Replication 2

Hung-Hsiang CHIEN

2020412

Problem 1

Table 1 below is the replication for Table 3 in the paper

```
weighted.var <- function(x, w, na.rm = FALSE) {
    if (na.rm) {
        w <- w[i <- !is.na(x)]
        x <- x[i]
    }
    sum.w <- sum(w)
    sum.w2 <- sum(w^2)
    mean.w <- sum(x * w) / sum(w)
    (sum.w / (sum.w^2 - sum.w2)) * sum(w * (x - mean.w)^2, na.rm =
na.rm)
}



table3_1 =transpose(df1 %>%
  filter(alt == 1, sample_af == 1) %>%
  dummy_columns(select_columns = "mode") %>%
  summarise(Helicopter = weighted.mean(mode_1,weight),
            "Water Taxi" = weighted.mean(mode_2,weight),
            Ferry =  weighted.mean(mode_3,weight),
            Hovercraft = weighted.mean(mode_4,weight),
            Helicopter_SD = sqrt(weighted.var(mode_1,weight)),
            WaterTaxi_SD = sqrt(weighted.var(mode_2,weight)),
            Ferry_SD =  sqrt(weighted.var(mode_3,weight)),
            Hovercrat_SD = sqrt(weighted.var(mode_4,weight))
            ) )

table3_1 =as.data.frame(cbind(table3_1[c(1:4),],table3_1[c(5:8),]))
colnames(table3_1) = c("Mean","SD")

table3_2 = transpose(df1 %>%
  filter(alt == 1, sample_noaf == 1) %>%
  dummy_columns(select_columns = "mode") %>%
  summarise(Helicopter = weighted.mean(mode_1,weight),
            "Water Taxi" = weighted.mean(mode_2,weight),
            Ferry =  weighted.mean(mode_3,weight),
            Hovercraft = weighted.mean(mode_4,weight),
            Helicopter_SD = sqrt(weighted.var(mode_1,weight)),
            WaterTaxi_SD = sqrt(weighted.var(mode_2,weight)),
            Ferry_SD =  sqrt(weighted.var(mode_3,weight)),
            Hovercrat_SD = sqrt(weighted.var(mode_4,weight))
            ) )

table3_2 =as.data.frame(cbind(table3_2[c(1:4),],table3_2[c(5:8),]))
```

```
colnames(table3_2) = c("Mean","SD")

table3_3 = transpose(df1 %>%
  filter(alt == 1, sample_all == 1) %>%
  dummy_columns(select_columns = "mode") %>%
  summarise(Helicopter = weighted.mean(mode_1,weight),
            "Water Taxi" = weighted.mean(mode_2,weight),
            Ferry =  weighted.mean(mode_3,weight),
            Hovercraft = weighted.mean(mode_4,weight),
            Helicopter_SD = sqrt(weighted.var(mode_1,weight)),
            WaterTaxi_SD = sqrt(weighted.var(mode_2,weight)),
            Ferry_SD =  sqrt(weighted.var(mode_3,weight)),
            Hovercrat_SD = sqrt(weighted.var(mode_4,weight))
            ) )


table3_3 =as.data.frame(cbind(table3_3[c(1:4),],table3_3[c(5:8),]))
colnames(table3_3) = c("Mean","SD")


table3 = bind_cols(table3_1,table3_2,table3_3) %>%
  mutate(`Transportation choices`
         = c("Helicopter","Water Taxi","Ferry","Hovercraft")) %>%
  arrange(desc(Mean))

table3 = table3[,c(7,1:6)]


panelb_1 = transpose(df2 %>%
  filter(nationality < 3 & sample_all ==1) %>%
  summarise(
    weighted.mean(gender,weight),
     weighted.mean(age,weight),
     weighted.mean(educ_low,weight),
     weighted.mean(educ_high,weight),
     weighted.mean(affected,weight),
    weighted.mean(children,weight),
