Replication

Kaley Burg

March 31, 2024

1. Paper Title and Information

- **Title:** Exclusion and Cooperation in Diverse Societies: Experimental Evidence from Israel
- **Authors:** RYAN D. ENOS Harvard University NOAM GIDRON Hebrew University of Jerusalem and Princeton University
- **Journal and Year:** American Political Science Review (2018)
- **Research Question and Design:**
  - Do exclusionary preferences stop individuals from cooperating with one another?
  - Does the lack of cooperation therefore prevent societies from providing public goods?
  - **Main question:** How do exclusionary attitudes among the Jewish majority predict discriminatory behaviour towards Palestinian citizens in Israel?
  - Authors also aim to understand how exclusionary attitudes are distinct from from behaviors.
  - **Data:**
  - Uses lab-in-the-field data to study cooperation and exclusionary attitudes of Jewish people towards Palestinian Citizens of Israel (PCI)
  - Measures cooperation (behaviours) with an economic decision making game
  - Measures exclusionary attitudes with a social distance scale which asks participants to choose the degree of proximity that they would accept an outgroup member. Responses range from family relative to no relationship.
  - **Variables:**
  - arab_accept, arab_accept_character, arab_reject_binary - all variables corresponding to the level of exclusionary attitudes towards PCI, either in categorical form, continuous numeric, or binary

- age - continuous age variable
- foreign_born - binary variable, 1 corresponding to foreign born
- sex - binary variable, 1 corresponding to male
- left_right - ordinal variable, 1 meaning far left to 7 meaning far right
- religiosity - categorical data with levels anti, orthodox, secular, traditional, ultra-orthodox (treated as ordinal in the manuscript but treated as categorical dummies in all analysis)
- education - ordinal variable with levels graduate, high school, primary, and undergraduate
- income - ordinal with levels average, high, low, very high, very low
- ethnicity
- arab_interactions - level of interaction with arabs with levels daily, monthly, never, week, and yearly

- **Hypothesis:**
  - **Note:** The authors do not explicitly state their hypotheses. Instead they state their contributions to the literature prior to their results section. I will be using this to infer their hypotheses
    1. Levels of exclusionary attitudes are high among the Jewish majority towards PCI
    2. These levels of exclusion are highest for lower-status Jews (low SES, uneducated, and the ultra-Orthodox population)
    3. Exclusionary attitudes are symbolic, meaning they are stable and affect other attitudes.
    4. Cooperation is informed/predicted by exclusionary attitiudes preferences for exclusion. In other words, those with greater exclusionary attitudes towards PCI are also less likely to cooperate with PCI.
    5. The attitudes-behaviour connection is not moderated by other factors such as perceptions of Arabs' trustworthiness.

