,data=data2,level=1, method='d',confsetsize=10,plotty=T)

*#Error too many*

glmulti(A\_per\_HH~.,data=data2,level=2, method='d',confsetsize=10,plotty=T)*#with interaction*

*#Error too many*

Boruta2 <- Boruta(A\_per\_HH ~ ., data=data2, doTrace = 2, ntree = 500)

print(Boruta2)

plot(Boruta2)

plot(Boruta2, xlab = "", xaxt = "n")

lz<-lapply(1:ncol(Boruta2$ImpHistory),**function**(i)

Boruta2$ImpHistory[is.finite(Boruta2$ImpHistory[,i]),i])

names(lz) <- colnames(Boruta2$ImpHistory)

Labels <- sort(sapply(lz,median))

axis(side = 1,las=2,labels = names(Labels),

at = 1:ncol(Boruta2$ImpHistory), cex.axis = 0.55)

myvars<-c("A\_per\_HH","job\_new","Local.authority", "Fussy",

"HHSize", "Q33\_Recoded", "Q31\_Recoded",

"children\_3\_to\_11","CupboardTinned","CupboardFish",

"CupboardMeat","age\_brackets","CupboardSalad",

"CupboardBread","CupboardFrozen","CupboardFruit")

data2<-data[myvars]

glmulti(A\_per\_HH~.,data=data2,level=1, method='d',confsetsize=10,plotty=T)

*#32768*

glmulti(A\_per\_HH~.,data=data2,level=2, method='d',confsetsize=10,plotty=T)*#with interaction*

*#1e9*

*#model selection*

fit2<-glm(A\_per\_HH~.,data=data2)

dd2<-dredge(fit2)

ds2<-subset(dd2,delta<2)

ds2

summary(fit2)

*###interaction*

glmulti(A\_per\_HH~Fussy+job\_new+Local.authority+HHSize+Q31\_Recoded+age\_brackets,data=data2,level=2, method='d',confsetsize=10,plotty=T)*#with interaction*

*#47850*

glmulti(A\_per\_HH~Fussy+job\_new+Local.authority+HHSize+Q31\_Recoded+age\_brackets,data=data2,level=2, method='h',confsetsize=10,plotty=T)*#with interaction*

modI<-glm(A\_per\_HH ~ job\_new + Local.authority + HHSize + Q31\_Recoded +

age\_brackets + Fussy + age\_brackets:job\_new + job\_new:Fussy +

Local.authority:Fussy,data=data2)

modII<-glm(A\_per\_HH ~ job\_new + Local.authority + HHSize + Q31\_Recoded +

age\_brackets + Fussy + age\_brackets:job\_new + job\_new:Fussy +

Local.authority:Fussy + age\_brackets:Fussy,data=data2)

sum<-model.avg(modI,modII)

summary(sum)

*##############remove LA*

names(data)

myvars<-c("A\_per\_HH", "Q36\_Recoded",

"job\_new","Q31\_Recoded",

"Fussy", "HHSize",

"Q33\_Recoded", "children\_3\_to\_11",

"age\_brackets","CupboardMeat",

"Discard\_Sellby", "Discard\_packs", "Discard\_cheese","Discard\_veg",

"CupboardMeat")

data3<-data[myvars]

Boruta4 <- Boruta(A\_per\_HH ~ ., data=data3, doTrace = 2, ntree = 500)

print(Boruta4)

plot(Boruta4)

plot(Boruta4, xlab = "", xaxt = "n")

lz<-lapply(1:ncol(Boruta4$ImpHistory),**function**(i)

Boruta4$ImpHistory[is.finite(Boruta4$ImpHistory[,i]),i])

names(lz) <- colnames(Boruta4$ImpHistory)

Labels <- sort(sapply(lz,median))

axis(side = 1,las=2,labels = names(Labels),

at = 1:ncol(Boruta4$ImpHistory), cex.axis = 0.55)

*#model selection*

glmulti(A\_per\_HH~.,data=data3,level=1, method='d',confsetsize=10,plotty=T)

*#16384*

glmulti(A\_per\_HH~.,data=data3,level=2, method='d',confsetsize=10,plotty=T)*#with interaction*

