kcca_replication.R

patrickhunnicutt

2020-06-08

```
#TO DO:
# all data files should source from single replication folder set by here(), rather than team-specific
# metadata with datafiles should be separate documents
# replication directory should have a ReadMe.txt file that describes each individual file and variables
# replication directory should have copies of all data collection instruments

################################################################################################
################################################################################################
#Replication code for:

#Mark T. Buntaine, Patrick Hunnicutt, and Polycarp Komakech
#The Challenges of Using Citizen Reporting to Improve Public Services: A Field Experiment on Solid Wast
#Journal of Public Administration Research & Theory

#Prepared by: Patrick Hunnicutt
#Patrick Hunnicutt contact (as of May 2020): hunnicutt@bren.ucsb.edu

#Compiled using R Version 3.6.1 (2019-07-05) (version "Action of the Toes") on x86_64-apple-darwin15.6.
################################################################################################
################################################################################################


##### PREAMBLE #####
rm(list=ls())

## (1) Load Packages -----
library(plyr) #Version X.XX
library(dplyr) #Version X.XX
library(readr)
library(stringr)
library(tidyr)
library(gdata)
library(gtools)
library(ggplot2)
library(grid)
library(gridExtra)
library(stringr)
library(xtable)
library(dummies)
library(devtools)
library(randomizr)
library(estimatr)
library(rgdal)
library(rgeos)
library(ri)
library(multiwayvcov)
library(lmtest)
```

```r
library(stats)
library(lfe)
library(stargazer)
library(raster)
library(here)
library(ggpubr)

## (2) Set Working Directory -----
setwd("~/Google Drive File Stream/My Drive/Kampala Solid Waste project/Phase 3/KCCA_SolidWaste_ph3repli
rmarkdown::render("kcca_replication.R", "pdf_document")
```

```
##
##
## processing file: kcca_replication.spin.Rmd
```

```
##
  |
  |                                                                      |   0%
  |
  |.....................                                                 |  33%
##   ordinary text without R code
##
##
  |
  |.....................................                                |  67%
## label: unnamed-chunk-4
```

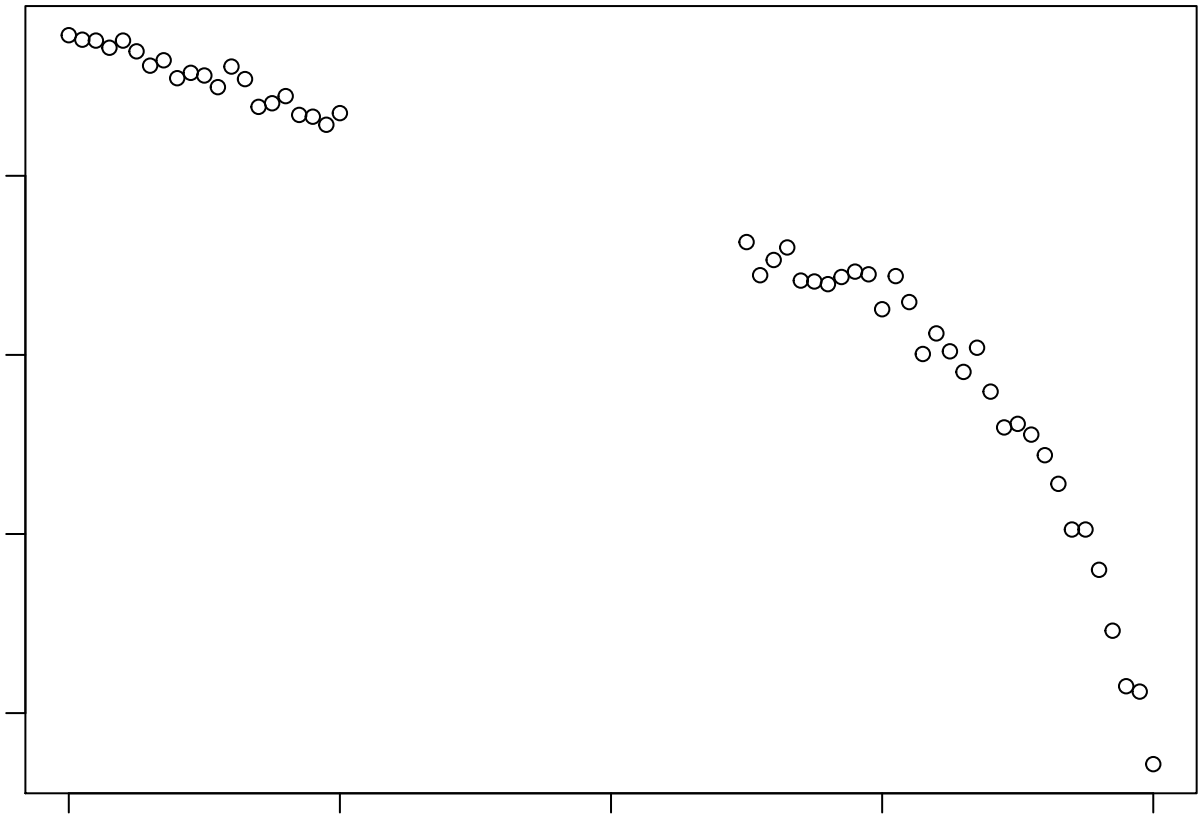
```
##
  |
  |......................................................................| 100%
##   ordinary text without R code
```

```
## output file: kcca_replication.knit.md
```

```
## /Applications/RStudio.app/Contents/MacOS/pandoc/pandoc +RTS -K512m -RTS kcca_replication.utf8.md --t
```

```
##
## Output created: kcca_replication.pdf
```

```r
# setwd(here())
getwd()
```



```
## [1] "/Volumes/GoogleDrive/My Drive/Kampala Solid Waste project/Phase 3/KCCA_SolidWaste_ph3replicatio
```

```r
## (3) Define Custom Functions -----
rcel <- function(n){
  out <- 10*ceiling(n/10)
  return(out)
}

balance.plot_f <- function(dta, var, treat, title){

  if (class(dta[,var])=="factor"){
    dta <- dta[!is.na(dta[,var]),]
    d <- table(dta[,treat], dta[,var])
    m <- max(d)
    n <- rcel(m)

    out <-  ggplot(data=dta, aes_string(x = var, fill = treat)) +
      geom_histogram(data=dta[dta[,treat]==1,], stat="count", col="grey20", alpha=0.75) +
      geom_histogram(data=dta[dta[,treat]==0,], stat="count", aes(y=..count..*(-1)), col="grey20", alpha
      scale_y_continuous(limits=c(-n,n), breaks=seq(-n,n,n),labels=abs(seq(-n,n,n))) +
      coord_flip() +
      # theme(legend.position="none") +
      ggtitle(title) +
      theme(plot.title = element_text(hjust = 0.5, size=12), axis.title.y = element_blank(), axis.title
    return(out)}
}

balance.plot_n <- function(dta, var, treat, title){
```

```
  dta <- dta[!is.na(dta[,var]),]
  # d <- density(dta[,var])
  # n <- round_any(max(d$y), 0.05)

  out <- ggplot(data=dta, aes_string(x = var, fill = treat)) +
    geom_histogram(data=dta[dta[,treat]==1,], col="grey20", alpha=0.75, bins=20) +
    geom_histogram(data=dta[dta[,treat]==0,], aes(y=..count..*(-1)), col="grey20", alpha=0.75, bins=20)
    # scale_y_continuous(limits=c(-n,n), breaks=seq(-n,n,n),labels=abs(seq(-n,n,n))) +
    coord_flip() +
    # theme(legend.position="none") +
    ggtitle(title) +
    theme(plot.title = element_text(hjust = 0.5, size=12), axis.title.y = element_blank(), axis.title.x
  return(out)
}

balance_plot <- function(dta, var, treat, title){
  if (class(dta[,var])=="factor"){
    dta <- dta[!is.na(dta[,var]),]
    d <- table(dta[,treat], dta[,var])
    m <- max(d)
    n <- rcel(m)

    out <-  ggplot(data=dta, aes_string(x = var, fill = treat)) +
      geom_histogram(data=dta[dta[,treat]==1,], stat="count", col="grey20", alpha=0.75) +
      geom_histogram(data=dta[dta[,treat]==0,], stat="count", aes(y=..count..*(-1)), col="grey20", alpha
      scale_y_continuous(limits=c(-n,n), breaks=seq(-n,n,n),labels=abs(seq(-n,n,n))) +
      coord_flip() +
      # theme(legend.position="none") +
      ggtitle(title) +
      theme(plot.title = element_text(hjust = 0.5, size=12), axis.title.y = element_blank(), axis.title
  } else if (class(dta[,var])=="numeric"|class(dta[,var])=="integer"){
    dta <- dta[!is.na(dta[,var]),]
    # d <- density(dta[,var])
    # n <- round_any(max(d$y), 0.05)

    out <- ggplot(data=dta, aes_string(x = var, fill = treat)) +
      geom_histogram(data=dta[dta[,treat]==1,], col="grey20", alpha=0.75, bins=20) +
      geom_histogram(data=dta[dta[,treat]==0,], aes(y=..count..*(-1)), col="grey20", alpha=0.75, bins=20
      # scale_y_continuous(limits=c(-n,n), breaks=seq(-n,n,n),labels=abs(seq(-n,n,n))) +
      coord_flip() +
      # theme(legend.position="none") +
      ggtitle(title) +
      theme(plot.title = element_text(hjust = 0.5, size=12), axis.title.y = element_blank(), axis.title
  }
  return(out)
}

lm.ri <- function(formula, dta, treat.var, sims, clust_var, m,...){ #treatment variable needs to be in
  require(lfe)

  ate <- coef(lm(formula, data=dta))[2]
  N <- nobs(lm(formula, data=dta))
  ate.samp.dist <- rep(NA,sims)
```

```
  for (i in 1:sims){
    dta[,treat.var] <- cluster_ra(clust_var, m)
    ate.samp.dist[i] <- coef(lm(formula, data=dta))[2]
  }

  p.two.way <- sum(abs(ate)<abs(ate.samp.dist))/sims
  p.one.way.greater <- sum(ate<ate.samp.dist)/sims
  p.one.way.lesser <- sum(ate>ate.samp.dist)/sims
  se <- sd(ate.samp.dist)

  coun <- list("ate" = ate, "ate.samp.dist" = ate.samp.dist, "se"=se, "p.two.way" = p.two.way, "p.one.wa
  return(coun)
}

did.ri <- function(formula, dta, treat.var, sims, clust_var, m,...){   require(lfe)

  ate <- coef(lm(formula, data=dta))[4]
  N <- nobs(lm(formula, data=dta))
  ate.samp.dist <- rep(NA,sims)

  for (i in 1:sims){
    dta[,treat.var] <- cluster_ra(clust_var, m)
    ate.samp.dist[i] <- coef(lm(formula, data=dta))[4]
  }

  p.two.way <- sum(abs(ate)<abs(ate.samp.dist))/sims
  p.one.way.greater <- sum(ate<ate.samp.dist)/sims
  p.one.way.lesser <- sum(ate>ate.samp.dist)/sims
  se <- sd(ate.samp.dist)

  coun <- list("ate" = ate, "ate.samp.dist" = ate.samp.dist, "se"=se, "p.two.way" = p.two.way, "p.one.wa
  return(coun)
}


#note: specify negative treatment effects
lm.ri.power.all3 <- function(formula.size, formula.dummy, formula.rank,
                             dta, treat.var, outcome.var, binary.var, rank.var,
                             sims, clust_var, m, te,
                             size.bar, dummy.bar, rank.bar, ...){
  require(randomizr)

  outcome.store <- dta[,outcome.var]
  binary.store <- dta[,binary.var]

  ate.samp.dist.size <- rep(NA,sims)
  ate.samp.dist.dummy <- rep(NA,sims)
  ate.samp.dist.rank <- rep(NA,sims)

  for (i in 1:sims){
    dta[,treat.var] <- cluster_ra(clust_var, m)

    total.treat <- -(sum(dta[,treat.var]==1)*te)
```

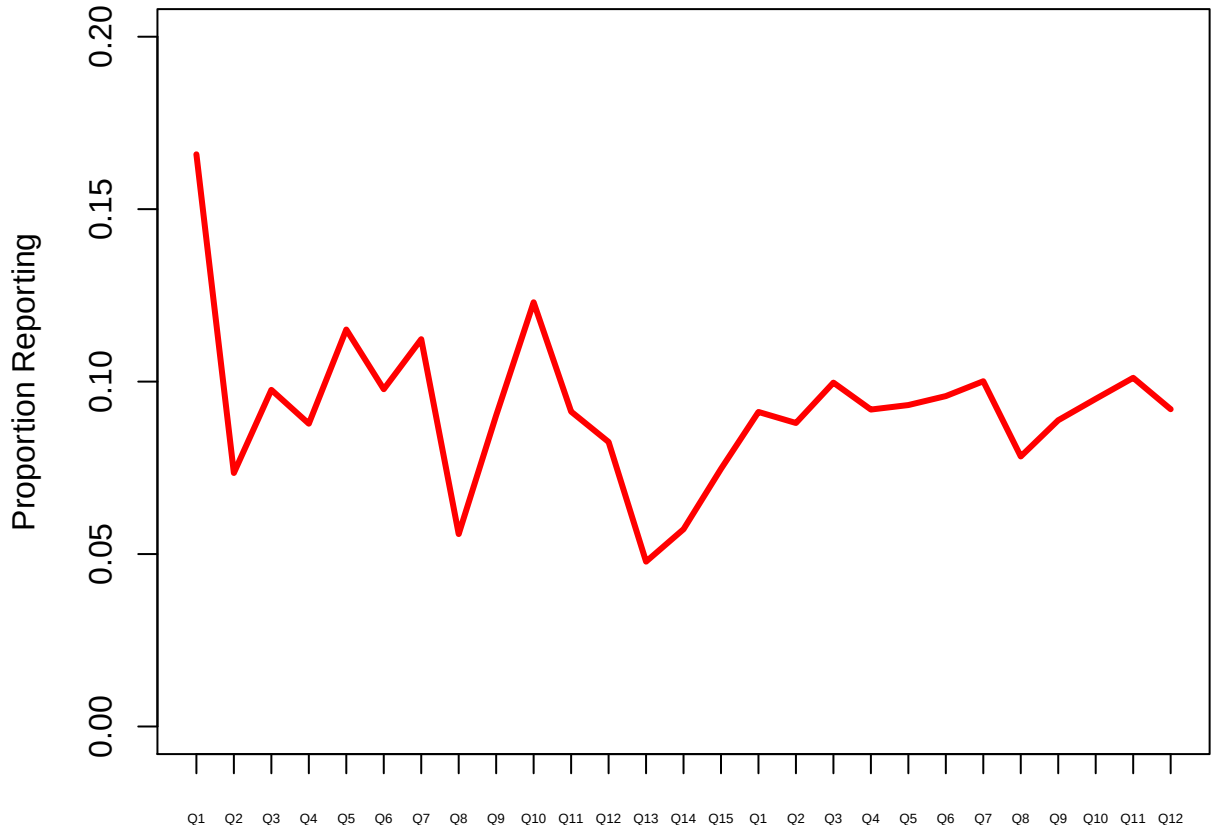
```
    #Size
    dta[,outcome.var] <- outcome.store

    while(total.treat>0){
      clean.it <- sample(row.names(dta[dta[,treat.var]==1 & dta[,outcome.var]>0,]), 1)

      if(total.treat>dta[row.names(dta)==clean.it, outcome.var]){
        total.treat <- total.treat - dta[row.names(dta)==clean.it, outcome.var]
        dta[row.names(dta)==clean.it, outcome.var] <- 0
      }

      if(total.treat<=dta[row.names(dta)==clean.it, outcome.var]){
        dta[row.names(dta)==clean.it, outcome.var] <- dta[row.names(dta)==clean.it, outcome.var] - total
        total.treat <- 0
      }
    }

    ate.samp.dist.size[i] <- coef(lm(formula.size, data=dta))[2]

    #Dummy
    #dta[,binary.var] <- binary.store
    dta[,binary.var] <- ifelse(dta[,outcome.var]==0, 0, 1)
    ate.samp.dist.dummy[i] <- coef(lm(formula.dummy, data=dta))[2]

    #Rank
    dta[,rank.var] <- rank(dta[,outcome.var], na.last = "keep")
    ate.samp.dist.rank[i] <- coef(lm(formula.rank, data=dta))[2]
  }

  power.size <- sum(ate.samp.dist.size<size.bar)/sims
  power.dummy <- sum(ate.samp.dist.dummy<dummy.bar)/sims
  power.rank <- sum(ate.samp.dist.rank<rank.bar)/sims

  coun <- list("ate.samp.dist.size" = ate.samp.dist.size, "power.size"=power.size,
               "ate.samp.dist.dummy" = ate.samp.dist.dummy, "power.dummy"=power.dummy,
               "ate.samp.dist.rank" = ate.samp.dist.rank, "power.rank"=power.rank)
  return(coun)
}


lm.cluster<-function(model, cluster)
{
  require(multiwayvcov)
  require(lmtest)
  vcovCL<-cluster.vcov(model, cluster)

  coef <- coeftest(model, vcovCL)
  w <- waldtest.default(model, vcov = vcovCL, test = "F")

  return(list(coef, w))
} #This function computes cluster standard errors using an lm() fit
```

```r
## (4) Read-in Data -----
main_dta <- read.csv("~/Google Drive File Stream/My Drive/Kampala Solid Waste project/Phase 3/KCCA_Solid
                     stringsAsFactors = FALSE, strip.white = TRUE)
#main_dta: data on the full sample of audited informal waste piles.

subsetA_dta <- read.csv("~/Google Drive File Stream/My Drive/Kampala Solid Waste project/Phase 3/KCCA_S
                        stringsAsFactors = FALSE, strip.white = TRUE)
#subsetA_dta: subsetted pile data; excludes piles due to discrepancies in the coordinates of the pile b

subsetB_dta <- read.csv("~/Google Drive File Stream/My Drive/Kampala Solid Waste project/Phase 3/KCCA_S
                        stringsAsFactors = FALSE, strip.white = TRUE)
#subsetB_dta: subsetted pile data; excludes piles due to discrepancies in the coordinates of the pile b
#size measurements as determined by comparing recorded pile sizes with photos of waste piles. Also cont
#analysis of spillover

hte_dta <- read.csv("~/Google Drive File Stream/My Drive/Kampala Solid Waste project/Phase 3/KCCA_SolidW
                    stringsAsFactors = FALSE, strip.white = TRUE)
#hte_dta: subsetted pile data (subsetB_dta) used for exploring HTE and sub-group effects.

lc1 <- readOGR(dsn="~/Google Drive File Stream/My Drive/Kampala Solid Waste project/Phase 3/KCCA_SolidWa
               layer="Kampala_villagesfff")
```

```
## OGR data source with driver: ESRI Shapefile
## Source: "/Volumes/GoogleDrive/My Drive/Kampala Solid Waste project/Phase 3/KCCA_SolidWaste_ph3replica
## with 755 features
## It has 7 fields
## Integer64 fields read as strings:  COUNT
```

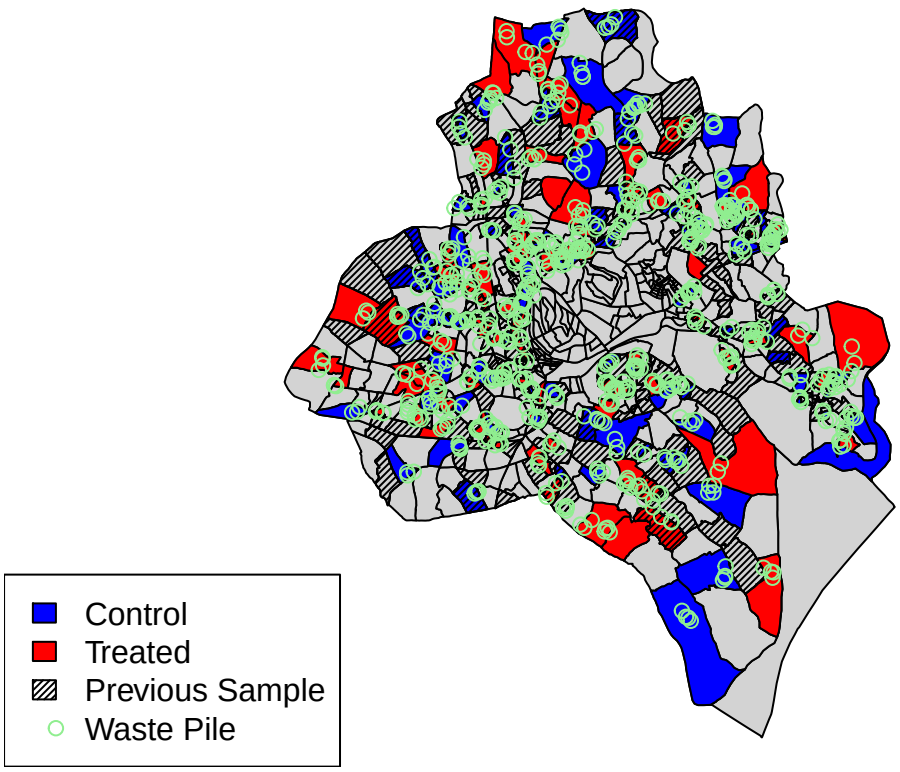
```r
#lc1: shapefile containing polygons of zones, parishes, and divisions in Kampala.

rr <- read.csv("~/Google Drive File Stream/My Drive/Kampala Solid Waste project/Phase 3/KCCA_SolidWaste_
               stringsAsFactors = FALSE, strip.white = TRUE)
##rr: data on reporting rates of citizen reporters over the study period.

p1.sample <- read.csv("~/Google Drive File Stream/My Drive/Kampala Solid Waste project/Phase 3/KCCA_Soli
                      stringsAsFactors = FALSE, strip.white = TRUE)
##p1.sample: assignment data on piles sampled in project's first phase (used for mapping experimental s

p2.sample <- read.csv("~/Google Drive File Stream/My Drive/Kampala Solid Waste project/Phase 3/KCCA_Soli
                      stringsAsFactors = FALSE, strip.white = TRUE)
##p2.sample: assignment data on piles sampled in project's second phase (used for mapping experimental

p3.sample <- read.csv("~/Google Drive File Stream/My Drive/Kampala Solid Waste project/Phase 3/KCCA_Soli
                      stringsAsFactors = FALSE, strip.white = TRUE)
##p3.sample: assignment data on piles sampled in project's third phase (used for mapping experimental s

##  (5) Clean Data -----
##Assign Unique ID to Zones in Shapefile
lc1@data$zone.id <- rownames(lc1@data)

lc1@data$zpd_uid <- NA
lc1@data$zpd_uid <- paste(lc1@data$VNAME, "_", lc1@data$Parishes, "_", lc1@data$Division)
lc1@data$zpd_uid <- tolower(lc1@data$zpd_uid)
lc1@data$zpd_uid <- str_replace_all(lc1@data$zpd_uid, " ", "")
```

```r
lc1@data$zpd_uid <- str_replace_all(lc1@data$zpd_uid, ",", "")
lc1@data$zpd_uid <- str_replace_all(lc1@data$zpd_uid, "-", ".")



##### FIGURES #####
## Figure 1: Reporting Rates -----
# tiff("./Figures/Figure1.tiff", width=5, height=3, units="in", res=300)
# par(mar = c(2.1, 4.1, 1, 1))
# plot(x=1:nrow(rr), y=rr$Usable.Response.Rate, type="l", col="red", lwd=3, xlab="Question Number", yla
# axis(1, at=1:nrow(rr), labels=rr$Question, cex.axis=0.45)
# dev.off()

# png("./Figures/Figure1.png", width=5, height=3, units="in", res=300)
par(mar = c(2.1, 4.1, 1, 1))
plot(x=1:nrow(rr), y=rr$Usable.Response.Rate, type="l", col="red", lwd=3, xlab="Question Number", ylab=
axis(1, at=1:nrow(rr), labels=rr$Question, cex.axis=0.45)
```




```r
# dev.off()

## Figure 2: Map of Experimental Sample -----
trt.zones <- unique(main_dta$zone.id[main_dta$treat==1])
trt.zones <- trt.zones[!is.na(trt.zones)]
ctl.zones <- unique(main_dta$zone.id[main_dta$treat==0])
ctl.zones <- ctl.zones[!is.na(ctl.zones)]

###Bring in Phase 1/2 zones that have continuing reports, from SolidWaste_Spillover_Phase3.R
```

```r
#Note, from SolidWaste_Phase1_Analysis: no subjects recruited from Ntinda, Village 7 (reason was never
p1.sample <- subset(p1.sample, X!=565)
p2.sample <- subset(p2.sample, InSample==1) #removing replacement zones not used, original zones not us
prev.zones <- c(p1.sample$X, p2.sample$zone.id.realized) #Note: some zones might not have active report

piles <- data.frame(main_dta$m2.pile.lon, main_dta$m2.pile.lat)
piles <- piles[complete.cases(piles),]
names(piles) <- c("lon","lat")

LongLatToUTM<-function(x,y,zone){
  xy <- data.frame(ID = 1:length(x), X = x, Y = y)
  coordinates(xy) <- c("X", "Y")
  proj4string(xy) <- CRS("+proj=longlat +datum=WGS84")  ## for example
  res <- spTransform(xy, CRS(paste("+proj=utm +zone=",zone," ellps=WGS84",sep='')))
  return(as.data.frame(res))
}

cords <- LongLatToUTM(x=piles$lon, y=piles$lat, zone=36)

# tiff("./Figures/Figure2.tiff", width=5, height=5, units="in", res=300)
# par(mar = c(1, 1, 1, 1))
# plot(lc1, col = "lightgrey")
# plot(lc1[row.names(lc1@data) %in% ctl.zones, ], col = "blue", add = TRUE)
# plot(lc1[row.names(lc1@data) %in% trt.zones, ], col = "red", add = TRUE)
# plot(lc1[row.names(lc1@data) %in% prev.zones, ], col = "black", density=40, add = TRUE)
# points(x=cords$X, y=cords$Y, col="lightgreen")
# legend("bottomleft", legend=c("Control","Treated","Previous Sample", "Waste Pile"), cex=1,
#        pch=c(NA,NA,NA,1), fill=c("blue","red","black","white"), density=c(NA,NA,40,NA), border=c(1,1,
# dev.off()

# png("./Figures/Figure2.png", width=5, height=5, units="in", res=300)
par(mar = c(1, 1, 1, 1))
plot(lc1, col = "lightgrey")
plot(lc1[row.names(lc1@data) %in% ctl.zones, ], col = "blue", add = TRUE)
plot(lc1[row.names(lc1@data) %in% trt.zones, ], col = "red", add = TRUE)
plot(lc1[row.names(lc1@data) %in% prev.zones, ], col = "black", density=40, add = TRUE)
points(x=cords$X, y=cords$Y, col="lightgreen")
legend("bottomleft", legend=c("Control","Treated","Previous Sample", "Waste Pile"), cex=1,
       pch=c(NA,NA,NA,1), fill=c("blue","red","black","white"), density=c(NA,NA,40,NA), border=c(1,1,1,
```




```
# dev.off()


## Figure 7: Dependent Variables, Midline 1 -----
#WASTE PILE DUMMY
table(subsetA_dta$treat, subsetA_dta$m1.waste.pile_d)
```

```
##
##       0   1
##   0  47 301
##   1  49 300
```

```
a<-data.frame((prop.table(table(subsetA_dta$treat, subsetA_dta$m1.waste.pile_d),1)))
names(a) <- c("treat", "desc", "prop")

a$desc <- ordered(factor(c(rep("Cleaned", 2), rep("Not Cleaned", 2))), levels=c("Cleaned", "Not Cleaned
a$desc.n <- as.numeric(a$desc)
a <- a[a$desc=="Cleaned",]

ctrl <- as.numeric(length(which(subsetA_dta$m1.waste.pile_d==0&subsetA_dta$treat==0)))
trt <- as.numeric(length(which(subsetA_dta$m1.waste.pile_d==0&subsetA_dta$treat==1)))
n.ctrl <- as.numeric(length(which(subsetA_dta$treat==0))) #count of units in control stays the same for
n.trt <- as.numeric(length(which(subsetA_dta$treat==1))) #count of units in treatment stays the same fo

ctrl.means <- rep(NA,1000)
trt.means <- rep(NA,1000)
set.seed(201)
for (i in 1:1000){
  ctrl.means[i] <- mean(sample(c(rep(0,ctrl),rep(1,n.ctrl-ctrl)), replace=T))
  trt.means[i] <- mean(sample(c(rep(0,trt),rep(1,n.trt-trt)), replace=T))
```

```
}

a$se <- c(sd(ctrl.means),sd(trt.means))
a$Assignment <- c("Control", "Treatment")

ctrl <- as.numeric(length(which(subsetA_dta$m1.waste.pile_d==0&subsetA_dta$treat==0)))
trt <- as.numeric(length(which(subsetA_dta$m1.waste.pile_d==0&subsetA_dta$treat==1)))
n.ctrl <- as.numeric(length(which(subsetA_dta$treat==0))) #count of units in control stays the same for
n.trt <- as.numeric(length(which(subsetA_dta$treat==1))) #count of units in treatment stays the same fo

prop.test(c(ctrl,trt), c(n.ctrl,n.trt), alternative = "l")
```

```
##
##  2-sample test for equality of proportions with continuity
##  correction
##
## data:  c(ctrl, trt) out of c(n.ctrl, n.trt)
## X-squared = 0.008982, df = 1, p-value = 0.4622
## alternative hypothesis: less
## 95 percent confidence interval:
##  -1.00000000  0.04046539
## sample estimates:
##    prop 1    prop 2
## 0.1350575 0.1404011
```

```
#WASTE PILE DUMMY ADJUSTED
table(subsetA_dta$treat, subsetA_dta$m1.waste.pile_d2)
```

```
##
##       0   1
##   0  55 293
##   1  67 282
```

```
a.adj <- data.frame((prop.table(table(subsetA_dta$treat, subsetA_dta$m1.waste.pile_d2),1)))
names(a.adj) <- c("treat", "desc", "prop")
a.adj$desc <- ordered(factor(c(rep("Cleaned", 2), rep("Not Cleaned", 2))), levels=c("Cleaned", "Not Clea
a.adj$desc.n <- as.numeric(a.adj$desc)
a.adj <- a.adj[a.adj$desc=="Cleaned",]


ctrl <- as.numeric(length(which(subsetA_dta$m1.waste.pile_d2==0&subsetA_dta$treat==0)))
trt <- as.numeric(length(which(subsetA_dta$m1.waste.pile_d2==0&subsetA_dta$treat==1)))

ctrl.means <- rep(NA,1000)
trt.means <- rep(NA,1000)
set.seed(201)
for (i in 1:1000){
  ctrl.means[i] <- mean(sample(c(rep(0,ctrl),rep(1,n.ctrl-ctrl)), replace=T))
  trt.means[i] <- mean(sample(c(rep(0,trt),rep(1,n.trt-trt)), replace=T))
}

a.adj$se <- c(sd(ctrl.means),sd(trt.means))
a.adj$Assignment <- c("Control", "Treatment")

prop.test(c(ctrl,trt), c(n.ctrl,n.trt), alternative="l")
```

```
##
##  2-sample test for equality of proportions with continuity
##  correction
##
## data:  c(ctrl, trt) out of c(n.ctrl, n.trt)
## X-squared = 1.1643, df = 1, p-value = 0.1403
## alternative hypothesis: less
## 95 percent confidence interval:
##  -1.0000000  0.0162362
## sample estimates:
##    prop 1    prop 2
## 0.1580460 0.1919771
```

```
##GENERAL PILE DESCRIPTION
table(subsetA_dta$treat, subsetA_dta$m1.pile.gdesc)
```

```
##
##
##   0  64
##   1  98
##
##     Large sack(s) or container(s) of rubbish that can easily be transported
##   0                                                                      12
##   1                                                                      22
##
##     Less than 10 pieces of non-organic waste
##   0                                       78
##   1                                       79
##
##     More than 10 pieces of non-organic waste
##   0                                      194
##   1                                      150
```

```
subsetA_dta$m1.pile.gdesc <- ifelse(subsetA_dta$m1.pile.gdesc==""&subsetA_dta$m1.waste.pile_d==1, "Miss
table(subsetA_dta$treat, subsetA_dta$m1.pile.gdesc)
```

```
##
##
##   0  44
##   1  47
##
##     Large sack(s) or container(s) of rubbish that can easily be transported
##   0                                                                      12
##   1                                                                      22
##
##     Less than 10 pieces of non-organic waste Missing
##   0                                       78      20
##   1                                       79      51
##
##     More than 10 pieces of non-organic waste
##   0                                      194
##   1                                      150
```

```
b <- data.frame(prop.table(table(subsetA_dta$treat, subsetA_dta$m1.pile.gdesc),1))
names(b) <- c("treat","desc","prop")
b$desc <- as.character(b$desc)
```

```r
#Removing missing responses from table
b <- b[!(b$desc=="Missing"),]
#Recoding possible answers -> groups
b$desc[b$desc==""] <- "Cleaned"
b$desc[b$desc=="More than 10 pieces of non-organic waste"] <- ">10 pcs.\nwaste"
b$desc[b$desc=="Less than 10 pieces of non-organic waste"] <- "<10 pcs.\nwaste"
b$desc[b$desc=="Large sack(s) or container(s) of rubbish that can easily be transported"] <- "In \nConta
b <- b[!is.na(b$desc),]
b$desc <- ordered(factor(b$desc,
                         levels=c("Cleaned","In \nContainers","<10 pcs.\nwaste", ">10 pcs.\nwaste")))
b$desc.n <- as.numeric(b$desc)

con.prop.transport <- rep(NA,1000)
trt.prop.transport <- rep(NA,1000)

con.prop.large <- rep(NA,1000)
trt.prop.large <- rep(NA,1000)

con.prop.small <- rep(NA,1000)
trt.prop.small <- rep(NA,1000)

con.prop.cleaned <- rep(NA,1000)
trt.prop.cleaned <- rep(NA,1000)

set.seed(209)
for (i in 1:1000){
  con.vec <- sample(subsetA_dta$m1.pile.gdesc[subsetA_dta$treat==0], replace=T, size=length(subsetA_dta
  trt.vec <- sample(subsetA_dta$m1.pile.gdesc[subsetA_dta$treat==1], replace=T, size=length(subsetA_dta

  con.vec <- con.vec[!is.na(con.vec)]
  trt.vec <- trt.vec[!is.na(trt.vec)]

  con.prop.transport[i] <- length(con.vec[con.vec=="Large sack(s) or container(s) of rubbish that can ea
  trt.prop.transport[i] <- length(trt.vec[trt.vec=="Large sack(s) or container(s) of rubbish that can ea

  con.prop.large[i] <- length(con.vec[con.vec=="More than 10 pieces of non-organic waste"]) / length(con
  trt.prop.large[i] <- length(trt.vec[trt.vec=="More than 10 pieces of non-organic waste"]) / length(tr

  con.prop.small[i] <- length(con.vec[con.vec=="Less than 10 pieces of non-organic waste"]) / length(con
  trt.prop.small[i] <- length(trt.vec[trt.vec=="Less than 10 pieces of non-organic waste"]) / length(tr

  con.prop.cleaned[i] <- length(con.vec[con.vec==""]) / length(con.vec)
  trt.prop.cleaned[i] <- length(trt.vec[trt.vec==""]) / length(trt.vec)
}

b$se[1:2] <- c(sd(con.prop.cleaned),sd(trt.prop.cleaned))
b$se[3:4] <- c(sd(con.prop.transport),sd(trt.prop.transport))
b$se[5:6] <- c(sd(con.prop.large),sd(trt.prop.large))
b$se[7:8] <- c(sd(con.prop.small),sd(trt.prop.small))

b$Assignment <- c("Control", "Treatment")

#T-tests
```

```r
ctrl.largesacks <- as.numeric(length(which(subsetA_dta$m1.pile.gdesc=="Large sack(s) or container(s) of
trt.largesacks <- as.numeric(length(which(subsetA_dta$m1.pile.gdesc=="Large sack(s) or container(s) of
ctrl.less10pcs <- as.numeric(length(which(subsetA_dta$m1.pile.gdesc=="Less than 10 pieces of non-organi
trt.less10pcs <- as.numeric(length(which(subsetA_dta$m1.pile.gdesc=="Less than 10 pieces of non-organic
ctrl.more10pcs <- as.numeric(length(which(subsetA_dta$m1.pile.gdesc=="More than 10 pieces of non-organi
trt.more10pcs <- as.numeric(length(which(subsetA_dta$m1.pile.gdesc=="More than 10 pieces of non-organic


prop.test(c(ctrl.largesacks,trt.largesacks), c(n.ctrl,n.trt), alternative = "l")
```

```
##
##  2-sample test for equality of proportions with continuity
##  correction
##
## data:  c(ctrl.largesacks, trt.largesacks) out of c(n.ctrl, n.trt)
## X-squared = 2.4775, df = 1, p-value = 0.05774
## alternative hypothesis: less
## 95 percent confidence interval:
##  -1.000000000  0.001086589
## sample estimates:
##     prop 1     prop 2
## 0.03448276 0.06303725
```

```r
prop.test(c(ctrl.less10pcs,trt.less10pcs), c(n.ctrl,n.trt), alternative = "l")
```

```
##
##  2-sample test for equality of proportions with continuity
##  correction
##
## data:  c(ctrl.less10pcs, trt.less10pcs) out of c(n.ctrl, n.trt)
## X-squared = 1.2716e-30, df = 1, p-value = 0.5
## alternative hypothesis: less
## 95 percent confidence interval:
##  -1.00000000  0.05205369
## sample estimates:
##    prop 1    prop 2
## 0.2241379 0.2263610
```

```r
prop.test(c(ctrl.more10pcs,trt.more10pcs), c(n.ctrl,n.trt), alternative = "g")
```

```
##
##  2-sample test for equality of proportions with continuity
##  correction
##
## data:  c(ctrl.more10pcs, trt.more10pcs) out of c(n.ctrl, n.trt)
## X-squared = 10.858, df = 1, p-value = 0.0004918
## alternative hypothesis: greater
## 95 percent confidence interval:
##  0.06301386 1.00000000
## sample estimates:
##    prop 1    prop 2
## 0.5574713 0.4297994
```

```r
##STORAGE
table(subsetA_dta$treat,subsetA_dta$m1.waste.stor)
```

```
##
##
##   0  44
##   1  47
##
##     All of the rubbish is neatly contained with sacks or other containers
##   0                                                                     10
##   1                                                                     18
##
##     Most of the rubbish is organized in sacks or other containers
##   0                                                           19
##   1                                                           26
##
##     No rubbish is contained in sacks or containers
##   0                                          248
##   1                                          238
##
##     Very little rubbish is contained within sacks or containers
##   0                                                         27
##   1                                                         20
```

```r
c <- data.frame(prop.table(table(subsetA_dta$treat,subsetA_dta$m1.waste.stor),1))
names(c) <- c("treat","desc","prop")
c$desc <- as.character(c$desc)
c$desc[c$desc==""] <- "Cleaned"
c$desc[c$desc=="No rubbish is contained in sacks or containers"] <- "None"
c$desc[c$desc=="All of the rubbish is neatly contained with sacks or other containers"] <- "Full \n"
c$desc[c$desc=="Most of the rubbish is organized in sacks or other containers"] <- "Most"
c$desc[c$desc=="Very little rubbish is contained within sacks or containers"] <- "Some"
c <- c[!is.na(c$desc),]
c$desc <- ordered(factor(c$desc,
                         levels=c("Cleaned","Full \n","Most","Some","None")))
c$desc.n <- as.numeric(c$desc)

con.prop.uc <- rep(NA,1000)
trt.prop.uc <- rep(NA,1000)

con.prop.pc <- rep(NA,1000)
trt.prop.pc <- rep(NA,1000)

con.prop.mc <- rep(NA,1000)
trt.prop.mc <- rep(NA,1000)

con.prop.fc <- rep(NA,1000)
trt.prop.fc <- rep(NA,1000)

con.prop.cc <- rep(NA,1000)
trt.prop.cc <- rep(NA,1000)

set.seed(201)
for (i in 1:1000){
  con.vec_c <- sample(subsetA_dta$m1.waste.stor[subsetA_dta$treat==0], replace=T, size=length(subsetA_d
  trt.vec_c <- sample(subsetA_dta$m1.waste.stor[subsetA_dta$treat==1], replace=T, size=length(subsetA_d
```

```r
  con.vec_c <- con.vec_c[!is.na(con.vec_c)]
  trt.vec_c <- trt.vec_c[!is.na(trt.vec_c)]

  con.prop.uc[i] <- length(con.vec_c[con.vec_c=="No rubbish is contained in sacks or containers"])/leng
  trt.prop.uc[i] <- length(trt.vec_c[trt.vec_c=="No rubbish is contained in sacks or containers"])/leng

  con.prop.pc[i] <- length(con.vec_c[con.vec_c=="Very little rubbish is contained within sacks or conta
  trt.prop.pc[i] <- length(trt.vec_c[trt.vec_c=="Very little rubbish is contained within sacks or conta

  con.prop.mc[i] <- length(con.vec_c[con.vec_c=="Most of the rubbish is organized in sacks or other con
  trt.prop.mc[i] <- length(trt.vec_c[trt.vec_c=="Most of the rubbish is organized in sacks or other con

  con.prop.fc[i] <- length(con.vec_c[con.vec_c=="All of the rubbish is neatly contained with sacks or o
  trt.prop.fc[i] <- length(trt.vec_c[trt.vec_c=="All of the rubbish is neatly contained with sacks or o

  con.prop.cc[i] <- length(con.vec_c[con.vec_c==""])/length(con.vec_c)
  trt.prop.cc[i] <- length(trt.vec_c[trt.vec_c==""])/length(trt.vec_c)

}

c$se[1:2] <- c(sd(con.prop.cc),sd(trt.prop.cc))
c$se[3:4] <- c(sd(con.prop.fc),sd(trt.prop.fc))
c$se[5:6] <- c(sd(con.prop.mc),sd(trt.prop.mc))
c$se[7:8] <- c(sd(con.prop.uc),sd(trt.prop.uc))
c$se[9:10] <- c(sd(con.prop.pc),sd(trt.prop.pc))
c$Assignment <- c("Control", "Treatment")

#T-tests
ctrl.full <- as.numeric(length(which(subsetA_dta$m1.waste.stor=="All of the rubbish is neatly contained
ctrl.most <- as.numeric(length(which(subsetA_dta$m1.waste.stor=="Most of the rubbish is organized in sa
ctrl.some <- as.numeric(length(which(subsetA_dta$m1.waste.stor=="Very little rubbish is contained withi
ctrl.none <- as.numeric(length(which(subsetA_dta$m1.waste.stor=="No rubbish is contained in sacks or co
trt.full <- as.numeric(length(which(subsetA_dta$m1.waste.stor=="All of the rubbish is neatly contained
trt.most <- as.numeric(length(which(subsetA_dta$m1.waste.stor=="Most of the rubbish is organized in sac
trt.some <- as.numeric(length(which(subsetA_dta$m1.waste.stor=="Very little rubbish is contained within
trt.none <- as.numeric(length(which(subsetA_dta$m1.waste.stor=="No rubbish is contained in sacks or con

prop.test(c(ctrl.full,trt.full), c(n.ctrl,n.trt), alternative = "l")
```

```
##
##  2-sample test for equality of proportions with continuity
##  correction
##
## data:  c(ctrl.full, trt.full) out of c(n.ctrl, n.trt)
## X-squared = 1.8024, df = 1, p-value = 0.08971
## alternative hypothesis: less
## 95 percent confidence interval:
##  -1.000000000  0.004446272
## sample estimates:
##     prop 1     prop 2
## 0.02873563 0.05157593
```

```r
prop.test(c(ctrl.most,trt.most), c(n.ctrl,n.trt), alternative = "l")
```

```
##
##  2-sample test for equality of proportions with continuity
##  correction
##
## data:  c(ctrl.most, trt.most) out of c(n.ctrl, n.trt)
## X-squared = 0.83691, df = 1, p-value = 0.1801
## alternative hypothesis: less
## 95 percent confidence interval:
##  -1.00000000  0.01355951
## sample estimates:
##     prop 1     prop 2
## 0.05459770 0.07449857
```

```r
prop.test(c(ctrl.some,trt.some), c(n.ctrl,n.trt), alternative = "g")
```

```
##
##  2-sample test for equality of proportions with continuity
##  correction
##
## data:  c(ctrl.some, trt.some) out of c(n.ctrl, n.trt)
## X-squared = 0.83991, df = 1, p-value = 0.1797
## alternative hypothesis: greater
## 95 percent confidence interval:
##  -0.01381795  1.00000000
## sample estimates:
##     prop 1     prop 2
## 0.07758621 0.05730659
```

```r
prop.test(c(ctrl.none,trt.none), c(n.ctrl,n.trt), alternative = "g")
```

```
##
##  2-sample test for equality of proportions with continuity
##  correction
##
## data:  c(ctrl.none, trt.none) out of c(n.ctrl, n.trt)
## X-squared = 0.63917, df = 1, p-value = 0.212
## alternative hypothesis: greater
## 95 percent confidence interval:
##  -0.02938888  1.00000000
## sample estimates:
##    prop 1    prop 2
## 0.7126437 0.6819484
```

```r
#ORGANIZATION
table(subsetA_dta$treat,subsetA_dta$m1.pile.org)
```

```
##
##         All of the rubbish is collected into a single pile
##   0  44                                                  34
##   1  47                                                  39
##
##     Most of the rubbish is organized around a single pile but other rubbish is spread nearby
##   0                                                                                       60
##   1                                                                                       69
```

```
##
##     Rubbish is spread all round no evidence of the rubbish being organized
##   0                                                                    210
##   1                                                                    194
```

```r
d <- data.frame(prop.table(table(subsetA_dta$treat,subsetA_dta$m1.pile.org),1))
names(d) <- c("treat","desc","prop")
d$desc <- as.character(d$desc)

#Recoding possible answers -> groups
d$desc[d$desc==""] <- "Cleaned"
d$desc[d$desc=="All of the rubbish is collected into a single pile"] <- "Pile"
d$desc[d$desc=="Most of the rubbish is organized around a single pile but other rubbish is spread nearby
d$desc[d$desc=="Rubbish is spread all round no evidence of the rubbish being organized"] <- "Dispersed"
d <- d[!is.na(d$desc),]
d$desc <- ordered(factor(d$desc, levels = c("Cleaned","Pile","Dispersed \naround Pile","Dispersed")))
d$desc.n <- as.numeric(d$desc)

con.prop.wopc <- rep(NA,1000)
trt.prop.wopc <- rep(NA,1000)

con.prop.sp <- rep(NA,1000)
trt.prop.sp <- rep(NA,1000)

con.prop.spd <- rep(NA,1000)
trt.prop.spd <- rep(NA,1000)

con.prop.npd <- rep(NA,1000)
trt.prop.npd <- rep(NA,1000)

set.seed(201)
for (i in 1:1000){
  con.vec_d <- sample(subsetA_dta$m1.pile.org[subsetA_dta$treat==0], replace=T, size=length(subsetA_dta
  trt.vec_d <- sample(subsetA_dta$m1.pile.org[subsetA_dta$treat==1], replace=T, size=length(subsetA_dta

  con.vec_d <- con.vec_d[!is.na(con.vec_d)]
  trt.vec_d <- trt.vec_d[!is.na(trt.vec_d)]

  con.prop.wopc[i] <- length(con.vec_d[con.vec_d==""])/length(con.vec_d)
  trt.prop.wopc[i] <- length(trt.vec_d[trt.vec_d==""])/length(trt.vec_d)

  con.prop.sp[i] <- length(con.vec_d[con.vec_d=="All of the rubbish is collected into a single pile"])/
  trt.prop.sp[i] <- length(trt.vec_d[trt.vec_d=="All of the rubbish is collected into a single pile"])/

  con.prop.spd[i] <- length(con.vec_d[con.vec_d=="Most of the rubbish is organized around a single pile
  trt.prop.spd[i] <- length(trt.vec_d[trt.vec_d=="Most of the rubbish is organized around a single pile

  con.prop.npd[i] <- length(con.vec_d[con.vec_d=="Rubbish is spread all round no evidence of the rubbis
  trt.prop.npd[i] <- length(trt.vec_d[trt.vec_d=="Rubbish is spread all round no evidence of the rubbis
}

d$se[1:2] <- c(sd(con.prop.wopc),sd(trt.prop.wopc))
d$se[3:4] <- c(sd(con.prop.sp),sd(trt.prop.sp))
d$se[5:6] <- c(sd(con.prop.spd),sd(trt.prop.spd))
```

```r
d$se[7:8] <- c(sd(con.prop.npd),sd(trt.prop.npd))
d$Assignment <- c("Control", "Treatment")

#T-tests
ctrl.single <- as.numeric(length(which(subsetA_dta$m1.pile.org=="All of the rubbish is collected into a
ctrl.partdispersed <- as.numeric(length(which(subsetA_dta$m1.pile.org=="Most of the rubbish is organize
ctrl.dispersed <- as.numeric(length(which(subsetA_dta$m1.pile.org=="Rubbish is spread all round no evid
trt.single <- as.numeric(length(which(subsetA_dta$m1.pile.org=="All of the rubbish is collected into a
trt.partdispersed <- as.numeric(length(which(subsetA_dta$m1.pile.org=="Most of the rubbish is organized
trt.dispersed <- as.numeric(length(which(subsetA_dta$m1.pile.org=="Rubbish is spread all round no evide

prop.test(c(ctrl.single,trt.single), c(n.ctrl,n.trt), alternative = "l")
```

```
##
##  2-sample test for equality of proportions with continuity
##  correction
##
## data:  c(ctrl.single, trt.single) out of c(n.ctrl, n.trt)
## X-squared = 0.23217, df = 1, p-value = 0.315
## alternative hypothesis: less
## 95 percent confidence interval:
##  -1.00000000  0.02696541
## sample estimates:
##     prop 1     prop 2
## 0.09770115 0.11174785
```

```r
prop.test(c(ctrl.partdispersed,trt.partdispersed), c(n.ctrl,n.trt), alternative = "l")
```

```
##
##  2-sample test for equality of proportions with continuity
##  correction
##
## data:  c(ctrl.partdispersed, trt.partdispersed) out of c(n.ctrl, n.trt)
## X-squared = 0.58096, df = 1, p-value = 0.223
## alternative hypothesis: less
## 95 percent confidence interval:
##  -1.00000000  0.02593861
## sample estimates:
##    prop 1    prop 2
## 0.1724138 0.1977077
```

```r
prop.test(c(ctrl.dispersed,trt.dispersed), c(n.ctrl,n.trt), alternative="g")
```

```
##
##  2-sample test for equality of proportions with continuity
##  correction
##
## data:  c(ctrl.dispersed, trt.dispersed) out of c(n.ctrl, n.trt)
## X-squared = 1.4292, df = 1, p-value = 0.1159
## alternative hypothesis: greater
## 95 percent confidence interval:
##  -0.01673046  1.00000000
## sample estimates:
##    prop 1    prop 2
## 0.6034483 0.5558739
```

```r
#BURNING
table(subsetA_dta$treat,subsetA_dta$m1.evi.burn)
```

```
##
##
##   0  44
##   1  47
##
##     Less than half of the area of the rubbish pile contains evidence of burning
##   0                                                                          60
##   1                                                                          50
##
##     More than half of the area of the rubbish pile contains evidence of burning
##   0                                                                          87
##   1                                                                          70
##
##     No evidence of burning
##   0                    157
##   1                    182
```

```r
e <- data.frame(prop.table(table(subsetA_dta$treat,subsetA_dta$m1.evi.burn),1))
names(e) <- c("treat","desc","prop")
e$desc <- as.character(e$desc)

con.prop.ec <- rep(NA,1000)
trt.prop.ec <- rep(NA,1000)

con.prop.bl <- rep(NA,1000)
trt.prop.bl <- rep(NA,1000)

con.prop.bm <- rep(NA,1000)
trt.prop.bm <- rep(NA,1000)

con.prop.nb <- rep(NA,1000)
trt.prop.nb <- rep(NA,1000)


#Recoding possible answers -> groups
e$desc[e$desc==""] <- "Cleaned"
e$desc[e$desc=="Less than half of the area of the rubbish pile contains evidence of burning"] <- "<50%"
e$desc[e$desc=="More than half of the area of the rubbish pile contains evidence of burning"] <- ">50%"
e$desc[e$desc=="No evidence of burning"] <- "None"
e <- e[!is.na(e$desc),]
e$desc <- ordered(factor(e$desc, levels = c("Cleaned","None", "<50%", ">50%")))
e$desc.n <- as.numeric(e$desc)

set.seed(201)

for (i in 1:1000){
  con.vec_e <- sample(subsetA_dta$m1.evi.burn[subsetA_dta$treat==0], replace=T, size=length(subsetA_dta
  trt.vec_e <- sample(subsetA_dta$m1.evi.burn[subsetA_dta$treat==1], replace=T, size=length(subsetA_dta

  con.vec_e <- con.vec_e[!is.na(con.vec_e)]
  trt.vec_e <- trt.vec_e[!is.na(trt.vec_e)]
```

```
  con.prop.ec[i] <- length(con.vec_e[con.vec_e==""])/length(con.vec_e)
  trt.prop.ec[i] <- length(trt.vec_e[trt.vec_e==""])/length(trt.vec_e)

  con.prop.nb[i] <- length(con.vec_e[con.vec_e=="No evidence of burning"])/length(con.vec_e)
  trt.prop.nb[i] <- length(trt.vec_e[trt.vec_e=="No evidence of burning"])/length(trt.vec_e)

  con.prop.bl[i] <- length(con.vec_e[con.vec_e=="Less than half of the area of the rubbish pile contains
  trt.prop.bl[i] <- length(trt.vec_e[trt.vec_e=="Less than half of the area of the rubbish pile contains

  con.prop.bm[i] <- length(con.vec_e[con.vec_e=="More than half of the area of the rubbish pile contains
  trt.prop.bm[i] <- length(trt.vec_e[trt.vec_e=="More than half of the area of the rubbish pile contains

}

e$se[1:2] <- c(sd(con.prop.ec),sd(trt.prop.ec))
e$se[3:4] <- c(sd(con.prop.bl),sd(trt.prop.bl))
e$se[5:6] <- c(sd(con.prop.bm),sd(trt.prop.bm))
e$se[7:8] <- c(sd(con.prop.nb),sd(trt.prop.nb))
e$Assignment <- c("Control", "Treatment")

#T-tests
ctrl.noburn <- as.numeric(length(which(subsetA_dta$m1.evi.burn=="No evidence of burning"&subsetA_dta$tr
ctrl.someburn <- as.numeric(length(which(subsetA_dta$m1.evi.burn=="Less than half of the area of the ru
ctrl.mostburn <- as.numeric(length(which(subsetA_dta$m1.evi.burn=="More than half of the area of the ru
trt.noburn <- as.numeric(length(which(subsetA_dta$m1.evi.burn=="No evidence of burning"&subsetA_dta$trea
trt.someburn <- as.numeric(length(which(subsetA_dta$m1.evi.burn=="Less than half of the area of the rubb
trt.mostburn <- as.numeric(length(which(subsetA_dta$m1.evi.burn=="More than half of the area of the rubb

prop.test(c(ctrl.noburn, trt.noburn), c(n.ctrl,n.trt), alternative = "l")
```

```
##
##  2-sample test for equality of proportions with continuity
##  correction
##
## data:  c(ctrl.noburn, trt.noburn) out of c(n.ctrl, n.trt)
## X-squared = 3.1753, df = 1, p-value = 0.03738
## alternative hypothesis: less
## 95 percent confidence interval:
##  -1.00000000 -0.00534564
## sample estimates:
##    prop 1    prop 2
## 0.4511494 0.5214900
```

```
prop.test(c(ctrl.someburn, trt.someburn), c(n.ctrl,n.trt), alternative = "g")
```