2. Replicate the figures and tables of the main findings found in the manuscript

- Below is all the code used to replicate these figures and tables

```
#load dataset
load("data.analysis.RData")
head(data.analysis)

#create subset of secular and unorthodox respondents
```

```
#for figures
data.secular <- data.analysis[data.analysis$secular==1,]
data.uo <- data.analysis[data.analysis$uo==1,]

######Figure 1######
uo.respondents_accept_secular <- table(data.uo$secular_accept.character)[c(6,3,4,2,1,7,5)]/sum(table(data.uo$secular_accept.character))
uo.respondents_accept_arabs <- table(data.uo$arab_accept_character)[c(6,3,4,2,1,7,5)]/sum(table(data.uo$arab_accept_character))

secular.respondents_accept_uo <- table(data.secular$uo_accept_secular)[c(6,3,4,2,1,7,5)]/sum(table(data.secular$uo_accept_secular))
secular.respondents_accept_arabs <- table(data.secular$arab_accept_character)[c(6,3,4,2,1,7,5)]/sum(table(data.secular$arab_accept_character))

secular.arab.matrix <- rbind(secular.respondents_accept_uo,secular.respondents_accept_arabs)

uo.arab.matrix <- rbind(uo.respondents_accept_secular,uo.respondents_accept_arabs)

png(file='distance_secular_arabs.png')
par(mar=c(6.5,5.5,4.1,2.1))
a<-barplot(secular.arab.matrix, beside=T, ylim=c(0,0.6), las=2, axes = FALSE, axisnames = FALSE,
            ylab="Share of respondents", cex.lab=2.2)
legend("topleft", c("Secular toward UO","Secular toward Arabs"), bty="n",
       fill=c("black", "gray"), cex=2.2)
labs=c("relative", "friend", "neighbor", "coworker", "citizen", "visitor", "none")
text(a[1,], par("usr")[3]- 0.02, labels = labs, srt = 45, adj = 1, xpd = TRUE, cex=2)
axis(2,cex.axis = 1.6)
dev.off()

png(file='distance_uo_arabs.png')
par(mar=c(6.5,5.5,4.1,2.1))
b <- barplot(uo.arab.matrix, beside=T, ylim=c(0,0.6), axes = FALSE, axisnames = FALSE,
             ylab="Share of respondents", cex.axis = 1.6, cex.lab=1.6)
legend("topleft", c("UO toward Secular","UO toward Arabs"), bty="n", fill=c("black", "gray"),
       cex=2)
labs=c("relative","friend", "neighbor", "coworker", "citizen", "visitor", "none")
text(b[1,], par("usr")[3]- 0.02, labels = labs, srt = 45, adj = 1, xpd = TRUE, cex=2)
axis(2,cex.axis = 1.6)
dev.off()
```

```r
######Table 1######
#without interaction with arabs factor
arab.accept.lm.religiosity <- lm(arab_accept ~ age + foreign_born+ sex + left_right +
                                    as.factor(religiosity)  + as.factor(education) + as.factor(income) +
                                    as.factor(ethnicity),
                                  data=data.analysis)
summary(arab.accept.lm.religiosity)

stargazer(arab.accept.lm.religiosity)


#with interactions factor
arab.accept.lm.religiosity.interactions <- lm(arab_accept ~ age + foreign_born + sex + left_right +
                                                 as.factor(religiosity) + as.factor(education) + as.factor(income) +
                                                 as.factor(ethnicity) + as.factor(arab_interactions),
                                               data=data.analysis)
summary(arab.accept.lm.religiosity.interactions)

stargazer(arab.accept.lm.religiosity,arab.accept.lm.religiosity.interactions,
          no.space=T, covariate.labels=c("Age", "Foreign Born", "Sex",  "Left-right",
                                          "Secular", "Traditional", "Ultra Orthodox",
                                          "High-school", "Primary-school", "Undergrad",
                                          "High income", "Low income", "Very high income", "Very low income",
                                          "Mixed", "Other", "Sephardic",
                                          "Month", "Never", "Week", "Year",
                                          "Constant"))



######Figure 2######
distance.x <- seq(min(as.numeric(data.analysis$left_right), na.rm=TRUE), max(as.numeric(data.analysis$left_right), na.rm=TRUE),1)

xnew.distance.left_right=list(left_right=distance.x,
                              religiosity=rep("secular", length(distance.x)),
                              age=rep(median(as.numeric(as.vector(data.analysis$age)), na.rm=TRUE), length(distance.x)),
                              foreign_born=rep("0", length(distance.x)),
                              sex=rep("1", length(distance.x)),
```

```
                                   education=rep("high", length(distance.x)),
                                   income=rep("average", length(distance.x)),
                                   ethnicity=rep("sep", length(distance.x))
)

pred.distance.left_right <-predict(arab.accept.lm.religiosity, newdata=xnew.distance.left_right,
                                         interval="confidence", level=0.95)

png(file='predict_left_right.png')
par(mar=c(7.1,9.5,4.1,2.1))
plot(distance.x, pred.distance.left_right[,1], type="n", ann=FALSE, ylim=c(1,7), axes=FALSE)
axis(2, at=1:7, lab=c("relative", "friend", "neighbor", "coworker", "citizen", "visitor", "none"),
      las=1, cex.axis=2.5)
axis(1, at=1:7, cex.axis=2.5, tick = FALSE)
lines(distance.x, pred.distance.left_right[,1], lwd=6)
lines(distance.x, pred.distance.left_right[,2], lwd=1, lty=2)
lines(distance.x, pred.distance.left_right[,3], lwd=1, lty=2)
box()
abline(h=c(1:7), lty=3)
dev.off()




#predicting social distance by education
distance.x <- c("primary", "high", "undergrad", "grad")

xnew.distance.education=list(left_right=rep(median(as.numeric(data.analysis$left_right), na.rm=TRUE), length(distance.x)),
                                  religiosity=rep("secular", length(distance.x)),
                                  age=rep(median(as.numeric(as.vector(data.analysis$age)), na.rm=TRUE), length(distance.x)),
                                  foreign_born=rep("0", length(distance.x)),
                                  sex=rep("1", length(distance.x)),
                                  education=distance.x,
                                  income=rep("average", length(distance.x)),
                                  ethnicity=rep("sep", length(distance.x))
)

pred.distance.education <-predict(arab.accept.lm.religiosity, newdata=xnew.distance.education, tryp="response",
                                        se=TRUE)

plotUpper.win <- pred.distance.education$fit + (1.96 * pred.distance.education$se.fit)
plotLower.win <- pred.distance.education$fit - (1.96 * pred.distance.education$se.fit)
```