*#1e9*

*#model selection*

fit3<-glm(A\_per\_HH~.,data=data3)

dd3<-dredge(fit3)

ds3<-subset(dd3,delta<2)

ds3

mod.avg(ds3)

*##interactions*

glmulti(A\_per\_HH~HHSize+job\_new+Q31\_Recoded+children\_3\_to\_11,data=data3,level=2, method='h',confsetsize=10,plotty=T)

*#113*

fit3I<-glm(A\_per\_HH~HHSize+job\_new+Q31\_Recoded+children\_3\_to\_11+

children\_3\_to\_11:HHSize +children\_3\_to\_11:job\_new + children\_3\_to\_11:Q31\_Recoded,data=data3)

dd3I<-dredge(fit3I)

ds3I<-subset(dd3I,delta<2)

ds3I

sum<-model.avg(ds3I)

summary(sum)

*####*

*##########remove LA & Discard*

names(data)

drops <- c("Local.authority","Discard\_fruit",

"Discard\_cooked","Discard\_meal",

"Discard\_unused", "Discard\_packs",

"Discard\_mould", "Discard\_Sellby",

"Discard\_veg", "Discard\_Bread",

"Discard\_meat","Discard\_cheese", "Discard\_milk",

"Discard\_ready", "Discard\_home","Discard\_parts" )

data4<-data[ , !(names(data) %in% drops)]

*#model selection*

*#model selection*

myvars<-c("A\_per\_HH","Fussy","HHSize",

"CupboardVeg",

"CupboardBread",

"CupboardMeat",

"CupboardFish",

"CupboardMilk",

"CupboardTinned",

"CupboardFrozen",

"CupboardSalad",

"age\_brackets",

"Q33\_Recoded",

"Q31\_Recoded",

"Q36\_Recoded",

"job\_new", "children\_3\_to\_11")

data4<-data[myvars]

glmulti(A\_per\_HH~HHSize+.,data=data4,level=1, method='d',confsetsize=10,plotty=T)

*#65536*

glmulti(A\_per\_HH~.,data=data4,level=2, method='d',confsetsize=10,plotty=T)*#with interaction*

*#1e9*

fit4<-glm(A\_per\_HH~.,data=data4)

dd4<-dredge(fit4)

ds4<-subset(dd4,delta<2)

ds4

summary(ds4)

model1<-glm(A\_per\_HH~age\_brackets+Fussy+HHSize+job\_new+Q31\_Recoded, data=data4)

model2<-glm(A\_per\_HH~age\_brackets+CupboardVeg+Fussy+HHSize+job\_new+Q31\_Recoded, data=data4)

model3<-glm(A\_per\_HH~age\_brackets+children\_3\_to\_11+Fussy+HHSize+job\_new+Q31\_Recoded, data=data4)

m4<-model.avg(model1,model2,model3)

summary(m4)

*#plot Figure 1 Random forests*

*# wrap1$Foodwaste<-wrap1$A\_per\_HH*

*# wrap1$A\_per\_HH<-NULL*

*# wrap1$Household<-wrap$HHSize*

*# wrap1$HHSize<-NULL*

*# wrap1$Vegetables<-wrap1$Discard\_veg*

*# wrap1$Composition<-wrap1$Q33\_Recoded*

*# wrap1$Ownership<-wrap1$Q31\_Recoded*

*# wrap1$Housetype<-wrap1$Q36\_Recoded*

*#*

*# wrap$Ownership<-wrap$Q31\_Recoded*

*# wrap$Housetype<-wrap$Q36\_Recoded*

par(mfrow = c(2,2))

plot(Boruta1, xlab = "", xaxt = "n", main="A")

lz<-lapply(1:ncol(Boruta1$ImpHistory),**function**(i)

Boruta1$ImpHistory[is.finite(Boruta1$ImpHistory[,i]),i])

colnames(Boruta1$ImpHistory)<-c("Local.authority", "Fussy", "HHSize", "ShoppingTrips",