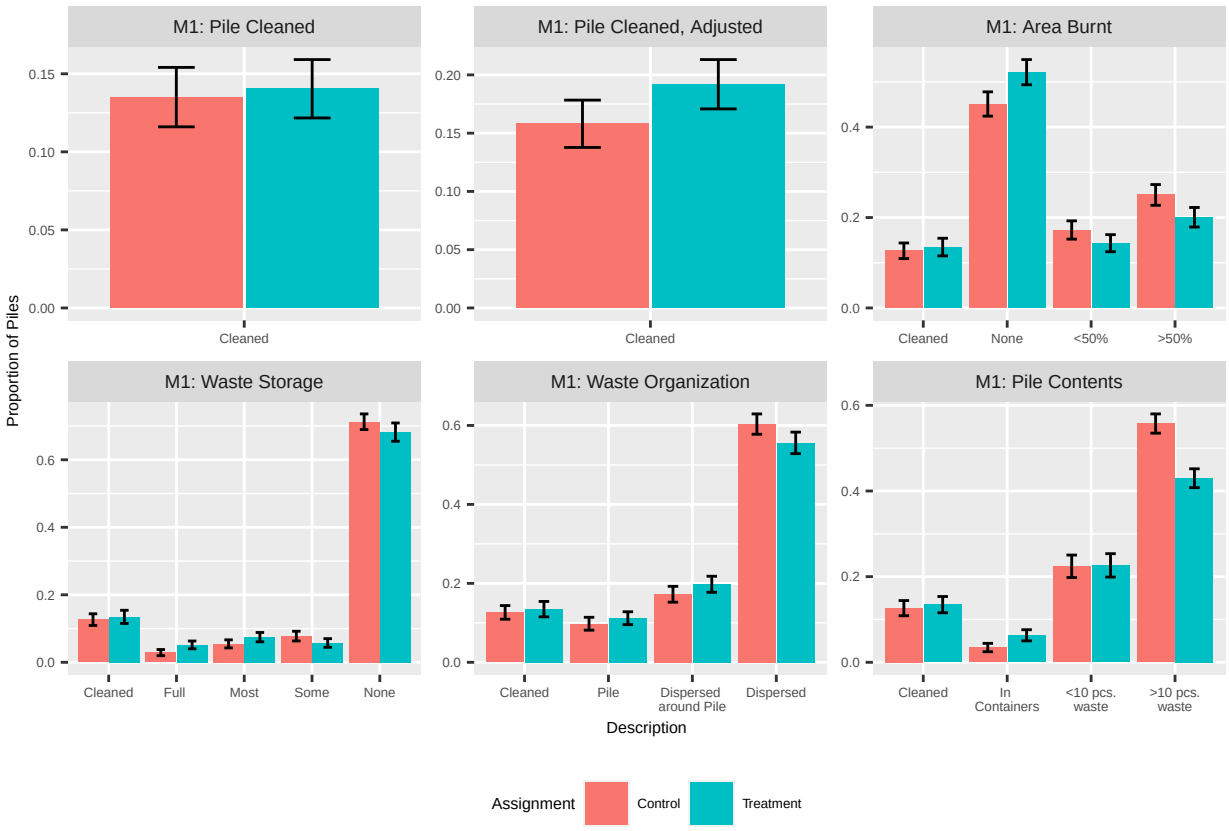
```
##
##  2-sample test for equality of proportions with continuity
##  correction
##
## data:  c(ctrl.someburn, trt.someburn) out of c(n.ctrl, n.trt)
## X-squared = 0.90529, df = 1, p-value = 0.1707
## alternative hypothesis: greater
## 95 percent confidence interval:
##  -0.01911874  1.00000000
```

```
## sample estimates:
##    prop 1    prop 2
## 0.1724138 0.1432665
```

```
prop.test(c(ctrl.mostburn, trt.mostburn), c(n.ctrl,n.trt), alternative = "g")
```

```
##
##  2-sample test for equality of proportions with continuity
##  correction
##
## data:  c(ctrl.mostburn, trt.mostburn) out of c(n.ctrl, n.trt)
## X-squared = 2.1643, df = 1, p-value = 0.07062
## alternative hypothesis: greater
## 95 percent confidence interval:
##  -0.0054113  1.0000000
## sample estimates:
##    prop 1    prop 2
## 0.2500000 0.2005731
```

```
#Final Plot
a$dv <- "M1: Pile Cleaned"
a.adj$dv <- "M1: Pile Cleaned, Adjusted"
b$dv <- "M1: Pile Contents"
c$dv <- "M1: Waste Storage"
d$dv <- "M1: Waste Organization"
e$dv <- "M1: Area Burnt"
m1_outcomes <- rbind(a,a.adj,b,c,d,e)
m1_outcomes$ci95.low <- m1_outcomes$prop - m1_outcomes$se
m1_outcomes$ci95.high <- m1_outcomes$prop + m1_outcomes$se

m1_outcomes$dv <- factor(ordered(m1_outcomes$dv, levels=c("M1: Pile Cleaned", "M1: Pile Cleaned, Adjust
                                                          "M1: Waste Storage", "M1: Waste Organization"
m1_outcomes_plot <- ggplot(m1_outcomes) +
  geom_bar(aes(desc, prop, group=Assignment, fill=Assignment), stat="identity", position=position_dodge
  geom_errorbar(aes(desc, ymin=ci95.low, ymax=ci95.high, group=Assignment), position=position_dodge(0.9
  scale_x_discrete("Description") +
  scale_y_continuous("Proportion of Piles") +
  facet_wrap(.~dv, scales="free") +
  theme(legend.position = "bottom",
        text=element_text(size=6),
        strip.text=element_text(size=7))
m1_outcomes_plot
```




```
# ggsave("./Figures/Figure7.tiff",
#        m1_outcomes_plot,
#        width=5, height=3.5, units="in")
# ggsave("./Figures/Figure7.png",
#        m1_outcomes_plot,
#        width=5, height=3.5, units="in")

## Figure 8: Descriptive Analyses, Midline 2 -----
table(subsetA_dta$treat, subsetA_dta$m2.waste.pile_d)
```

```
##
##       0   1
##   0  78 270
##   1  74 275
```

```
a2<-data.frame((prop.table(table(subsetA_dta$treat, subsetA_dta$m2.waste.pile_d),1)))
names(a2) <- c("treat", "desc", "prop")

a2$desc <- ordered(factor(c(rep("Cleaned", 2), rep("Not Cleaned", 2))), levels=c("Cleaned", "Not Cleaned
a2$desc.n <- as.numeric(a2$desc)
a2 <- a2[a2$desc=="Cleaned",]

ctrl <- as.numeric(length(which(subsetA_dta$m2.waste.pile_d==0&subsetA_dta$treat==0)))
trt <- as.numeric(length(which(subsetA_dta$m2.waste.pile_d==0&subsetA_dta$treat==1)))
n.ctrl <- as.numeric(length(which(subsetA_dta$treat==0))) #count of units in control stays the same for
n.trt <- as.numeric(length(which(subsetA_dta$treat==1))) #count of units in treatment stays the same fo

ctrl.means <- rep(NA,1000)
```

```r
trt.means <- rep(NA,1000)
set.seed(201)
for (i in 1:1000){
  ctrl.means[i] <- mean(sample(c(rep(0,ctrl),rep(1,n.ctrl-ctrl)), replace=T))
  trt.means[i] <- mean(sample(c(rep(0,trt),rep(1,n.trt-trt)), replace=T))
}

a2$se <- c(sd(ctrl.means),sd(trt.means))
a2$Assignment <- c("Control", "Treatment")

ctrl <- as.numeric(length(which(subsetA_dta$m2.waste.pile_d==0&subsetA_dta$treat==0)))
trt <- as.numeric(length(which(subsetA_dta$m2.waste.pile_d==0&subsetA_dta$treat==1)))
prop.test(c(ctrl,trt), c(n.ctrl,n.trt), alternative = "l")
```

```
##
##  2-sample test for equality of proportions with continuity
##  correction
##
## data:  c(ctrl, trt) out of c(n.ctrl, n.trt)
## X-squared = 0.087134, df = 1, p-value = 0.6161
## alternative hypothesis: less
## 95 percent confidence interval:
##  -1.00000000  0.06642406
## sample estimates:
##    prop 1    prop 2
## 0.2241379 0.2120344
```

```r
##WASTE PILE DUMMY ADJUSTED
table(subsetA_dta$treat, subsetA_dta$m2.waste.pile_d2)
```

```
##
##       0   1
##   0  90 258
##   1  92 257
```

```r
a2.adj <- data.frame((prop.table(table(subsetA_dta$treat, subsetA_dta$m2.waste.pile_d2),1)))
names(a2.adj) <- c("treat", "desc", "prop")
a2.adj$desc <- ordered(factor(c(rep("Cleaned", 2), rep("Not Cleaned", 2))), levels=c("Cleaned", "Not Cl
a2.adj$desc.n <- as.numeric(a2.adj$desc)
a2.adj <- a2.adj[a2.adj$desc=="Cleaned",]

ctrl <- as.numeric(length(which(subsetA_dta$m2.waste.pile_d2==0&subsetA_dta$treat==0)))
trt <- as.numeric(length(which(subsetA_dta$m2.waste.pile_d2==0&subsetA_dta$treat==1)))

ctrl.means <- rep(NA,1000)
trt.means <- rep(NA,1000)
set.seed(201)
for (i in 1:1000){
  ctrl.means[i] <- mean(sample(c(rep(0,ctrl),rep(1,n.ctrl-ctrl)), replace=T))
  trt.means[i] <- mean(sample(c(rep(0,trt),rep(1,n.trt-trt)), replace=T))
}

a2.adj$se <- c(sd(ctrl.means),sd(trt.means))
a2.adj$Assignment <- c("Control", "Treatment")
```

```r
prop.test(c(ctrl,trt), c(n.ctrl,n.trt), alternative="l")
```

```
##
##  2-sample test for equality of proportions with continuity
##  correction
##
## data:  c(ctrl, trt) out of c(n.ctrl, n.trt)
## X-squared = 0.0040598, df = 1, p-value = 0.4746
## alternative hypothesis: less
## 95 percent confidence interval:
##  -1.00000000  0.05261127
## sample estimates:
##    prop 1    prop 2
## 0.2586207 0.2636103
```

```r
##GENERAL PILE DESCRIPTION
table(subsetA_dta$treat, subsetA_dta$m2.pile.gdesc)
```

```
##
##
##   0  78
##   1  73
##
##     Large sack(s) or container(s) of rubbish that can easily be transported
##   0                                                                       9
##   1                                                                      18
##
##     Less than 10 pieces of non-organic waste
##   0                                       48
##   1                                       48
##
##     More than 10 pieces of non-organic waste
##   0                                      213
##   1                                      210
```

```r
subsetA_dta$m2.pile.gdesc <- ifelse(subsetA_dta$m2.pile.gdesc==""&subsetA_dta$m2.waste.pile_d==1, "Miss
table(subsetA_dta$treat, subsetA_dta$m2.pile.gdesc)
```

```
##
##
##   0  78
##   1  73
##
##     Large sack(s) or container(s) of rubbish that can easily be transported
##   0                                                                       9
##   1                                                                      18
##
##     Less than 10 pieces of non-organic waste
##   0                                       48
##   1                                       48
##
##     More than 10 pieces of non-organic waste
##   0                                      213
##   1                                      210
```

```r
b2 <- data.frame(prop.table(table(subsetA_dta$treat, subsetA_dta$m2.pile.gdesc),1))
names(b2) <- c("treat","desc","prop")
b2$desc <- as.character(b2$desc)
#Removing missing responses from table
b2 <- b2[!(b2$desc=="Missing"),]
#Recoding possible answers -> groups
b2$desc[b2$desc==""] <- "Cleaned"
b2$desc[b2$desc=="More than 10 pieces of non-organic waste"] <- ">10 pcs.\nwaste"
b2$desc[b2$desc=="Less than 10 pieces of non-organic waste"] <- "<10 pcs.\nwaste"
b2$desc[b2$desc=="Large sack(s) or container(s) of rubbish that can easily be transported"] <- "In \nCo
b2 <- b2[!is.na(b2$desc),]
b2$desc <- ordered(factor(b2$desc,
                          levels=c("Cleaned","In \nContainers","<10 pcs.\nwaste", ">10 pcs.\nwaste")))
b2$desc.n <- as.numeric(b2$desc)

con.prop.transport <- rep(NA,1000)
trt.prop.transport <- rep(NA,1000)

con.prop.large <- rep(NA,1000)
trt.prop.large <- rep(NA,1000)

con.prop.small <- rep(NA,1000)
trt.prop.small <- rep(NA,1000)

con.prop.cleaned <- rep(NA,1000)
trt.prop.cleaned <- rep(NA,1000)

set.seed(209)
for (i in 1:1000){
  con.vec <- sample(subsetA_dta$m2.pile.gdesc[subsetA_dta$treat==0], replace=T, size=length(subsetA_dta
  trt.vec <- sample(subsetA_dta$m2.pile.gdesc[subsetA_dta$treat==1], replace=T, size=length(subsetA_dta

  con.vec <- con.vec[!is.na(con.vec)]
  trt.vec <- trt.vec[!is.na(trt.vec)]

  con.prop.transport[i] <- length(con.vec[con.vec=="Large sack(s) or container(s) of rubbish that can e
  trt.prop.transport[i] <- length(trt.vec[trt.vec=="Large sack(s) or container(s) of rubbish that can e

  con.prop.large[i] <- length(con.vec[con.vec=="More than 10 pieces of non-organic waste"]) / length(co
  trt.prop.large[i] <- length(trt.vec[trt.vec=="More than 10 pieces of non-organic waste"]) / length(tr

  con.prop.small[i] <- length(con.vec[con.vec=="Less than 10 pieces of non-organic waste"]) / length(co
  trt.prop.small[i] <- length(trt.vec[trt.vec=="Less than 10 pieces of non-organic waste"]) / length(tr

  con.prop.cleaned[i] <- length(con.vec[con.vec==""]) / length(con.vec)
  trt.prop.cleaned[i] <- length(trt.vec[trt.vec==""]) / length(trt.vec)
}

b2$se[1:2] <- c(sd(con.prop.cleaned),sd(trt.prop.cleaned))
b2$se[3:4] <- c(sd(con.prop.transport),sd(trt.prop.transport))
b2$se[5:6] <- c(sd(con.prop.large),sd(trt.prop.large))
b2$se[7:8] <- c(sd(con.prop.small),sd(trt.prop.small))
```

```r
b2$Assignment <- c("Control", "Treatment")

#T-tests
ctrl.largesacks <- as.numeric(length(which(subsetA_dta$m2.pile.gdesc=="Large sack(s) or container(s) of
trt.largesacks <- as.numeric(length(which(subsetA_dta$m2.pile.gdesc=="Large sack(s) or container(s) of
ctrl.less10pcs <- as.numeric(length(which(subsetA_dta$m2.pile.gdesc=="Less than 10 pieces of non-organi
trt.less10pcs <- as.numeric(length(which(subsetA_dta$m2.pile.gdesc=="Less than 10 pieces of non-organic
ctrl.more10pcs <- as.numeric(length(which(subsetA_dta$m2.pile.gdesc=="More than 10 pieces of non-organi
trt.more10pcs <- as.numeric(length(which(subsetA_dta$m2.pile.gdesc=="More than 10 pieces of non-organic


prop.test(c(ctrl.largesacks,trt.largesacks), c(n.ctrl,n.trt), alternative = "l")
```

```
##
##  2-sample test for equality of proportions with continuity
##  correction
##
## data:  c(ctrl.largesacks, trt.largesacks) out of c(n.ctrl, n.trt)
## X-squared = 2.4421, df = 1, p-value = 0.05906
## alternative hypothesis: less
## 95 percent confidence interval:
##  -1.000000000  0.001136307
## sample estimates:
##     prop 1     prop 2
## 0.02586207 0.05157593
```

```r
prop.test(c(ctrl.less10pcs,trt.less10pcs), c(n.ctrl,n.trt), alternative = "l")
```

```
##
##  2-sample test for equality of proportions with continuity
##  correction
##
## data:  c(ctrl.less10pcs, trt.less10pcs) out of c(n.ctrl, n.trt)
## X-squared = 7.7821e-31, df = 1, p-value = 0.5
## alternative hypothesis: less
## 95 percent confidence interval:
##  -1.00000000  0.04373242
## sample estimates:
##    prop 1    prop 2
## 0.1379310 0.1375358
```

```r
prop.test(c(ctrl.more10pcs,trt.more10pcs), c(n.ctrl,n.trt), alternative = "g")
```

```
##
##  2-sample test for equality of proportions with continuity
##  correction
##
## data:  c(ctrl.more10pcs, trt.more10pcs) out of c(n.ctrl, n.trt)
## X-squared = 0.040868, df = 1, p-value = 0.4199
## alternative hypothesis: greater
## 95 percent confidence interval:
##  -0.05337895  1.00000000
## sample estimates:
##    prop 1    prop 2
## 0.6120690 0.6017192
```

```
##STORAGE
table(subsetA_dta$treat,subsetA_dta$m2.waste.stor)
```

```
##
##
##   0  78
##   1  73
##
##     All of the rubbish is neatly contained with sacks or other containers
##   0                                                                    12
##   1                                                                    21
##
##     Most of the rubbish is organized in sacks or other containers
##   0                                                             19
##   1                                                             16
##
##     No rubbish is contained in sacks or containers
##   0                                           222
##   1                                           215
##
##     Very little rubbish is contained within sacks or containers
##   0                                                       17
##   1                                                       24
```

```
c2 <- data.frame(prop.table(table(subsetA_dta$treat,subsetA_dta$m2.waste.stor),1))
names(c2) <- c("treat","desc","prop")
c2$desc <- as.character(c2$desc)
c2$desc[c2$desc==""] <- "Cleaned"
c2$desc[c2$desc=="No rubbish is contained in sacks or containers"] <- "None"
c2$desc[c2$desc=="All of the rubbish is neatly contained with sacks or other containers"] <- "Full \n"
c2$desc[c2$desc=="Most of the rubbish is organized in sacks or other containers"] <- "Most"
c2$desc[c2$desc=="Very little rubbish is contained within sacks or containers"] <- "Some"
c2 <- c2[!is.na(c2$desc),]
c2$desc <- ordered(factor(c2$desc,
                          levels=c("Cleaned","Full \n","Most","Some","None")))
c2$desc.n <- as.numeric(c$desc)

con.prop.uc <- rep(NA,1000)
trt.prop.uc <- rep(NA,1000)

con.prop.pc <- rep(NA,1000)
trt.prop.pc <- rep(NA,1000)

con.prop.mc <- rep(NA,1000)
trt.prop.mc <- rep(NA,1000)

con.prop.fc <- rep(NA,1000)
trt.prop.fc <- rep(NA,1000)

con.prop.cc <- rep(NA,1000)
trt.prop.cc <- rep(NA,1000)

set.seed(201)
for (i in 1:1000){
```

```r
  con.vec_c <- sample(subsetA_dta$m2.waste.stor[subsetA_dta$treat==0], replace=T, size=length(subsetA_dt
  trt.vec_c <- sample(subsetA_dta$m2.waste.stor[subsetA_dta$treat==1], replace=T, size=length(subsetA_dt

  con.vec_c <- con.vec_c[!is.na(con.vec_c)]
  trt.vec_c <- trt.vec_c[!is.na(trt.vec_c)]

  con.prop.uc[i] <- length(con.vec_c[con.vec_c=="No rubbish is contained in sacks or containers"])/leng
  trt.prop.uc[i] <- length(trt.vec_c[trt.vec_c=="No rubbish is contained in sacks or containers"])/leng

  con.prop.pc[i] <- length(con.vec_c[con.vec_c=="Very little rubbish is contained within sacks or conta
  trt.prop.pc[i] <- length(trt.vec_c[trt.vec_c=="Very little rubbish is contained within sacks or conta

  con.prop.mc[i] <- length(con.vec_c[con.vec_c=="Most of the rubbish is organized in sacks or other con
  trt.prop.mc[i] <- length(trt.vec_c[trt.vec_c=="Most of the rubbish is organized in sacks or other con

  con.prop.fc[i] <- length(con.vec_c[con.vec_c=="All of the rubbish is neatly contained with sacks or o
  trt.prop.fc[i] <- length(trt.vec_c[trt.vec_c=="All of the rubbish is neatly contained with sacks or o

  con.prop.cc[i] <- length(con.vec_c[con.vec_c==""])/length(con.vec_c)
  trt.prop.cc[i] <- length(trt.vec_c[trt.vec_c==""])/length(trt.vec_c)

}

c2$se[1:2] <- c(sd(con.prop.cc),sd(trt.prop.cc))
c2$se[3:4] <- c(sd(con.prop.fc),sd(trt.prop.fc))
c2$se[5:6] <- c(sd(con.prop.mc),sd(trt.prop.mc))
c2$se[7:8] <- c(sd(con.prop.uc),sd(trt.prop.uc))
c2$se[9:10] <- c(sd(con.prop.pc),sd(trt.prop.pc))
c2$Assignment <- c("Control", "Treatment")

#T-tests
ctrl.full <- as.numeric(length(which(subsetA_dta$m2.waste.stor=="All of the rubbish is neatly contained
ctrl.most <- as.numeric(length(which(subsetA_dta$m2.waste.stor=="Most of the rubbish is organized in sa
ctrl.some <- as.numeric(length(which(subsetA_dta$m2.waste.stor=="Very little rubbish is contained withi
ctrl.none <- as.numeric(length(which(subsetA_dta$m2.waste.stor=="No rubbish is contained in sacks or co
trt.full <- as.numeric(length(which(subsetA_dta$m2.waste.stor=="All of the rubbish is neatly contained
trt.most <- as.numeric(length(which(subsetA_dta$m2.waste.stor=="Most of the rubbish is organized in sac
trt.some <- as.numeric(length(which(subsetA_dta$m2.waste.stor=="Very little rubbish is contained within
trt.none <- as.numeric(length(which(subsetA_dta$m2.waste.stor=="No rubbish is contained in sacks or con

prop.test(c(ctrl.full,trt.full), c(n.ctrl,n.trt), alternative = "l")
```

```
##
##  2-sample test for equality of proportions with continuity
##  correction
##
## data:  c(ctrl.full, trt.full) out of c(n.ctrl, n.trt)
## X-squared = 2.0118, df = 1, p-value = 0.07804
## alternative hypothesis: less
## 95 percent confidence interval:
##  -1.00000000  0.00358567
## sample estimates:
##     prop 1     prop 2
## 0.03448276 0.06017192
```

```r
prop.test(c(ctrl.most,trt.most), c(n.ctrl,n.trt), alternative = "l")
```

```
##
##  2-sample test for equality of proportions with continuity
##  correction
##
## data:  c(ctrl.most, trt.most) out of c(n.ctrl, n.trt)
## X-squared = 0.12645, df = 1, p-value = 0.6389
## alternative hypothesis: less
## 95 percent confidence interval:
##  -1.00000000  0.03883233
## sample estimates:
##     prop 1     prop 2
## 0.05459770 0.04584527
```

```r
prop.test(c(ctrl.some,trt.some), c(n.ctrl,n.trt), alternative = "g")
```

```
##
##  2-sample test for equality of proportions with continuity
##  correction
##
## data:  c(ctrl.some, trt.some) out of c(n.ctrl, n.trt)
## X-squared = 0.91473, df = 1, p-value = 0.8306
## alternative hypothesis: greater
## 95 percent confidence interval:
##  -0.05207304  1.00000000
## sample estimates:
##     prop 1     prop 2
## 0.04885057 0.06876791
```

```r
prop.test(c(ctrl.none,trt.none), c(n.ctrl,n.trt), alternative = "g")
```

```
##
##  2-sample test for equality of proportions with continuity
##  correction
##
## data:  c(ctrl.none, trt.none) out of c(n.ctrl, n.trt)
## X-squared = 0.26941, df = 1, p-value = 0.3019
## alternative hypothesis: greater
## 95 percent confidence interval:
##  -0.04122874  1.00000000
## sample estimates:
##    prop 1    prop 2
## 0.6379310 0.6160458
```

```r
#ORGANIZATION
table(subsetA_dta$treat,subsetA_dta$m2.pile.org)
```

```
##
##         All of the rubbish is collected into a single pile
##   0  78                                                  40
##   1  73                                                  41
##
##     Most of the rubbish is organized around a single pile but other rubbish is spread nearby
##   0                                                                                       81
##   1                                                                                       80
```

```
##
##     Rubbish is spread all round no evidence of the rubbish being organized
##   0                                                                    149
##   1                                                                    155
```

```r
d2 <- data.frame(prop.table(table(subsetA_dta$treat,subsetA_dta$m2.pile.org),1))
names(d2) <- c("treat","desc","prop")
d2$desc <- as.character(d2$desc)

#Recoding possible answers -> groups
d2$desc[d2$desc==""] <- "Cleaned"
d2$desc[d2$desc=="All of the rubbish is collected into a single pile"] <- "Pile"
d2$desc[d2$desc=="Most of the rubbish is organized around a single pile but other rubbish is spread nea
d2$desc[d2$desc=="Rubbish is spread all round no evidence of the rubbish being organized"] <- "Dispersed
d2 <- d2[!is.na(d2$desc),]
d2$desc <- ordered(factor(d2$desc, levels = c("Cleaned","Pile","Dispersed \naround Pile","Dispersed")))
d2$desc.n <- as.numeric(d2$desc)

con.prop.wopc <- rep(NA,1000)
trt.prop.wopc <- rep(NA,1000)

con.prop.sp <- rep(NA,1000)
trt.prop.sp <- rep(NA,1000)

con.prop.spd <- rep(NA,1000)
trt.prop.spd <- rep(NA,1000)

con.prop.npd <- rep(NA,1000)
trt.prop.npd <- rep(NA,1000)

set.seed(201)
for (i in 1:1000){
  con.vec_d <- sample(subsetA_dta$m2.pile.org[subsetA_dta$treat==0], replace=T, size=length(subsetA_dta
  trt.vec_d <- sample(subsetA_dta$m2.pile.org[subsetA_dta$treat==1], replace=T, size=length(subsetA_dta

  con.vec_d <- con.vec_d[!is.na(con.vec_d)]
  trt.vec_d <- trt.vec_d[!is.na(trt.vec_d)]

  con.prop.wopc[i] <- length(con.vec_d[con.vec_d==""])/length(con.vec_d)
  trt.prop.wopc[i] <- length(trt.vec_d[trt.vec_d==""])/length(trt.vec_d)

  con.prop.sp[i] <- length(con.vec_d[con.vec_d=="All of the rubbish is collected into a single pile"])/
  trt.prop.sp[i] <- length(trt.vec_d[trt.vec_d=="All of the rubbish is collected into a single pile"])/

  con.prop.spd[i] <- length(con.vec_d[con.vec_d=="Most of the rubbish is organized around a single pile
  trt.prop.spd[i] <- length(trt.vec_d[trt.vec_d=="Most of the rubbish is organized around a single pile

  con.prop.npd[i] <- length(con.vec_d[con.vec_d=="Rubbish is spread all round no evidence of the rubbis
  trt.prop.npd[i] <- length(trt.vec_d[trt.vec_d=="Rubbish is spread all round no evidence of the rubbis
}

d2$se[1:2] <- c(sd(con.prop.wopc),sd(trt.prop.wopc))
d2$se[3:4] <- c(sd(con.prop.sp),sd(trt.prop.sp))
d2$se[5:6] <- c(sd(con.prop.spd),sd(trt.prop.spd))
```

```r
d2$se[7:8] <- c(sd(con.prop.npd),sd(trt.prop.npd))
d2$Assignment <- c("Control", "Treatment")

#T-tests
ctrl.single <- as.numeric(length(which(subsetA_dta$m2.pile.org=="All of the rubbish is collected into a
ctrl.partdispersed <- as.numeric(length(which(subsetA_dta$m2.pile.org=="Most of the rubbish is organize
ctrl.dispersed <- as.numeric(length(which(subsetA_dta$m2.pile.org=="Rubbish is spread all round no evid
trt.single <- as.numeric(length(which(subsetA_dta$m2.pile.org=="All of the rubbish is collected into a
trt.partdispersed <- as.numeric(length(which(subsetA_dta$m2.pile.org=="Most of the rubbish is organized
trt.dispersed <- as.numeric(length(which(subsetA_dta$m2.pile.org=="Rubbish is spread all round no evide

prop.test(c(ctrl.single,trt.single), c(n.ctrl,n.trt), alternative = "l")#p=0.315; single pile
```

```
##
##  2-sample test for equality of proportions with continuity
##  correction
##
## data:  c(ctrl.single, trt.single) out of c(n.ctrl, n.trt)
## X-squared = 7.9904e-31, df = 1, p-value = 0.5
## alternative hypothesis: less
## 95 percent confidence interval:
##  -1.00000000  0.03993301
## sample estimates:
##    prop 1    prop 2
## 0.1149425 0.1174785
```

```r
prop.test(c(ctrl.partdispersed,trt.partdispersed), c(n.ctrl,n.trt), alternative = "l")#p=0.223; most in
```

```
##
##  2-sample test for equality of proportions with continuity
##  correction
##
## data:  c(ctrl.partdispersed, trt.partdispersed) out of c(n.ctrl, n.trt)
## X-squared = 0.00043095, df = 1, p-value = 0.5083
## alternative hypothesis: less
## 95 percent confidence interval:
##  -1.00000000  0.05891913
## sample estimates:
##    prop 1    prop 2
## 0.2327586 0.2292264
```

```r
prop.test(c(ctrl.dispersed,trt.dispersed), c(n.ctrl,n.trt), alternative="g")#p=0.116; no pile
```

```
##
##  2-sample test for equality of proportions with continuity
##  correction
##
## data:  c(ctrl.dispersed, trt.dispersed) out of c(n.ctrl, n.trt)
## X-squared = 0.12151, df = 1, p-value = 0.6363
## alternative hypothesis: greater
## 95 percent confidence interval:
##  -0.08061956  1.00000000
## sample estimates:
##    prop 1    prop 2
## 0.4281609 0.4441261
```

```r
#BURNING
table(subsetA_dta$treat,subsetA_dta$m2.evi.burn)
```

```
##
##
##   0  78
##   1  73
##
##     Less than half of the area of the rubbish pile contains evidence of burning
##   0                                                                          72
##   1                                                                          58
##
##     More than half of the area of the rubbish pile contains evidence of burning
##   0                                                                          64
##   1                                                                          72
##
##     No evidence of burning
##   0                    134
##   1                    146
```

```r
e2 <- data.frame(prop.table(table(subsetA_dta$treat,subsetA_dta$m2.evi.burn),1))
names(e2) <- c("treat","desc","prop")
e2$desc <- as.character(e2$desc)

con.prop.ec <- rep(NA,1000)
trt.prop.ec <- rep(NA,1000)

con.prop.bl <- rep(NA,1000)
trt.prop.bl <- rep(NA,1000)

con.prop.bm <- rep(NA,1000)
trt.prop.bm <- rep(NA,1000)

con.prop.nb <- rep(NA,1000)
trt.prop.nb <- rep(NA,1000)


#Recoding possible answers -> groups
e2$desc[e2$desc==""] <- "Cleaned"
e2$desc[e2$desc=="Less than half of the area of the rubbish pile contains evidence of burning"] <- "<50
e2$desc[e2$desc=="More than half of the area of the rubbish pile contains evidence of burning"] <- ">50
e2$desc[e2$desc=="No evidence of burning"] <- "None"
e2 <- e2[!is.na(e2$desc),]
e2$desc <- ordered(factor(e2$desc, levels = c("Cleaned","None", "<50%", ">50%")))
e2$desc.n <- as.numeric(e2$desc)

set.seed(201)

for (i in 1:1000){
  con.vec_e <- sample(subsetA_dta$m2.evi.burn[subsetA_dta$treat==0], replace=T, size=length(subsetA_dta
  trt.vec_e <- sample(subsetA_dta$m2.evi.burn[subsetA_dta$treat==1], replace=T, size=length(subsetA_dta

  con.vec_e <- con.vec_e[!is.na(con.vec_e)]
  trt.vec_e <- trt.vec_e[!is.na(trt.vec_e)]
```

```
  con.prop.ec[i] <- length(con.vec_e[con.vec_e==""])/length(con.vec_e)
  trt.prop.ec[i] <- length(trt.vec_e[trt.vec_e==""])/length(trt.vec_e)

  con.prop.nb[i] <- length(con.vec_e[con.vec_e=="No evidence of burning"])/length(con.vec_e)
  trt.prop.nb[i] <- length(trt.vec_e[trt.vec_e=="No evidence of burning"])/length(trt.vec_e)

  con.prop.bl[i] <- length(con.vec_e[con.vec_e=="Less than half of the area of the rubbish pile contains
  trt.prop.bl[i] <- length(trt.vec_e[trt.vec_e=="Less than half of the area of the rubbish pile contains

  con.prop.bm[i] <- length(con.vec_e[con.vec_e=="More than half of the area of the rubbish pile contains
  trt.prop.bm[i] <- length(trt.vec_e[trt.vec_e=="More than half of the area of the rubbish pile contains

}

e2$se[1:2] <- c(sd(con.prop.ec),sd(trt.prop.ec))
e2$se[3:4] <- c(sd(con.prop.bl),sd(trt.prop.bl))
e2$se[5:6] <- c(sd(con.prop.bm),sd(trt.prop.bm))
e2$se[7:8] <- c(sd(con.prop.nb),sd(trt.prop.nb))
e2$Assignment <- c("Control", "Treatment")

#T-tests
ctrl.noburn <- as.numeric(length(which(subsetA_dta$m2.evi.burn=="No evidence of burning"&subsetA_dta$tr
ctrl.someburn <- as.numeric(length(which(subsetA_dta$m2.evi.burn=="Less than half of the area of the ru
ctrl.mostburn <- as.numeric(length(which(subsetA_dta$m2.evi.burn=="More than half of the area of the ru
trt.noburn <- as.numeric(length(which(subsetA_dta$m2.evi.burn=="No evidence of burning"&subsetA_dta$trea
trt.someburn <- as.numeric(length(which(subsetA_dta$m2.evi.burn=="Less than half of the area of the rubb
trt.mostburn <- as.numeric(length(which(subsetA_dta$m2.evi.burn=="More than half of the area of the rubb

prop.test(c(ctrl.noburn, trt.noburn), c(n.ctrl,n.trt), alternative = "l")#p=0.037; no evidence of burni
```

```
##
##  2-sample test for equality of proportions with continuity
##  correction
##
## data:  c(ctrl.noburn, trt.noburn) out of c(n.ctrl, n.trt)
## X-squared = 0.67052, df = 1, p-value = 0.2064
## alternative hypothesis: less
## 95 percent confidence interval:
##  -1.00000000  0.03064037
## sample estimates:
##    prop 1    prop 2
## 0.3850575 0.4183381
```

```
prop.test(c(ctrl.someburn, trt.someburn), c(n.ctrl,n.trt), alternative = "g")#p=0.171; some evidence of
```