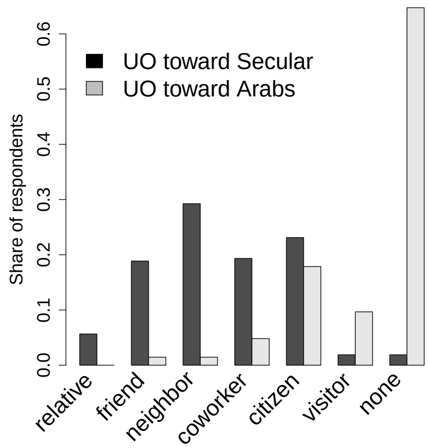
```
png("predict_education.png")
par(mar=c(7.1,9.5,4.1,2.1))
plot(1:4, pred.distance.education$fit,ann=FALSE, ylim=c(1,7), axes=FALSE, xlim=c(0.5,4.5),
     pch=19, lwd=6)
axis(2, at=1:7, lab=c("relative", "friend", "neighbor", "coworker", "citizen", "visitor", "none"),
     las=1, cex.axis=2.5)
text(axTicks(1), par("usr")[3]-0.25, srt=45, adj=1,
     labels=c("prim.", "high", "under.", "grad"),
     xpd=T, cex=2.5)
box()
for(x in 1:length(pred.distance.education$fit))
  lines(c(x,x),c(plotUpper.win[x],plotLower.win[x]), lwd=6)
abline(h=c(1:7), lty=3)
dev.off()

#predicting social distance by religiosity
distance.x <- c("secular", "trad", "orthodox", "u_orthodox")

xnew.distance.rel=list(left_right=rep(median(as.numeric(data.analysis$left_right), na.rm=TRUE), length(distance.x)), #5
                       religiosity=distance.x,
                       income=rep("average", length(distance.x)),
                       age=rep(median(as.numeric(as.vector(data.analysis$age)), na.rm=TRUE), length(distance.x)), #22
                       sex=rep("1", length(distance.x)), #male
                       foreign_born=rep("0", length(distance.x)),
                       education=rep("high", length(distance.x)),
                       income=rep("average", length(distance.x)),
                       ethnicity=rep("sep", length(distance.x))
)

pred.distance.rel <-predict(arab.accept.lm.religiosity, newdata=xnew.distance.rel, tryp="response",
                            se=TRUE)

plotUpper.win <- pred.distance.rel$fit + (1.96 * pred.distance.rel$se.fit)
plotLower.win <- pred.distance.rel$fit - (1.96 * pred.distance.rel$se.fit)

png("predict_religiosity.png")
par(mar=c(7.1,9.5,4.1,2.1))
plot(1:4, pred.distance.rel$fit, ann=FALSE, ylim=c(1,7), axes=FALSE, xlim=c(0.5,4.5),
     pch=19, lwd=6)
axis(2, at=1:7, lab=c("relative", "friend", "neighbor", "coworker", "citizen", "visitor", "none"),
     las=1, cex.axis=2.5)
```

```r
#axis(1, at=1:4, lab=c("secular", "trad.", "religious", "UO"), cex.axis=2)
text(axTicks(1), par("usr")[3]-0.25, srt=45, adj=1,
      labels=c("secular", "trad.", "religious", "UO"),
      xpd=T, cex=2.5)
box()
for(x in 1:length(pred.distance.rel$fit)) lines(c(x,x),c(plotUpper.win[x],plotLower.win[x]), lwd=6)
abline(h=c(1:7), lty=3)
dev.off()




###########Table 2##########
arab.coop.glm.distance <- glm(arab_coop ~ arab_reject_binary,
                              data=data.analysis, family=binomial)
summary(arab.coop.glm.distance)

arab.coop.glm.distance.covariates <- glm(arab_coop ~ arab_reject_binary + age +
                                           foreign_born + sex + left_right +
                                           as.factor(religiosity) + as.factor(education) + as.factor(income) + as.factor(ethnicity),
                                         data=data.analysis, family=binomial)
summary(arab.coop.glm.distance.covariates)

arab.coop.trust <- glm(arab_coop ~ arab_reject_binary + trust.b,
                       data=data.analysis, family=binomial)
summary(arab.coop.trust)

arab.coop.trust.covariates <- glm(arab_coop ~ arab_reject_binary + trust.b +
                                    age + foreign_born +
                                    sex + left_right +
                                    as.factor(religiosity) + as.factor(education) + as.factor(income) + as.factor(ethnicity),
                                  data=data.analysis, family=binomial)
summary(arab.coop.trust.covariates)


stargazer(arab.coop.glm.distance, arab.coop.glm.distance.covariates, arab.coop.trust, arab.coop.trust.covariates,
          no.space=T,  covariate.labels=c("Social Distance (binary)", "Age", "Foreign Born", "Sex", "Left-right",
                                          "Secular", "Traditional", "Ultra Orthodox", "High-school", "Primary school", "Undergrad",
                                          "High", "Low", "Very high", "Very low", "Mixed", "Other","Sephardic", "Trust in Arabs", "Constant")
```