     weighted.mean(swim,weight),
     weighted.mean(nationality_1,weight),
    weighted.mean(wage_hr_ppp,weight,na.rm = TRUE),
     weighted.mean(wage_imput_ppp,weight),
     weighted.mean(life_exp,weight),
    weighted.mean(fatalism,weight),
    sgender = sqrt(weighted.var(gender,weight,na.rm = TRUE)),
    sage = sqrt(weighted.var(age,weight,na.rm = TRUE)),
    seduc_low = sqrt(weighted.var(educ_low,weight,na.rm = TRUE)),
    seduc_high = sqrt(weighted.var(educ_high,weight,na.rm = TRUE)),
    saffected  = sqrt(weighted.var(affected,weight,na.rm = TRUE)),
    schildren = sqrt(weighted.var(children,weight,na.rm = TRUE)),
    sswim = sqrt(weighted.var(swim,weight,na.rm = TRUE)),
    snationality_1  = sqrt(weighted.var(nationality_1,weight,na.rm = TRUE)),
    swage_hr_ppp = sqrt(weighted.var(wage_hr_ppp,weight,na.rm = TRUE)),
    swage_imput_ppp = sqrt(weighted.var(wage_imput_ppp,weight,na.rm = TRUE)),
```

```
    slife_exp = sqrt(weighted.var(life_exp,weight,na.rm = TRUE)),
    sfatalism = sqrt(weighted.var(fatalism,weight,na.rm = TRUE))
  ))



panelb_1 = as.data.frame(cbind(panelb_1[c(1:12),],panelb_1[c(13:24),]))
colnames(panelb_1) = c("Mean","SD")

panelb_2 = transpose(df2 %>%
  filter(nationality == 3 & sample_all ==1) %>%
  summarise(
    weighted.mean(gender,weight),
     weighted.mean(age,weight),
     weighted.mean(educ_low,weight),
     weighted.mean(educ_high,weight),
     weighted.mean(affected,weight),
    weighted.mean(children,weight),
     weighted.mean(swim,weight),
     weighted.mean(nationality_1,weight),
    weighted.mean(wage_hr_ppp,weight,na.rm = TRUE),
     weighted.mean(wage_imput_ppp,weight),
     weighted.mean(life_exp,weight),
    weighted.mean(fatalism,weight),
    sgender = sqrt(weighted.var(gender,weight,na.rm = TRUE)),
    sage = sqrt(weighted.var(age,weight,na.rm = TRUE)),
    seduc_low = sqrt(weighted.var(educ_low,weight,na.rm = TRUE)),
    seduc_high = sqrt(weighted.var(educ_high,weight,na.rm = TRUE)),
    saffected  = sqrt(weighted.var(affected,weight,na.rm = TRUE)),
    schildren = sqrt(weighted.var(children,weight,na.rm = TRUE)),
    sswim = sqrt(weighted.var(swim,weight,na.rm = TRUE)),
    snationality_1  = sqrt(weighted.var(nationality_1,weight,na.rm = TRUE)),
    swage_hr_ppp = sqrt(weighted.var(wage_hr_ppp,weight,na.rm = TRUE)),
    swage_imput_ppp = sqrt(weighted.var(wage_imput_ppp,weight,na.rm = TRUE)),
    slife_exp = sqrt(weighted.var(life_exp,weight,na.rm = TRUE)),
    sfatalism = sqrt(weighted.var(fatalism,weight,na.rm = TRUE))
  ))



panelb_2 = as.data.frame(cbind(panelb_2[c(1:12),],panelb_2[c(13:24),]))
colnames(panelb_2) = c("Mean","SD")

panelb_3 = transpose(df2 %>%
  filter( sample_all ==1) %>%
  summarise(
    weighted.mean(gender,weight),
     weighted.mean(age,weight),
     weighted.mean(educ_low,weight),
     weighted.mean(educ_high,weight),
     weighted.mean(affected,weight),
```

```
    weighted.mean(children,weight),
     weighted.mean(swim,weight),
     weighted.mean(nationality_1,weight),
    weighted.mean(wage_hr_ppp,