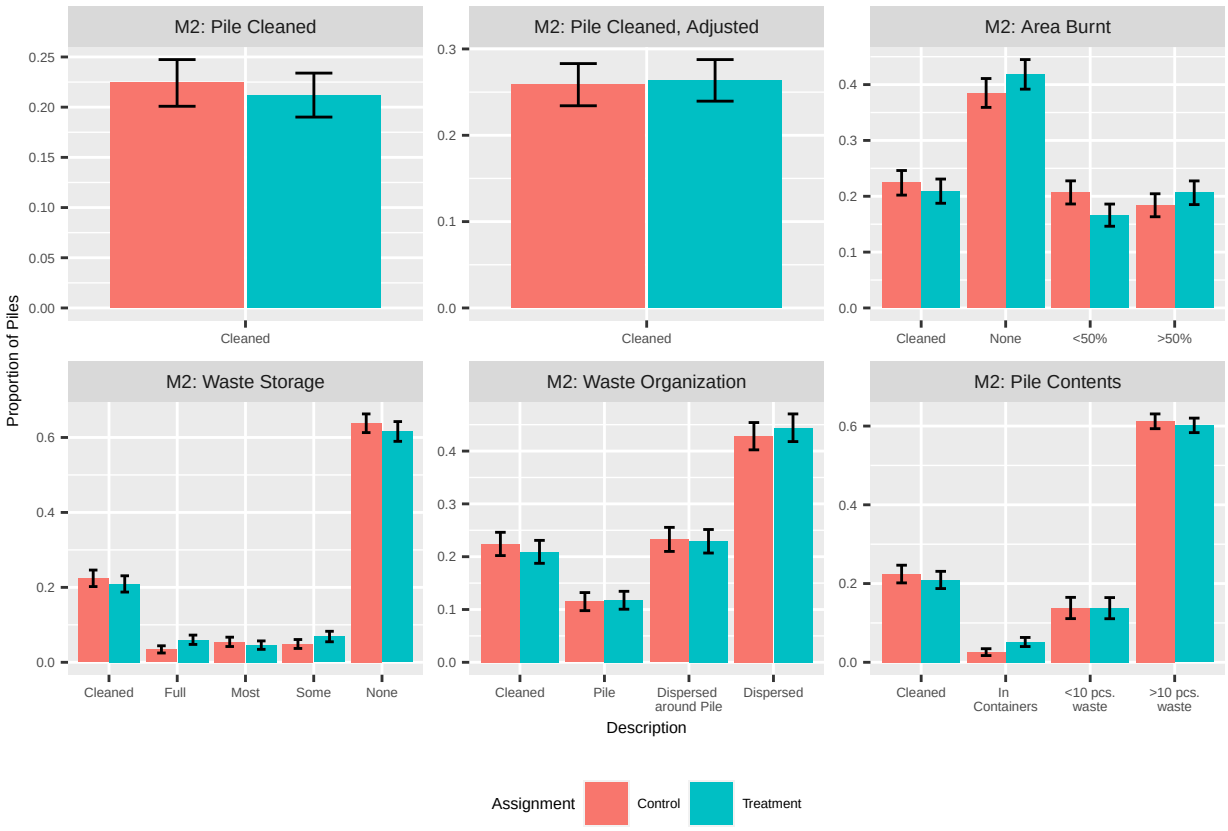
```
##
##  2-sample test for equality of proportions with continuity
##  correction
##
## data:  c(ctrl.someburn, trt.someburn) out of c(n.ctrl, n.trt)
## X-squared = 1.6442, df = 1, p-value = 0.09987
## alternative hypothesis: greater
## 95 percent confidence interval:
##  -0.01063843  1.00000000
```

```
## sample estimates:
##    prop 1    prop 2
## 0.2068966 0.1661891
```

```
prop.test(c(ctrl.mostburn, trt.mostburn), c(n.ctrl,n.trt), alternative = "g")#p=0.071; significant evid
```

```
##
##  2-sample test for equality of proportions with continuity
##  correction
##
## data:  c(ctrl.mostburn, trt.mostburn) out of c(n.ctrl, n.trt)
## X-squared = 0.42303, df = 1, p-value = 0.7423
## alternative hypothesis: greater
## 95 percent confidence interval:
##  -0.07462327  1.00000000
## sample estimates:
##    prop 1    prop 2
## 0.1839080 0.2063037
```

```
#Final Plot
a2$dv <- "M2: Pile Cleaned"
a2.adj$dv <- "M2: Pile Cleaned, Adjusted"
b2$dv <- "M2: Pile Contents"
c2$dv <- "M2: Waste Storage"
d2$dv <- "M2: Waste Organization"
e2$dv <- "M2: Area Burnt"
m2_outcomes <- rbind(a2,a2.adj,b2,c2,d2,e2)
m2_outcomes$ci95.low <- m2_outcomes$prop - m2_outcomes$se
m2_outcomes$ci95.high <- m2_outcomes$prop + m2_outcomes$se

m2_outcomes$dv <- factor(ordered(m2_outcomes$dv, levels=c("M2: Pile Cleaned", "M2: Pile Cleaned, Adjust
                                                          "M2: Waste Storage", "M2: Waste Organization"
m2_outcomes_plot <- ggplot(m2_outcomes) +
  geom_bar(aes(desc, prop, group=Assignment, fill=Assignment), stat="identity", position=position_dodge
  geom_errorbar(aes(desc, ymin=ci95.low, ymax=ci95.high, group=Assignment), position=position_dodge(0.9
  scale_x_discrete("Description") +
  scale_y_continuous("Proportion of Piles") +
  facet_wrap(.~dv, scales="free") +
  theme(legend.position = "bottom",
        text=element_text(size=6),
        strip.text=element_text(size=7))
m2_outcomes_plot
```




```
# ggsave("./Figures/Figure8.tiff",
#        m2_outcomes_plot,
#        width=5, height=3.5, units="in")
# ggsave("./Figures/Figure8.png",
#        m2_outcomes_plot,
#        width=5, height=3.5, units="in")

## Figure 9: Reporting Consistency -----
plot.df <- subset(main_dta, select=c(zone.id,bad.overall,prop.deviant_overall))
plot.df <- plot.df[!duplicated(plot.df),]

main_dta$bad.overall
```

```
##   [1]  0  0  0  0  0  0  0  0 NA NA NA NA  0  0  0  0  0  0  0  0 NA NA NA
##  [24] NA  0  0  0  0  0  0  0  0  0  0  0  0 NA NA NA NA  0  0  0  0  0  0
##  [47]  0  0 NA NA NA NA  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0
##  [70]  0  0  0  0  0  0  0 NA NA NA NA  0  0  0  0  1  1  1  1 NA NA NA NA
##  [93] NA NA NA NA  1  1  1  1  1  1  1  1  1  1  1  1  0  0  0  0  0  0  0
## [116]  0  1  1  1  1 NA NA NA NA NA NA NA NA  0  0  0  0  0  0  0  0 NA NA
## [139] NA NA NA NA NA NA NA NA NA NA  1  1  1  1 NA NA NA NA  0  0  0  0  0
## [162]  0  0  0 NA NA NA NA NA NA NA NA NA NA NA NA  1  1  1  1  0  0  0  0
## [185]  1  1  1  1 NA NA NA NA  1  1  1  1 NA NA NA NA NA NA NA NA  0  0  0
## [208]  0  0  0  0  0 NA NA NA NA  0  0  0  0  0  0  0  0  0  0  0  0  0  0
## [231]  0  0 NA NA NA NA NA NA NA NA  0  0  0  0  1  1  1  1  0  0  0  0  0
## [254]  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  1  1  1  1 NA NA NA NA
## [277] NA NA NA NA  0  0  0  0  1  1  1  1 NA NA NA NA  0  0  0  0  0  0  0
## [300]  0  0  0  0  0 NA NA NA NA  0  0  0  0  1  1  1  1 NA NA NA NA  0  0
## [323]  0  0  0  0  0  0  0  0  0  0 NA NA NA NA  0  0  0  0 NA NA NA NA  0
```

```
## [346]  0  0  0  0  0  0  0 NA NA NA NA NA NA NA NA  0  0  0  0 NA NA NA NA
## [369]  0  0  0  0 NA NA NA NA  0  0  0  0 NA NA NA NA NA NA NA NA  0  0  0
## [392]  0 NA NA NA NA NA NA NA NA NA NA NA NA  1  1  1  1  0  0  0  0  1  1
## [415]  1  1  0  0  0  0  1  1  1  1  0  0  0  0  1  1  1  1  0  0  0  0  0
## [438]  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0 NA NA NA NA NA NA NA NA
## [461]  1  1  1  1 NA NA NA NA  0  0  0  0  0  0  0  0  0  0  0  0 NA NA NA
## [484] NA  0  0  0  0 NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA  0  0
## [507]  0  0  0  0  0  0 NA NA NA NA  0  0  0  0 NA NA NA NA  1  1  1  1  0
## [530]  0  0  0  0  0  0  0 NA NA NA NA  0  0  0  0  0  0  0  0  0  0  0  0
## [553]  1  1  1  1 NA NA NA NA  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0
## [576]  0 NA NA NA NA NA NA NA NA  0  0  0  0 NA NA NA NA  0  0  0  0 NA NA
## [599] NA NA  0  0  0  0  0  0  0  0 NA NA NA NA NA NA NA NA  1  1  1  1  0
## [622]  0  0  0  1  1  1  1 NA NA NA NA NA NA NA NA  1  1  1  1 NA NA NA NA
## [645]  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  1  1  1  1  0  0  0
## [668]  0  1  1  1  1  0  0  0  0  0  0  0  0  1  1  1  1  0  0  0  0  1  1
## [691]  1  1  0  0  0  0  0  0  0  0 NA NA NA NA NA NA NA NA NA NA NA NA  0
## [714]  0  0  0 NA NA NA NA  0  0  0  0 NA NA NA NA  0  0  0  0  1  1  1  1
## [737]  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0 NA NA NA NA  0  0  0
## [760]  0  1  1  1  1  0  0  0  0 NA NA NA NA  0  0  0  0 NA NA NA NA NA NA
## [783] NA NA  0  0  0  0 NA NA NA NA  0  0  0  0  1  1  1  1
```

```
zid <- plot.df$zone.id

lc1@data$prop.bad.overall <- NA
for(i in 1:length(zid)) {
  lc1@data$prop.deviant_overall[lc1@data$zone.id==zid[i]] <- unique(plot.df$prop.deviant_overall[plot.d
}

summary(lc1@data$prop.deviant_overall)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.    NA's
##  0.0000  0.1523  0.2308  0.2104  0.2899  0.4022     625
```

```
summary(plot.df$prop.deviant_overall)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.    NA's
##  0.0000  0.1523  0.2308  0.2104  0.2899  0.4022      70
```

```
lc1.f <- fortify(lc1, region="zone.id")
lc1.df <- left_join(lc1.f, lc1@data, by=c("id"="zone.id"))

bad.service <- as.character(plot.df$zone.id[plot.df$bad.overall==1])
bad.service <- bad.service[!is.na(bad.service)]

#Basemap of Kampala
basemap <- ggplot() +
  geom_polygon(data=lc1, aes(long,lat,group=group), fill="white") +
  geom_path(data=lc1, aes(long,lat,group=group), color="grey65", size=0.15) +
  theme_void() +
  coord_equal()
```

```
## Regions defined for each Polygons
```

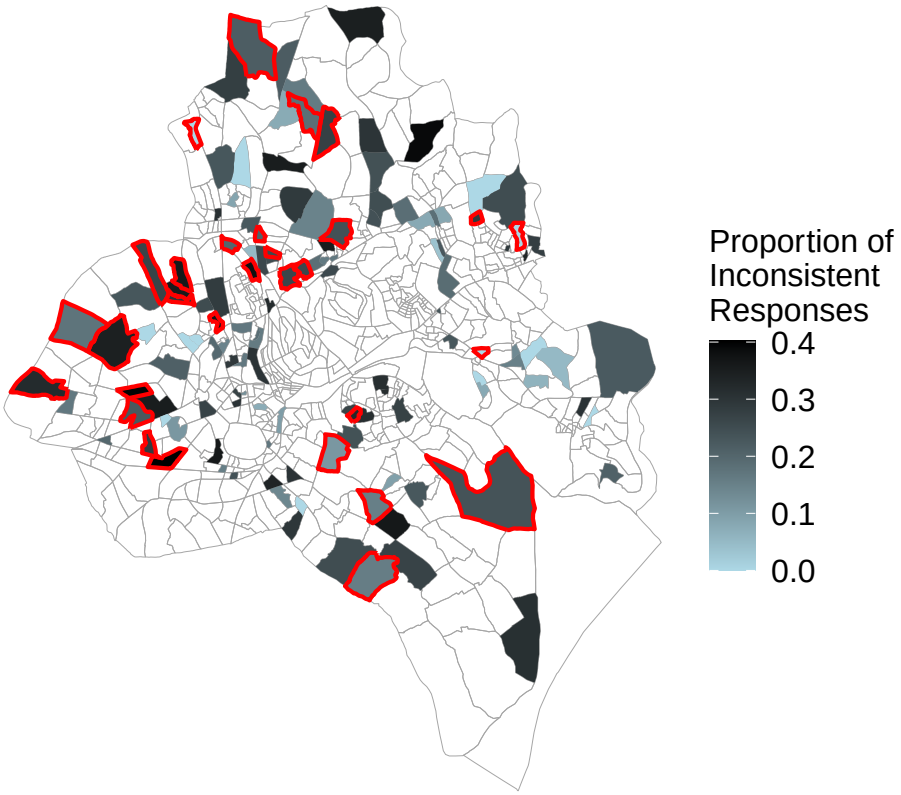
```
## Regions defined for each Polygons
```

```
#Final Figure
bad.overall_map <- basemap +
```

```
  geom_polygon(data=lc1.df[is.na(lc1.df$prop.deviant_overall)==FALSE,], aes(long,lat,group=group,fill=pr
  geom_path(data=lc1.df[c(lc1.df$id %in% bad.service & is.na(lc1.df$prop.deviant_overall)==FALSE),], aes
  theme_void() +
  coord_equal() +
  scale_fill_gradient("Proportion of\nInconsistent\nResponses", high="black", low="light blue") +
  theme(legend.title = element_text(size=12), legend.text = element_text(size=12), legend.position = "ri
```

```
## Coordinate system already present. Adding new coordinate system, which will replace the existing one
```

```
bad.overall_map
```




```
# tiff("./Figures/Figure9.tiff", width=5, height=5, units="in", res=300)
# bad.overall_map
# dev.off()
# png("./Figures/Figure9.png", width=5, height=5, units="in", res=300)
# bad.overall_map
# dev.off()

## Figure J1: Balance Table -----
##Making Pile Size Declies for better visualization
main_dta$b.pile.area_dec <- NA
main_dta$b.pile.area_dec <- ntile(main_dta$b.pile.area_m, 10)

table(main_dta$b.pile.area_dec, useNA="always")
```

```
##
##    1    2    3    4    5    6    7    8    9   10 <NA>
##   78   77   78   77   77   78   77   78   77   77   26
```

```
pre.ph3 <- subset(main_dta, select=c(unique.id, zone.id, treat, div,
                                     p1.p2.monitoring, b.type.site, b.pile.area_dec,
                                     lights.mean, area.km2, road.density,ls.pop_2016))
table(pre.ph3$b.type.site)
```

```
##
##
##                                          45
##              Household sack or bin for collection
##                                          17
##                        Littering in public place
##                                          33
##                 Official dumping site or continer
##                                          97
##                                       Other
##                                          10
##                 Pile within household for burning
##                                          39
##                                    Pit latrine
##                                           1
##                 Pit within household for burrying
##                                           3
##                      Small pile outside household
##                                          29
## Unofficial dumping site (used by many households)
##                                         500
```

```
##Change Variable Names for Plotting
var.names <- c(names(pre.ph3)[1:3], "Division",
               "Previous Phase Reporting", "Pile Type", "Baseline Pile Area, Verified (Decile)",
               "Zone-Level NTL (Mean)", "Zone-Level Area (Mean)", "Zone-Level Road Density (Mean)", "Zo

let <- str_replace_all(paste(LETTERS[seq(1,9)],"."),"  ", "")

var.names[4:12] <- paste(let, var.names[4:12])
var.names
```

```
##  [1] "unique.id"
##  [2] "zone.id"
##  [3] "treat"
##  [4] "A. Division"
##  [5] "B. Previous Phase Reporting"
##  [6] "C. Pile Type"
##  [7] "D. Baseline Pile Area, Verified (Decile)"
##  [8] "E. Zone-Level NTL (Mean)"
##  [9] "F. Zone-Level Area (Mean)"
## [10] "G. Zone-Level Road Density (Mean)"
## [11] "H. Zone-Level Population (2016)"
## [12] "I. NA"
```

```
##Changing Labels w/in Variables
c.cols <- 4:7
pre.ph3[,c.cols] = apply(pre.ph3[,c.cols], 2, function(x) (as.character(x)))
n.cols <- 8:ncol(pre.ph3)
pre.ph3[,n.cols] = apply(pre.ph3[,n.cols], 2, function(x) (as.numeric(x)))
```

```
##Changing labels within variables
table(pre.ph3$p1.p2.monitoring, useNA="always")
```

```
##
##    0    1 <NA>
##  536  264    0
```

```
pre.ph3$p1.p2.monitoring <- ifelse(pre.ph3$p1.p2.monitoring=="1","yes",pre.ph3$p1.p2.monitoring)
pre.ph3$p1.p2.monitoring <- ifelse(pre.ph3$p1.p2.monitoring=="0","no",pre.ph3$p1.p2.monitoring)
table(pre.ph3$p1.p2.monitoring, useNA="always")
```

```
##
##   no  yes <NA>
##  536  264    0
```

```
table(pre.ph3$b.type.site, useNA="always")
```

```
##
##
##                                            45
##              Household sack or bin for collection
##                                            17
##                         Littering in public place
##                                            33
##                Official dumping site or continer
##                                            97
##                                         Other
##                                            10
##                Pile within household for burning
##                                            39
##                                     Pit latrine
##                                             1
##                Pit within household for burrying
##                                             3
##                     Small pile outside household
##                                            29
## Unofficial dumping site (used by many households)
##                                           500
##                                          <NA>
##                                            26
```

```
pre.ph3$b.type.site <- ifelse(pre.ph3$b.type.site=="Pile within household for burning","Household pile"
pre.ph3$b.type.site <- ifelse(pre.ph3$b.type.site=="Pit within household for burrying","Household pile"
pre.ph3$b.type.site <- ifelse(pre.ph3$b.type.site=="Household sack or bin for collection","Household sa
pre.ph3$b.type.site <- ifelse(pre.ph3$b.type.site=="Littering in public place","Littering",pre.ph3$b.ty
pre.ph3$b.type.site <- ifelse(pre.ph3$b.type.site=="Official dumping site or continer","Official dumpin
pre.ph3$b.type.site <- ifelse(pre.ph3$b.type.site=="Unofficial dumping site (used by many households)",
pre.ph3$b.type.site <- ifelse(pre.ph3$b.type.site=="Small pile outside household","Small pile",pre.ph3$
pre.ph3$b.type.site <- ifelse(pre.ph3$b.type.site=="Pit latrine","Other",pre.ph3$b.type.site)
pre.ph3$b.type.site <- ifelse(pre.ph3$b.type.site=="","Missing",pre.ph3$b.type.site)
table(pre.ph3$b.type.site, useNA="always")
```

```
##
##          Household pile   Household sack or bin               Littering
##                      42                      17                      33
```

```
##                 Missing   Official dumping site                   Other
##                      45                      97                      11
##              Small pile Unofficial dumping site                    <NA>
##                      29                     500                      26
```

```
pre.ph3$b.pile.area_dec <- str_replace_all(pre.ph3$b.pile.area_dec, " ", "")

##Creating factors with appropriate levels
table(pre.ph3$treat, useNA = "always")
```

```
##
##    0    1 <NA>
##  400  400    0
```

```
pre.ph3$treat <- factor(pre.ph3$treat, levels=c("0", "1"))
table(pre.ph3$treat, useNA="always")
```

```
##
##    0    1 <NA>
##  400  400    0
```

```
table(pre.ph3$div, useNA="always")
```

```
##
##  central  kawempe makindye   nakawa   rubaga     <NA>
##       88      156      152      236      168        0
```

```
pre.ph3$div <- factor(pre.ph3$div, levels=c("central","kawempe","makindye","rubaga","nakawa"))
table(pre.ph3$div, useNA="always")
```

```
##
##  central  kawempe makindye   rubaga   nakawa     <NA>
##       88      156      152      168      236        0
```

```
table(pre.ph3$p1.p2.monitoring, useNA="always")
```

```
##
##   no  yes <NA>
##  536  264    0
```

```
pre.ph3$p1.p2.monitoring <- factor(pre.ph3$p1.p2.monitoring)
table(pre.ph3$p1.p2.monitoring, useNA="always")
```

```
##
##   no  yes <NA>
##  536  264    0
```

```
table(pre.ph3$b.type.site, useNA="always")
```

```
##
##          Household pile   Household sack or bin               Littering
##                      42                      17                      33
##                 Missing   Official dumping site                   Other
##                      45                      97                      11
##              Small pile Unofficial dumping site                    <NA>
##                      29                     500                      26
```

```
pre.ph3$b.type.site <- ifelse(is.na(pre.ph3$b.type.site),"Missing",pre.ph3$b.type.site)
pre.ph3$b.type.site <- factor(pre.ph3$b.type.site,
```

```
                                levels=c("Missing","Other","Household sack or bin","Household pile","Small
                                         "Littering","Official dumping site","Unofficial dumping site"))
table(pre.ph3$b.type.site, useNA="always")
```

```
##
##                 Missing                   Other   Household sack or bin
##                      71                      11                      17
##          Household pile              Small pile               Littering
##                      42                      29                      33
##   Official dumping site Unofficial dumping site                    <NA>
##                      97                     500                       0
```

```
table(pre.ph3$b.pile.area_dec, useNA="always")
```

```
##
##    1   10    2    3    4    5    6    7    8    9 <NA>
##   78   77   77   78   77   77   78   77   78   77   26
```

```
pre.ph3$b.pile.area_dec <- factor(pre.ph3$b.pile.area_dec, levels=c("1", "2", "3", "4", "5", "6", "7",
table(pre.ph3$b.pile.area_dec, useNA="always")
```

```
##
##    1    2    3    4    5    6    7    8    9   10 <NA>
##   78   77   78   77   77   78   77   78   77   77   26
```