"LeftOvers", "RiceLeft", "RiceLeft2", "PastaLeft" ,

"PastaLeft2", "Discard\_fruit","Discard\_cooked", "Discard\_meal" ,

"Discard\_unused", "Discard\_packs","Discard\_mould" , "Discard\_Sellby" ,

"Discard\_veg", "Discard\_Bread","Discard\_meat" , "Discard\_cheese" ,

"Discard\_milk", "Discard\_ready","Discard\_home" , "Discard\_parts" ,

"HowDisposed", "ShoppingHabits","Preplan" , "CupboardFruit" ,

"CupboardVeg", "CupboardBread","CupboardMeat" , "CupboardFish" ,

"CupboardMilk", "CupboardReady","CupboardTinned" , "CupboardFrozen",

"CupboardSalad", "PreplanList","Cheese" , "Apple" ,

"Carrot", "gender","age\_brackets" , "Age\_over65" ,

"Ownership", "Housetype","Composition", "children\_3\_to\_11",

"job\_new", "social\_grouping","shadowMax","shadowMean","shadowMin" )

names(lz) <- colnames(Boruta1$ImpHistory)

Labels <- sort(sapply(lz,median))

axis(side = 1,las=2,labels = names(Labels),

at = 1:ncol(Boruta1$ImpHistory), cex.axis = 0.7)

plot(Boruta2, xlab = "", xaxt = "n", main="B")

lz<-lapply(1:ncol(Boruta2$ImpHistory),**function**(i)

Boruta2$ImpHistory[is.finite(Boruta2$ImpHistory[,i]),i])

colnames(Boruta2$ImpHistory)<-c("Local.authority", "Fussy", "HHSize", "ShoppingTrips",

"LeftOvers", "RiceLeft", "RiceLeft2", "PastaLeft" ,

"PastaLeft2", "HowDisposed", "ShoppingHabits","Preplan" , "CupboardFruit" ,

"CupboardVeg", "CupboardBread","CupboardMeat" , "CupboardFish" ,

"CupboardMilk", "CupboardReady","CupboardTinned" , "CupboardFrozen",

"CupboardSalad", "PreplanList","Cheese" , "Apple" ,

"Carrot", "gender","age\_brackets" , "Age\_over65" ,

"Ownership", "Housetype","Composition", "children\_3\_to\_11",

"job\_new", "social\_grouping","shadowMax","shadowMean","shadowMin" )

names(lz) <- colnames(Boruta2$ImpHistory)

Labels <- sort(sapply(lz,median))

axis(side = 1,las=2,labels = names(Labels),

at = 1:ncol(Boruta2$ImpHistory), cex.axis = 0.7)

*#Boruta3 <- Boruta(A\_per\_HH ~ ., data=data4, doTrace = 2, ntree = 500)*

*#print(Boruta3)*

*#plot(Boruta3)*

plot(Boruta3, xlab = "", xaxt = "n", main="C")

lz<-lapply(1:ncol(Boruta3$ImpHistory),**function**(i)

Boruta3$ImpHistory[is.finite(Boruta3$ImpHistory[,i]),i])

colnames(Boruta3$ImpHistory)<-c("Fussy", "HHSize", "CupboardVeg", "CupboardBread","CupboardMeat" , "CupboardFish" ,

"CupboardMilk", "CupboardTinned" , "CupboardFrozen",

"CupboardSalad", "age\_brackets" , "Ownership", "Housetype","Composition","job\_new",

"children\_3\_to\_11","shadowMax","shadowMean","shadowMin" )

names(lz) <- colnames(Boruta3$ImpHistory)

Labels <- sort(sapply(lz,median))

axis(side = 1,las=2,labels = names(Labels),

at = 1:ncol(Boruta3$ImpHistory), cex.axis = 0.55)

plot(Boruta4, xlab = "", xaxt = "n", main="D")

lz<-lapply(1:ncol(Boruta4$ImpHistory),**function**(i)

Boruta4$ImpHistory[is.finite(Boruta4$ImpHistory[,i]),i])

colnames(Boruta4$ImpHistory)<-c("Housetype", "job\_new", "Ownership", "Fussy", "HHSize", "Composition",

"children\_3\_to\_11", "age\_brackets", "CupboardMeat", "Discard\_Sellby",

"Discard\_packs","Discard\_cheese", "Discard\_veg",

"shadowMax","shadowMean","shadowMin" )

names(lz) <- colnames(Boruta4$ImpHistory)

Labels <- sort(sapply(lz,median))

axis(side = 1,las=2,labels = names(Labels),

at = 1:ncol(Boruta4$ImpHistory), cex.axis = 0.55)