                  weight,na.rm = TRUE),
     weighted.mean(wage_imput_ppp,
                   weight),
     weighted.mean(life_exp,
                   weight),
    weighted.mean(fatalism,
                  weight),
    sgender = sqrt(weighted.var(gender,
                                weight,na.rm = TRUE)),
    sage = sqrt(weighted.var(age,
                             weight,na.rm = TRUE)),
    seduc_low = sqrt(weighted.var(educ_low,
                                  weight,na.rm = TRUE)),
    seduc_high = sqrt(weighted.var(educ_high,
                                   weight,na.rm = TRUE)),
    saffected  = sqrt(weighted.var(affected,
                                   weight,na.rm = TRUE)),
    schildren = sqrt(weighted.var(children,
                                  weight,na.rm = TRUE)),
    sswim = sqrt(weighted.var(swim,
                              weight,na.rm = TRUE)),
    snationality_1  = sqrt(weighted.var(nationality_1,
                                        weight,na.rm = TRUE)),
    swage_hr_ppp = sqrt(weighted.var(wage_hr_ppp,
                                     weight,na.rm = TRUE)),
    swage_imput_ppp = sqrt(weighted.var(wage_imput_ppp,
                                        weight,na.rm = TRUE)),
    slife_exp = sqrt(weighted.var(life_exp,
                                  weight,na.rm = TRUE)),
    sfatalism = sqrt(weighted.var(fatalism,
                                  weight,na.rm = TRUE))
  ))


panelb_3 = as.data.frame(cbind(panelb_3[c(1:12),],panelb_3[c(13:24),]))
colnames(panelb_3) = c("Mean","SD")

a = c("Gender (1 = Male)",
      "Age",
      "Educational Level: less than completed university",
      "Educational Level: complete university or more",
      "Personally affected by civil conflict (Yes = 1)",
      "Have Children? (Yes = 1)",
      "Knows how to swim?",
      "Sierrra Leonean",
      "Hourly wage (USD,PPP)- Measured",
      "Hourly wage (USD,PPP)- imputed",
      "Self-reported life expectancy",
      "Self-reported fatalism")
```

```
panelb = bind_cols(panelb_1,panelb_2,panelb_3) %>%
  mutate(Variable = a)


panelb = panelb[,c(7,1:6)]

a = capture.output(stargazer(table3, header = F,
                             summary = F,digits = 2,
                             type = "latex",rownames = F)  )
b = capture.output(stargazer(panelb, header = F,
                             summary = F,digits = 2,
                             type = "latex",rownames = F) )

c = star_panel( a,b, panel.names =
                  c("Transportation choices",
                                  " Respondent characteristics and attitudes"),
                reg = FALSE)

c[8] =  "  & Mean & SD & Mean & SD & Mean & SD \\\\ "
c[15]=""

c=c(c[1:7],
    " & \\multicolumn{2}{c}{Africans}
    & \\multicolumn{2}{c}{Non-Africans} & \\multicolumn{2}{c}{All} \\\\ ",
    " & \\multicolumn{2}{c}{(N=336)} & \\multicolumn{2}{c}{(N=225)}
    & \\multicolumn{2}{c}{(N=561)} \\\\ ",
    "\\hline \\\\[-1.8ex] " ,c[8:31]  )
cat(c,sep = "\n")
```