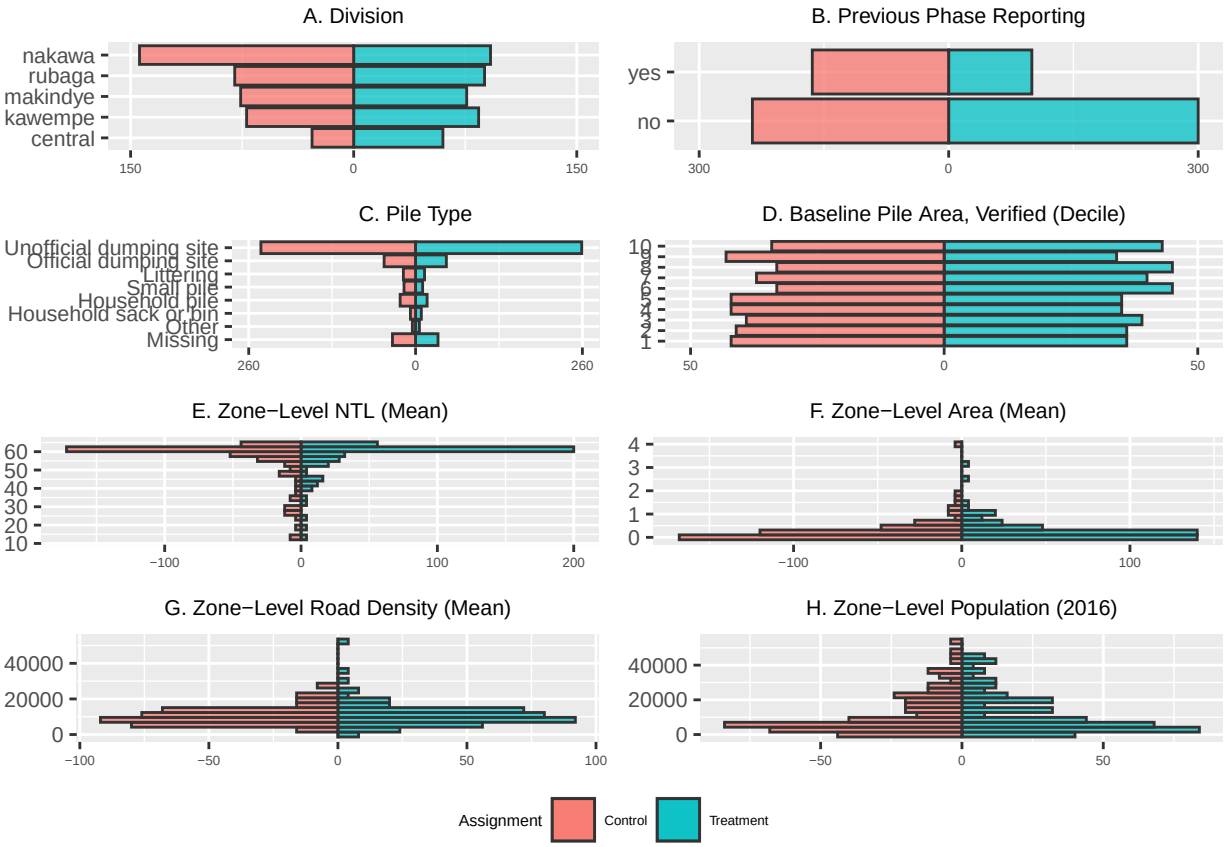
```
table(pre.ph3$treat,pre.ph3$div)
```

```
##
##     central kawempe makindye rubaga nakawa
##   0      28      72       76     80    144
##   1      60      84       76     88     92
```

```
plot.list <- list()
for(i in 4:7){
  p <- balance.plot_f(dta=pre.ph3, var=names(pre.ph3)[i], treat="treat", title=var.names[i])
  plot.list[[i]] <- p +
    scale_fill_discrete("Assignment", labels = c("Control", "Treatment")) +
    theme(text=element_text(size=6),
          plot.title = element_text(size=8))
}
```

```
## Warning: Ignoring unknown parameters: binwidth, bins, pad

## Warning: Ignoring unknown parameters: binwidth, bins, pad

## Warning: Ignoring unknown parameters: binwidth, bins, pad

## Warning: Ignoring unknown parameters: binwidth, bins, pad

## Warning: Ignoring unknown parameters: binwidth, bins, pad

## Warning: Ignoring unknown parameters: binwidth, bins, pad

## Warning: Ignoring unknown parameters: binwidth, bins, pad

## Warning: Ignoring unknown parameters: binwidth, bins, pad
```

```
for(i in 8:11){
  p <- balance.plot_n(dta=pre.ph3, var=names(pre.ph3)[i], treat="treat", title=var.names[i])
  plot.list[[i]] <-  p +
    scale_fill_discrete("Assignment", labels = c("Control", "Treatment")) +
    theme(text=element_text(size=6),
          plot.title = element_text(size=8))
}

plot.list <- plot.list[4:11]
balance_out <- ggarrange(plot.list[[1]], plot.list[[2]], plot.list[[3]], plot.list[[4]], plot.list[[5]]
                         plot.list[[6]], plot.list[[7]], plot.list[[8]], ncol=2, nrow=4, common.legend =
balance_out
```




```
# ggsave("./SI Figures/FigureJ1.tiff",
#        balance_out,
#        width=5, height=5, units="in")
# ggsave("./SI Figures/FigureJ1.png",
#        balance_out,
#        width=5, height=5, units="in")


##### TABLES ######
## Table 1 & J1: RI Results, Primary Outcomes ----
ri.cph3.analysis <- subsetB_dta

#Midline 1 Waste Accumulation
ri.m1.size <- lm.ri(formula = m1.pile.area_m~treat+p1.p2.monitoring+b.pile.area_m+lights.mean+road.dens
```

```
                    dta = ri.cph3.analysis,
                    treat.var = "treat",
                    clust_var = ri.cph3.analysis$zone.id,
                    m=length(unique(ri.cph3.analysis$zone.id))/2,
                    sims=10000)#This runs with NAs removed from the treat and cluster variables

ri.m1.size_new <- lm.ri(formula = m1.size_final~treat+p1.p2.monitoring+b.pile.area_m+lights.mean+road.d
                        dta = ri.cph3.analysis,
                        treat.var = "treat",
                        clust_var = ri.cph3.analysis$zone.id,
                        m=length(unique(ri.cph3.analysis$zone.id))/2,
                        sims=10000)#This runs with NAs removed from the treat and cluster variables


ri.m2.size <- lm.ri(formula = m2.pile.area_m~treat+p1.p2.monitoring+b.pile.area_m+lights.mean+road.densi
                    dta = ri.cph3.analysis,
                    treat.var = "treat",
                    clust_var = ri.cph3.analysis$zone.id,
                    m=length(unique(ri.cph3.analysis$zone.id))/2,
                    sims=10000)

ri.m2.size_new <- lm.ri(formula = m2.size_final~treat+p1.p2.monitoring+b.pile.area_m+lights.mean+road.d
                        dta = ri.cph3.analysis,
                        treat.var = "treat",
                        clust_var = ri.cph3.analysis$zone.id,
                        m=length(unique(ri.cph3.analysis$zone.id))/2,
                        sims=10000)


#Rank Test

ri.rank_m1 <- lm.ri(formula=rank.m1~treat+p1.p2.monitoring+rank.b+lights.mean+road.density+area.km2+div
                    dta=ri.cph3.analysis,
                    treat.var="treat",
                    clust_var = ri.cph3.analysis$zone.id,
                    m=length(unique(ri.cph3.analysis$zone.id))/2,
                    sims=10000)

ri.rank_m2 <- lm.ri(formula=rank.m2~treat+p1.p2.monitoring+rank.b+lights.mean+road.density+area.km2+div
                    dta=ri.cph3.analysis,
                    treat.var="treat",
                    clust_var = ri.cph3.analysis$zone.id,
                    m=length(unique(ri.cph3.analysis$zone.id))/2,
                    sims=10000)

ri.rank_m1_new <- lm.ri(formula=rank.m1_new~treat+p1.p2.monitoring+rank.b_new+lights.mean+road.density+
                        dta=ri.cph3.analysis,
                        treat.var="treat",
                        clust_var = ri.cph3.analysis$zone.id,
                        m=length(unique(ri.cph3.analysis$zone.id))/2,
                        sims=10000)

ri.rank_m2_new <- lm.ri(formula=rank.m2_new~treat+p1.p2.monitoring+rank.b_new+lights.mean+road.density+
```

```
                        dta=ri.cph3.analysis,
                        treat.var="treat",
                        clust_var = ri.cph3.analysis$zone.id,
                        m=length(unique(ri.cph3.analysis$zone.id))/2,
                        sims=10000)

ri.wpd1 <- lm.ri(formula = m1.waste.pile_d~treat+p1.p2.monitoring+lights.mean+road.density+div+area.km2
                 dta = ri.cph3.analysis,
                 treat.var = "treat",
                 clust_var = ri.cph3.analysis$zone.id,
                 m=length(unique(ri.cph3.analysis$zone.id))/2,
                 sims=10000)

ri.cph3.analysis$m1.waste.pile_d2 <- ifelse(ri.cph3.analysis$m1.waste.pile_d2=="Yes", 1, ri.cph3.analysi
ri.cph3.analysis$m1.waste.pile_d2 <- ifelse(ri.cph3.analysis$m1.waste.pile_d2=="No", 0, ri.cph3.analysis
ri.wpd1_adj <- lm.ri(formula = m1.waste.pile_d2~treat+p1.p2.monitoring+lights.mean+road.density+div+area
                     dta = ri.cph3.analysis,
                     treat.var = "treat",
                     clust_var = ri.cph3.analysis$zone.id,
                     m=length(unique(ri.cph3.analysis$zone.id))/2,
                     sims=10000)

ri.wpd2 <- lm.ri(formula = m2.waste.pile_d~treat+p1.p2.monitoring+lights.mean+road.density+div+area.km2
                 dta = ri.cph3.analysis,
                 treat.var = "treat",
                 clust_var = ri.cph3.analysis$zone.id,
                 m=length(unique(ri.cph3.analysis$zone.id))/2,
                 sims=10000)

ri.cph3.analysis$m2.waste.pile_d2 <- ifelse(ri.cph3.analysis$m2.waste.pile_d2=="Yes", 1, ri.cph3.analysi
ri.cph3.analysis$m2.waste.pile_d2 <- ifelse(ri.cph3.analysis$m2.waste.pile_d2=="No", 0, ri.cph3.analysis
ri.wpd2_adj <- lm.ri(formula = m2.waste.pile_d2~treat+p1.p2.monitoring+lights.mean+road.density+div+area
                     dta = ri.cph3.analysis,
                     treat.var = "treat",
                     clust_var = ri.cph3.analysis$zone.id,
                     m=length(unique(ri.cph3.analysis$zone.id))/2,
                     sims=10000)



reg.adj_ri.raw <- data.frame(matrix(nrow=5, ncol=6))
rownames(reg.adj_ri.raw) <- c("Audit","Treatment Effect","Standard Error", "p-value", "N")
colnames(reg.adj_ri.raw) <- c("M1.1", "M2.1", "M1.2", "M2.2", "M1.3", "M2.3")

reg.adj_ri.raw$M1.1[1] <- "M1"
reg.adj_ri.raw$M2.1[1] <- "M2"
reg.adj_ri.raw$M1.2[1] <- "M1"
reg.adj_ri.raw$M2.2[1] <- "M2"
reg.adj_ri.raw$M1.3[1] <- "M1"
reg.adj_ri.raw$M2.3[1] <- "M2"


reg.adj_ri.raw$M1.1[2] <- ri.m1.size$ate
```

```
reg.adj_ri.raw$M1.1[3] <- ri.m1.size$se
reg.adj_ri.raw$M1.1[4] <- ri.m1.size$p.one.way.lesser
reg.adj_ri.raw$M1.1[5] <- ri.m1.size$N

reg.adj_ri.raw$M2.1[2] <- ri.m2.size$ate
reg.adj_ri.raw$M2.1[3] <- ri.m2.size$se
reg.adj_ri.raw$M2.1[4] <- ri.m2.size$p.one.way.lesser
reg.adj_ri.raw$M2.1[5] <- ri.m2.size$N

reg.adj_ri.raw$M1.2[2] <- ri.wpd1$ate
reg.adj_ri.raw$M1.2[3] <- ri.wpd1$se
reg.adj_ri.raw$M1.2[4] <- ri.wpd1$p.one.way.lesser
reg.adj_ri.raw$M1.2[5] <- ri.wpd1$N

reg.adj_ri.raw$M2.2[2] <- ri.wpd2$ate
reg.adj_ri.raw$M2.2[3] <- ri.wpd2$se
reg.adj_ri.raw$M2.2[4] <- ri.wpd2$p.one.way.lesser
reg.adj_ri.raw$M2.2[5] <- ri.wpd2$N

reg.adj_ri.raw$M1.3[2] <- ri.rank_m1$ate
reg.adj_ri.raw$M1.3[3] <- ri.rank_m1$se
reg.adj_ri.raw$M1.3[4] <- ri.rank_m1$p.one.way.lesser
reg.adj_ri.raw$M1.3[5] <- ri.rank_m1$N

reg.adj_ri.raw$M2.3[2] <- ri.rank_m2$ate
reg.adj_ri.raw$M2.3[3] <- ri.rank_m2$se
reg.adj_ri.raw$M2.3[4] <- ri.rank_m2$p.one.way.lesser
reg.adj_ri.raw$M2.3[5] <- ri.rank_m2$N

reg.adj_ri.raw[2,] <- round(as.numeric(reg.adj_ri.raw[2,]), 3)
reg.adj_ri.raw[3,] <- round(as.numeric(reg.adj_ri.raw[3,]), 3)
reg.adj_ri.raw[4,] <- round(as.numeric(reg.adj_ri.raw[4,]), 3)
reg.adj_ri.raw[5,] <- round(as.numeric(reg.adj_ri.raw[5,]), 3)


colnames(reg.adj_ri.raw) <- c("Pile Size", "Pile Size",
                              "Pile Cleaned", "Pile Cleaned",
                              "Pile Rank", "Pile Rank")

reg.adj_ri.cln <- data.frame(matrix(nrow=5, ncol=6))
rownames(reg.adj_ri.cln) <- c("Audit", "Treatment Effect","Standard Error", "p-value", "N")
colnames(reg.adj_ri.cln) <- c("M1.1", "M2.1", "M1.2", "M2.2", "M1.3", "M2.3")

reg.adj_ri.cln$M1.1[1] <- "M1"
reg.adj_ri.cln$M1.2[1] <- "M1"
reg.adj_ri.cln$M1.3[1] <- "M1"
reg.adj_ri.cln$M2.1[1] <- "M2"
reg.adj_ri.cln$M2.2[1] <- "M2"
reg.adj_ri.cln$M2.3[1] <- "M2"

reg.adj_ri.cln$M1.1[2] <- ri.m1.size_new$ate
reg.adj_ri.cln$M1.1[3] <- ri.m1.size_new$se
reg.adj_ri.cln$M1.1[4] <- ri.m1.size_new$p.one.way.lesser
```

```r
reg.adj_ri.cln$M1.1[5] <- ri.m1.size_new$N

reg.adj_ri.cln$M2.1[2] <- ri.m2.size_new$ate
reg.adj_ri.cln$M2.1[3] <- ri.m2.size_new$se
reg.adj_ri.cln$M2.1[4] <- ri.m2.size_new$p.one.way.lesser
reg.adj_ri.cln$M2.1[5] <- ri.m2.size_new$N

reg.adj_ri.cln$M1.2[2] <- ri.wpd1_adj$ate
reg.adj_ri.cln$M1.2[3] <- ri.wpd1_adj$se
reg.adj_ri.cln$M1.2[4] <- ri.wpd1_adj$p.one.way.lesser
reg.adj_ri.cln$M1.2[5] <- ri.wpd1_adj$N

reg.adj_ri.cln$M2.2[2] <- ri.wpd2_adj$ate
reg.adj_ri.cln$M2.2[3] <- ri.wpd2_adj$se
reg.adj_ri.cln$M2.2[4] <- ri.wpd2_adj$p.one.way.lesser
reg.adj_ri.cln$M2.2[5] <- ri.wpd2_adj$N

reg.adj_ri.cln$M1.3[2] <- ri.rank_m1_new$ate
reg.adj_ri.cln$M1.3[3] <- ri.rank_m1_new$se
reg.adj_ri.cln$M1.3[4] <- ri.rank_m1_new$p.one.way.lesser
reg.adj_ri.cln$M1.3[5] <- ri.rank_m1_new$N

reg.adj_ri.cln$M2.3[2] <- ri.rank_m2_new$ate
reg.adj_ri.cln$M2.3[3] <- ri.rank_m2_new$se
reg.adj_ri.cln$M2.3[4] <- ri.rank_m2_new$p.one.way.lesser
reg.adj_ri.cln$M2.3[5] <- ri.rank_m2_new$N

reg.adj_ri.cln[2,] <- round(as.numeric(reg.adj_ri.cln[2,]), 3)
reg.adj_ri.cln[3,] <- round(as.numeric(reg.adj_ri.cln[3,]), 3)
reg.adj_ri.cln[4,] <- round(as.numeric(reg.adj_ri.cln[4,]), 3)
reg.adj_ri.cln[5,] <- round(as.numeric(reg.adj_ri.cln[5,]), 3)


colnames(reg.adj_ri.cln) <- c("Pile Size", "Pile Size",
                              "Pile Cleaned", "Pile Cleaned",
                              "Pile Rank", "Pile Rank")
#Table J1
out1 <- list(reg.adj_ri.raw)
attr(out1, "message") <- c("Note: results calculated using raw waste pile size measurements.")
print(xtableList(out1, caption="RI Results, Primary Dependent Variables (Raw)"), caption.placement="top
```

```
## % latex table generated in R 3.6.1 by xtable 1.8-4 package
## % Tue Jun  9 13:05:39 2020
## \begin{table}[ht]
## \centering
## \caption{RI Results, Primary Dependent Variables (Raw)}
## \begin{tabular}{rllllll}
##   \hline
##  & Pile Size & Pile Size & Pile Cleaned & Pile Cleaned & Pile Rank & Pile Rank \\
##   \hline
## Audit & M1 & M2 & M1 & M2 & M1 & M2 \\
##   Treatment Effect & 16.572 & -0.678 & -0.001 & 0.001 & -10.567 & -4.246 \\
##   Standard Error & 16.172 & 16.316 & 0.029 & 0.036 & 17.913 & 16.676 \\
##   p-value & 0.802 & 0.484 & 0.493 & 0.508 & 0.274 & 0.396 \\
```

```
##   N & 679 & 679 & 679 & 679 & 679 & 679 \\
##    \hline
## \multicolumn{7}{l}{Note: results calculated using raw waste pile size measurements.}\\
## \end{tabular}
## \end{table}
```

```
#Table 1
out2 <- list(reg.adj_ri.cln)
attr(out2, "message") <- c("Note: results calculated using cleaned waste pile size measurements.")
print(xtableList(out2, caption="RI Results, Primary Dependent Variables (Cleaned)"), caption.placement=
```

```
## % latex table generated in R 3.6.1 by xtable 1.8-4 package
## % Tue Jun  9 13:05:39 2020
## \begin{table}[ht]
## \centering
## \caption{RI Results, Primary Dependent Variables (Cleaned)}
## \begin{tabular}{rllllll}
##   \hline
##  & Pile Size & Pile Size & Pile Cleaned & Pile Cleaned & Pile Rank & Pile Rank \\
##   \hline
## Audit & M1 & M2 & M1 & M2 & M1 & M2 \\
##   Treatment Effect & -4.234 & -7.781 & -0.032 & -0.011 & -11.109 & -6.544 \\
##   Standard Error & 3.467 & 12.746 & 0.033 & 0.039 & 16.673 & 16.264 \\
##   p-value & 0.112 & 0.303 & 0.176 & 0.389 & 0.262 & 0.342 \\
##   N & 679 & 679 & 679 & 679 & 679 & 679 \\
##    \hline
## \multicolumn{7}{l}{Note: results calculated using cleaned waste pile size measurements.}\\
## \end{tabular}
## \end{table}
```

```
## Table 2 & J2: RI, Secondary Outcomes  -----
ri.cph3.analysis <- subsetB_dta

ri.uw1_m1 <- lm.ri(formula = m1.ea.uwa1~treat+p1.p2.monitoring+b.pile.area_m+lights.mean+road.density+di
                   dta = ri.cph3.analysis,
                   treat.var = "treat",
                   clust_var = ri.cph3.analysis$zone.id,
                   m=length(unique(ri.cph3.analysis$zone.id))/2,
                   sims=10000)
ri.uw1_m2 <- lm.ri(formula = m1.ea.uwa2~treat+p1.p2.monitoring+b.pile.area_m+lights.mean+road.density+di
                   dta = ri.cph3.analysis,
                   treat.var = "treat",
                   clust_var = ri.cph3.analysis$zone.id,
                   m=length(unique(ri.cph3.analysis$zone.id))/2,
                   sims=10000)

ri.uw2_m1 <- lm.ri(formula = m1.ea.uwa2~treat+p1.p2.monitoring+b.pile.area_m+lights.mean+road.density+di
                   dta = ri.cph3.analysis,
                   treat.var = "treat",
                   clust_var = ri.cph3.analysis$zone.id,
                   m=length(unique(ri.cph3.analysis$zone.id))/2,
                   sims=10000)
ri.uw2_m2 <- lm.ri(formula = m2.ea.uwa2~treat+p1.p2.monitoring+b.pile.area_m+lights.mean+road.density+di
                   dta = ri.cph3.analysis,
                   treat.var = "treat",
```

```
                    clust_var = ri.cph3.analysis$zone.id,
                    m=length(unique(ri.cph3.analysis$zone.id))/2,
                    sims=10000)

ri.po1_m1 <- lm.ri(formula = m1.ea.po1~treat+p1.p2.monitoring+b.pile.area_m+lights.mean+road.density+di
                    dta = ri.cph3.analysis,
                    treat.var = "treat",
                    clust_var = ri.cph3.analysis$zone.id,
                    m=length(unique(ri.cph3.analysis$zone.id))/2,
                    sims=10000)
ri.po1_m2 <- lm.ri(formula = m2.ea.po1~treat+p1.p2.monitoring+b.pile.area_m+lights.mean+road.density+di
                    dta = ri.cph3.analysis,
                    treat.var = "treat",
                    clust_var = ri.cph3.analysis$zone.id,
                    m=length(unique(ri.cph3.analysis$zone.id))/2,
                    sims=10000)

ri.po2_m1 <- lm.ri(formula = m1.ea.po2~treat+p1.p2.monitoring+b.pile.area_m+lights.mean+road.density+di
                    dta = ri.cph3.analysis,
                    treat.var = "treat",
                    clust_var = ri.cph3.analysis$zone.id,
                    m=length(unique(ri.cph3.analysis$zone.id))/2,
                    sims=10000)
ri.po2_m2 <- lm.ri(formula = m2.ea.po2~treat+p1.p2.monitoring+b.pile.area_m+lights.mean+road.density+di
                    dta = ri.cph3.analysis,
                    treat.var = "treat",
                    clust_var = ri.cph3.analysis$zone.id,
                    m=length(unique(ri.cph3.analysis$zone.id))/2,
                    sims=10000)

ri.brn1_m1 <- lm.ri(formula = m1.ea.brn1~treat+p1.p2.monitoring+b.pile.area_m+lights.mean+road.density+
                     dta = ri.cph3.analysis,
                     treat.var = "treat",
                     clust_var = ri.cph3.analysis$zone.id,
                     m=length(unique(ri.cph3.analysis$zone.id))/2,
                     sims=10000)
ri.brn1_m2 <- lm.ri(formula = m2.ea.brn1~treat+p1.p2.monitoring+b.pile.area_m+lights.mean+road.density+
                     dta = ri.cph3.analysis,
                     treat.var = "treat",
                     clust_var = ri.cph3.analysis$zone.id,
                     m=length(unique(ri.cph3.analysis$zone.id))/2,
                     sims=10000)

ri.brn2_m1 <- lm.ri(formula = m1.ea.brn2~treat+p1.p2.monitoring+b.pile.area_m+lights.mean+road.density+
                     dta = ri.cph3.analysis,
                     treat.var = "treat",
                     clust_var = ri.cph3.analysis$zone.id,
                     m=length(unique(ri.cph3.analysis$zone.id))/2,
                     sims=10000)
ri.brn2_m2 <- lm.ri(formula = m2.ea.brn2~treat+p1.p2.monitoring+b.pile.area_m+lights.mean+road.density+
                     dta = ri.cph3.analysis,
                     treat.var = "treat",
                     clust_var = ri.cph3.analysis$zone.id,
```

```
                    m=length(unique(ri.cph3.analysis$zone.id))/2,
                    sims=10000)

ri.uw1_m1.cln <- lm.ri(formula = m1.ea.uwa1.cln~treat+p1.p2.monitoring+b.pile.area_m+lights.mean+road.d
                       dta = ri.cph3.analysis,
                       treat.var = "treat",
                       clust_var = ri.cph3.analysis$zone.id,
                       m=length(unique(ri.cph3.analysis$zone.id))/2,
                       sims=10000)
ri.uw1_m2.cln <- lm.ri(formula = m1.ea.uwa2.cln~treat+p1.p2.monitoring+b.pile.area_m+lights.mean+road.d
                       dta = ri.cph3.analysis,
                       treat.var = "treat",
                       clust_var = ri.cph3.analysis$zone.id,
                       m=length(unique(ri.cph3.analysis$zone.id))/2,
                       sims=10000)

ri.uw2_m1.cln <- lm.ri(formula = m1.ea.uwa2.cln~treat+p1.p2.monitoring+b.pile.area_m+lights.mean+road.d
                       dta = ri.cph3.analysis,
                       treat.var = "treat",
                       clust_var = ri.cph3.analysis$zone.id,
                       m=length(unique(ri.cph3.analysis$zone.id))/2,
                       sims=10000)
ri.uw2_m2.cln <- lm.ri(formula = m2.ea.uwa2.cln~treat+p1.p2.monitoring+b.pile.area_m+lights.mean+road.d
                       dta = ri.cph3.analysis,
                       treat.var = "treat",
                       clust_var = ri.cph3.analysis$zone.id,
                       m=length(unique(ri.cph3.analysis$zone.id))/2,
                       sims=10000)

ri.po1_m1.cln <- lm.ri(formula = m1.ea.po1.cln~treat+p1.p2.monitoring+b.pile.area_m+lights.mean+road.de
                       dta = ri.cph3.analysis,
                       treat.var = "treat",
                       clust_var = ri.cph3.analysis$zone.id,
                       m=length(unique(ri.cph3.analysis$zone.id))/2,
                       sims=10000)
ri.po1_m2.cln <- lm.ri(formula = m2.ea.po1.cln~treat+p1.p2.monitoring+b.pile.area_m+lights.mean+road.de
                       dta = ri.cph3.analysis,
                       treat.var = "treat",
                       clust_var = ri.cph3.analysis$zone.id,
                       m=length(unique(ri.cph3.analysis$zone.id))/2,
                       sims=10000)

ri.po2_m1.cln <- lm.ri(formula = m1.ea.po2.cln~treat+p1.p2.monitoring+b.pile.area_m+lights.mean+road.de
                       dta = ri.cph3.analysis,
                       treat.var = "treat",
                       clust_var = ri.cph3.analysis$zone.id,
                       m=length(unique(ri.cph3.analysis$zone.id))/2,
                       sims=10000)
ri.po2_m2.cln <- lm.ri(formula = m2.ea.po2.cln~treat+p1.p2.monitoring+b.pile.area_m+lights.mean+road.de
                       dta = ri.cph3.analysis,
                       treat.var = "treat",
                       clust_var = ri.cph3.analysis$zone.id,
                       m=length(unique(ri.cph3.analysis$zone.id))/2,
```

```
                        sims=10000)

ri.brn1_m1.cln <- lm.ri(formula = m1.ea.brn1.cln~treat+p1.p2.monitoring+b.pile.area_m+lights.mean+road.
                        dta = ri.cph3.analysis,
                        treat.var = "treat",
                        clust_var = ri.cph3.analysis$zone.id,
                        m=length(unique(ri.cph3.analysis$zone.id))/2,
                        sims=10000)
ri.brn1_m2.cln <- lm.ri(formula = m2.ea.brn1.cln~treat+p1.p2.monitoring+b.pile.area_m+lights.mean+road.
                        dta = ri.cph3.analysis,
                        treat.var = "treat",
                        clust_var = ri.cph3.analysis$zone.id,
                        m=length(unique(ri.cph3.analysis$zone.id))/2,
                        sims=10000)

ri.brn2_m1.cln <- lm.ri(formula = m1.ea.brn2.cln~treat+p1.p2.monitoring+b.pile.area_m+lights.mean+road.
                        dta = ri.cph3.analysis,
                        treat.var = "treat",
                        clust_var = ri.cph3.analysis$zone.id,
                        m=length(unique(ri.cph3.analysis$zone.id))/2,
                        sims=10000)
ri.brn2_m2.cln <- lm.ri(formula = m2.ea.brn2.cln~treat+p1.p2.monitoring+b.pile.area_m+lights.mean+road.
                        dta = ri.cph3.analysis,
                        treat.var = "treat",
                        clust_var = ri.cph3.analysis$zone.id,
                        m=length(unique(ri.cph3.analysis$zone.id))/2,
                        sims=10000)

reg.adj_ri.alt.dvs_m1 <- data.frame(matrix(nrow=5, ncol=6))
rownames(reg.adj_ri.alt.dvs_m1) <- c("Variable Specification", "Treatment Effect","Standard Error", "p-
colnames(reg.adj_ri.alt.dvs_m1) <- c("M1.1", "M2.1", "M1.2", "M2.2", "M1.3", "M2.3")

reg.adj_ri.alt.dvs_m1$M1.1[1] <- "A"
reg.adj_ri.alt.dvs_m1$M1.2[1] <- "A"
reg.adj_ri.alt.dvs_m1$M1.3[1] <- "A"
reg.adj_ri.alt.dvs_m1$M2.1[1] <- "B"
reg.adj_ri.alt.dvs_m1$M2.2[1] <- "B"
reg.adj_ri.alt.dvs_m1$M2.3[1] <- "B"

reg.adj_ri.alt.dvs_m1$M1.1[2] <- ri.uw1_m1$ate
reg.adj_ri.alt.dvs_m1$M1.1[3] <- ri.uw1_m1$se
reg.adj_ri.alt.dvs_m1$M1.1[4] <- ri.uw1_m1$p.one.way.lesser
reg.adj_ri.alt.dvs_m1$M1.1[5] <- ri.uw1_m1$N

reg.adj_ri.alt.dvs_m1$M2.1[2] <- ri.uw2_m1$ate
reg.adj_ri.alt.dvs_m1$M2.1[3] <- ri.uw2_m1$se
reg.adj_ri.alt.dvs_m1$M2.1[4] <- ri.uw2_m1$p.one.way.lesser
reg.adj_ri.alt.dvs_m1$M2.1[5] <- ri.uw2_m1$N

reg.adj_ri.alt.dvs_m1$M1.2[2] <- ri.po1_m1$ate
reg.adj_ri.alt.dvs_m1$M1.2[3] <- ri.po1_m1$se
reg.adj_ri.alt.dvs_m1$M1.2[4] <- ri.po1_m1$p.one.way.lesser
reg.adj_ri.alt.dvs_m1$M1.2[5] <- ri.po1_m1$N
```

```
reg.adj_ri.alt.dvs_m1$M2.2[2] <- ri.po2_m1$ate
reg.adj_ri.alt.dvs_m1$M2.2[3] <- ri.po2_m1$se
reg.adj_ri.alt.dvs_m1$M2.2[4] <- ri.po2_m1$p.one.way.lesser
reg.adj_ri.alt.dvs_m1$M2.2[5] <- ri.po2_m1$N

reg.adj_ri.alt.dvs_m1$M1.3[2] <- ri.brn1_m1$ate
reg.adj_ri.alt.dvs_m1$M1.3[3] <- ri.brn1_m1$se
reg.adj_ri.alt.dvs_m1$M1.3[4] <- ri.brn1_m1$p.one.way.lesser
reg.adj_ri.alt.dvs_m1$M1.3[5] <- ri.brn1_m1$N

reg.adj_ri.alt.dvs_m1$M2.3[2] <- ri.brn2_m1$ate
reg.adj_ri.alt.dvs_m1$M2.3[3] <- ri.brn2_m1$se
reg.adj_ri.alt.dvs_m1$M2.3[4] <- ri.brn2_m1$p.one.way.lesser
reg.adj_ri.alt.dvs_m1$M2.3[5] <- ri.brn2_m1$N

reg.adj_ri.alt.dvs_m1[2,] <- round(as.numeric(reg.adj_ri.alt.dvs_m1[2,]), 3)
reg.adj_ri.alt.dvs_m1[3,] <- round(as.numeric(reg.adj_ri.alt.dvs_m1[3,]), 3)
reg.adj_ri.alt.dvs_m1[4,] <- round(as.numeric(reg.adj_ri.alt.dvs_m1[4,]), 3)
reg.adj_ri.alt.dvs_m1[5,] <- round(as.numeric(reg.adj_ri.alt.dvs_m1[5,]), 3)


colnames(reg.adj_ri.alt.dvs_m1) <- c("Uncontained", "Uncontained",
                                     "Disorganized", "Disorganized",
                                     "Burnt", "Burnt")

reg.adj_ri.alt.dvs.cln_m1 <- data.frame(matrix(nrow=5, ncol=6))
rownames(reg.adj_ri.alt.dvs.cln_m1) <- c("Variable Specification", "Treatment Effect","Standard Error",
colnames(reg.adj_ri.alt.dvs.cln_m1) <- c("M1.1", "M2.1", "M1.2", "M2.2", "M1.3", "M2.3")

reg.adj_ri.alt.dvs.cln_m1$M1.1[1] <- "A"
reg.adj_ri.alt.dvs.cln_m1$M1.2[1] <- "A"
reg.adj_ri.alt.dvs.cln_m1$M1.3[1] <- "A"
reg.adj_ri.alt.dvs.cln_m1$M2.1[1] <- "B"
reg.adj_ri.alt.dvs.cln_m1$M2.2[1] <- "B"
reg.adj_ri.alt.dvs.cln_m1$M2.3[1] <- "B"

reg.adj_ri.alt.dvs.cln_m1$M1.1[2] <- ri.uw1_m1.cln$ate
reg.adj_ri.alt.dvs.cln_m1$M1.1[3] <- ri.uw1_m1.cln$se
reg.adj_ri.alt.dvs.cln_m1$M1.1[4] <- ri.uw1_m1.cln$p.one.way.lesser
reg.adj_ri.alt.dvs.cln_m1$M1.1[5] <- ri.uw1_m1.cln$N

reg.adj_ri.alt.dvs.cln_m1$M2.1[2] <- ri.uw2_m1.cln$ate
reg.adj_ri.alt.dvs.cln_m1$M2.1[3] <- ri.uw2_m1.cln$se
reg.adj_ri.alt.dvs.cln_m1$M2.1[4] <- ri.uw2_m1.cln$p.one.way.lesser
reg.adj_ri.alt.dvs.cln_m1$M2.1[5] <- ri.uw2_m1.cln$N

reg.adj_ri.alt.dvs.cln_m1$M1.2[2] <- ri.po1_m1.cln$ate
reg.adj_ri.alt.dvs.cln_m1$M1.2[3] <- ri.po1_m1.cln$se
reg.adj_ri.alt.dvs.cln_m1$M1.2[4] <- ri.po1_m1.cln$p.one.way.lesser
reg.adj_ri.alt.dvs.cln_m1$M1.2[5] <- ri.po1_m1.cln$N

reg.adj_ri.alt.dvs.cln_m1$M2.2[2] <- ri.po2_m1.cln$ate
reg.adj_ri.alt.dvs.cln_m1$M2.2[3] <- ri.po2_m1.cln$se
```

```
reg.adj_ri.alt.dvs.cln_m1$M2.2[4] <- ri.po2_m1.cln$p.one.way.lesser
reg.adj_ri.alt.dvs.cln_m1$M2.2[5] <- ri.po2_m1.cln$N

reg.adj_ri.alt.dvs.cln_m1$M1.3[2] <- ri.brn1_m1.cln$ate
reg.adj_ri.alt.dvs.cln_m1$M1.3[3] <- ri.brn1_m1.cln$se
reg.adj_ri.alt.dvs.cln_m1$M1.3[4] <- ri.brn1_m1.cln$p.one.way.lesser
reg.adj_ri.alt.dvs.cln_m1$M1.3[5] <- ri.brn1_m1.cln$N

reg.adj_ri.alt.dvs.cln_m1$M2.3[2] <- ri.brn2_m1.cln$ate
reg.adj_ri.alt.dvs.cln_m1$M2.3[3] <- ri.brn2_m1.cln$se
reg.adj_ri.alt.dvs.cln_m1$M2.3[4] <- ri.brn2_m1.cln$p.one.way.lesser
reg.adj_ri.alt.dvs.cln_m1$M2.3[5] <- ri.brn2_m1.cln$N

reg.adj_ri.alt.dvs.cln_m1[2,] <- round(as.numeric(reg.adj_ri.alt.dvs.cln_m1[2,]),3)
reg.adj_ri.alt.dvs.cln_m1[3,] <- round(as.numeric(reg.adj_ri.alt.dvs.cln_m1[3,]),3)
reg.adj_ri.alt.dvs.cln_m1[4,] <- round(as.numeric(reg.adj_ri.alt.dvs.cln_m1[4,]),3)
reg.adj_ri.alt.dvs.cln_m1[5,] <- round(as.numeric(reg.adj_ri.alt.dvs.cln_m1[5,]),3)

#Benjamini-Hochberg-Yekutieli correction, as pre-specified
reg.adj_ri.alt.dvs.cln_m1[6,] <- round(p.adjust(reg.adj_ri.alt.dvs.cln_m1[4,], method="BH"), 3)

reg.adj_ri.alt.dvs.cln_m1 <- reg.adj_ri.alt.dvs.cln_m1[c(1,2,3,4,6,5),]
row.names(reg.adj_ri.alt.dvs.cln_m1)[5] <- "BHY p-value"

colnames(reg.adj_ri.alt.dvs.cln_m1) <- c("Uncontained", "Uncontained",
                                         "Disorganized", "Disorganized",
                                         "Burnt", "Burnt")

#Table J2a
out3 <- list(reg.adj_ri.alt.dvs_m1)
attr(out3, "message") <- c("Note: results calculated using raw waste pile size measurements.")
print(xtableList(out3, caption="RI Results, Midline 1, Secondary Dependent Variables (Raw)"), digits=3,

## % latex table generated in R 3.6.1 by xtable 1.8-4 package
## % Tue Jun  9 13:15:04 2020
## \begin{table}[ht]
## \centering
## \caption{RI Results, Midline 1, Secondary Dependent Variables (Raw)}
## \begin{tabular}{rllllll}
##   \hline
##  & Uncontained & Uncontained & Disorganized & Disorganized & Burnt & Burnt \\
##   \hline
## Variable Specification & A & B & A & B & A & B \\
##   Treatment Effect & 16.347 & 16.296 & 16.355 & 16.344 & -2.211 & -2.276 \\
##   Standard Error & 16.067 & 16.122 & 16.13 & 16.243 & 0.863 & 0.999 \\
##   p-value & 0.796 & 0.79 & 0.794 & 0.79 & 0.005 & 0.011 \\
##   N & 679 & 679 & 679 & 679 & 679 & 679 \\
##    \hline
## \multicolumn{7}{l}{Note: results calculated using raw waste pile size measurements.}\\
## \end{tabular}
## \end{table}
#Table 2a
out4 <- list(reg.adj_ri.alt.dvs.cln_m1)
```

```r
attr(out4, "message") <- c("Note: results calculated using cleaned waste pile size measurements.")
print(xtableList(out4, caption="RI Results, Midline 1, Secondary Dependent Variables (Cleaned)"), digits
```

```
## % latex table generated in R 3.6.1 by xtable 1.8-4 package
## % Tue Jun  9 13:15:04 2020
## \begin{table}[ht]
## \centering
## \caption{RI Results, Midline 1, Secondary Dependent Variables (Cleaned)}
## \begin{tabular}{rllllll}
##   \hline
##  & Uncontained & Uncontained & Disorganized & Disorganized & Burnt & Burnt \\
##   \hline
## Variable Specification & A & B & A & B & A & B \\
##   Treatment Effect & -4.444 & -4.521 & -4.536 & -4.546 & -2.298 & -2.495 \\
##   Standard Error & 3.338 & 3.368 & 3.428 & 3.415 & 0.768 & 0.859 \\
##   p-value & 0.094 & 0.092 & 0.096 & 0.094 & 0.001 & 0.001 \\
##   BHY p-value & 0.096 & 0.096 & 0.096 & 0.096 & 0.003 & 0.003 \\
##   N & 679 & 679 & 679 & 679 & 679 & 679 \\
##    \hline
## \multicolumn{7}{l}{Note: results calculated using cleaned waste pile size measurements.}\\
## \end{tabular}
## \end{table}
```

```r
reg.adj_ri.alt.dvs_m2 <- data.frame(matrix(nrow=5, ncol=6))
rownames(reg.adj_ri.alt.dvs_m2) <- c("Variable Specification", "Treatment Effect","Standard Error", "p-
colnames(reg.adj_ri.alt.dvs_m2) <- c("M1.1", "M2.1", "M1.2", "M2.2", "M1.3", "M2.3")

reg.adj_ri.alt.dvs_m2$M1.1[1] <- "A"
reg.adj_ri.alt.dvs_m2$M1.2[1] <- "A"
reg.adj_ri.alt.dvs_m2$M1.3[1] <- "A"
reg.adj_ri.alt.dvs_m2$M2.1[1] <- "B"
reg.adj_ri.alt.dvs_m2$M2.2[1] <- "B"
reg.adj_ri.alt.dvs_m2$M2.3[1] <- "B"

reg.adj_ri.alt.dvs_m2$M1.1[2] <- ri.uw1_m2$ate
reg.adj_ri.alt.dvs_m2$M1.1[3] <- ri.uw1_m2$se
reg.adj_ri.alt.dvs_m2$M1.1[4] <- ri.uw1_m2$p.one.way.lesser
reg.adj_ri.alt.dvs_m2$M1.1[5] <- ri.uw1_m2$N

reg.adj_ri.alt.dvs_m2$M2.1[2] <- ri.uw2_m2$ate
reg.adj_ri.alt.dvs_m2$M2.1[3] <- ri.uw2_m2$se
reg.adj_ri.alt.dvs_m2$M2.1[4] <- ri.uw2_m2$p.one.way.lesser
reg.adj_ri.alt.dvs_m2$M2.1[5] <- ri.uw2_m2$N

reg.adj_ri.alt.dvs_m2$M1.2[2] <- ri.po1_m2$ate
reg.adj_ri.alt.dvs_m2$M1.2[3] <- ri.po1_m2$se
reg.adj_ri.alt.dvs_m2$M1.2[4] <- ri.po1_m2$p.one.way.lesser
reg.adj_ri.alt.dvs_m2$M1.2[5] <- ri.po1_m2$N

reg.adj_ri.alt.dvs_m2$M2.2[2] <- ri.po2_m2$ate
reg.adj_ri.alt.dvs_m2$M2.2[3] <- ri.po2_m2$se
reg.adj_ri.alt.dvs_m2$M2.2[4] <- ri.po2_m2$p.one.way.lesser
reg.adj_ri.alt.dvs_m2$M2.2[5] <- ri.po2_m2$N
```

```
reg.adj_ri.alt.dvs_m2$M1.3[2] <- ri.brn1_m2$ate
reg.adj_ri.alt.dvs_m2$M1.3[3] <- ri.brn1_m2$se
reg.adj_ri.alt.dvs_m2$M1.3[4] <- ri.brn1_m2$p.one.way.lesser
reg.adj_ri.alt.dvs_m2$M1.3[5] <- ri.brn1_m2$N

reg.adj_ri.alt.dvs_m2$M2.3[2] <- ri.brn2_m2$ate
reg.adj_ri.alt.dvs_m2$M2.3[3] <- ri.brn2_m2$se
reg.adj_ri.alt.dvs_m2$M2.3[4] <- ri.brn2_m2$p.one.way.lesser
reg.adj_ri.alt.dvs_m2$M2.3[5] <- ri.brn2_m2$N

reg.adj_ri.alt.dvs_m2[2,] <- round(as.numeric(reg.adj_ri.alt.dvs_m2[2,]), 3)
reg.adj_ri.alt.dvs_m2[3,] <- round(as.numeric(reg.adj_ri.alt.dvs_m2[3,]), 3)
reg.adj_ri.alt.dvs_m2[4,] <- round(as.numeric(reg.adj_ri.alt.dvs_m2[4,]), 3)
reg.adj_ri.alt.dvs_m2[5,] <- round(as.numeric(reg.adj_ri.alt.dvs_m2[5,]), 3)


colnames(reg.adj_ri.alt.dvs_m2) <- c("Uncontained", "Uncontained",
                                     "Disorganized", "Disorganized",
                                     "Burnt", "Burnt")

reg.adj_ri.alt.dvs.cln_m2 <- data.frame(matrix(nrow=5, ncol=6))
rownames(reg.adj_ri.alt.dvs.cln_m2) <- c("Variable Specification", "Treatment Effect","Standard Error",
colnames(reg.adj_ri.alt.dvs.cln_m2) <- c("M1.1", "M2.1", "M1.2", "M2.2", "M1.3", "M2.3")

reg.adj_ri.alt.dvs.cln_m2$M1.1[1] <- "A"
reg.adj_ri.alt.dvs.cln_m2$M1.2[1] <- "A"
reg.adj_ri.alt.dvs.cln_m2$M1.3[1] <- "A"
reg.adj_ri.alt.dvs.cln_m2$M2.1[1] <- "B"
reg.adj_ri.alt.dvs.cln_m2$M2.2[1] <- "B"
reg.adj_ri.alt.dvs.cln_m2$M2.3[1] <- "B"

reg.adj_ri.alt.dvs.cln_m2$M1.1[2] <- ri.uw1_m2.cln$ate
reg.adj_ri.alt.dvs.cln_m2$M1.1[3] <- ri.uw1_m2.cln$se
reg.adj_ri.alt.dvs.cln_m2$M1.1[4] <- ri.uw1_m2.cln$p.one.way.lesser
reg.adj_ri.alt.dvs.cln_m2$M1.1[5] <- ri.uw1_m2.cln$N

reg.adj_ri.alt.dvs.cln_m2$M2.1[2] <- ri.uw2_m2.cln$ate
reg.adj_ri.alt.dvs.cln_m2$M2.1[3] <- ri.uw2_m2.cln$se
reg.adj_ri.alt.dvs.cln_m2$M2.1[4] <- ri.uw2_m2.cln$p.one.way.lesser
reg.adj_ri.alt.dvs.cln_m2$M2.1[5] <- ri.uw2_m2.cln$N

reg.adj_ri.alt.dvs.cln_m2$M1.2[2] <- ri.po1_m2.cln$ate
reg.adj_ri.alt.dvs.cln_m2$M1.2[3] <- ri.po1_m2.cln$se
reg.adj_ri.alt.dvs.cln_m2$M1.2[4] <- ri.po1_m2.cln$p.one.way.lesser
reg.adj_ri.alt.dvs.cln_m2$M1.2[5] <- ri.po1_m2.cln$N

reg.adj_ri.alt.dvs.cln_m2$M2.2[2] <- ri.po2_m2.cln$ate
reg.adj_ri.alt.dvs.cln_m2$M2.2[3] <- ri.po2_m2.cln$se
reg.adj_ri.alt.dvs.cln_m2$M2.2[4] <- ri.po2_m2.cln$p.one.way.lesser
reg.adj_ri.alt.dvs.cln_m2$M2.2[5] <- ri.po2_m2.cln$N

reg.adj_ri.alt.dvs.cln_m2$M1.3[2] <- ri.brn1_m2.cln$ate
reg.adj_ri.alt.dvs.cln_m2$M1.3[3] <- ri.brn1_m2.cln$se
```

```
reg.adj_ri.alt.dvs.cln_m2$M1.3[4] <- ri.brn1_m2.cln$p.one.way.lesser
reg.adj_ri.alt.dvs.cln_m2$M1.3[5] <- ri.brn1_m2.cln$N

reg.adj_ri.alt.dvs.cln_m2$M2.3[2] <- ri.brn2_m2.cln$ate
reg.adj_ri.alt.dvs.cln_m2$M2.3[3] <- ri.brn2_m2.cln$se
reg.adj_ri.alt.dvs.cln_m2$M2.3[4] <- ri.brn2_m2.cln$p.one.way.lesser
reg.adj_ri.alt.dvs.cln_m2$M2.3[5] <- ri.brn2_m2.cln$N

reg.adj_ri.alt.dvs.cln_m2[2,] <- round(as.numeric(reg.adj_ri.alt.dvs.cln_m2[2,]),3)
reg.adj_ri.alt.dvs.cln_m2[3,] <- round(as.numeric(reg.adj_ri.alt.dvs.cln_m2[3,]),3)
reg.adj_ri.alt.dvs.cln_m2[4,] <- round(as.numeric(reg.adj_ri.alt.dvs.cln_m2[4,]),3)
reg.adj_ri.alt.dvs.cln_m2[5,] <- round(as.numeric(reg.adj_ri.alt.dvs.cln_m2[5,]),3)

#Benjamini-Hochberg-Yekutieli correction, as pre-specified
reg.adj_ri.alt.dvs.cln_m2[6,] <- round(p.adjust(reg.adj_ri.alt.dvs.cln_m2[4,], method="BH"),3)

reg.adj_ri.alt.dvs.cln_m2 <- reg.adj_ri.alt.dvs.cln_m2[c(1,2,3,4,6,5),]
row.names(reg.adj_ri.alt.dvs.cln_m2)[5] <- "BHY p-value"

colnames(reg.adj_ri.alt.dvs.cln_m2) <- c("Uncontained", "Uncontained",
                                         "Disorganized", "Disorganized",
                                         "Burnt", "Burnt")

#Table J2b
out3 <- list(reg.adj_ri.alt.dvs_m2)
attr(out3, "message") <- c("Note: results calculated using raw waste pile size measurements.")
print(xtableList(out3, caption="RI Results, Midline 2, Secondary Dependent Variables (Raw)"), digits=3,
```

```
## % latex table generated in R 3.6.1 by xtable 1.8-4 package
## % Tue Jun  9 13:15:05 2020
## \begin{table}[ht]
## \centering
## \caption{RI Results, Midline 2, Secondary Dependent Variables (Raw)}
## \begin{tabular}{rllllll}
##   \hline
##  & Uncontained & Uncontained & Disorganized & Disorganized & Burnt & Burnt \\
##   \hline
## Variable Specification & A & B & A & B & A & B \\
##   Treatment Effect & 16.296 & -0.683 & -0.337 & -0.267 & 0.479 & -1.23 \\
##   Standard Error & 16.141 & 14.534 & 16.403 & 16.469 & 3.253 & 3.259 \\
##   p-value & 0.792 & 0.49 & 0.492 & 0.493 & 0.549 & 0.366 \\
##   N & 679 & 679 & 679 & 679 & 679 & 679 \\
##    \hline
## \multicolumn{7}{l}{Note: results calculated using raw waste pile size measurements.}\\
## \end{tabular}
## \end{table}
```

```
#Table 2b
out4 <- list(reg.adj_ri.alt.dvs.cln_m2)
attr(out4, "message") <- c("Note: results calculated using cleaned waste pile size measurements.")
print(xtableList(out4, caption="RI Results, Midline 2, Secondary Dependent Variables (Cleaned)"), digits
```

```
## % latex table generated in R 3.6.1 by xtable 1.8-4 package
## % Tue Jun  9 13:15:05 2020
```

```
## \begin{table}[ht]
## \centering
## \caption{RI Results, Midline 2, Secondary Dependent Variables (Cleaned)}
## \begin{tabular}{rllllll}
##   \hline
##  & Uncontained & Uncontained & Disorganized & Disorganized & Burnt & Burnt \\
##   \hline
## Variable Specification & A & B & A & B & A & B \\
##   Treatment Effect & -4.521 & -7.662 & -7.337 & -7.216 & -2.002 & -3.097 \\
##   Standard Error & 3.354 & 10.499 & 12.753 & 12.798 & 1.931 & 2.418 \\
##   p-value & 0.094 & 0.284 & 0.316 & 0.319 & 0.174 & 0.123 \\
##   BHY p-value & 0.319 & 0.319 & 0.319 & 0.319 & 0.319 & 0.319 \\
##   N & 679 & 679 & 679 & 679 & 679 & 679 \\
##    \hline
## \multicolumn{7}{l}{Note: results calculated using cleaned waste pile size measurements.}\\
## \end{tabular}
## \end{table}
```

```
## Table F1: Scalar Values, Alternative Dependent Variables -----
main_dta$m1.ea.uwa1.p <- NA
main_dta$m1.ea.uwa1.p <- ifelse(main_dta$m1.waste.pile_d == "No",
                                0.0, main_dta$m1.ea.uwa1)
main_dta$m1.ea.uwa1.p <- ifelse(main_dta$m1.waste.stor == "All of the rubbish is neatly contained with s
                                0.0, main_dta$m1.ea.uwa1)
main_dta$m1.ea.uwa1.p <- ifelse(main_dta$m1.waste.stor == "Most of the rubbish is organized in sacks or
                                0.333, main_dta$m1.ea.uwa1)
main_dta$m1.ea.uwa1.p <- ifelse(main_dta$m1.waste.stor == "Very little rubbish is contained within sacks
                                0.667, main_dta$m1.ea.uwa1)
main_dta$m1.ea.uwa1.p <- ifelse(main_dta$m1.waste.stor == "No rubbish is contained in sacks or containe
                                1.0, main_dta$m1.ea.uwa1)
tab_a <- data.frame(Characteristic=c(rep("Waste Containment", 5), rep("Waste Organization", 5), rep("Was
                    `Abbreviated Response`=c("Pile Cleaned", "Fully Contained", "Mostly Contained", "Mos
                                             "Pile Cleaned", "Fully Organized", "Mostly Organized", "Mos
                                             "Pile Cleaned", "No Burning", "Less than 50% burnt", "More
                    Scalar=c(0, 0, 0.333, 0.667, 1,
                             0, 0, 0.333, 0.667, 1,
                             0, 0, 0.45, 0.55))

tab_b <- data.frame(Characteristic=c(rep("Waste Containment", 5), rep("Waste Organization", 5), rep("Was
                    `Abbreviated Response`=c("Pile Cleaned", "Fully Contained", "Mostly Contained", "Mos
                                             "Pile Cleaned", "Fully Organized", "Mostly Organized", "Mos
                                             "Pile Cleaned", "No Burning", "Less than 50% burnt", "More
                    Scalar=c(0, 0, 0.25, 0.75, 1,
                             0, 0, 0.25, 0.75, 1,
                             0, 0, 0.33, 0.667))

print(xtable(tab_a, "Scalar Values, Variable Specification A"), include.rownames=F)
```

```
## % latex table generated in R 3.6.1 by xtable 1.8-4 package
## % Tue Jun  9 13:15:05 2020
## \begin{table}[ht]
## \centering
## \begin{tabular}{llr}
##   \hline
## Characteristic & Abbreviated.Response & Scalar \\
```

```
##   \hline
## Waste Containment & Pile Cleaned & 0.00 \\
##   Waste Containment & Fully Contained & 0.00 \\
##   Waste Containment & Mostly Contained & 0.33 \\
##   Waste Containment & Mostly Uncontained & 0.67 \\
##   Waste Containment & Uncontained & 1.00 \\
##   Waste Organization & Pile Cleaned & 0.00 \\
##   Waste Organization & Fully Organized & 0.00 \\
##   Waste Organization & Mostly Organized & 0.33 \\
##   Waste Organization & Mostly Unorganized & 0.67 \\
##   Waste Organization & Unorganized & 1.00 \\
##   Waste Burning & Pile Cleaned & 0.00 \\
##   Waste Burning & No Burning & 0.00 \\
##   Waste Burning & Less than 50\% burnt & 0.45 \\
##   Waste Burning & More than 50\% burnt & 0.55 \\
##    \hline
## \end{tabular}
## \caption{Scalar Values, Variable Specification A}
## \end{table}
```

```
print(xtable(tab_b, "Scalar Values, Variable Specification B"), include.rownames=F)
```

```
## % latex table generated in R 3.6.1 by xtable 1.8-4 package
## % Tue Jun  9 13:15:05 2020
## \begin{table}[ht]
## \centering
## \begin{tabular}{llr}
##   \hline
## Characteristic & Abbreviated.Response & Scalar \\
##   \hline
## Waste Containment & Pile Cleaned & 0.00 \\
##   Waste Containment & Fully Contained & 0.00 \\
##   Waste Containment & Mostly Contained & 0.25 \\
##   Waste Containment & Mostly Uncontained & 0.75 \\
##   Waste Containment & Uncontained & 1.00 \\
##   Waste Organization & Pile Cleaned & 0.00 \\
##   Waste Organization & Fully Organized & 0.00 \\
##   Waste Organization & Mostly Organized & 0.25 \\
##   Waste Organization & Mostly Unorganized & 0.75 \\
##   Waste Organization & Unorganized & 1.00 \\
##   Waste Burning & Pile Cleaned & 0.00 \\
##   Waste Burning & No Burning & 0.00 \\
##   Waste Burning & Less than 50\% burnt & 0.33 \\
##   Waste Burning & More than 50\% burnt & 0.67 \\
##    \hline
## \end{tabular}
## \caption{Scalar Values, Variable Specification B}
## \end{table}
```

```
## Table I1: Spillover Results, Primary Dependent Variables (Cleaned) -----
cph3.analysis_spill <- subset(subsetB_dta, indirect.prob!=0) #All zones potentially affected by indirec

cph3.analysis_spill$realized.prob <- NA
cph3.analysis_spill$realized.prob <- ifelse(cph3.analysis_spill$treat==1 & cph3.analysis_spill$indirect
                                            cph3.analysis_spill$d1i1, cph3.analysis_spill$realized.prob
```

```
cph3.analysis_spill$realized.prob <- ifelse(cph3.analysis_spill$treat==1 & cph3.analysis_spill$indirect=
                                            cph3.analysis_spill$d1i0, cph3.analysis_spill$realized.prob)
cph3.analysis_spill$realized.prob <- ifelse(cph3.analysis_spill$treat==0 & cph3.analysis_spill$indirect=
                                            cph3.analysis_spill$d0i1, cph3.analysis_spill$realized.prob)
cph3.analysis_spill$realized.prob <- ifelse(cph3.analysis_spill$treat==0 & cph3.analysis_spill$indirect=
                                            cph3.analysis_spill$d0i0, cph3.analysis_spill$realized.prob)

cph3.analysis_spill$spillover <- NA
cph3.analysis_spill$spillover <- paste("d", cph3.analysis_spill$treat, "i", cph3.analysis_spill$indirect
table(cph3.analysis_spill$spillover, useNA="always")
```

```
##
## d0i0 d0i1 d1i0 d1i1 <NA>
##   97  177  120  193    0
```

```
cph3.analysis_spill$spillover <- ordered(cph3.analysis_spill$spillover, levels=c("d0i0", "d0i1", "d1i0"
table(cph3.analysis_spill$spillover, useNA="always")
```

```
##
## d0i0 d0i1 d1i0 d1i1 <NA>
##   97  177  120  193    0
```

```
#Size
spill.m1_area <- lm(m1.size_final~spillover+p1.p2.monitoring+b.pile.area_m+lights.mean+road.density+area
                    data = cph3.analysis_spill, weights=1/realized.prob)
summary(spill.m1_area)
```

```
##
## Call:
## lm(formula = m1.size_final ~ spillover + p1.p2.monitoring + b.pile.area_m +
##     lights.mean + road.density + area.km2 + div + ls.pop_2016,
##     data = cph3.analysis_spill, weights = 1/realized.prob)
##
## Weighted Residuals:
##     Min      1Q  Median      3Q     Max
## -843.54  -35.32  -14.12    5.18 1303.78
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)      -8.4658247 21.9299659  -0.386  0.69961
## spillover.L      -8.9293340  4.6281743  -1.929  0.05419 .
## spillover.Q       1.4219577  4.6586928   0.305  0.76031
## spillover.C       2.4612364  4.8189352   0.511  0.60973
## p1.p2.monitoring  2.5021088  4.8595635   0.515  0.60684
## b.pile.area_m     0.0554866  0.0100527   5.520  5.2e-08 ***
## lights.mean       0.6152552  0.3408568   1.805  0.07161 .
## road.density     -0.0011891  0.0004043  -2.941  0.00341 **
## area.km2          0.8043834  5.5447322   0.145  0.88471
## divkawempe        5.2734065  8.0246484   0.657  0.51135
## divmakindye       0.4760138  9.2740649   0.051  0.95908
## divnakawa         8.2165573  8.8855634   0.925  0.35552
## divrubaga         5.6190485  8.5233012   0.659  0.51000
## ls.pop_2016      -0.0001410  0.0001859  -0.758  0.44848
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
##
## Residual standard error: 126.5 on 561 degrees of freedom
##   (12 observations deleted due to missingness)
## Multiple R-squared:  0.1102, Adjusted R-squared:  0.08953
## F-statistic: 5.342 on 13 and 561 DF,  p-value: 4.148e-09
```

```
spill.m1_area.c <- lm.cluster(spill.m1_area, cph3.analysis_spill$zone.id)


spill.m2_area <- lm(m2.size_final~spillover+p1.p2.monitoring+b.pile.area_m+lights.mean+road.density+area
                    data = cph3.analysis_spill, weights=1/realized.prob)
summary(spill.m2_area)
```

```
##
## Call:
## lm(formula = m2.size_final ~ spillover + p1.p2.monitoring + b.pile.area_m +
##     lights.mean + road.density + area.km2 + div + ls.pop_2016,
##     data = cph3.analysis_spill, weights = 1/realized.prob)
##
## Weighted Residuals:
##     Min      1Q  Median      3Q     Max
## -3513.7  -131.6   -40.8    60.0 12580.5
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)      -18.548143 128.146557  -0.145  0.88497
## spillover.L      -68.330386  27.044483  -2.527  0.01179 *
## spillover.Q       35.495913  27.222817   1.304  0.19280
## spillover.C      -32.014762  28.159184  -1.137  0.25606
## p1.p2.monitoring -49.937854  28.396594  -1.759  0.07919 .
## b.pile.area_m      0.187744   0.058742   3.196  0.00147 **
## lights.mean        2.729657   1.991778   1.370  0.17109
## road.density      -0.004937   0.002363  -2.090  0.03711 *
## area.km2          -5.565688  32.400339  -0.172  0.86367
## divkawempe       -19.698746  46.891594  -0.420  0.67458
## divmakindye       66.084603  54.192491   1.219  0.22319
## divnakawa        -56.153330  51.922304  -1.081  0.27994
## divrubaga         54.480382  49.805445   1.094  0.27448
## ls.pop_2016       -0.002982   0.001086  -2.745  0.00624 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 739.3 on 561 degrees of freedom
##   (12 observations deleted due to missingness)
## Multiple R-squared:  0.104,  Adjusted R-squared:  0.08324
## F-statistic: 5.009 on 13 and 561 DF,  p-value: 2.112e-08
```

```
spill.m2_area.c <- lm.cluster(spill.m2_area, cph3.analysis_spill$zone.id)

#Dummy
spill.m1_d <- lm(m1.waste.pile_d~spillover+p1.p2.monitoring+lights.mean+road.density+area.km2+div+ls.pop
                 data = cph3.analysis_spill, weights=1/realized.prob)
spill.m1_d.c <- lm.cluster(spill.m1_d, cph3.analysis_spill$zone.id)

spill.m2_d <- lm(m2.waste.pile_d~spillover+p1.p2.monitoring+lights.mean+road.density+area.km2+div+ls.pop
```

```
                 data = cph3.analysis_spill, weights=1/realized.prob)
spill.m2_d.c <- lm.cluster(spill.m2_d, cph3.analysis_spill$zone.id)

#Rank, re-ranked for subset
cph3.analysis_spill$rank.b_new2 <- rank(cph3.analysis_spill$b.pile.area_m, na.last = "keep")
cph3.analysis_spill$rank.m1_new2 <- rank(cph3.analysis_spill$m1.size_final, na.last ="keep")
cph3.analysis_spill$rank.m2_new2 <- rank(cph3.analysis_spill$m2.size_final, na.last ="keep")

spill.m1_rank <- lm(rank.m1_new2~spillover+p1.p2.monitoring+rank.b_new2+lights.mean+road.density+area.k
                    data = cph3.analysis_spill, weights=1/realized.prob)
spill.m1_rank.c <- lm.cluster(spill.m1_rank, cph3.analysis_spill$zone.id)

spill.m2_rank <- lm(rank.m2_new2~spillover+p1.p2.monitoring+rank.b_new2+lights.mean+road.density+area.k
                    data = cph3.analysis_spill, weights=1/realized.prob)
spill.m2_rank.c <- lm.cluster(spill.m2_rank, cph3.analysis_spill$zone.id)

stargazer(spill.m1_d, spill.m1_area, spill.m1_rank, spill.m2_d, spill.m2_area, spill.m2_rank,
          dep.var.caption = "DV: Cleaned (0: Yes, 1: No), Pile Size ($m^{2}$), Pile Size Rank",
          dep.var.labels = c("M1 Cleaned","M1 Size","M1 Rank","M2 Cleaned","M2 Size","M2 Rank"),
          covariate.labels = c("Control, Indirect", "Treated, No Indirect","Treated, Indirect"),
          omit = c("p1.p2.monitoring", "b.pile.area_m", "rank.b", "lights.mean", "road.density", "area.
          p = list(spill.m1_d.c[[1]][c(2:4,1),4], spill.m1_area.c[[1]][c(2:4,1),4], spill.m1_rank.c[[1]]
                   spill.m2_d.c[[1]][c(2:4,1),4], spill.m2_area.c[[1]][c(2:4,1),4], spill.m2_rank.c[[1]]
          se = list(spill.m1_d.c[[1]][c(2:4,1),2], spill.m1_area.c[[1]][c(2:4,1),2], spill.m1_rank.c[[1
                    spill.m2_d.c[[1]][c(2:4,1),2], spill.m2_area.c[[1]][c(2:4,1),2], spill.m2_rank.c[[1
          add.lines = list(c("Covariates", "Yes", "Yes", "Yes", "Yes", "Yes", "Yes")),
          df = FALSE,
          notes.label = "Notes: two-tailed tests; weighted by inverse probability of assignment to expos
          omit.stat = c("rsq","ser","adj.rsq"),
          intercept.bottom = TRUE)
```

```
##
## % Table created by stargazer v.5.2.2 by Marek Hlavac, Harvard University. E-mail: hlavac at fas.harva
## % Date and time: Tue, Jun 09, 2020 - 13:15:05
## \begin{table}[!htbp] \centering
##   \caption{}
##   \label{}
## \begin{tabular}{@{\extracolsep{5pt}}lcccccc}
## \\[-1.8ex]\hline
## \hline \\[-1.8ex]
##  & \multicolumn{6}{c}{DV: Cleaned (0: Yes, 1: No), Pile Size ($m^{2}$), Pile Size Rank} \\
## \cline{2-7}
## \\[-1.8ex] & M1 Cleaned & M1 Size & M1 Rank & M2 Cleaned & M2 Size & M2 Rank \\
## \\[-1.8ex] & (1) & (2) & (3) & (4) & (5) & (6)\\
## \hline \\[-1.8ex]
##  Control, Indirect & 0.001 & $-$8.929 & 10.280 & $-$0.003 & $-$68.330 & 3.391 \\
##   & (0.029) & (9.014) & (16.978) & (0.042) & (57.618) & (14.049) \\
##   & & & & & & \\
##  Treated, No Indirect & 0.011 & 1.422 & 6.689 & 0.040 & 35.496 & 18.014 \\
##   & (0.032) & (7.440) & (16.091) & (0.042) & (42.806) & (12.987) \\
##   & & & & & & \\
##  Treated, Indirect & $-$0.020 & 2.461 & $-$3.356 & $-$0.047 & $-$32.015 & $-$26.016 \\
##   & (0.033) & (4.667) & (15.679) & (0.046) & (24.755) & (16.075) \\
##   & & & & & & \\
```

```
##  Constant & 1.110$^{***}$ & $-$8.466 & 177.610$^{**}$ & 1.045$^{***}$ & $-$18.548 & 218.074$^{***}$ \
##   & (0.136) & (21.319) & (73.886) & (0.186) & (106.761) & (58.661) \\
##   & & & & & & \\
## \hline \\[-1.8ex]
## Covariates & Yes & Yes & Yes & Yes & Yes & Yes \\
## Observations & 575 & 575 & 575 & 575 & 575 & 575 \\
## F Statistic & 1.746$^{*}$ & 5.342$^{***}$ & 9.497$^{***}$ & 1.726$^{*}$ & 5.009$^{***}$ & 8.691$^{***
## \hline
## \hline \\[-1.8ex]
## Notes: two-tailed tests; weighted by inverse probability of assignment to exposure type; baseline is
ef{tab:main-ate.cln}. & \multicolumn{6}{r}{$^{*}$p$<$0.1; $^{**}$p$<$0.05; $^{***}$p$<$0.01} \\
## \end{tabular}
## \end{table}
```

```
## Table H2: No Spillover Results, Primary Dependent Variables (Cleaned) -----
cph3.analysis_nspill <- subset(subsetB_dta, indirect.prob==0) #All zones not affected by indirect Ph3 s
nspill.m1_area <- lm(m1.size_final~treat+p1.p2.monitoring+b.pile.area_m+lights.mean+road.density+area.k
                     data = cph3.analysis_nspill)
summary(nspill.m1_area)
```

```
##
## Call:
## lm(formula = m1.size_final ~ treat + p1.p2.monitoring + b.pile.area_m +
##     lights.mean + road.density + area.km2 + div + ls.pop_2016,
##     data = cph3.analysis_nspill)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -118.373   -8.444   -0.600    8.119  125.926
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)      -9.556e+00  1.796e+01  -0.532    0.596
## treat             3.393e+00  6.031e+00   0.563    0.575
## p1.p2.monitoring -2.588e+00  6.083e+00  -0.425    0.672
## b.pile.area_m     3.962e-01  2.441e-02  16.228   <2e-16 ***
## lights.mean      -1.349e-02  2.927e-01  -0.046    0.963
## road.density      7.592e-04  6.813e-04   1.114    0.268
## area.km2          5.096e+00  8.990e+00   0.567    0.572
## divkawempe       -6.550e-01  9.954e+00  -0.066    0.948
## divmakindye       1.209e+00  1.002e+01   0.121    0.904
## divnakawa         6.193e+00  1.013e+01   0.611    0.543
## divrubaga        -1.290e+01  1.163e+01  -1.109    0.270
## ls.pop_2016       3.276e-04  3.455e-04   0.948    0.345
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 25.54 on 92 degrees of freedom
## Multiple R-squared:  0.7619, Adjusted R-squared:  0.7334
## F-statistic: 26.76 on 11 and 92 DF,  p-value: < 2.2e-16
```

```
nspill.m1_area.c <- lm.cluster(nspill.m1_area, cph3.analysis_nspill$zone.id)
```

```
nspill.m2_area <- lm(m2.size_final~treat+p1.p2.monitoring+b.pile.area_m+lights.mean+road.density+area.k
```

```
                    data = cph3.analysis_nspill)
summary(nspill.m2_area)
```

```
##
## Call:
## lm(formula = m2.size_final ~ treat + p1.p2.monitoring + b.pile.area_m +
##     lights.mean + road.density + area.km2 + div + ls.pop_2016,
##     data = cph3.analysis_nspill)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -81.905 -12.973  -4.978   5.803 172.678
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)      -2.875e+01  2.202e+01  -1.306   0.1950
## treat            -2.578e+00  7.394e+00  -0.349   0.7282
## p1.p2.monitoring -1.713e-01  7.458e+00  -0.023   0.9817
## b.pile.area_m     1.501e-01  2.993e-02   5.017 2.55e-06 ***
## lights.mean       7.512e-01  3.589e-01   2.093   0.0391 *
## road.density     -5.124e-04  8.353e-04  -0.614   0.5411
## area.km2          6.327e+00  1.102e+01   0.574   0.5674
## divkawempe       -4.790e-01  1.220e+01  -0.039   0.9688
## divmakindye       1.806e+01  1.229e+01   1.470   0.1450
## divnakawa         1.800e+01  1.242e+01   1.449   0.1506
## divrubaga         7.837e-01  1.426e+01   0.055   0.9563
## ls.pop_2016      -3.501e-04  4.236e-04  -0.826   0.4107
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 31.32 on 92 degrees of freedom
## Multiple R-squared:  0.2839, Adjusted R-squared:  0.1983
## F-statistic: 3.316 on 11 and 92 DF,  p-value: 0.0007173
```

```
nspill.m2_area.c <- lm.cluster(nspill.m2_area, cph3.analysis_nspill$zone.id)

#Dummy
nspill.m1_d <- lm(m1.waste.pile_d~treat+p1.p2.monitoring+lights.mean+road.density+area.km2+div+ls.pop_2
                  data = cph3.analysis_nspill)
nspill.m1_d.c <- lm.cluster(nspill.m1_d, cph3.analysis_nspill$zone.id)

nspill.m2_d <- lm(m2.waste.pile_d~treat+p1.p2.monitoring+lights.mean+road.density+area.km2+div+ls.pop_2
                  data = cph3.analysis_nspill)
nspill.m2_d.c <- lm.cluster(nspill.m2_d, cph3.analysis_nspill$zone.id)

#Rank, re-ranked for subset
cph3.analysis_nspill$rank.b_new2 <- rank(cph3.analysis_nspill$b.pile.area_m, na.last = "keep")
cph3.analysis_nspill$rank.m1_new2 <- rank(cph3.analysis_nspill$m1.size_final, na.last ="keep")
cph3.analysis_nspill$rank.m2_new2 <- rank(cph3.analysis_nspill$m2.size_final, na.last ="keep")

nspill.m1_rank <- lm(rank.m1_new2~treat+p1.p2.monitoring+rank.b_new2+lights.mean+road.density+area.km2+
                     data = cph3.analysis_nspill)
nspill.m1_rank.c <- lm.cluster(nspill.m1_rank, cph3.analysis_nspill$zone.id)
```

```
nspill.m2_rank <- lm(rank.m2_new2~treat+p1.p2.monitoring+rank.b_new2+lights.mean+road.density+area.km2+
                    data = cph3.analysis_nspill)
nspill.m2_rank.c <- lm.cluster(nspill.m2_rank, cph3.analysis_nspill$zone.id)

stargazer(nspill.m1_d, nspill.m1_area, nspill.m1_rank, nspill.m2_d, nspill.m2_area, nspill.m2_rank,
          dep.var.caption = "DV: Cleaned (0: Yes, 1: No), Pile Size ($m^{2}$), Pile Size Rank",
          dep.var.labels = c("M1 Cleaned","M1 Size","M1 Rank","M2 Cleaned","M2 Size","M2 Rank"),
          covariate.labels = c("Treatment", "P1/P2 Monitoring"),
          omit = c("p1.p2.monitoring", "b.pile.area_m", "rank.b", "lights.mean", "road.density", "area.
          p = list(nspill.m1_d.c[[1]][c(2:4,1),4], nspill.m1_area.c[[1]][c(2:4,1),4], nspill.m1_rank.c[
                   nspill.m2_d.c[[1]][c(2:4,1),4], nspill.m2_area.c[[1]][c(2:4,1),4], nspill.m2_rank.c[
          se = list(nspill.m1_d.c[[1]][c(2:4,1),2], nspill.m1_area.c[[1]][c(2:4,1),2], nspill.m2_rank.c
                    nspill.m2_d.c[[1]][c(2:4,1),2], nspill.m2_area.c[[1]][c(2:4,1),2], nspill.m2_rank.c
          add.lines = list(c("Covariates", "Yes", "Yes", "Yes", "Yes", "Yes", "Yes")),
          df = FALSE,
          notes.label = "Notes: two-tailed tests; rank variables are specific to the subset. For a full
          omit.stat = c("rsq","ser","adj.rsq"),
          intercept.bottom = TRUE)
```

```
##
## % Table created by stargazer v.5.2.2 by Marek Hlavac, Harvard University. E-mail: hlavac at fas.harva
## % Date and time: Tue, Jun 09, 2020 - 13:15:06
## \begin{table}[!htbp] \centering
##   \caption{}
##   \label{}
## \begin{tabular}{@{\extracolsep{5pt}}lcccccc}
## \\[-1.8ex]\hline
## \hline \\[-1.8ex]
##  & \multicolumn{6}{c}{DV: Cleaned (0: Yes, 1: No), Pile Size ($m^{2}$), Pile Size Rank} \\
## \cline{2-7}
## \\[-1.8ex] & M1 Cleaned & M1 Size & M1 Rank & M2 Cleaned & M2 Size & M2 Rank \\
## \\[-1.8ex] & (1) & (2) & (3) & (4) & (5) & (6)\\
## \hline \\[-1.8ex]
##  Treatment & $-$0.137 & 3.393 & $-$18.615$^{***}$ & $-$0.047 & $-$2.578 & $-$15.123$^{***}$ \\
##   & (0.094) & (5.295) & (5.047) & (0.071) & (6.748) & (5.047) \\
##   & & & & & & \\
##  P1/P2 Monitoring & 0.769$^{**}$ & $-$9.556 & 23.420 & 0.562$^{*}$ & $-$28.749 & $-$7.821 \\
##   & (0.321) & (10.861) & (18.205) & (0.294) & (19.090) & (18.205) \\
##   & & & & & & \\
## \hline \\[-1.8ex]
## Covariates & Yes & Yes & Yes & Yes & Yes & Yes \\
## Observations & 104 & 104 & 104 & 104 & 104 & 104 \\
## F Statistic & 1.067 & 26.758$^{***}$ & 3.079$^{***}$ & 0.786 & 3.316$^{***}$ & 3.073$^{***}$ \\
## \hline
## \hline \\[-1.8ex]
## Notes: two-tailed tests; rank variables are specific to the subset. For a full set of variable descri
ef{tab:main-ate.cln}. & \multicolumn{6}{r}{$^{*}$p$<$0.1; $^{**}$p$<$0.05; $^{***}$p$<$0.01} \\
## \end{tabular}
## \end{table}
```

```
## Table J3: HTE, Party ----
subsetB_dta$Party <- as.character(subsetB_dta$Party)
subsetB_dta$party.type <- ifelse(subsetB_dta$Party!="NRM" & subsetB_dta$Party!="INDEPENDENT", "OPPOSITI
subsetB_dta$party.type <- factor(subsetB_dta$party.type, levels=c("NRM","INDEPENDENT","OPPOSITION"))
```

```r
##Binary
lm.wpd1_p <- lm(m1.waste.pile_d~treat*party.type+p1.p2.monitoring+lights.mean+road.density+area.km2+div
summary(lm.wpd1_p)
```

```
##
## Call:
## lm(formula = m1.waste.pile_d ~ treat * party.type + p1.p2.monitoring +
##     lights.mean + road.density + area.km2 + div, data = subsetB_dta)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.00278  0.03076  0.10050  0.15724  0.35352
##
## Coefficients:
##                               Estimate Std. Error t value Pr(>|t|)
## (Intercept)                  1.217e+00  1.272e-01   9.564  < 2e-16 ***
## treat                       -2.899e-02  4.740e-02  -0.612  0.54112
## party.typeINDEPENDENT       -8.143e-02  5.487e-02  -1.484  0.13857
## party.typeOPPOSITION        -7.468e-02  5.993e-02  -1.246  0.21351
## p1.p2.monitoring             5.398e-02  3.717e-02   1.452  0.14723
## lights.mean                 -2.070e-03  1.706e-03  -1.213  0.22580
## road.density                -7.423e-06  2.877e-06  -2.580  0.01024 *
## area.km2                     2.316e-02  3.406e-02   0.680  0.49693
## divkawempe                  -7.209e-02  6.338e-02  -1.137  0.25603
## divmakindye                 -1.386e-01  6.704e-02  -2.067  0.03941 *
## divnakawa                   -1.761e-01  5.855e-02  -3.007  0.00281 **
## divrubaga                    2.519e-02  9.522e-02   0.265  0.79148
## treat:party.typeINDEPENDENT  4.636e-02  8.131e-02   0.570  0.56887
## treat:party.typeOPPOSITION   2.113e-02  7.830e-02   0.270  0.78746
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.3178 on 393 degrees of freedom
##   (284 observations deleted due to missingness)
## Multiple R-squared:  0.0627, Adjusted R-squared:  0.03169
## F-statistic: 2.022 on 13 and 393 DF,  p-value: 0.01809
```

```r
lm.wpd1_pc <- lm.cluster(lm.wpd1_p, subsetB_dta$zone.id)

lm.wpd2_p <- lm(m2.waste.pile_d~treat*party.type+p1.p2.monitoring+lights.mean+road.density+area.km2+div
summary(lm.wpd2_p)
```

```
##
## Call:
## lm(formula = m2.waste.pile_d ~ treat * party.type + p1.p2.monitoring +
##     lights.mean + road.density + area.km2 + div, data = subsetB_dta)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.95079  0.06672  0.19593  0.27007  0.62689
##
## Coefficients:
##                               Estimate Std. Error t value Pr(>|t|)
## (Intercept)                  1.071e+00  1.700e-01   6.303 7.85e-10 ***
```

```
## treat                        -1.771e-02  6.333e-02  -0.280   0.7799
## party.typeINDEPENDENT        -1.589e-02  7.330e-02  -0.217   0.8285
## party.typeOPPOSITION          2.878e-02  8.007e-02   0.359   0.7195
## p1.p2.monitoring              1.318e-02  4.966e-02   0.265   0.7909
## lights.mean                  -1.513e-03  2.280e-03  -0.664   0.5072
## road.density                 -6.251e-06  3.844e-06  -1.626   0.1047
## area.km2                      2.721e-02  4.550e-02   0.598   0.5502
## divkawempe                   -1.457e-01  8.467e-02  -1.721   0.0861 .
## divmakindye                  -2.062e-01  8.956e-02  -2.303   0.0218 *
## divnakawa                    -1.824e-01  7.822e-02  -2.332   0.0202 *
## divrubaga                    -1.721e-02  1.272e-01  -0.135   0.8925
## treat:party.typeINDEPENDENT -1.347e-02  1.086e-01  -0.124   0.9014
## treat:party.typeOPPOSITION  -7.791e-02  1.046e-01  -0.745   0.4569
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.4245 on 393 degrees of freedom
##   (284 observations deleted due to missingness)
## Multiple R-squared:  0.03444,    Adjusted R-squared:  0.002501
## F-statistic: 1.078 on 13 and 393 DF,  p-value: 0.3763
```

```
lm.wpd2_pc <- lm.cluster(lm.wpd2_p, subsetB_dta$zone.id)

## Pile Size
lm.area1_p <- lm(m1.size_final~treat*party.type+p1.p2.monitoring+lights.mean+road.density+area.km2+b.pi
summary(lm.area1_p)
```

```
##
## Call:
## lm(formula = m1.size_final ~ treat * party.type + p1.p2.monitoring +
##     lights.mean + road.density + area.km2 + b.pile.area_m + div,
##     data = subsetB_dta)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -111.92  -13.66   -6.38    5.61  555.73
##
## Coefficients:
##                               Estimate Std. Error t value Pr(>|t|)
## (Intercept)                  -1.446e-01  1.638e+01  -0.009   0.9930
## treat                        -8.101e-01  6.102e+00  -0.133   0.8945
## party.typeINDEPENDENT         4.340e-02  7.068e+00   0.006   0.9951
## party.typeOPPOSITION          1.701e+01  7.745e+00   2.196   0.0287 *
## p1.p2.monitoring             -1.216e+00  4.813e+00  -0.253   0.8006
## lights.mean                   3.153e-01  2.198e-01   1.434   0.1523
## road.density                 -8.099e-04  3.707e-04  -2.185   0.0295 *
## area.km2                     -3.566e+00  4.385e+00  -0.813   0.4166
## b.pile.area_m                 1.943e-01  1.951e-02   9.958   <2e-16 ***
## divkawempe                    6.942e+00  8.161e+00   0.851   0.3955
## divmakindye                   2.021e+00  8.631e+00   0.234   0.8150
## divnakawa                     7.449e+00  7.538e+00   0.988   0.3237
## divrubaga                    -9.199e+00  1.227e+01  -0.750   0.4537
## treat:party.typeINDEPENDENT -1.713e+00  1.047e+01  -0.164   0.8701
## treat:party.typeOPPOSITION  -1.404e+01  1.008e+01  -1.392   0.1647
## ---
```

```
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 40.91 on 392 degrees of freedom
##   (284 observations deleted due to missingness)
## Multiple R-squared:  0.2404, Adjusted R-squared:  0.2132
## F-statistic: 8.859 on 14 and 392 DF,  p-value: < 2.2e-16
```

```
lm.area1_pc <- lm.cluster(lm.area1_p, subsetB_dta$zone.id)

lm.area2_p <- lm(m2.size_final~treat*party.type+p1.p2.monitoring+lights.mean+road.density+area.km2+b.pil
summary(lm.area2_p)
```

```
##
## Call:
## lm(formula = m2.size_final ~ treat * party.type + p1.p2.monitoring +
##     lights.mean + road.density + area.km2 + b.pile.area_m + div,
##     data = subsetB_dta)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
##  -97.63  -20.08   -8.94    4.78 1934.24
##
## Coefficients:
##                                Estimate Std. Error t value Pr(>|t|)
## (Intercept)                  44.444189  42.590489   1.044    0.297
## treat                        11.626293  15.868858   0.733    0.464
## party.typeINDEPENDENT         3.015865  18.380712   0.164    0.870
## party.typeOPPOSITION          4.222171  20.140561   0.210    0.834
## p1.p2.monitoring             -6.972694  12.516424  -0.557    0.578
## lights.mean                  -0.298521   0.571545  -0.522    0.602
## road.density                 -0.001182   0.000964  -1.226    0.221
## area.km2                     -8.872983  11.401888  -0.778    0.437
## b.pile.area_m                 0.061987   0.050745   1.222    0.223
## divkawempe                    3.715895  21.222693   0.175    0.861
## divmakindye                  12.362540  22.442932   0.551    0.582
## divnakawa                    -6.007113  19.602603  -0.306    0.759
## divrubaga                   -31.647583  31.897817  -0.992    0.322
## treat:party.typeINDEPENDENT -13.613280  27.221093  -0.500    0.617
## treat:party.typeOPPOSITION   24.703964  26.223213   0.942    0.347
##
## Residual standard error: 106.4 on 392 degrees of freedom
##   (284 observations deleted due to missingness)
## Multiple R-squared:  0.02614,    Adjusted R-squared:  -0.00864
## F-statistic: 0.7516 on 14 and 392 DF,  p-value: 0.7216
```

```
lm.area2_pc <- lm.cluster(lm.area2_p, subsetB_dta$zone.id)

##Output Table
keepA <- c("treat", "party.typeINDEPENDENT", "party.typeOPPOSITION", "treat:party.typeINDEPENDENT", "tr
keepB <- c("treat", "party.typeINDEPENDENT", "party.typeOPPOSITION", "b.pile.area_m", "treat:party.type
stargazer(lm.wpd1_p, lm.area1_p, lm.wpd2_p, lm.area2_p, type = "latex",
          dep.var.caption = "DV: Cleaned (0: Yes, 1: No) or Waste Pile Size ($m^{2}$)",
          dep.var.labels = c("M1 Cleaned","M1 Size","M2 Cleaned","M2 Size"),
          keep =c("treat","party.typeINDEPENDENT","party.typeOPPOSITION","b.pile.area_m","treat:party.ty
          covariate.labels = c("Treatment","Indepedent","Opposition","Baseline Pile Area",
```

```
                                "Treatment X Indepedent","Treatment X Opposition"),
          add.lines = list(c("Covariates", "Yes", "Yes","Yes","Yes")),
          p = list(lm.wpd1_pc[[1]][keepA, "Pr(>|t|)"], lm.area1_pc[[1]][keepB, "Pr(>|t|)"],
                   lm.wpd2_pc[[1]][keepA,"Pr(>|t|)"], lm.area2_pc[[1]][keepB,"Pr(>|t|)"]),
          se = list(lm.wpd1_pc[[1]][keepA, "Std. Error"], lm.area1_pc[[1]][keepB, "Std. Error"],
                    lm.wpd2_pc[[1]][keepA,"Std. Error"], lm.area2_pc[[1]][keepB,"Std. Error"]),
          notes.label = "<em>Note: </em> two-tailed tests",
          df = FALSE,intercept.bottom = TRUE)
```

```
##
## % Table created by stargazer v.5.2.2 by Marek Hlavac, Harvard University. E-mail: hlavac at fas.harva
## % Date and time: Tue, Jun 09, 2020 - 13:15:06
## \begin{table}[!htbp] \centering
##   \caption{}
##   \label{}
## \begin{tabular}{@{\extracolsep{5pt}}lcccc}
## \\[-1.8ex]\hline
## \hline \\[-1.8ex]
##  & \multicolumn{4}{c}{DV: Cleaned (0: Yes, 1: No) or Waste Pile Size ($m^{2}$)} \\
## \cline{2-5}
## \\[-1.8ex] & M1 Cleaned & M1 Size & M2 Cleaned & M2 Size \\
## \\[-1.8ex] & (1) & (2) & (3) & (4)\\
## \hline \\[-1.8ex]
##  Treatment & $-$0.029 & $-$0.810 & $-$0.018 & 11.626 \\
##   & (0.059) & (5.702) & (0.077) & (10.122) \\
##   & & & & \\
##  Indepedent & $-$0.081 & 0.043 & $-$0.016 & 3.016 \\
##   & (0.063) & (5.000) & (0.080) & (5.845) \\
##   & & & & \\
##  Opposition & $-$0.075 & 17.006 & 0.029 & 4.222 \\
##   & (0.063) & (14.922) & (0.085) & (10.676) \\
##   & & & & \\
##  Baseline Pile Area &  & 0.194$^{***}$ &  & 0.062 \\
##   &  & (0.050) &  & (0.040) \\
##   & & & & \\
##  Treatment X Indepedent & 0.046 & $-$1.713 & $-$0.013 & $-$13.613 \\
##   & (0.094) & (7.433) & (0.113) & (14.759) \\
##   & & & & \\
##  Treatment X Opposition & 0.021 & $-$14.037 & $-$0.078 & 24.704 \\
##   & (0.085) & (14.166) & (0.111) & (32.656) \\
##   & & & & \\
## \hline \\[-1.8ex]
## Covariates & Yes & Yes & Yes & Yes \\
## Observations & 407 & 407 & 407 & 407 \\
## R$^{2}$ & 0.063 & 0.240 & 0.034 & 0.026 \\
## Adjusted R$^{2}$ & 0.032 & 0.213 & 0.003 & $-$0.009 \\
## Residual Std. Error & 0.318 & 40.909 & 0.425 & 106.380 \\
## F Statistic & 2.022$^{**}$ & 8.859$^{***}$ & 1.078 & 0.752 \\
## \hline
## \hline \\[-1.8ex]
## <em>Note: </em> two-tailed tests & \multicolumn{4}{r}{$^{*}$p$<$0.1; $^{**}$p$<$0.05; $^{***}$p$<$0.0
## \end{tabular}
## \end{table}
```

```r
## Table J4: Subgroup Effects; Response Rates -----
rrate_m1 <- lm(m1.size_final~mean.rrate+b.pile.area_m+p1.p2.monitoring+lights.mean+road.density+div+area
summary(rrate_m1)
```

```
##
## Call:
## lm(formula = m1.size_final ~ mean.rrate + b.pile.area_m + p1.p2.monitoring +
##     lights.mean + road.density + div + area.km2 + ls.pop_2016,
##     data = hte_dta)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -170.996  -10.452   -4.245    2.607  299.645
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)       2.304e+01  1.646e+01   1.400    0.162
## mean.rrate       -1.283e+01  2.057e+01  -0.624    0.533
## b.pile.area_m     1.653e-01  1.818e-02   9.092   <2e-16 ***
## p1.p2.monitoring -4.335e-01  4.196e+00  -0.103    0.918
## lights.mean      -1.050e-01  2.451e-01  -0.429    0.669
## road.density     -1.709e-04  2.907e-04  -0.588    0.557
## divkawempe        9.196e+00  6.165e+00   1.492    0.137
## divmakindye       2.451e+00  6.423e+00   0.382    0.703
## divnakawa         5.297e+00  6.375e+00   0.831    0.407
## divrubaga        -3.874e-02  6.107e+00  -0.006    0.995
## area.km2         -7.305e+00  4.834e+00  -1.511    0.132
## ls.pop_2016      -1.815e-04  1.760e-04  -1.032    0.303
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 30.91 on 332 degrees of freedom
##   (4 observations deleted due to missingness)
## Multiple R-squared:  0.2184, Adjusted R-squared:  0.1925
## F-statistic: 8.433 on 11 and 332 DF,  p-value: 4.069e-13
```

```r
rrate_m1c <- lm.cluster(rrate_m1, cluster=hte_dta$zone.id)

rrate_m2 <- lm(m2.size_final~mean.rrate+b.pile.area_m+p1.p2.monitoring+lights.mean+road.density+div+area
summary(rrate_m2)
```

```
##
## Call:
## lm(formula = m2.size_final ~ mean.rrate + b.pile.area_m + p1.p2.monitoring +
##     lights.mean + road.density + div + area.km2 + ls.pop_2016,
##     data = hte_dta)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
##  -89.50  -24.58  -12.72    2.22 1945.47
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)       5.513e+01  6.236e+01   0.884    0.377
```

```
## mean.rrate       -3.155e+01  7.796e+01  -0.405    0.686
## b.pile.area_m     8.003e-02  6.890e-02   1.162    0.246
## p1.p2.monitoring -8.463e+00  1.590e+01  -0.532    0.595
## lights.mean       5.980e-02  9.288e-01   0.064    0.949
## road.density     -8.787e-04  1.102e-03  -0.798    0.426
## divkawempe        1.172e+00  2.336e+01   0.050    0.960
## divmakindye       2.246e+01  2.434e+01   0.923    0.357
## divnakawa        -1.371e+01  2.416e+01  -0.567    0.571
## divrubaga        -6.125e+00  2.314e+01  -0.265    0.791
## area.km2         -1.005e+01  1.832e+01  -0.548    0.584
## ls.pop_2016      -1.003e-03  6.669e-04  -1.505    0.133
##
## Residual standard error: 117.1 on 332 degrees of freedom
##   (4 observations deleted due to missingness)
## Multiple R-squared:  0.01801,    Adjusted R-squared:  -0.01452
## F-statistic: 0.5536 on 11 and 332 DF,  p-value: 0.8655
```

```
rrate_m2c <- lm.cluster(rrate_m2, cluster=hte_dta$zone.id)

rrate_m1.d <- lm(m1.waste.pile_d~mean.rrate+p1.p2.monitoring+lights.mean+road.density+div+area.km2+ls.p
summary(rrate_m1.d)
```

```
##
## Call:
## lm(formula = m1.waste.pile_d ~ mean.rrate + p1.p2.monitoring +
##     lights.mean + road.density + div + area.km2 + ls.pop_2016,
##     data = hte_dta)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.96442  0.03796  0.10514  0.17219  0.31987
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)       8.970e-01  1.785e-01   5.025 8.22e-07 ***
## mean.rrate        2.831e-01  2.233e-01   1.268   0.2058
## p1.p2.monitoring  1.002e-01  4.556e-02   2.200   0.0285 *
## lights.mean      -1.448e-03  2.660e-03  -0.544   0.5866
## road.density     -4.665e-06  3.156e-06  -1.478   0.1403
## divkawempe       -1.137e-02  6.686e-02  -0.170   0.8651
## divmakindye      -1.685e-01  6.967e-02  -2.419   0.0161 *
## divnakawa        -9.407e-02  6.918e-02  -1.360   0.1749
## divrubaga        -2.005e-02  6.622e-02  -0.303   0.7622
## area.km2          6.331e-02  5.238e-02   1.209   0.2276
## ls.pop_2016       2.389e-06  1.904e-06   1.254   0.2106
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.3356 on 333 degrees of freedom
##   (4 observations deleted due to missingness)
## Multiple R-squared:  0.05865,    Adjusted R-squared:  0.03038
## F-statistic: 2.075 on 10 and 333 DF,  p-value: 0.02598
```

```
rrate_m1.dc <- lm.cluster(rrate_m1.d, cluster=hte_dta$zone.id)
```

```r
rrate_m2.d <- lm(m2.waste.pile_d~mean.rrate+p1.p2.monitoring+lights.mean+road.density+div+area.km2+ls.p
summary(rrate_m2.d)
```

```
##
## Call:
## lm(formula = m2.waste.pile_d ~ mean.rrate + p1.p2.monitoring +
##     lights.mean + road.density + div + area.km2 + ls.pop_2016,
##     data = hte_dta)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.91775  0.09724  0.18477  0.23164  0.43933
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)        1.224e+00  2.163e-01   5.660 3.27e-08 ***
## mean.rrate        -4.336e-02  2.706e-01  -0.160   0.8728
## p1.p2.monitoring -1.926e-02  5.519e-02  -0.349   0.7273
## lights.mean       -4.898e-03  3.223e-03  -1.520   0.1295
## road.density      -3.570e-06  3.824e-06  -0.934   0.3512
## divkawempe        -9.733e-02  8.100e-02  -1.202   0.2304
## divmakindye       -1.766e-01  8.441e-02  -2.093   0.0371 *
## divnakawa         -1.518e-01  8.382e-02  -1.811   0.0710 .
## divrubaga         -7.554e-02  8.023e-02  -0.941   0.3471
## area.km2          -2.592e-02  6.346e-02  -0.408   0.6832
## ls.pop_2016        1.701e-06  2.307e-06   0.737   0.4615
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.4066 on 333 degrees of freedom
##   (4 observations deleted due to missingness)
## Multiple R-squared:  0.02281,    Adjusted R-squared:  -0.006533
## F-statistic: 0.7774 on 10 and 333 DF,  p-value: 0.6507
```

```r
rrate_m2.dc <- lm.cluster(rrate_m2.d, cluster=hte_dta$zone.id)

stargazer(rrate_m1.d, rrate_m1, rrate_m2.d, rrate_m2,
          type="latex",
          title="Estimated Effects of Treatment Conditional on Zone-Level Response Rate",
          label="squeaky-wheel2",
          dep.var.caption = "DV: Cleaned (0: Yes, 1: No) or Waste Pile Size ($m^{2}$)",
          dep.var.labels = c("M1 Cleaned","M1 Size","M2 Cleaned","M2 Size"),
          covariate.labels = c("Response Rate","Baseline Pile Area","P1/P2 Monitoring"),
          keep = c("mean.rrate", "b.pile.area_m", "p1.p2.monitoring"),
          add.lines = list(c("Covariates", "Yes", "Yes","Yes","Yes")),
          p = list(rrate_m1.dc[[1]][2:4,4], rrate_m1c[[1]][2:4,4], rrate_m2.dc[[1]][2:4,4], rrate_m2c[[
          se = list(rrate_m1.dc[[1]][2:4,2], rrate_m1c[[1]][2:4,2], rrate_m2.dc[[1]][2:4,2], rrate_m2c[
          notes.label = "Note: two-tailed tests",
          column.sep.width = "1pt",
          df = FALSE,intercept.bottom = TRUE)
```

```
##
## % Table created by stargazer v.5.2.2 by Marek Hlavac, Harvard University. E-mail: hlavac at fas.harva
## % Date and time: Tue, Jun 09, 2020 - 13:15:06
```

```
## \begin{table}[!htbp] \centering
##   \caption{Estimated Effects of Treatment Conditional on Zone-Level Response Rate}
##   \label{squeaky-wheel2}
## \begin{tabular}{@{\extracolsep{1pt}}lcccc}
## \\[-1.8ex]\hline
## \hline \\[-1.8ex]
##  & \multicolumn{4}{c}{DV: Cleaned (0: Yes, 1: No) or Waste Pile Size ($m^{2}$)} \\
## \cline{2-5}
## \\[-1.8ex] & M1 Cleaned & M1 Size & M2 Cleaned & M2 Size \\
## \\[-1.8ex] & (1) & (2) & (3) & (4)\\
## \hline \\[-1.8ex]
##  Response Rate & 0.283 & $-$12.831 & $-$0.043 & $-$31.552 \\
##   & (0.272) & (22.332) & (0.300) & (55.359) \\
##   & & & & \\
##  Baseline Pile Area &  & 0.165 &  & 0.080$^{*}$ \\
##   &  & (0.101) &  & (0.041) \\
##   & & & & \\
##  P1/P2 Monitoring & 0.100$^{*}$ & $-$0.433 & $-$0.019 & $-$8.463 \\
##   & (0.053) & (2.958) & (0.068) & (10.114) \\
##   & & & & \\
## \hline \\[-1.8ex]
## Covariates & Yes & Yes & Yes & Yes \\
## Observations & 344 & 344 & 344 & 344 \\
## R$^{2}$ & 0.059 & 0.218 & 0.023 & 0.018 \\
## Adjusted R$^{2}$ & 0.030 & 0.192 & $-$0.007 & $-$0.015 \\
## Residual Std. Error & 0.336 & 30.914 & 0.407 & 117.148 \\
## F Statistic & 2.075$^{**}$ & 8.433$^{***}$ & 0.777 & 0.554 \\
## \hline
## \hline \\[-1.8ex]
## Note: two-tailed tests & \multicolumn{4}{r}{$^{*}$p$<$0.1; $^{**}$p$<$0.05; $^{***}$p$<$0.01} \\
## \end{tabular}
## \end{table}
```

```
## Table J5: Subgroup Effects; Reported Quality of Initial Waste Services -----
cons.service_m1 <- lm(m1.size_final~waste.service_modal+b.pile.area_m+p1.p2.monitoring+lights.mean+road
summary(cons.service_m1)
```

```
##
## Call:
## lm(formula = m1.size_final ~ waste.service_modal + b.pile.area_m +
##     p1.p2.monitoring + lights.mean + road.density + div + area.km2 +
##     ls.pop_2016, data = hte_dta)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -171.742  -11.101   -4.301    2.455  299.996
##
## Coefficients:
##                       Estimate Std. Error t value Pr(>|t|)
## (Intercept)         19.0201667 15.4116435   1.234    0.218
## waste.service_modal -0.8575907  3.7798825  -0.227    0.821
## b.pile.area_m        0.1653955  0.0181926   9.091   <2e-16 ***
## p1.p2.monitoring    -0.3155760  4.2001443  -0.075    0.940
## lights.mean         -0.1139279  0.2450535  -0.465    0.642
## road.density        -0.0001392  0.0002930  -0.475    0.635
```

```
## divkawempe           9.8335829  6.3000111   1.561    0.120
## divmakindye          2.8118129  6.6227000   0.425    0.671
## divnakawa            5.5066465  6.3896976   0.862    0.389
## divrubaga            0.9232416  6.2355617   0.148    0.882
## area.km2            -7.3900422  4.8818381  -1.514    0.131
## ls.pop_2016         -0.0001769  0.0001759  -1.006    0.315
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 30.93 on 332 degrees of freedom
##   (4 observations deleted due to missingness)
## Multiple R-squared:  0.2176, Adjusted R-squared:  0.1917
## F-statistic: 8.394 on 11 and 332 DF,  p-value: 4.74e-13
```

```
cons.service_m1c <- lm.cluster(cons.service_m1, hte_dta$zone.id)

cons.service_m2  <- lm(m2.size_final~waste.service_modal+b.pile.area_m+p1.p2.monitoring+lights.mean+roa
summary(cons.service_m2)
```

```
##
## Call:
## lm(formula = m2.size_final ~ waste.service_modal + b.pile.area_m +
##     p1.p2.monitoring + lights.mean + road.density + div + area.km2 +
##     ls.pop_2016, data = hte_dta)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
##  -81.80  -26.82  -12.22    2.04 1936.45
##
## Coefficients:
##                       Estimate Std. Error t value Pr(>|t|)
## (Intercept)          5.386e+01  5.818e+01   0.926    0.355
## waste.service_modal  2.176e+01  1.427e+01   1.525    0.128
## b.pile.area_m        8.258e-02  6.868e-02   1.202    0.230
## p1.p2.monitoring    -6.044e+00  1.586e+01  -0.381    0.703
## lights.mean         -8.452e-02  9.251e-01  -0.091    0.927
## road.density        -1.107e-03  1.106e-03  -1.000    0.318
## divkawempe          -5.997e+00  2.378e+01  -0.252    0.801
## divmakindye          1.323e+01  2.500e+01   0.529    0.597
## divnakawa           -1.585e+01  2.412e+01  -0.657    0.512
## divrubaga           -1.372e+01  2.354e+01  -0.583    0.560
## area.km2            -1.521e+01  1.843e+01  -0.825    0.410
## ls.pop_2016         -9.610e-04  6.641e-04  -1.447    0.149
##
## Residual standard error: 116.8 on 332 degrees of freedom
##   (4 observations deleted due to missingness)
## Multiple R-squared:  0.02436,    Adjusted R-squared:  -0.007962
## F-statistic: 0.7537 on 11 and 332 DF,  p-value: 0.6863
```

```
cons.service_m2c <- lm.cluster(cons.service_m2, hte_dta$zone.id)

cons.service_m1.d <- lm(m1.waste.pile_d~waste.service_modal+p1.p2.monitoring+lights.mean+road.density+di
summary(cons.service_m1.d)
```

```
##
```

```
## Call:
## lm(formula = m1.waste.pile_d ~ waste.service_modal + p1.p2.monitoring +
##     lights.mean + road.density + div + area.km2 + ls.pop_2016,
##     data = hte_dta)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.96527  0.04833  0.11500  0.16686  0.32481
##
## Coefficients:
##                       Estimate Std. Error t value Pr(>|t|)
## (Intercept)          9.944e-01  1.672e-01   5.946 6.95e-09 ***
## waste.service_modal  4.328e-02  4.104e-02   1.055  0.29236
## p1.p2.monitoring     9.980e-02  4.561e-02   2.188  0.02936 *
## lights.mean         -1.377e-03  2.660e-03  -0.517  0.60517
## road.density        -5.676e-06  3.182e-06  -1.784  0.07534 .
## divkawempe          -3.436e-02  6.832e-02  -0.503  0.61540
## divmakindye         -1.868e-01  7.187e-02  -2.599  0.00975 **
## divnakawa           -1.014e-01  6.935e-02  -1.462  0.14470
## divrubaga           -5.144e-02  6.761e-02  -0.761  0.44732
## area.km2             6.015e-02  5.290e-02   1.137  0.25636
## ls.pop_2016          2.318e-06  1.904e-06   1.218  0.22422
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.3359 on 333 degrees of freedom
##   (4 observations deleted due to missingness)
## Multiple R-squared:  0.05725,    Adjusted R-squared:  0.02894
## F-statistic: 2.022 on 10 and 333 DF,  p-value: 0.03053
```

```
cons.service_m1.dc <- lm.cluster(cons.service_m1.d, hte_dta$zone.id)

cons.service_m2.d <- lm(m2.waste.pile_d~waste.service_modal+p1.p2.monitoring+lights.mean+road.density+di
summary(cons.service_m2.d)
```

```
##
## Call:
## lm(formula = m2.waste.pile_d ~ waste.service_modal + p1.p2.monitoring +
##     lights.mean + road.density + div + area.km2 + ls.pop_2016,
##     data = hte_dta)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.92015  0.09964  0.18369  0.23883  0.46275
##
## Coefficients:
##                       Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.221e+00  2.024e-01   6.034 4.25e-09 ***
## waste.service_modal  2.720e-02  4.966e-02   0.548   0.5842
## p1.p2.monitoring    -1.618e-02  5.519e-02  -0.293   0.7695
## lights.mean         -5.084e-03  3.219e-03  -1.579   0.1152
## road.density        -3.848e-06  3.850e-06  -0.999   0.3183
## divkawempe          -1.061e-01  8.268e-02  -1.284   0.2001
## divmakindye         -1.882e-01  8.697e-02  -2.164   0.0312 *
## divnakawa           -1.544e-01  8.393e-02  -1.840   0.0666 .
```

```
## divrubaga           -8.480e-02  8.182e-02  -1.036   0.3008
## area.km2            -3.240e-02  6.402e-02  -0.506   0.6131
## ls.pop_2016          1.758e-06  2.304e-06   0.763   0.4459
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.4065 on 333 degrees of freedom
##   (4 observations deleted due to missingness)
## Multiple R-squared:  0.02362,    Adjusted R-squared:  -0.005705
## F-statistic: 0.8054 on 10 and 333 DF,  p-value: 0.6236
```

```
cons.service_m2.dc <- lm.cluster(cons.service_m2.d, hte_dta$zone.id)

stargazer(cons.service_m1.d, cons.service_m1, cons.service_m2.d, cons.service_m2, type = "latex",
          title="Estimated Effects of Treatment Conditional on Baseline Quality of Service Provision",
          label="overall_table",
          dep.var.caption = "DV: Cleaned (0: Yes, 1: No) or Waste Pile Size ($m^{2}$)",
          dep.var.labels = c("M1 Cleaned","M1 Size","M2 Cleaned","M2 Size"),
          covariate.labels = c("Reported Service Quality","Baseline Pile Area","P1/P2 Monitoring"),
          keep = c("waste.service_modal", "b.pile.area_m", "p1.p2.monitoring"),
          add.lines = list(c("Covariates", "Yes", "Yes","Yes","Yes")),
          p = list(cons.service_m1.dc[[1]][2:4,4], cons.service_m1c[[1]][2:4,4], cons.service_m2.dc[[1]]
          se = list(cons.service_m1.dc[[1]][2:4,2], cons.service_m1c[[1]][2:4,2], cons.service_m2.dc[[1
          notes.label = "Note: two-tailed tests",
          column.sep.width = "1pt",
          df = FALSE,intercept.bottom = TRUE)
```

```
##
## % Table created by stargazer v.5.2.2 by Marek Hlavac, Harvard University. E-mail: hlavac at fas.harva
## % Date and time: Tue, Jun 09, 2020 - 13:15:06
## \begin{table}[!htbp] \centering
##   \caption{Estimated Effects of Treatment Conditional on Baseline Quality of Service Provision}
##   \label{overall_table}
## \begin{tabular}{@{\extracolsep{1pt}}lcccc}
## \\[-1.8ex]\hline
## \hline \\[-1.8ex]
##  & \multicolumn{4}{c}{DV: Cleaned (0: Yes, 1: No) or Waste Pile Size ($m^{2}$)} \\
## \cline{2-5}
## \\[-1.8ex] & M1 Cleaned & M1 Size & M2 Cleaned & M2 Size \\
## \\[-1.8ex] & (1) & (2) & (3) & (4)\\
## \hline \\[-1.8ex]
##  Reported Service Quality & 0.043 & $-$0.858 & 0.027 & 21.762 \\
##   & (0.045) & (3.905) & (0.056) & (14.978) \\
##   & & & & \\
##  Baseline Pile Area &  & 0.165 &  & 0.083$^{**}$ \\
##   &  & (0.101) &  & (0.039) \\
##   & & & & \\
##  P1/P2 Monitoring & 0.100$^{**}$ & $-$0.316 & $-$0.016 & $-$6.044 \\
##   & (0.051) & (2.973) & (0.068) & (9.805) \\
##   & & & & \\
## \hline \\[-1.8ex]
## Covariates & Yes & Yes & Yes & Yes \\
## Observations & 344 & 344 & 344 & 344 \\
## R$^{2}$ & 0.057 & 0.218 & 0.024 & 0.024 \\
## Adjusted R$^{2}$ & 0.029 & 0.192 & $-$0.006 & $-$0.008 \\
```

```
## Residual Std. Error & 0.336 & 30.930 & 0.406 & 116.769 \\
## F Statistic & 2.022$^{**}$ & 8.394$^{***}$ & 0.805 & 0.754 \\
## \hline
## \hline \\[-1.8ex]
## Note: two-tailed tests & \multicolumn{4}{r}{$^{*}$p$<$0.1; $^{**}$p$<$0.05; $^{***}$p$<$0.01} \\
## \end{tabular}
## \end{table}
```

```
## Table J6: Subgroup Effects; Reported Zone-Level Dissatisfaction -----
cons.dis_m1 <- lm(m1.size_final~dissatisfaction_modal+b.pile.area_m+p1.p2.monitoring+lights.mean+road.d
summary(cons.dis_m1)
```

```
##
## Call:
## lm(formula = m1.size_final ~ dissatisfaction_modal + b.pile.area_m +
##     p1.p2.monitoring + lights.mean + road.density + div + area.km2 +
##     ls.pop_2016, data = hte_dta)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -168.894   -9.205   -3.947    2.997  306.544
##
## Coefficients:
##                          Estimate Std. Error t value Pr(>|t|)
## (Intercept)             6.244e+00  1.440e+01   0.434    0.665
## dissatisfaction_modal -5.358e-01  3.815e+00  -0.140    0.888
## b.pile.area_m          1.626e-01  1.674e-02   9.718   <2e-16 ***
## p1.p2.monitoring      -5.942e-01  3.951e+00  -0.150    0.881
## lights.mean            4.382e-02  2.277e-01   0.192    0.848
## road.density          -1.002e-05  2.717e-04  -0.037    0.971
## divkawempe             5.914e+00  5.906e+00   1.001    0.317
## divmakindye            3.433e+00  6.043e+00   0.568    0.570
## divnakawa              8.571e+00  6.095e+00   1.406    0.161
## divrubaga              1.398e+00  5.786e+00   0.242    0.809
## area.km2              -4.828e+00  4.475e+00  -1.079    0.281
## ls.pop_2016           -1.198e-04  1.642e-04  -0.730    0.466
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 28.37 on 316 degrees of freedom
##   (20 observations deleted due to missingness)
## Multiple R-squared:  0.2427, Adjusted R-squared:  0.2163
## F-statistic: 9.207 on 11 and 316 DF,  p-value: 2.541e-14
```

```
cons.dis_m1c <- lm.cluster(cons.dis_m1, hte_dta$zone.id)

cons.dis_m2 <- lm(m2.size_final~dissatisfaction_modal+b.pile.area_m+p1.p2.monitoring+lights.mean+road.d
summary(cons.dis_m2)
```

```
##
## Call:
## lm(formula = m2.size_final ~ dissatisfaction_modal + b.pile.area_m +
##     p1.p2.monitoring + lights.mean + road.density + div + area.km2 +
##     ls.pop_2016, data = hte_dta)
##
```

```
## Residuals:
##     Min      1Q  Median      3Q     Max
## -101.38  -25.49   -8.26    4.39 1922.13
##
## Coefficients:
##                          Estimate Std. Error t value Pr(>|t|)
## (Intercept)             2.367e+01  5.801e+01   0.408   0.6836
## dissatisfaction_modal   3.207e+01  1.537e+01   2.086   0.0378 *
## b.pile.area_m           6.716e-02  6.743e-02   0.996   0.3200
## p1.p2.monitoring       -1.376e+01  1.592e+01  -0.864   0.3881
## lights.mean             3.053e-01  9.174e-01   0.333   0.7395
## road.density           -8.052e-04  1.095e-03  -0.735   0.4626
## divkawempe             -1.544e+01  2.380e+01  -0.649   0.5170
## divmakindye             1.879e+01  2.435e+01   0.772   0.4409
## divnakawa              -1.298e+01  2.456e+01  -0.528   0.5975
## divrubaga              -1.469e+01  2.331e+01  -0.630   0.5290
## area.km2               -6.633e+00  1.803e+01  -0.368   0.7132
## ls.pop_2016            -7.151e-04  6.617e-04  -1.081   0.2807
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 114.3 on 316 degrees of freedom
##   (20 observations deleted due to missingness)
## Multiple R-squared:  0.02853,    Adjusted R-squared:  -0.005289
## F-statistic: 0.8436 on 11 and 316 DF,  p-value: 0.5965
```

```
cons.dis_m2c <- lm.cluster(cons.dis_m2, hte_dta$zone.id)

cons.dis_m1.d <- lm(m1.waste.pile_d~dissatisfaction_modal+p1.p2.monitoring+lights.mean+road.density+div
summary(cons.dis_m1.d)
```

```
##
## Call:
## lm(formula = m1.waste.pile_d ~ dissatisfaction_modal + p1.p2.monitoring +
##     lights.mean + road.density + div + area.km2 + ls.pop_2016,
##     data = hte_dta)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.93915  0.05983  0.10284  0.15679  0.32237
##
## Coefficients:
##                          Estimate Std. Error t value Pr(>|t|)
## (Intercept)             1.002e+00  1.689e-01   5.933 7.77e-09 ***
## dissatisfaction_modal   5.843e-02  4.476e-02   1.305   0.1927
## p1.p2.monitoring        7.190e-02  4.639e-02   1.550   0.1221
## lights.mean            -1.028e-03  2.672e-03  -0.385   0.7007
## road.density           -6.788e-06  3.191e-06  -2.127   0.0342 *
## divkawempe             -2.761e-02  6.929e-02  -0.399   0.6905
## divmakindye            -1.719e-01  7.087e-02  -2.426   0.0158 *
## divnakawa              -9.715e-02  7.154e-02  -1.358   0.1755
## divrubaga              -5.598e-02  6.789e-02  -0.824   0.4103
## area.km2                5.640e-02  5.246e-02   1.075   0.2832
## ls.pop_2016             1.798e-06  1.921e-06   0.936   0.3499
## ---
```

```
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.3331 on 317 degrees of freedom
##   (20 observations deleted due to missingness)
## Multiple R-squared:  0.0586, Adjusted R-squared:  0.0289
## F-statistic: 1.973 on 10 and 317 DF,  p-value: 0.03566
```

```
cons.dis_m1.dc <- lm.cluster(cons.dis_m1.d, hte_dta$zone.id)

cons.dis_m2.d <- lm(m2.waste.pile_d~dissatisfaction_modal+p1.p2.monitoring+lights.mean+road.density+div
summary(cons.dis_m2.d)
```

```
##
## Call:
## lm(formula = m2.waste.pile_d ~ dissatisfaction_modal + p1.p2.monitoring +
##     lights.mean + road.density + div + area.km2 + ls.pop_2016,
##     data = hte_dta)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -0.9244  0.1006  0.1880  0.2279  0.4944
##
## Coefficients:
##                          Estimate Std. Error t value Pr(>|t|)
## (Intercept)             1.222e+00  2.053e-01   5.951 7.07e-09 ***
## dissatisfaction_modal   1.900e-02  5.439e-02   0.349   0.7271
## p1.p2.monitoring       -4.296e-02  5.638e-02  -0.762   0.4466
## lights.mean            -4.589e-03  3.248e-03  -1.413   0.1586
## road.density           -5.161e-06  3.878e-06  -1.331   0.1842
## divkawempe             -9.146e-02  8.421e-02  -1.086   0.2782
## divmakindye            -1.755e-01  8.613e-02  -2.037   0.0424 *
## divnakawa              -1.537e-01  8.695e-02  -1.767   0.0781 .
## divrubaga              -8.372e-02  8.251e-02  -1.015   0.3110
## area.km2               -3.365e-02  6.375e-02  -0.528   0.5980
## ls.pop_2016             1.348e-06  2.334e-06   0.578   0.5640
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.4048 on 317 degrees of freedom
##   (20 observations deleted due to missingness)
## Multiple R-squared:  0.02556,    Adjusted R-squared:  -0.005179
## F-statistic: 0.8315 on 10 and 317 DF,  p-value: 0.5985
```

```
cons.dis_m2.dc <- lm.cluster(cons.dis_m2.d, hte_dta$zone.id)

stargazer(cons.dis_m1.d, cons.dis_m1, cons.dis_m2.d, cons.dis_m2, type = "latex",
          title="Estimated Effects of Treatment Conditional on Zone-Level Dissatisfaction at Baseline",
          label="dissatisfaction",
          dep.var.caption = "DV: Cleaned (0: Yes, 1: No) or Waste Pile Size ($m^{2}$)",
          dep.var.labels = c("M1 Cleaned","M1 Size","M2 Cleaned","M2 Size"),
          covariate.labels = c("Dissatisfaction","Baseline Pile Area","P1/P2 Monitoring"),
          keep = c("dissatisfaction_modal", "b.pile.area_m", "p1.p2.monitoring"),
          p = list(cons.dis_m1.dc[[1]][2:4,4], cons.dis_m1c[[1]][2:4,4], cons.dis_m2.dc[[1]][2:4,4], co
          se = list(cons.dis_m1.dc[[1]][2:4,2], cons.dis_m1c[[1]][2:4,2], cons.dis_m2.dc[[1]][2:4,2], c
          add.lines = list(c("Covariates", "Yes", "Yes","Yes","Yes")),
```

```
          notes.label = "Note: two-tailed tests",
          column.sep.width = "1pt",
          df = FALSE,intercept.bottom = TRUE)
```

```
##
## % Table created by stargazer v.5.2.2 by Marek Hlavac, Harvard University. E-mail: hlavac at fas.harva
## % Date and time: Tue, Jun 09, 2020 - 13:15:07
## \begin{table}[!htbp] \centering
##   \caption{Estimated Effects of Treatment Conditional on Zone-Level Dissatisfaction at Baseline}
##   \label{dissatisfaction}
## \begin{tabular}{@{\extracolsep{1pt}}lcccc}
## \\[-1.8ex]\hline
## \hline \\[-1.8ex]
##  & \multicolumn{4}{c}{DV: Cleaned (0: Yes, 1: No) or Waste Pile Size ($m^{2}$)} \\
## \cline{2-5}
## \\[-1.8ex] & M1 Cleaned & M1 Size & M2 Cleaned & M2 Size \\
## \\[-1.8ex] & (1) & (2) & (3) & (4)\\
## \hline \\[-1.8ex]
##  Dissatisfaction & 0.058 & $-$0.536 & 0.019 & 32.067 \\
##   & (0.044) & (4.397) & (0.062) & (26.648) \\
##   & & & & \\
##  Baseline Pile Area &  & 0.163 &  & 0.067 \\
##   &  & (0.103) &  & (0.042) \\
##   & & & & \\
##  P1/P2 Monitoring & 0.072 & $-$0.594 & $-$0.043 & $-$13.759 \\
##   & (0.051) & (3.078) & (0.069) & (13.123) \\
##   & & & & \\
## \hline \\[-1.8ex]
## Covariates & Yes & Yes & Yes & Yes \\
## Observations & 328 & 328 & 328 & 328 \\
## R$^{2}$ & 0.059 & 0.243 & 0.026 & 0.029 \\
## Adjusted R$^{2}$ & 0.029 & 0.216 & $-$0.005 & $-$0.005 \\
## Residual Std. Error & 0.333 & 28.368 & 0.405 & 114.301 \\
## F Statistic & 1.973$^{**}$ & 9.207$^{***}$ & 0.832 & 0.844 \\
## \hline
## \hline \\[-1.8ex]
## Note: two-tailed tests & \multicolumn{4}{r}{$^{*}$p$<$0.1; $^{**}$p$<$0.05; $^{***}$p$<$0.01} \\
## \end{tabular}
## \end{table}
```

```
## Table J7: Subgroup Effects; Reported Zone-Level Consistency -----
over.bad_m1 <- lm(m1.size_final~prop.deviant_overall+b.pile.area_m+p1.p2.monitoring+lights.mean+road.de
summary(over.bad_m1)
```

```
##
## Call:
## lm(formula = m1.size_final ~ prop.deviant_overall + b.pile.area_m +
##     p1.p2.monitoring + lights.mean + road.density + div + area.km2 +
##     ls.pop_2016, data = hte_dta)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -171.363  -11.020   -4.351    2.692  300.381
##
## Coefficients:
```

```
##                         Estimate Std. Error t value Pr(>|t|)
## (Intercept)           20.7153474 15.7779046   1.313    0.190
## prop.deviant_overall -8.3101482 21.9075152  -0.379    0.705
## b.pile.area_m         0.1662168  0.0182901   9.088   <2e-16 ***
## p1.p2.monitoring     -0.1234781  4.1970894  -0.029    0.977
## lights.mean          -0.1056540  0.2465255  -0.429    0.669
## road.density         -0.0001746  0.0002960  -0.590    0.556
## divkawempe            9.7225603  6.1685871   1.576    0.116
## divmakindye           2.9014514  6.5353414   0.444    0.657
## divnakawa             5.2849917  6.3836917   0.828    0.408
## divrubaga             0.9527995  6.1176940   0.156    0.876
## area.km2             -7.5627990  4.8176723  -1.570    0.117
## ls.pop_2016          -0.0001799  0.0001761  -1.021    0.308
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 30.93 on 332 degrees of freedom
##   (4 observations deleted due to missingness)
## Multiple R-squared:  0.2178, Adjusted R-squared:  0.1919
## F-statistic: 8.405 on 11 and 332 DF,  p-value: 4.546e-13
```

```
over.bad_m1c <- lm.cluster(over.bad_m1, hte_dta$zone.id)

over.bad_m2 <- lm(m2.size_final~prop.deviant_overall+b.pile.area_m+p1.p2.monitoring+lights.mean+road.de
summary(over.bad_m2)
```

```
##
## Call:
## lm(formula = m2.size_final ~ prop.deviant_overall + b.pile.area_m +
##     p1.p2.monitoring + lights.mean + road.density + div + area.km2 +
##     ls.pop_2016, data = hte_dta[hte_dta$delt.b_m2pile.area_new <
##     6000, ])
##
## Residuals:
##     Min      1Q  Median      3Q     Max
##  -89.91  -25.27  -12.48    1.51 1930.46
##
## Coefficients:
##                         Estimate Std. Error t value Pr(>|t|)
## (Intercept)           6.793e+01  5.956e+01   1.141    0.255
## prop.deviant_overall -1.314e+02  8.269e+01  -1.589    0.113
## b.pile.area_m         9.214e-02  6.904e-02   1.335    0.183
## p1.p2.monitoring     -6.156e+00  1.584e+01  -0.389    0.698
## lights.mean           2.276e-01  9.305e-01   0.245    0.807
## road.density         -1.213e-03  1.117e-03  -1.086    0.278
## divkawempe            5.178e+00  2.328e+01   0.222    0.824
## divmakindye           2.961e+01  2.467e+01   1.200    0.231
## divnakawa            -1.543e+01  2.410e+01  -0.640    0.523
## divrubaga             1.490e+00  2.309e+01   0.065    0.949
## area.km2             -1.061e+01  1.819e+01  -0.583    0.560
## ls.pop_2016          -1.054e-03  6.649e-04  -1.586    0.114
##
## Residual standard error: 116.7 on 332 degrees of freedom
##   (4 observations deleted due to missingness)
## Multiple R-squared:  0.02495,    Adjusted R-squared:  -0.007357
```

```
## F-statistic: 0.7723 on 11 and 332 DF,  p-value: 0.6678
```

```
over.bad_m2c <- lm.cluster(over.bad_m2, hte_dta$zone.id)

over.bad_m1.d <- lm(m1.waste.pile_d~prop.deviant_overall+p1.p2.monitoring+lights.mean+road.density+div+a
summary(over.bad_m1.d)
```

```
##
## Call:
## lm(formula = m1.waste.pile_d ~ prop.deviant_overall + p1.p2.monitoring +
##     lights.mean + road.density + div + area.km2 + ls.pop_2016,
##     data = hte_dta)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.96766  0.04977  0.11124  0.16382  0.31673
##
## Coefficients:
##                          Estimate Std. Error t value Pr(>|t|)
## (Intercept)           9.899e-01  1.716e-01   5.768 1.83e-08 ***
## prop.deviant_overall -6.632e-02  2.370e-01  -0.280   0.7797
## p1.p2.monitoring      9.686e-02  4.565e-02   2.122   0.0346 *
## lights.mean          -1.054e-03  2.679e-03  -0.393   0.6943
## road.density         -5.316e-06  3.218e-06  -1.652   0.0996 .
## divkawempe           -1.687e-02  6.705e-02  -0.252   0.8015
## divmakindye          -1.649e-01  7.096e-02  -2.324   0.0207 *
## divnakawa            -9.757e-02  6.940e-02  -1.406   0.1606
## divrubaga            -3.025e-02  6.651e-02  -0.455   0.6495
## area.km2              6.921e-02  5.230e-02   1.323   0.1867
## ls.pop_2016           2.231e-06  1.909e-06   1.169   0.2434
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.3364 on 333 degrees of freedom
##   (4 observations deleted due to missingness)
## Multiple R-squared:  0.05433,    Adjusted R-squared:  0.02593
## F-statistic: 1.913 on 10 and 333 DF,  p-value: 0.04257
```

```
over.bad_m1.dc <- lm.cluster(over.bad_m1.d, hte_dta$zone.id)

over.bad_m2.d <- lm(m2.waste.pile_d~prop.deviant_overall+p1.p2.monitoring+lights.mean+road.density+div+a
summary(over.bad_m2.d)
```

```
##
## Call:
## lm(formula = m2.waste.pile_d ~ prop.deviant_overall + p1.p2.monitoring +
##     lights.mean + road.density + div + area.km2 + ls.pop_2016,
##     data = hte_dta)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.93959  0.09918  0.18463  0.23627  0.43911
##
## Coefficients:
##                          Estimate Std. Error t value Pr(>|t|)
```

```
## (Intercept)           1.269e+00  2.070e-01   6.130 2.49e-09 ***
## prop.deviant_overall -3.388e-01  2.858e-01  -1.185   0.2367
## p1.p2.monitoring     -1.392e-02  5.507e-02  -0.253   0.8007
## lights.mean          -4.438e-03  3.232e-03  -1.373   0.1707
## road.density         -4.492e-06  3.882e-06  -1.157   0.2481
## divkawempe           -8.746e-02  8.088e-02  -1.081   0.2803
## divmakindye          -1.586e-01  8.559e-02  -1.853   0.0647 .
## divnakawa            -1.563e-01  8.371e-02  -1.867   0.0628 .
## divrubaga            -5.718e-02  8.023e-02  -0.713   0.4766
## area.km2             -2.608e-02  6.309e-02  -0.413   0.6796
## ls.pop_2016           1.577e-06  2.302e-06   0.685   0.4938
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.4058 on 333 degrees of freedom
##   (4 observations deleted due to missingness)
## Multiple R-squared:  0.02684,    Adjusted R-squared:  -0.002382
## F-statistic: 0.9185 on 10 and 333 DF,  p-value: 0.5161
```

```
over.bad_m2.dc <- lm.cluster(over.bad_m2.d, hte_dta$zone.id)

stargazer(over.bad_m1.d, over.bad_m1, over.bad_m2.d, over.bad_m2, type = "latex",
          title="Treatment Effect of Citizen Reporting Conditional on Consistency of Zone-Level Reports
          label="overall_table",
          dep.var.caption = "DV: Cleaned (0: Yes, 1: No) or Waste Pile Size ($m^{2}$)",
          dep.var.labels = c("M1 Cleaned","M1 Size","M2 Cleaned","M2 Size"),
          covariate.labels = c("Inconsistency","Baseline Pile Area","P1/P2 Monitoring"),
          keep = c("prop.deviant_overall", "b.pile.area_m", "p1.p2.monitoring"),
          p = list(over.bad_m1.dc[[1]][2:4,4], over.bad_m1c[[1]][2:4,4], over.bad_m2.dc[[1]][2:4,4], ov
          se = list(over.bad_m1.dc[[1]][2:4,2], over.bad_m1c[[1]][2:4,2], over.bad_m2.dc[[1]][2:4,2], o
          add.lines = list(c("Covariates", "Yes", "Yes","Yes","Yes")),
          notes.label = "Note: two-tailed tests",
          column.sep.width = "1pt",
          df = FALSE,intercept.bottom = TRUE)
```

```
##
## % Table created by stargazer v.5.2.2 by Marek Hlavac, Harvard University. E-mail: hlavac at fas.harva
## % Date and time: Tue, Jun 09, 2020 - 13:15:07
## \begin{table}[!htbp] \centering
##   \caption{Treatment Effect of Citizen Reporting Conditional on Consistency of Zone-Level Reports on
##   \label{overall_table}
## \begin{tabular}{@{\extracolsep{1pt}}lcccc}
## \\[-1.8ex]\hline
## \hline \\[-1.8ex]
##  & \multicolumn{4}{c}{DV: Cleaned (0: Yes, 1: No) or Waste Pile Size ($m^{2}$)} \\
## \cline{2-5}
## \\[-1.8ex] & M1 Cleaned & M1 Size & M2 Cleaned & M2 Size \\
## \\[-1.8ex] & (1) & (2) & (3) & (4)\\
## \hline \\[-1.8ex]
##  Inconsistency & $-$0.066 & $-$8.310 & $-$0.339 & $-$131.438 \\
##   & (0.221) & (19.395) & (0.313) & (120.071) \\
##   & & & & \\
##  Baseline Pile Area &  & 0.166 &  & 0.092$^{**}$ \\
##   &  & (0.101) &  & (0.043) \\
##   & & & & \\
```

```
##  P1/P2 Monitoring & 0.097$^{*}$ & $-$0.123 & $-$0.014 & $-$6.156 \\
##   & (0.051) & (3.002) & (0.067) & (8.908) \\
##   & & & & \\
## \hline \\[-1.8ex]
## Covariates & Yes & Yes & Yes & Yes \\
## Observations & 344 & 344 & 344 & 344 \\
## R$^{2}$ & 0.054 & 0.218 & 0.027 & 0.025 \\
## Adjusted R$^{2}$ & 0.026 & 0.192 & $-$0.002 & $-$0.007 \\
## Residual Std. Error & 0.336 & 30.926 & 0.406 & 116.734 \\
## F Statistic & 1.913$^{**}$ & 8.405$^{***}$ & 0.919 & 0.772 \\
## \hline
## \hline \\[-1.8ex]
## Note: two-tailed tests & \multicolumn{4}{r}{$^{*}$p$<$0.1; $^{**}$p$<$0.05; $^{***}$p$<$0.01} \\
## \end{tabular}
## \end{table}
```

```
## Tables J8 & J9: HTE; Distance to Division HQ -----
disthq_m1 <- lm(m1.size_final~treat*pile_divhq_distKM+b.pile.area_m+p1.p2.monitoring+lights.mean+road.d
summary(disthq_m1)
```

```
##
## Call:
## lm(formula = m1.size_final ~ treat * pile_divhq_distKM + b.pile.area_m +
##     p1.p2.monitoring + lights.mean + road.density + div + area.km2 +
##     ls.pop_2016, data = subsetB_dta)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -276.10  -14.96   -7.25    2.89  564.00
##
## Coefficients:
##                            Estimate Std. Error t value Pr(>|t|)
## (Intercept)             -3.824e+01  2.058e+01  -1.858  0.06364 .
## treat                   -1.408e-01  6.584e+00  -0.021  0.98294
## pile_divhq_distKM        5.426e+00  2.234e+00   2.429  0.01542 *
## b.pile.area_m            1.279e-01  1.262e-02  10.137  < 2e-16 ***
## p1.p2.monitoring        -1.081e+00  3.612e+00  -0.299  0.76481
## lights.mean              7.975e-01  2.755e-01   2.894  0.00392 **
## road.density            -6.233e-04  3.093e-04  -2.015  0.04427 *
## divkawempe               5.352e+00  6.614e+00   0.809  0.41867
## divmakindye              3.292e-01  6.769e+00   0.049  0.96123
## divnakawa                9.115e+00  6.793e+00   1.342  0.18009
## divrubaga               -4.563e-02  6.698e+00  -0.007  0.99457
## area.km2                -2.782e+00  4.184e+00  -0.665  0.50627
## ls.pop_2016             -1.387e-05  1.683e-04  -0.082  0.93439
## treat:pile_divhq_distKM -1.660e+00  2.231e+00  -0.744  0.45716
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 42.59 on 665 degrees of freedom
##   (12 observations deleted due to missingness)
## Multiple R-squared:  0.1586, Adjusted R-squared:  0.1422
## F-statistic: 9.645 on 13 and 665 DF,  p-value: < 2.2e-16
```

```r
disthq_m1c <- lm.cluster(disthq_m1, subsetB_dta$zone.id)

disthq_m2 <- lm(m2.size_final~treat*pile_divhq_distKM+b.pile.area_m+p1.p2.monitoring+lights.mean+road.d
summary(disthq_m2)
```

```
##
## Call:
## lm(formula = m2.size_final ~ treat * pile_divhq_distKM + b.pile.area_m +
##     p1.p2.monitoring + lights.mean + road.density + div + area.km2 +
##     ls.pop_2016, data = subsetB_dta)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -378.78  -27.01  -12.83    4.42 2349.66
##
## Coefficients:
##                            Estimate Std. Error t value Pr(>|t|)
## (Intercept)               4.896e+00  6.698e+01   0.073 0.941748
## treat                    -5.962e+00  2.142e+01  -0.278 0.780867
## pile_divhq_distKM         3.307e+00  7.269e+00   0.455 0.649296
## b.pile.area_m             1.575e-01  4.106e-02   3.837 0.000137 ***
## p1.p2.monitoring         -1.435e+01  1.175e+01  -1.221 0.222378
## lights.mean               8.087e-01  8.965e-01   0.902 0.367369
## road.density             -1.350e-03  1.006e-03  -1.342 0.180116
## divkawempe               -6.950e+00  2.152e+01  -0.323 0.746821
## divmakindye               2.815e+00  2.202e+01   0.128 0.898341
## divnakawa                -1.739e+01  2.210e+01  -0.787 0.431645
## divrubaga                 1.269e+01  2.179e+01   0.582 0.560669
## area.km2                 -1.092e+00  1.361e+01  -0.080 0.936077
## ls.pop_2016              -9.699e-04  5.478e-04  -1.771 0.077086 .
## treat:pile_divhq_distKM -7.458e-01  7.261e+00  -0.103 0.918213
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 138.6 on 665 degrees of freedom
##   (12 observations deleted due to missingness)
## Multiple R-squared:  0.04146,    Adjusted R-squared:  0.02272
## F-statistic: 2.213 on 13 and 665 DF,  p-value: 0.007976
```

```r
disthq_m2c <- lm.cluster(disthq_m2, subsetB_dta$zone.id)

tthq_m1 <- lm(m1.size_final~treat*pile_divhq_tt+b.pile.area_m+p1.p2.monitoring+lights.mean+road.density
summary(tthq_m1)
```

```
##
## Call:
## lm(formula = m1.size_final ~ treat * pile_divhq_tt + b.pile.area_m +
##     p1.p2.monitoring + lights.mean + road.density + div + area.km2 +
##     ls.pop_2016, data = subsetB_dta)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -274.37  -14.67   -7.15    2.39  565.34
##
```

```
## Coefficients:
##                       Estimate Std. Error t value Pr(>|t|)
## (Intercept)         -3.042e+01  2.010e+01  -1.514  0.13059
## treat               -2.055e+00  7.313e+00  -0.281  0.77875
## pile_divhq_tt        3.453e+00  1.771e+00   1.950  0.05162 .
## b.pile.area_m        1.278e-01  1.262e-02  10.126  < 2e-16 ***
## p1.p2.monitoring    -1.620e+00  3.614e+00  -0.448  0.65418
## lights.mean          7.022e-01  2.660e-01   2.640  0.00848 **
## road.density        -6.509e-04  3.092e-04  -2.105  0.03568 *
## divkawempe           5.756e+00  6.597e+00   0.873  0.38325
## divmakindye          8.092e-01  6.760e+00   0.120  0.90476
## divnakawa            8.124e+00  6.854e+00   1.185  0.23635
## divrubaga            8.333e-01  6.660e+00   0.125  0.90047
## area.km2            -2.461e+00  4.192e+00  -0.587  0.55734
## ls.pop_2016         -5.740e-05  1.674e-04  -0.343  0.73181
## treat:pile_divhq_tt -7.693e-01  1.814e+00  -0.424  0.67164
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 42.65 on 665 degrees of freedom
##   (12 observations deleted due to missingness)
## Multiple R-squared:  0.1565, Adjusted R-squared:   0.14
## F-statistic: 9.491 on 13 and 665 DF,  p-value: < 2.2e-16
```

```
tthq_m1c <- lm.cluster(tthq_m1, subsetB_dta$zone.id)

tthq_m2 <- lm(m2.size_final~treat*pile_divhq_tt+b.pile.area_m+p1.p2.monitoring+lights.mean+road.density
summary(tthq_m2)
```

```
##
## Call:
## lm(formula = m2.size_final ~ treat * pile_divhq_tt + b.pile.area_m +
##     p1.p2.monitoring + lights.mean + road.density + div + area.km2 +
##     ls.pop_2016, data = subsetB_dta)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -377.89  -27.41  -13.15    4.61 2350.51
##
## Coefficients:
##                       Estimate Std. Error t value Pr(>|t|)
## (Intercept)          2.144e+01  6.531e+01   0.328 0.742819
## treat               -1.200e+01  2.377e+01  -0.505 0.613857
## pile_divhq_tt        5.750e-01  5.755e+00   0.100 0.920440
## b.pile.area_m        1.575e-01  4.103e-02   3.839 0.000135 ***
## p1.p2.monitoring    -1.465e+01  1.175e+01  -1.247 0.212673
## lights.mean          6.310e-01  8.644e-01   0.730 0.465645
## road.density        -1.356e-03  1.005e-03  -1.349 0.177641
## divkawempe          -6.313e+00  2.144e+01  -0.294 0.768495
## divmakindye          3.525e+00  2.197e+01   0.160 0.872572
## divnakawa           -1.769e+01  2.228e+01  -0.794 0.427329
## divrubaga            1.355e+01  2.165e+01   0.626 0.531667
## area.km2            -8.535e-01  1.362e+01  -0.063 0.950061
## ls.pop_2016         -1.005e-03  5.441e-04  -1.847 0.065190 .
## treat:pile_divhq_tt  1.119e+00  5.895e+00   0.190 0.849498
```

```
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 138.6 on 665 degrees of freedom
##   (12 observations deleted due to missingness)
## Multiple R-squared:  0.04128,    Adjusted R-squared:  0.02254
## F-statistic: 2.202 on 13 and 665 DF,  p-value: 0.008314
```

```
tthq_m2c <- lm.cluster(tthq_m2, subsetB_dta$zone.id)

disthq_m1.d <- lm(m1.waste.pile_d~treat*pile_divhq_distKM+b.pile.area_m+p1.p2.monitoring+lights.mean+roa
summary(disthq_m1.d)
```

```
##
## Call:
## lm(formula = m1.waste.pile_d ~ treat * pile_divhq_distKM + b.pile.area_m +
##     p1.p2.monitoring + lights.mean + road.density + div + area.km2 +
##     ls.pop_2016, data = subsetB_dta)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.98555  0.06396  0.11472  0.16012  0.28398
##
## Coefficients:
##                            Estimate Std. Error t value Pr(>|t|)
## (Intercept)              9.717e-01  1.620e-01   5.997 3.29e-09 ***
## treat                   -2.518e-02  5.182e-02  -0.486   0.6272
## pile_divhq_distKM       -1.604e-04  1.758e-02  -0.009   0.9927
## b.pile.area_m           -9.808e-05  9.933e-05  -0.987   0.3238
## p1.p2.monitoring         5.048e-02  2.843e-02   1.776   0.0763 .
## lights.mean             -1.023e-03  2.169e-03  -0.472   0.6372
## road.density            -3.803e-06  2.434e-06  -1.562   0.1187
## divkawempe              -2.571e-02  5.206e-02  -0.494   0.6216
## divmakindye             -1.348e-01  5.328e-02  -2.530   0.0117 *
## divnakawa               -8.137e-02  5.346e-02  -1.522   0.1285
## divrubaga               -5.348e-02  5.272e-02  -1.014   0.3108
## area.km2                 5.512e-02  3.293e-02   1.674   0.0946 .
## ls.pop_2016              2.535e-06  1.325e-06   1.913   0.0561 .
## treat:pile_divhq_distKM  9.302e-03  1.756e-02   0.530   0.5966
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.3352 on 665 degrees of freedom
##   (12 observations deleted due to missingness)
## Multiple R-squared:  0.03357,    Adjusted R-squared:  0.01467
## F-statistic: 1.777 on 13 and 665 DF,  p-value: 0.043
```

```
disthq_m1.dc <- lm.cluster(disthq_m1.d, subsetB_dta$zone.id)

disthq_m2.d <- lm(m2.waste.pile_d~treat*pile_divhq_distKM+b.pile.area_m+p1.p2.monitoring+lights.mean+roa
summary(disthq_m2.d)
```

```
##
## Call:
## lm(formula = m2.waste.pile_d ~ treat * pile_divhq_distKM + b.pile.area_m +
```

```
##     p1.p2.monitoring + lights.mean + road.density + div + area.km2 +
##     ls.pop_2016, data = subsetB_dta)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -0.9641  0.1224  0.1812  0.2351  0.5066
##
## Coefficients:
##                            Estimate Std. Error t value Pr(>|t|)
## (Intercept)              9.672e-01  1.967e-01   4.916 1.11e-06 ***
## treat                   -3.969e-02  6.292e-02  -0.631   0.5284
## pile_divhq_distKM       -1.141e-02  2.135e-02  -0.535   0.5931
## b.pile.area_m           -2.242e-04  1.206e-04  -1.859   0.0635 .
## p1.p2.monitoring        -6.240e-03  3.452e-02  -0.181   0.8566
## lights.mean             -7.537e-04  2.633e-03  -0.286   0.7748
## road.density            -5.189e-06  2.956e-06  -1.756   0.0796 .
## divkawempe              -4.081e-02  6.321e-02  -0.646   0.5187
## divmakindye             -1.435e-01  6.469e-02  -2.218   0.0269 *
## divnakawa               -9.243e-02  6.492e-02  -1.424   0.1550
## divrubaga               -6.948e-02  6.402e-02  -1.085   0.2781
## area.km2                 4.353e-02  3.998e-02   1.089   0.2766
## ls.pop_2016              1.942e-06  1.609e-06   1.207   0.2279
## treat:pile_divhq_distKM  1.520e-02  2.133e-02   0.713   0.4762
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.4071 on 665 degrees of freedom
##   (12 observations deleted due to missingness)
## Multiple R-squared:  0.02388,    Adjusted R-squared:  0.004794
## F-statistic: 1.251 on 13 and 665 DF,  p-value: 0.2384
```

```
disthq_m2.dc <- lm.cluster(disthq_m2.d, subsetB_dta$zone.id)

tthq_m1.d <- lm(m1.waste.pile_d~treat*pile_divhq_tt+b.pile.area_m+p1.p2.monitoring+lights.mean+road.dens
summary(tthq_m1.d)
```

```
##
## Call:
## lm(formula = m1.waste.pile_d ~ treat * pile_divhq_tt + b.pile.area_m +
##     p1.p2.monitoring + lights.mean + road.density + div + area.km2 +
##     ls.pop_2016, data = subsetB_dta)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.98407  0.06494  0.11444  0.15834  0.28714
##
## Coefficients:
##                       Estimate Std. Error t value Pr(>|t|)
## (Intercept)          9.422e-01  1.580e-01   5.965 3.98e-09 ***
## treat               -2.334e-02  5.748e-02  -0.406   0.6848
## pile_divhq_tt        3.575e-03  1.392e-02   0.257   0.7974
## b.pile.area_m       -9.982e-05  9.923e-05  -1.006   0.3148
## p1.p2.monitoring     4.978e-02  2.841e-02   1.752   0.0802 .
## lights.mean         -6.646e-04  2.091e-03  -0.318   0.7507
## road.density        -3.877e-06  2.431e-06  -1.595   0.1111
```

```
## divkawempe          -2.664e-02  5.185e-02  -0.514   0.6076
## divmakindye         -1.344e-01  5.314e-02  -2.529   0.0117 *
## divnakawa           -8.443e-02  5.387e-02  -1.567   0.1176
## divrubaga           -5.366e-02  5.235e-02  -1.025   0.3057
## area.km2             5.341e-02  3.295e-02   1.621   0.1055
## ls.pop_2016          2.474e-06  1.316e-06   1.880   0.0605 .
## treat:pile_divhq_tt  5.754e-03  1.426e-02   0.404   0.6866
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.3352 on 665 degrees of freedom
##   (12 observations deleted due to missingness)
## Multiple R-squared:  0.03377,    Adjusted R-squared:  0.01488
## F-statistic: 1.788 on 13 and 665 DF,  p-value: 0.04126
```

```
tthq_m1.dc <- lm.cluster(tthq_m1.d, subsetB_dta$zone.id)

tthq_m2.d <- lm(m2.waste.pile_d~treat*pile_divhq_tt+b.pile.area_m+p1.p2.monitoring+lights.mean+road.dens
summary(tthq_m2.d)
```

```
##
## Call:
## lm(formula = m2.waste.pile_d ~ treat * pile_divhq_tt + b.pile.area_m +
##     p1.p2.monitoring + lights.mean + road.density + div + area.km2 +
##     ls.pop_2016, data = subsetB_dta)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -0.9449  0.1243  0.1810  0.2345  0.5177
##
## Coefficients:
##                       Estimate Std. Error t value Pr(>|t|)
## (Intercept)          8.196e-01  1.916e-01   4.276 2.18e-05 ***
## treat               -5.617e-02  6.974e-02  -0.805   0.4209
## pile_divhq_tt        6.260e-03  1.689e-02   0.371   0.7110
## b.pile.area_m       -2.249e-04  1.204e-04  -1.868   0.0622 .
## p1.p2.monitoring    -6.606e-03  3.446e-02  -0.192   0.8480
## lights.mean          1.182e-03  2.536e-03   0.466   0.6414
## road.density        -5.207e-06  2.949e-06  -1.766   0.0779 .
## divkawempe          -5.059e-02  6.291e-02  -0.804   0.4216
## divmakindye         -1.474e-01  6.447e-02  -2.287   0.0225 *
## divnakawa           -1.032e-01  6.536e-02  -1.579   0.1149
## divrubaga           -7.744e-02  6.352e-02  -1.219   0.2232
## area.km2             3.463e-02  3.997e-02   0.866   0.3866
## ls.pop_2016          1.946e-06  1.596e-06   1.219   0.2234
## treat:pile_divhq_tt  1.462e-02  1.730e-02   0.845   0.3983
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.4067 on 665 degrees of freedom
##   (12 observations deleted due to missingness)
## Multiple R-squared:  0.02566,    Adjusted R-squared:  0.006608
## F-statistic: 1.347 on 13 and 665 DF,  p-value: 0.1806
```

```r
tthq_m2.dc <- lm.cluster(tthq_m2.d, subsetB_dta$zone.id)

#Table I8
vars <- c("treat", "pile_divhq_distKM", "treat:pile_divhq_distKM", "b.pile.area_m", "p1.p2.monitoring")
stargazer(disthq_m1, disthq_m1.d, disthq_m2, disthq_m2.d, type = "latex",
          title="Treatment Effect of Citizen Reporting Conditional on Distance (km) to Nearest KCCA Divi
          label="table:hte_distance",
          dep.var.caption = "DV: Cleaned (0: Yes, 1: No) or Waste Pile Size ($m^{2}$)",
          dep.var.labels = c("M1 Cleaned","M1 Size","M2 Cleaned","M2 Size"),
          covariate.labels = c("Treatment", "Distance", "Baseline Pile Area","P1/P2 Monitoring", "Treatm
          keep = vars,
          p = list(disthq_m1c[[1]][vars,"Pr(>|t|)"], disthq_m1.dc[[1]][vars,"Pr(>|t|)"], disthq_m2c[[1]]
          se = list(disthq_m1c[[1]][vars,"Std. Error"], disthq_m1.dc[[1]][vars,"Std. Error"], disthq_m2
          add.lines = list(c("Covariates", "Yes", "Yes","Yes","Yes")),
          notes.label = "Note: two-tailed tests",
          column.sep.width = "1pt",
          df = FALSE,intercept.bottom = TRUE)
```

```
##
## % Table created by stargazer v.5.2.2 by Marek Hlavac, Harvard University. E-mail: hlavac at fas.harva
## % Date and time: Tue, Jun 09, 2020 - 13:15:07
## \begin{table}[!htbp] \centering
##   \caption{Treatment Effect of Citizen Reporting Conditional on Distance (km) to Nearest KCCA Divisio
##   \label{table:hte_distance}
## \begin{tabular}{@{\extracolsep{1pt}}lcccc}
## \\[-1.8ex]\hline
## \hline \\[-1.8ex]
##  & \multicolumn{4}{c}{DV: Cleaned (0: Yes, 1: No) or Waste Pile Size ($m^{2}$)} \\
## \cline{2-5}
## \\[-1.8ex] & M1 Cleaned & M1 Size & M2 Cleaned & M2 Size \\
## \\[-1.8ex] & (1) & (2) & (3) & (4)\\
## \hline \\[-1.8ex]
##  Treatment & $-$0.141 & $-$0.025 & $-$5.962 & $-$0.040 \\
##   & (5.490) & (0.060) & (13.625) & (0.075) \\
##   & & & & \\
##  Distance & 5.426$^{**}$ & $-$0.0002 & 3.307 & $-$0.011 \\
##   & (2.753) & (0.019) & (4.852) & (0.022) \\
##   & & & & \\
##  Baseline Pile Area & 0.128$^{**}$ & $-$0.0001 & 0.158$^{**}$ & $-$0.0002$^{*}$ \\
##   & (0.055) & (0.0001) & (0.064) & (0.0001) \\
##   & & & & \\
##  P1/P2 Monitoring & $-$1.081 & 0.050$^{*}$ & $-$14.355 & $-$0.006 \\
##   & (3.262) & (0.030) & (10.894) & (0.038) \\
##   & & & & \\
##  Treatment*Distance & $-$1.660 & 0.009 & $-$0.746 & 0.015 \\
##   & (2.051) & (0.021) & (3.264) & (0.023) \\
##   & & & & \\
## \hline \\[-1.8ex]
## Covariates & Yes & Yes & Yes & Yes \\
## Observations & 679 & 679 & 679 & 679 \\
## R$^{2}$ & 0.159 & 0.034 & 0.041 & 0.024 \\
## Adjusted R$^{2}$ & 0.142 & 0.015 & 0.023 & 0.005 \\
## Residual Std. Error & 42.592 & 0.335 & 138.586 & 0.407 \\
## F Statistic & 9.645$^{***}$ & 1.777$^{**}$ & 2.213$^{***}$ & 1.251 \\
```

```
## \hline
## \hline \\[-1.8ex]
## Note: two-tailed tests & \multicolumn{4}{r}{$^{*}$p$<$0.1; $^{**}$p$<$0.05; $^{***}$p$<$0.01} \\
## \end{tabular}
## \end{table}
```

```
#Table I9
vars <- c("treat", "pile_divhq_tt", "treat:pile_divhq_tt", "b.pile.area_m", "p1.p2.monitoring")
stargazer(tthq_m1, tthq_m1.d, tthq_m2, tthq_m2.d, type = "latex",
          title="Treatment Effect of Citizen Reporting Conditional on Travel Time (min.) to Nearest KCC
          label="table:hte_traveltime",
          dep.var.caption = "DV: Cleaned (0: Yes, 1: No) or Waste Pile Size ($m^{2}$)",
          dep.var.labels = c("M1 Cleaned","M1 Size","M2 Cleaned","M2 Size"),
          covariate.labels = c("Treatment", "Travel Time", "Baseline Pile Area","P1/P2 Monitoring", "Tr
          keep = vars,
          p = list(tthq_m1c[[1]][vars,"Pr(>|t|)"], tthq_m1.dc[[1]][vars,"Pr(>|t|)"], tthq_m2c[[1]][vars
          se = list(tthq_m1c[[1]][vars,"Std. Error"], tthq_m1.dc[[1]][vars,"Std. Error"], tthq_m2c[[1]]
          add.lines = list(c("Covariates", "Yes", "Yes","Yes","Yes")),
          notes.label = "Note: two-tailed tests",
          column.sep.width = "1pt",
          df = FALSE,intercept.bottom = TRUE)
```

```
##
## % Table created by stargazer v.5.2.2 by Marek Hlavac, Harvard University. E-mail: hlavac at fas.harva
## % Date and time: Tue, Jun 09, 2020 - 13:15:07
## \begin{table}[!htbp] \centering
##   \caption{Treatment Effect of Citizen Reporting Conditional on Travel Time (min.) to Nearest KCCA D
##   \label{table:hte_traveltime}
## \begin{tabular}{@{\extracolsep{1pt}}lcccc}
## \\[-1.8ex]\hline
## \hline \\[-1.8ex]
##  & \multicolumn{4}{c}{DV: Cleaned (0: Yes, 1: No) or Waste Pile Size ($m^{2}$)} \\
## \cline{2-5}
## \\[-1.8ex] & M1 Cleaned & M1 Size & M2 Cleaned & M2 Size \\
## \\[-1.8ex] & (1) & (2) & (3) & (4)\\
## \hline \\[-1.8ex]
##  Treatment & $-$2.055 & $-$0.023 & $-$11.997 & $-$0.056 \\
##   & (5.216) & (0.065) & (19.969) & (0.083) \\
##   & & & & \\
##  Travel Time & 3.453$^{*}$ & 0.004 & 0.575 & 0.006 \\
##   & (1.894) & (0.016) & (2.874) & (0.019) \\
##   & & & & \\
##  Baseline Pile Area & 0.128$^{**}$ & $-$0.0001 & 0.158$^{**}$ & $-$0.0002$^{*}$ \\
##   & (0.055) & (0.0001) & (0.064) & (0.0001) \\
##   & & & & \\
##  P1/P2 Monitoring & $-$1.620 & 0.050$^{*}$ & $-$14.651 & $-$0.007 \\
##   & (3.216) & (0.030) & (11.015) & (0.038) \\
##   & & & & \\
##  Treatment*Travel Time & $-$0.769 & 0.006 & 1.119 & 0.015 \\
##   & (1.478) & (0.016) & (2.576) & (0.019) \\
##   & & & & \\
## \hline \\[-1.8ex]
## Covariates & Yes & Yes & Yes & Yes \\
## Observations & 679 & 679 & 679 & 679 \\
## R$^{2}$ & 0.156 & 0.034 & 0.041 & 0.026 \\
```

```
## Adjusted R$^{2}$ & 0.140 & 0.015 & 0.023 & 0.007 \\
## Residual Std. Error & 42.646 & 0.335 & 138.599 & 0.407 \\
## F Statistic & 9.491$^{***}$ & 1.788$^{**}$ & 2.202$^{***}$ & 1.347 \\
## \hline
## \hline \\[-1.8ex]
## Note: two-tailed tests & \multicolumn{4}{r}{$^{*}$p$<$0.1; $^{**}$p$<$0.05; $^{***}$p$<$0.01} \\
## \end{tabular}
## \end{table}
```

```
## Table J10: Summary Statistics for Core Regressions -----
vars_d <- c("unique.id", "zone.id", "treat",
            "b.waste.pile_d", "m1.waste.pile_d", "m2.waste.pile_d",
            "p1.p2.monitoring", "lights.mean", "road.density", "area.km2", "ls.pop_2016", "div")
vars_s <- c("unique.id", "zone.id", "treat",
            "b.pile.area_m", "m1.size_final", "m2.size_final",
            "m1.ea.uwa1", "m1.ea.uwa2", "m2.ea.uwa1", "m2.ea.uwa2",
            "m1.ea.brn1", "m1.ea.brn2", "m2.ea.brn1", "m2.ea.brn2",
            "m1.ea.po1", "m1.ea.po2", "m2.ea.po1", "m2.ea.po2",
            "p1.p2.monitoring", "lights.mean", "road.density", "area.km2", "ls.pop_2016", "div")
summ.df_d <- subsetA_dta[,names(subsetA_dta)%in%vars_d]
summ.df_s <- subsetB_dta[,names(subsetB_dta)%in%vars_s]

out_s <- c("b.pile.area_m", "m1.size_final", "m2.size_final",
           "m1.ea.uwa1", "m1.ea.uwa2", "m2.ea.uwa1", "m2.ea.uwa2",
           "m1.ea.brn1", "m1.ea.brn2", "m2.ea.brn1", "m2.ea.brn2",
           "m1.ea.po1", "m1.ea.po2", "m2.ea.po1", "m2.ea.po2")
out_d <- c("b.waste.pile_d", "m1.waste.pile_d", "m2.waste.pile_d")
cov_c <- c("lights.mean", "road.density", "area.km2", "ls.pop_2016")
cov_f <- c("p1.p2.monitoring", "div")

#Columns should be mean, min, max
outcome_size <- data.frame(Type=c("Outcome", rep("", length(out_s)-1)),
                           Mean=sapply(summ.df_s[,names(summ.df_s)%in%out_s], mean, na.rm=T),
                           Min=sapply(summ.df_s[,names(summ.df_s)%in%out_s], min, na.rm=T),
                           Max=sapply(summ.df_s[,names(summ.df_s)%in%out_s], max, na.rm=T),
                           Modal="-")

outcome_pile <- data.frame(Type=c("Outcome", rep("", length(out_d)-1)),
                           Mean="-",
                           Min="-",
                           Max="-",
                           Modal=sapply(summ.df_d[,names(summ.df_d)%in%out_d], modal, na.rm=T))

covariate_c <- data.frame(Type=c("Covariate", rep("", length(cov_c)-1)),
                          Mean=sapply(summ.df_s[,names(summ.df_s)%in%cov_c], mean, na.rm=T),
                          Min=sapply(summ.df_s[,names(summ.df_s)%in%cov_c], min, na.rm=T),
                          Max=sapply(summ.df_s[,names(summ.df_s)%in%cov_c], max, na.rm=T),
                          Modal="-")

covariate_f <- data.frame(Type=c("Covariate", rep("", length(cov_f)-1)),
                          Mean="-",
                          Min="-",
                          Max="-",
                          Modal=sapply(summ.df_s[,names(summ.df_s)%in%cov_f], modal, na.rm=T))
```

```
master_out <- rbind(outcome_size, outcome_pile, covariate_c, covariate_f)

master_out["m1.waste.pile_d", "Mean"] <- mean(as.numeric(summ.df_d$m1.waste.pile_d), na.rm=T)
master_out["m2.waste.pile_d", "Mean"] <- mean(as.numeric(summ.df_d$m2.waste.pile_d), na.rm=T)
master_out["p1.p2.monitoring", "Mean"] <- mean(as.numeric(summ.df_s$p1.p2.monitoring), na.rm=T)

master_out <- master_out[c("m1.size_final", "m2.size_final",
                           "m1.waste.pile_d", "m2.waste.pile_d",
                           #  "m1.ea.uwa1", "m1.ea.uwa2", "m2.ea.uwa1", "m2.ea.uwa2",
                           #  "m1.ea.brn1", "m1.ea.brn2", "m2.ea.brn1", "m2.ea.brn2",
                           # "m1.ea.po1", "m1.ea.po2", "m2.ea.po1", "m2.ea.po2",
                           "div", "p1.p2.monitoring", "b.pile.area_m",
                           "lights.mean", "road.density", "area.km2", "ls.pop_2016"),]
row.names(master_out) <- c("Pile Size, M1 $(m^2)$", "Pile Size, M2 $(m^2)$",
                           "Pile Dummy, M1", "Pile Dummy, M2",
                           "Division", "P1/P2 Monitoring", "Pile Area, Baseline $(m^2)$",
                           "Zone-NTL $(nW cm^{???2} sr^{???1})$", "Zone-Road Density", "Zone-Area $(km^

xtable(master_out)
```

```
## % latex table generated in R 3.6.1 by xtable 1.8-4 package
## % Tue Jun  9 13:15:08 2020
## \begin{table}[ht]
## \centering
## \begin{tabular}{rllllll}
##   \hline
##  & Type & Mean & Min & Max & Modal \\
##   \hline
## Pile Size, M1 \$(m\verb|^|2)\$ &  & 18.2898885672938 & 0 & 600 & - \\
##   Pile Size, M2 \$(m\verb|^|2)\$ &  & 25.6546309696093 & 0 & 2500 & - \\
##   Pile Dummy, M1 &  & 0.862266857962697 & - & - & 1 \\
##   Pile Dummy, M2 &  & 0.781922525107604 & - & - & 1 \\
##   Division & Covariate & - & - & - & nakawa \\
##   P1/P2 Monitoring &  & 0.331403762662808 & - & - & 0 \\
##   Pile Area, Baseline \$(m\verb|^|2)\$ & Outcome & 38.0982272069465 & 0.5 & 2000 & - \\
##   Zone-NTL \$(nW cm\verb|^|\{???2\} sr\verb|^|\{???1\})\$ & Covariate & 55.6619302995723 & 12.333333
##   Zone-Road Density &  & 10776.9395288664 & 833.8710321534 & 53072.9150024269 & - \\
##   Zone-Area \$(km\verb|^|2)\$ &  & 0.338752532561505 & 0.0085 & 4.0021 & - \\
##   Zone-Population &  & 12668.8335021708 & 45 & 53662 & - \\
##    \hline
## \end{tabular}
## \end{table}
```

```
## Table J11: RI, Diff-in-Diff Specification -----
did.dta <- data.frame(pile.id = rep(as.character(unique(subsetB_dta$unique.id)), each=3),
                      audit = rep(c("b", "m1", "m2"), length(unique(subsetB_dta$unique.id))),
                      post.trt = NA,
                      zone.id = NA,
                      treat = NA,
                      p1.p2.monitoring = NA,
                      pile.size = NA,
                      pile.d = NA,
                      lights.mean = NA,
                      road.density = NA,
```

```
                      div = NA,
                      area.km2 = NA,
                      ls.pop_2016 = NA)

id <- as.character(unique(did.dta$pile.id))

for(i in 1:length(id)) {
  did.dta$pile.size[did.dta$audit=="b"&did.dta$pile.id==id[i]] <- subsetB_dta$b.pile.area_m[subsetB_dta$
  did.dta$pile.size[did.dta$audit=="m1"&did.dta$pile.id==id[i]] <- subsetB_dta$m1.size_final[subsetB_dta
  did.dta$pile.size[did.dta$audit=="m2"&did.dta$pile.id==id[i]] <- subsetB_dta$m2.size_final[subsetB_dta

  did.dta$pile.d[did.dta$audit=="b"&did.dta$pile.id==id[i]] <- subsetB_dta$b.waste.pile_d[subsetB_dta$un
  did.dta$pile.d[did.dta$audit=="m1"&did.dta$pile.id==id[i]] <- subsetB_dta$m1.waste.pile_d[subsetB_dta$
  did.dta$pile.d[did.dta$audit=="m2"&did.dta$pile.id==id[i]] <- subsetB_dta$m2.waste.pile_d[subsetB_dta$

  did.dta$zone.id[did.dta$pile.id==id[i]] <- subsetB_dta$zone.id[subsetB_dta$unique.id==id[i]]
  did.dta$treat[did.dta$pile.id==id[i]] <- subsetB_dta$treat[subsetB_dta$unique.id==id[i]]
  did.dta$p1.p2.monitoring[did.dta$pile.id==id[i]] <- subsetB_dta$p1.p2.monitoring[subsetB_dta$unique.id
  did.dta$lights.mean[did.dta$pile.id==id[i]] <- subsetB_dta$lights.mean[subsetB_dta$unique.id==id[i]]
  did.dta$road.density[did.dta$pile.id==id[i]] <- subsetB_dta$road.density[subsetB_dta$unique.id==id[i]]
  did.dta$div[did.dta$pile.id==id[i]] <- subsetB_dta$div[subsetB_dta$unique.id==id[i]]
  did.dta$area.km2[did.dta$pile.id==id[i]] <- subsetB_dta$area.km2[subsetB_dta$unique.id==id[i]]
  did.dta$ls.pop_2016[did.dta$pile.id==id[i]] <- subsetB_dta$ls.pop_2016[subsetB_dta$unique.id==id[i]]
}

did.dta$post.trt <- ifelse(did.dta$audit=="b", "p0", "p1")
summary(did.dta)
```

```
##                    pile.id     audit      post.trt            zone.id
##  agip_makindye_kibuli_a:   3   b :691   Length:2073        Min.   :  0.0
##  agip_makindye_kibuli_b:   3   m1:691   Class :character   1st Qu.:144.0
##  agip_makindye_kibuli_c:   3   m2:691   Mode  :character   Median :341.0
##  agip_makindye_kibuli_d:   3                               Mean   :344.4
##  b4_nakawa_banda_a     :   3                               3rd Qu.:518.0
##  b4_nakawa_banda_b     :   3                               Max.   :750.0
##  (Other)               :2055
##      treat        p1.p2.monitoring   pile.size          pile.d
##  Min.   :0.0000   Min.   :0.0000   Min.   :   0.00   Length:2073
##  1st Qu.:0.0000   1st Qu.:0.0000   1st Qu.:   3.00   Class :character
##  Median :1.0000   Median :0.0000   Median :   8.00   Mode  :character
##  Mean   :0.5036   Mean   :0.3314   Mean   :  27.35
##  3rd Qu.:1.0000   3rd Qu.:1.0000   3rd Qu.:  18.00
##  Max.   :1.0000   Max.   :1.0000   Max.   :2500.00
##
##   lights.mean     road.density         div              area.km2
##  Min.   :12.33   Min.   :  833.9   Length:2073        Min.   :0.0085
##  1st Qu.:55.00   1st Qu.: 6891.0   Class :character   1st Qu.:0.0691
##  Median :61.39   Median : 9634.8   Mode  :character   Median :0.1535
##  Mean   :55.66   Mean   :10776.9                      Mean   :0.3388
##  3rd Qu.:62.00   3rd Qu.:13578.4                      3rd Qu.:0.4407
##  Max.   :63.00   Max.   :53072.9                      Max.   :4.0021
##                  NA's   :36
##   ls.pop_2016
##  Min.   :   45
```

```
##  1st Qu.: 3652
##  Median : 8419
##  Mean   :12669
##  3rd Qu.:19828
##  Max.   :53662
##
```

```
did_size <- did.ri(formula = pile.size~treat+post.trt+treat*post.trt+p1.p2.monitoring+lights.mean+road.d
                   dta = did.dta,
                   treat.var = "treat",
                   clust_var = did.dta$zone.id,
                   m=length(unique(did.dta$zone.id))/2,
                   sims=10000)


did_table <- data.frame(matrix(nrow=4, ncol=1))
rownames(did_table) <- c("Treatment*Post-Treatment","Standard Error", "p-value", "N")
colnames(did_table) <- c("Pile Size")

did_table[1,1] <- did_size$ate[[1]]
did_table[2,1] <- did_size$se[[1]]
did_table[3,1] <- did_size$p.one.way.lesser[[1]]
did_table[4,1] <- did_size$N[[1]]

out <- list(did_table)
attr(out, "message") <- c("Note: results calculated using cleaned waste pile size measurements.")
print(xtableList(out, caption="RI Results: Diff-in-Diff, Pile Size (Cleaned)"), caption.placement="top"
```