Table 1:

	Africans (N=336)		Non-Africans (N=225)		All (N=561)	
	Mean	SD	Mean	SD	Mean	SD
Panel A: Transportation choices						
Ferry	0.67	0.47	0.36	0.48	0.57	0.50
Water Taxi	0.20	0.40	0.36	0.48	0.25	0.43
Hovercraft	0.11	0.32	0.25	0.43	0.16	0.37
Helicopter	0.02	0.13	0.03	0.16	0.02	0.14
Panel B: Respondent characteristics and attitudes						
Gender (1 = Male)	0.78	0.42	0.76	0.43	0.77	0.42
Age	39.87	10.91	41.17	11.98	40.34	11.31
Educational Level: less than completed university	0.23	0.42	0.13	0.34	0.19	0.40
Educational Level: complete university or more	0.77	0.42	0.87	0.34	0.81	0.40
Personally affected by civil conflict (Yes = 1)	0.58	0.49	0.15	0.36	0.43	0.50
Have Children? (Yes = 1)	0.81	0.39	0.69	0.46	0.77	0.42
Knows how to swim?	0.36	0.48	0.74	0.44	0.50	0.50
Sierrra Leonean	0.58	0.50	0	0	0.37	0.48
Hourly wage (USD,PPP)- Measured	25.68	28.09	50.77	57.03	34.38	42.19
Hourly wage (USD,PPP)- imputed	29.05	27.66	47.60	51.39	35.64	38.81
Self-reported life expectancy	42.75	11.90	39.77	12.27	41.69	12.10
Self-reported fatalism	4.21	3.05	3.27	2.58	3.87	2.92

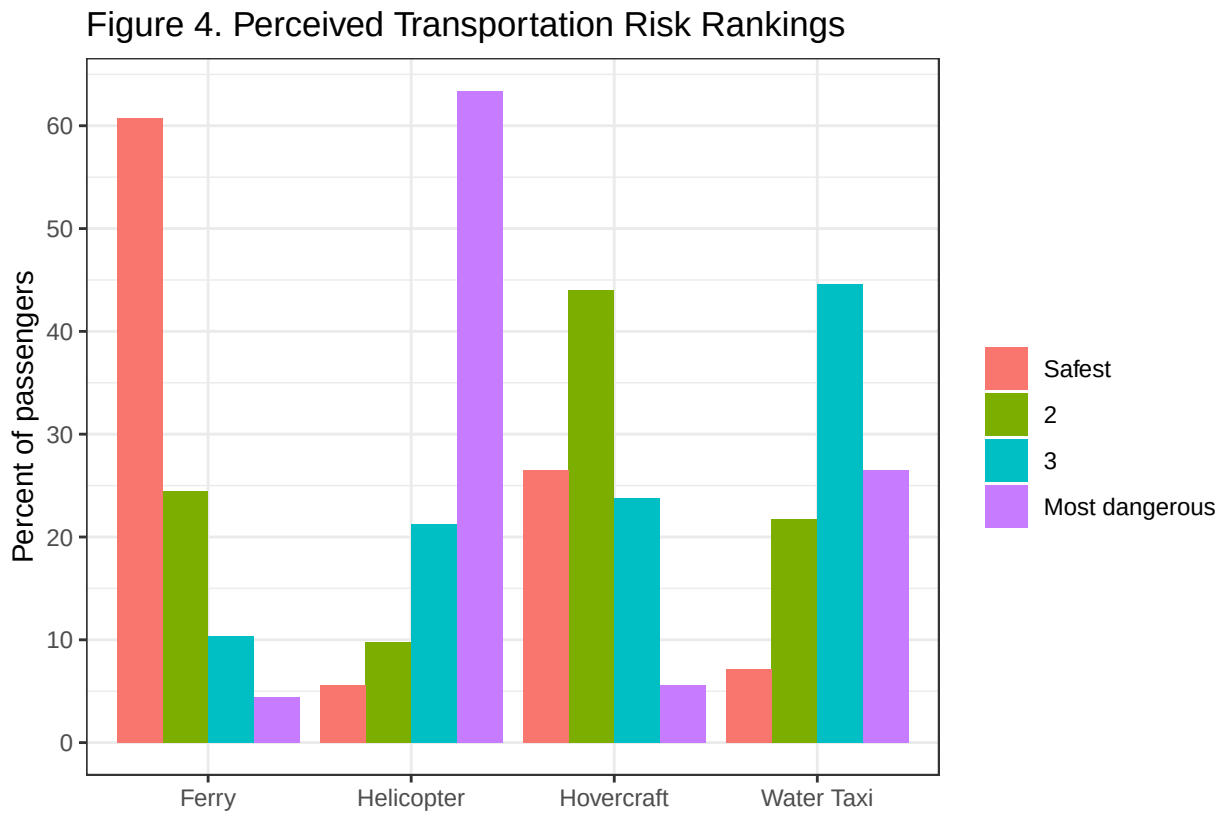
Problem2: Figure 4

```
df4 = df2 %>% filter(sample_all == 1) %>%