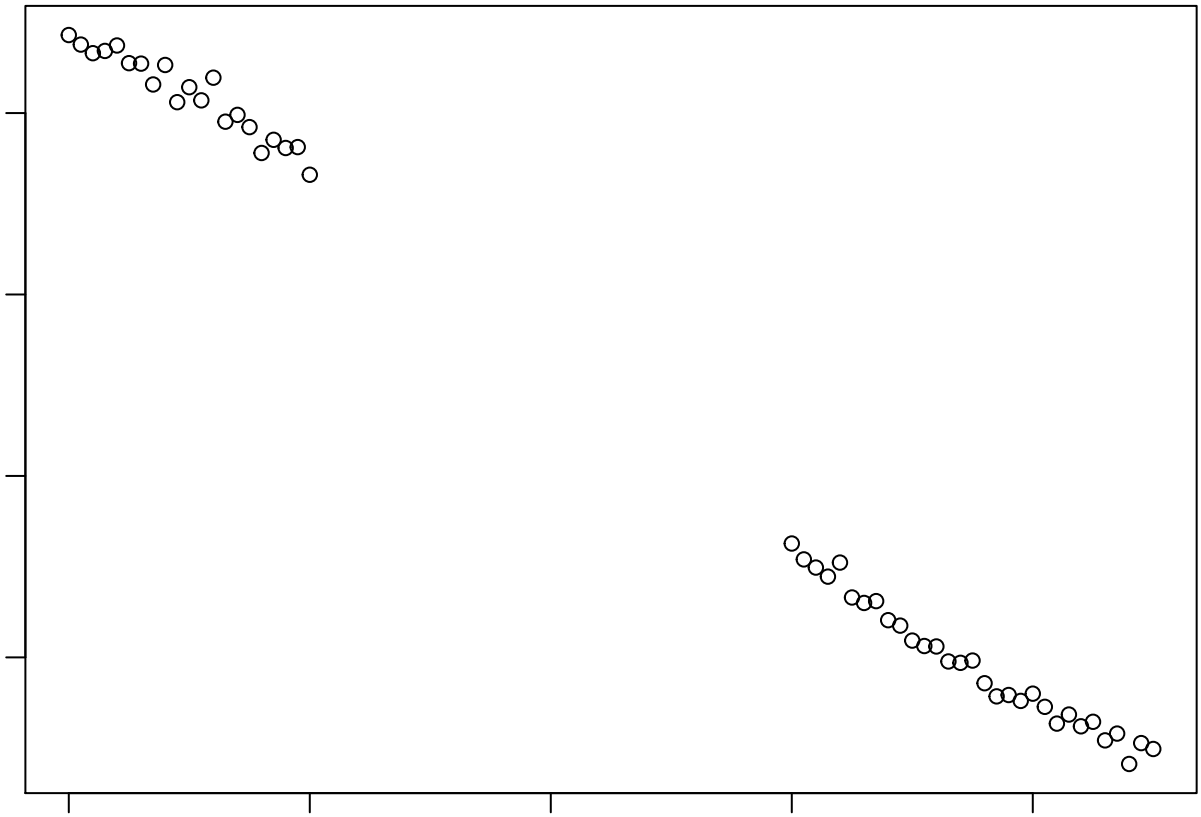
```
## % latex table generated in R 3.6.1 by xtable 1.8-4 package
## % Tue Jun  9 13:15:45 2020
## \begin{table}[ht]
## \centering
## \caption{RI Results: Diff-in-Diff, Pile Size (Cleaned)}
## \begin{tabular}{rr}
##   \hline
##  & Pile Size \\
##   \hline
## Treatment*Post-Treatment & -5.34 \\
##   Standard Error & 0.69 \\
##   p-value & 0.02 \\
##   N & 2037.00 \\
##    \hline
## \multicolumn{2}{l}{Note: results calculated using cleaned waste pile size measurements.}\\
## \end{tabular}
## \end{table}
```

```
## Table K1: Minimum Detectable Effect ----
ri.cph3.analysis <- subsetB_dta

set.seed(202)
#Waste Accumulation
ri.m1.size_new <- lm.ri(formula = m1.size_final~treat+p1.p2.monitoring+b.pile.area_m+lights.mean+road.d
                        dta = ri.cph3.analysis,
                        treat.var = "treat",
                        clust_var = ri.cph3.analysis$zone.id,
```