  group_by(weight,safest1_rain)

a = c(rep(1,4),rep(2,4),rep(3,4),rep(4,4))
df4 = as.data.frame(wtd.table(df4$safest1_rain,weights = df4$weight)) %>%
  bind_rows(as.data.frame(wtd.table(df4$safest2_rain,weights = df4$weight))) %>%
  bind_rows(as.data.frame(wtd.table(df4$safest3_rain,weights = df4$weight))) %>%
  bind_rows(as.data.frame(wtd.table(df4$safest4_rain,weights = df4$weight))) %>%
  bind_cols(as.data.frame(a)) %>%
  mutate(Var1 = case_when(
    Var1 == 1 ~ "Helicopter",
    Var1 == 2 ~ "Water Taxi",
    Var1 == 3 ~ "Ferry",
    Var1 == 4 ~"Hovercraft"
  ),
    a = case_when(
      a == 4 ~ "Most dangerous",
      a == 3 ~ "3",
      a == 2 ~ "2",
      a == 1 ~ "Safest"
    )) %>%
  group_by(a) %>%
  mutate(total = sum(Freq)) %>%
  ungroup()
```

```
df4$a = factor(df4$a, levels = c("Safest","2","3","Most dangerous"))


ggplot(data = df4, aes(x = Var1, y = Freq/total*100))+
  geom_bar(aes( fill = a),stat = "identity",position = "dodge") +
  labs(x = "", y = "Percent of passengers",
       title = "Figure 4. Perceived Transportation Risk Rankings", fill = "")+
  scale_y_continuous(breaks = seq(0,70,10))+
  theme_bw()
```



Problem 3

Table 2 is the replication for Table A3 in online appendix. I added a line of code: resizebox*textwidth*80mm to adjust the width of the table.

```
model1  = clogit (
  alt ~ risk_obs + transp_cost+
    strata(choice_situation)+cluster(serial),
  weights = weight,
  data = (df1 %>% filter(sample_af == 1)),method = "approximate"
)

model2  = clogit (
  alt ~ risk_obs+transp_cost +
    seats+noise+crowdness+convloc+clientele+
    strata(choice_situation)+cluster(serial),
```

```
  weights = weight,
  data = (df1 %>% filter(sample_af == 1)),method = "approximate")

model3  = clogit (
  alt ~ risk_obs + transp_cost+
    strata(choice_situation)+cluster(serial),
  weights = weight,
  data = (df1 %>% filter(sample_noaf == 1)),method = "approximate"
)

model4  = clogit (
  alt ~ risk_obs+transp_cost +
    seats+noise+crowdness+convloc+clientele+
    strata(choice_situation)+cluster(serial),
  weights = weight,
  data = (df1 %>% filter(sample_noaf == 1)),method = "approximate")

model5  = clogit (
  alt ~ risk_obs + transp_cost+
    strata(choice_situation)+cluster(serial),
  weights = weight,
  data = (df1 %>% filter(sample_all == 1)),method = "approximate"
)

model6  = clogit (
  alt ~ risk_obs+transp_cost +
    seats+noise+crowdness+convloc+clientele+
    strata(choice_situation)+cluster(serial), weights = weight,
  data = (df1 %>% filter(sample_all == 1)),method = "approximate")



df3 = df1 %>%
    filter(alt==1, sample_af == 1 | sample_noaf == 1) %>%
    drop_na(alt,risk_obs,transp_cost,choice_situation,serial,weight) %>%
    group_by( sample_af) %>%
    summarise(trips = n(), travelers = length(unique(serial)))

trips = c("Number of trips",
          rep(df3$trips[1],2),
          rep(df3$trips[2],2),
          rep(sum(df3$trips),2))
travelers = c("Number of decision makers",
              rep(df3$travelers[1],2),
              rep(df3$travelers[2],2),
              rep(sum(df3$travelers),2))



model = list(model1,model2,model3,model4,model5,model6)

mean_vsl = c("Mean VSL",1:6)
low_vsl = c("2.5,percentile",1:6)
```

```
high_vsl = c("97.5 percentile",1:6)

for (i in 1:6) {
  sd = deltamethod(~x1/x2, coef(model[[i]]),vcov(model[[i]]))
  mean_vsl[i+1] = round(model[[i]]$coefficients[1]/-model[[i]]$coefficients[2],3)
  low_vsl[i+1] = round(qnorm(0.025,as.numeric(mean_vsl[i+1]),sd),3)
  high_vsl[i+1] = round(qnorm(0.975,as.numeric(mean_vsl[i+1]),sd),3)
}



a =capture.output( stargazer(model,type = "latex", omit = c(3:10),
          column.labels = c("Africans","Non-Africans","All"),
          column.separate = c(2,2,2),
          dep.var.labels   = "",
          covariate.labels =
            c("Prob. of completing the trip (1-p$\\_i$)",
                              "Total transportation cost (Cost\\_ij)"),
          digits = 3,
          keep.stat = c("n","ll"),
          add.lines =
          list(
            c("Controls for Perceived attributes", "No" ,"Yes", "No","Yes","No" , "Yes" ),
                    trips,travelers,mean_vsl,low_vsl,high_vsl ),
          model.numbers = TRUE,
          header = FALSE,
          style = "qje",
          table.layout = "-c#-t-s-a-n"


))
a = c(a[c(1:16,20,17,21,22,18,19)],"\\hline \\\\[-1.8ex] ",a[23:31])
a[18] =  "Observations (respondent-alternative options) &
          3,281 & 3,281 & 2,124 & 2,124 & 5,405 & 5,405 \\\\ "

cat(a,sep  = "\n")
```

Table 2:

	Africans		Non-Africans		All	
	(1)	(2)	(3)	(4)	(5)	(6)
Prob. of completing the trip (1-p_i)	6.668***	8.996***	10.408***	10.524***	7.641***	9.391***
	(1.275)	(1.471)	(1.636)	(1.718)	(0.990)	(1.110)
Total transportation cost (Cost_ij)	−0.021***	−0.012***	−0.004	−0.004	−0.016***	−0.010***
	(0.001)	(0.001)	(0.002)	(0.002)	(0.001)	(0.001)
Controls for Perceived attributes	No	Yes	No	Yes	No	Yes
Observations (respondent-alternative options)	3,281	3,281	2,124	2,124	5,405	5,405
Number of trips	1083	1083	710	710	1793	1793
Number of decision makers	336	336	225	225	561	561
Log Likelihood	−1,216.268	−1,160.398	−618.982	−612.805	−1,869.394	−1,818.358
Mean VSL	319.985	778.496	2586.682	2960.805	482.681	984.267
2.5,percentile	156.025	235.992	-3644.366	-4657.212	243.193	199.131
97.5 percentile	483.945	1321	8817.73	10578.822	722.169	1769.403

Notes:

***Significant at the 1 percent level.
**Significant at the 5 percent level.
*Significant at the 10 percent level.