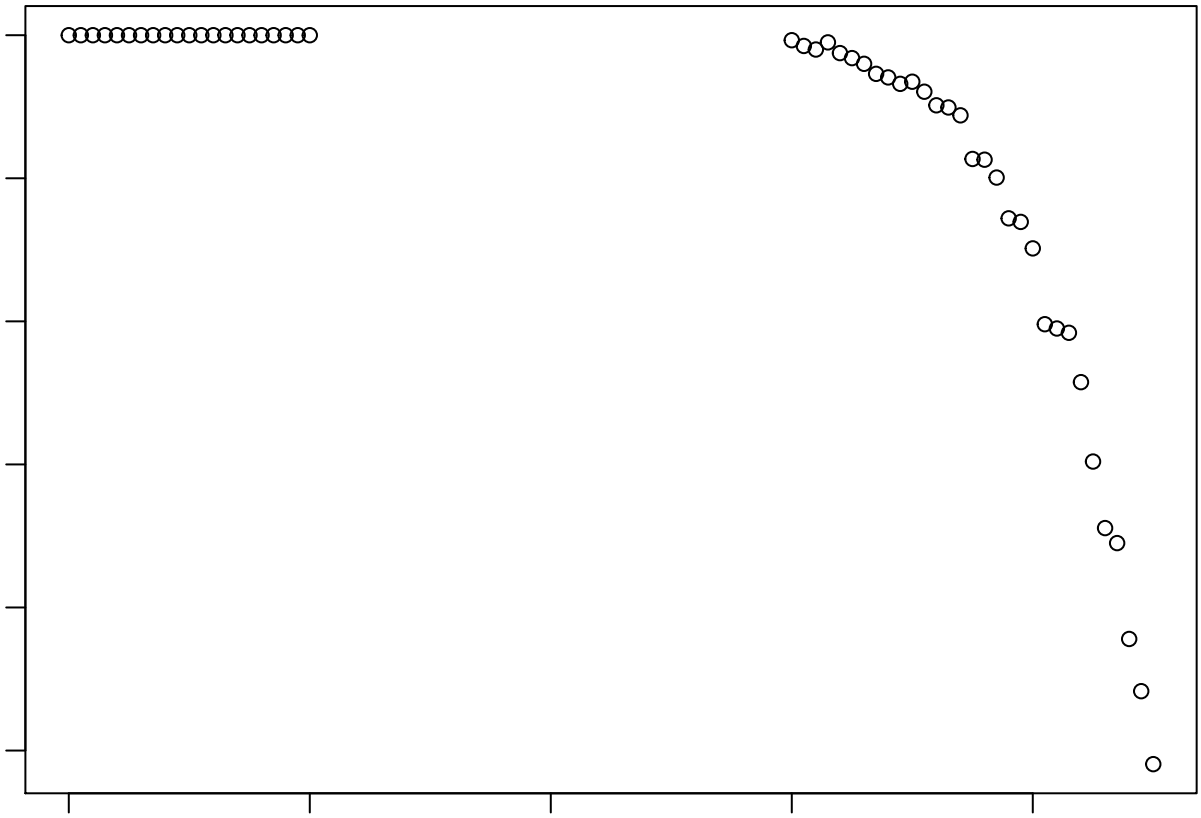
```
                        m=length(unique(ri.cph3.analysis$zone.id))/2,
                        sims=10000) #This runs with NAs removed from the treat and cluster variables

ri.m2.size_new <- lm.ri(formula = m2.size_final~treat+p1.p2.monitoring+b.pile.area_m+lights.mean+road.d
                        dta = ri.cph3.analysis,
                        treat.var = "treat",
                        clust_var = ri.cph3.analysis$zone.id,
                        m=length(unique(ri.cph3.analysis$zone.id))/2,
                        sims=10000)


#Pile Dummies
ri.cph3.analysis$m1.waste.pile_d2 <- ifelse(ri.cph3.analysis$m1.waste.pile_d2=="Yes", 1, ri.cph3.analysi
ri.cph3.analysis$m1.waste.pile_d2 <- ifelse(ri.cph3.analysis$m1.waste.pile_d2=="No", 0, ri.cph3.analysis
ri.wpd1_adj <- lm.ri(formula = m1.waste.pile_d2~treat+p1.p2.monitoring+lights.mean+road.density+div+area
                     dta = ri.cph3.analysis,
                     treat.var = "treat",
                     clust_var = ri.cph3.analysis$zone.id,
                     m=length(unique(ri.cph3.analysis$zone.id))/2,
                     sims=10000)

ri.cph3.analysis$m2.waste.pile_d2 <- ifelse(ri.cph3.analysis$m2.waste.pile_d2=="Yes", 1, ri.cph3.analysi
ri.cph3.analysis$m2.waste.pile_d2 <- ifelse(ri.cph3.analysis$m2.waste.pile_d2=="No", 0, ri.cph3.analysis
ri.wpd2_adj <- lm.ri(formula = m2.waste.pile_d2~treat+p1.p2.monitoring+lights.mean+road.density+div+area
                     dta = ri.cph3.analysis,
                     treat.var = "treat",
                     clust_var = ri.cph3.analysis$zone.id,
                     m=length(unique(ri.cph3.analysis$zone.id))/2,
                     sims=10000)

#Rank Test
ri.rank_m1_new <- lm.ri(formula=rank.m1_new~treat+p1.p2.monitoring+rank.b_new+lights.mean+road.density+a
                        dta=ri.cph3.analysis,
                        treat.var="treat",
                        clust_var = ri.cph3.analysis$zone.id,
                        m=length(unique(ri.cph3.analysis$zone.id))/2,
                        sims=10000)

ri.rank_m2_new <- lm.ri(formula=rank.m2_new~treat+p1.p2.monitoring+rank.b_new+lights.mean+road.density+a
                        dta=ri.cph3.analysis,
                        treat.var="treat",
                        clust_var = ri.cph3.analysis$zone.id,
                        m=length(unique(ri.cph3.analysis$zone.id))/2,
                        sims=10000)


te.seq <- c(seq(from=-1, to=-4, by=-0.1),seq(from=-8, to=-10, by=-0.1))
power.store.size <- rep(NA, length(te.seq))
power.store.dummy <- rep(NA, length(te.seq))
power.store.rank <- rep(NA, length(te.seq))


for (i in 1:length(te.seq)){
```

```
  ri.reverse.m1 <- lm.ri.power.all3(formula.size = m1.size_final~treat+p1.p2.monitoring+b.pile.area_m+li
                                    formula.dummy = m1.waste.pile_d2~treat+p1.p2.monitoring+lights.mean
                                    formula.rank = rank.m1_new~treat+p1.p2.monitoring+rank.b_new+lights
                                    size.bar = quantile(sort(ri.m1.size_new$ate.samp.dist), probs = 0.05
                                    dummy.bar = quantile(sort(ri.wpd1_adj$ate.samp.dist), probs = 0.05)
                                    rank.bar = quantile(sort(ri.rank_m1_new$ate.samp.dist), probs = 0.05
                                    dta = ri.cph3.analysis,
                                    treat.var = "treat",
                                    outcome.var = "m1.size_final",
                                    binary.var = "m1.waste.pile_d2",
                                    rank.var = "rank.m1_new",
                                    sims = 2000,
                                    clust_var = ri.cph3.analysis$zone.id,
                                    m=length(unique(ri.cph3.analysis$zone.id))/2,
                                    te = te.seq[i]
  )

  power.store.size[i] <- ri.reverse.m1$power.size
  power.store.dummy[i] <- ri.reverse.m1$power.dummy
  power.store.rank[i] <- ri.reverse.m1$power.rank
}


te.seq2 <- c(seq(from=-2, to=-5, by=-0.1),seq(from=-8, to=-10, by=-0.1))

power.store.size2 <- rep(NA, length(te.seq2))
power.store.dummy2 <- rep(NA, length(te.seq2))
power.store.rank2 <- rep(NA, length(te.seq2))


for (i in 1:length(te.seq2)){

  ri.reverse.m2 <- lm.ri.power.all3(formula.size = m2.size_final~treat+p1.p2.monitoring+b.pile.area_m+li
                                    formula.dummy = m2.waste.pile_d2~treat+p1.p2.monitoring+lights.mean
                                    formula.rank = rank.m2_new~treat+p1.p2.monitoring+rank.b_new+lights
                                    size.bar = quantile(sort(ri.m2.size_new$ate.samp.dist), probs = 0.05
                                    dummy.bar = quantile(sort(ri.wpd2_adj$ate.samp.dist), probs = 0.05)
                                    rank.bar = quantile(sort(ri.rank_m2_new$ate.samp.dist), probs = 0.05
                                    dta = ri.cph3.analysis,
                                    treat.var = "treat",
                                    outcome.var = "m2.size_final",
                                    binary.var = "m2.waste.pile_d2",
                                    rank.var = "rank.m2_new",
                                    sims = 2000,
                                    clust_var = ri.cph3.analysis$zone.id,
                                    m=length(unique(ri.cph3.analysis$zone.id))/2,
                                    te = te.seq2[i]
  )

  power.store.size2[i] <- ri.reverse.m2$power.size
  power.store.dummy2[i] <- ri.reverse.m2$power.dummy
  power.store.rank2[i] <- ri.reverse.m2$power.rank
}
```

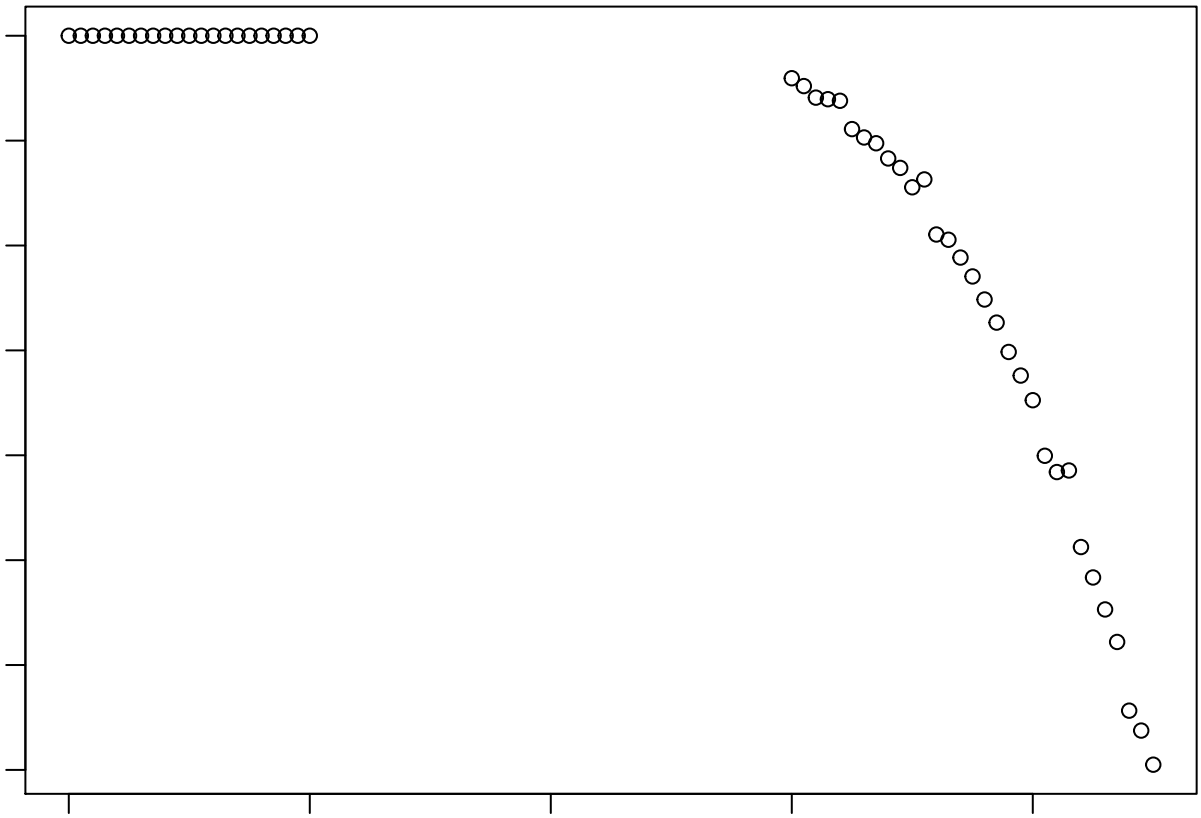
```
#Plots
plot(te.seq,power.store.size)
```



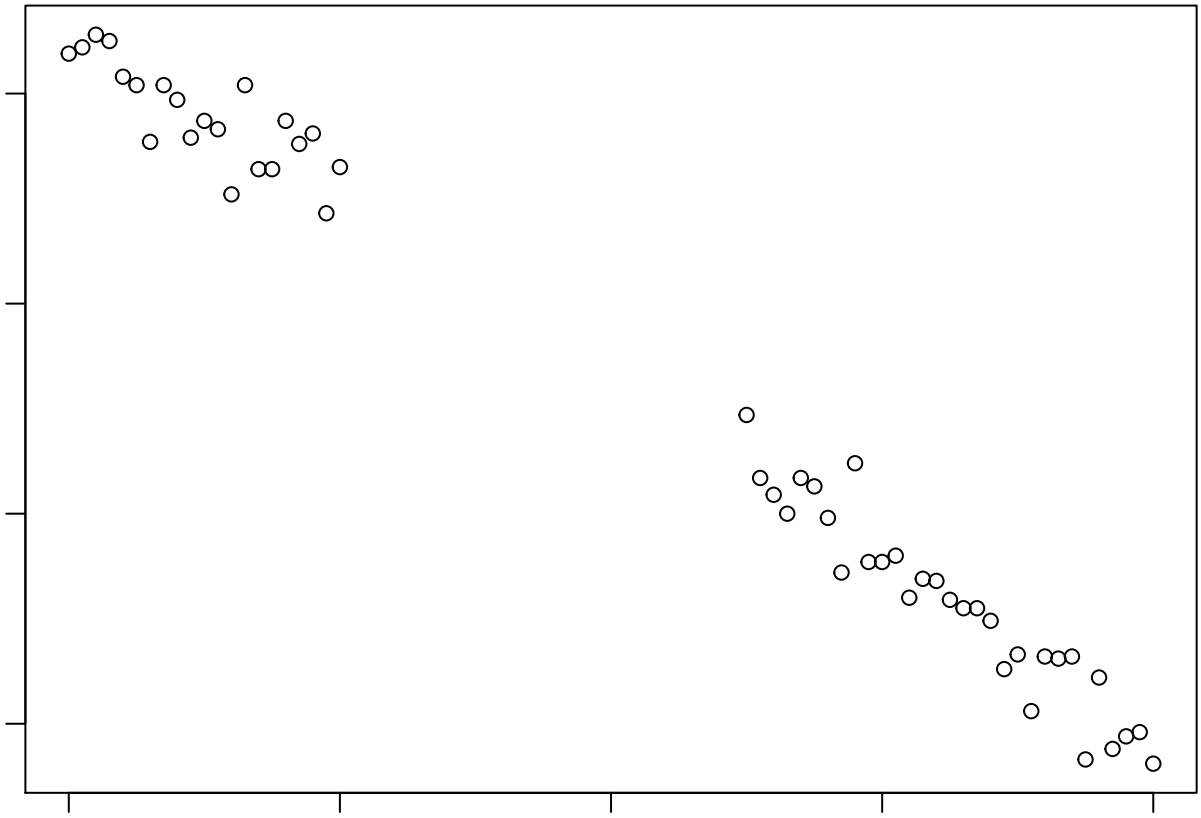
```
plot(te.seq,power.store.dummy)
```



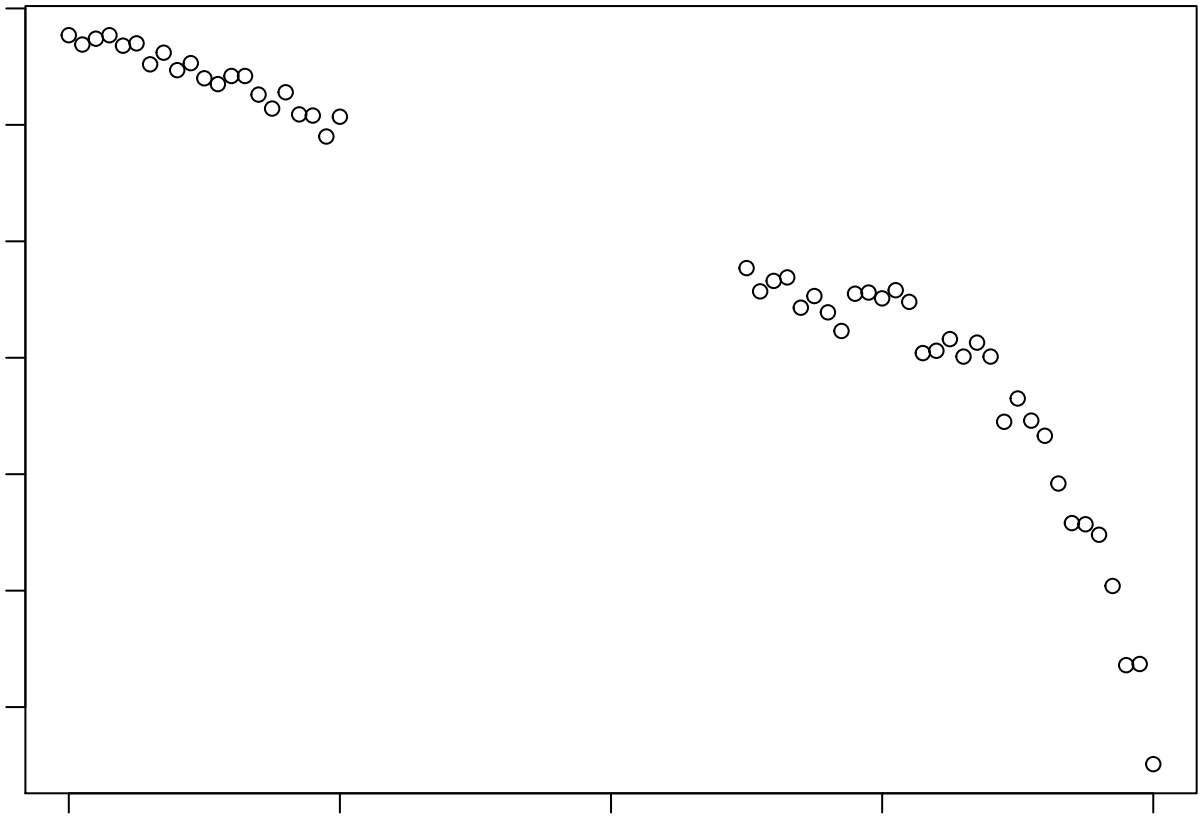
```
plot(te.seq,power.store.rank)
```



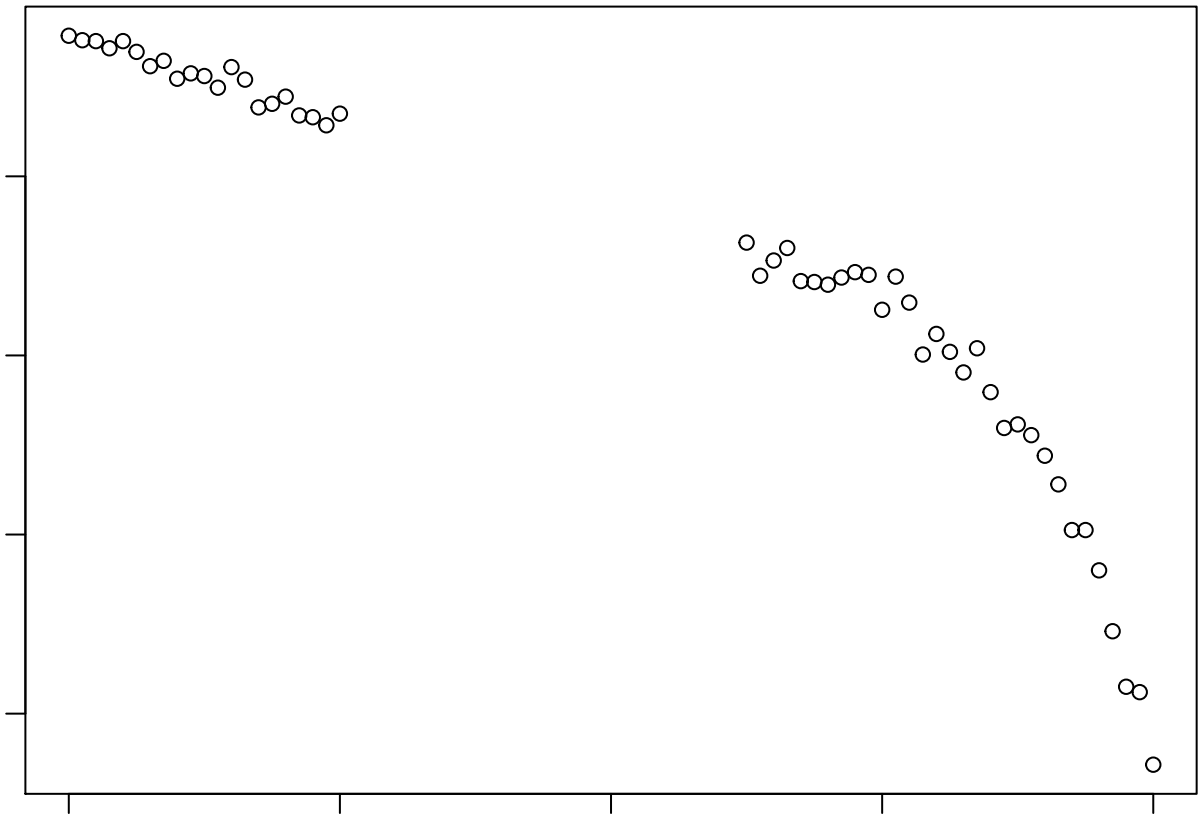
```
plot(te.seq2,power.store.size2)
```



```
plot(te.seq2,power.store.dummy2)
```



```
plot(te.seq2,power.store.rank2)
```



```r
#Table
mid1 <- c(max(te.seq[power.store.size>0.8]),
          max(te.seq[power.store.dummy>0.8]),
          max(te.seq[power.store.rank>0.8])
)

mid1.sd.c <- sd(ri.cph3.analysis$m1.size_final[ri.cph3.analysis$treat==0])

mid1.ses <- -round(mid1/mid1.sd.c,2)

mid2 <- c("--", #max(te.seq[power.store.size2>0.8]),
          max(te.seq[power.store.dummy2>0.8]),
          max(te.seq[power.store.rank2>0.8])
)

mid2.sd.c <- sd(ri.cph3.analysis$m2.size_final[ri.cph3.analysis$treat==0])

mid2.ses <- c("--",-round(as.numeric(mid2[2:3])/mid2.sd.c,2))

#Table
tab.dta <- data.matrix(rbind(mid1,mid1.ses,mid2,mid2.ses))
rownames(tab.dta) <- c("Midline 1","", "Midline 2","")
colnames(tab.dta) <- c("Pile Size","Pile Cleaned","Pile Rank")
xtable(tab.dta, summary=FALSE,
       label = "SI-tab-MDE",
       caption = "Minimum Detectable Effects $(m^2)$",
       align = "lccc") # Tue Dec  4 08:55:21 2018
```

```
## % latex table generated in R 3.6.1 by xtable 1.8-4 package
## % Wed Jun 10 08:12:47 2020
## \begin{table}[ht]
## \centering
## \begin{tabular}{lccc}
##   \hline
##  & Pile Size & Pile Cleaned & Pile Rank \\
##   \hline
## Midline.1 & -8.8 & -2 & -2.7 \\
##   X & 0.16 & 0.04 & 0.05 \\
##   Midline.2 & -- & -1.8 & -2.3 \\
##   X.1 & -- & 0.01 & 0.01 \\
##    \hline
## \end{tabular}
## \caption{Minimum Detectable Effects $(m^2)$}
## \label{SI-tab-MDE}
## \end{table}
```