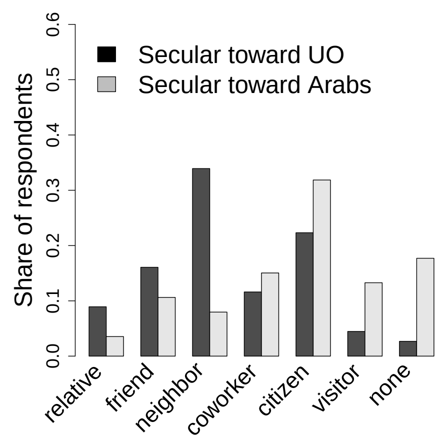
```
    )
```

0.0.1 Figures

Figure 1: Social Distance by Group

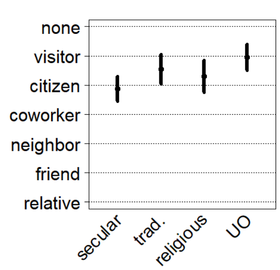


(a) Distance UO Arabs

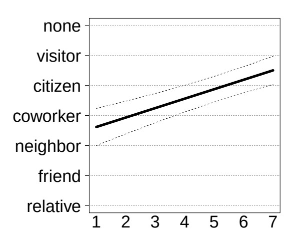


(b) Distance Secular Arabs

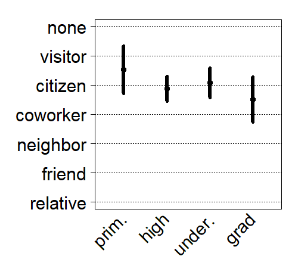
Figure 2: Religiosity, ideology, education, and social distance



(a) Religiosity



(b) Political Ideology



(c) Education

0.0.2 Tables

Table 1

	Dependent variable:	
	Social Distance with PCI	
	(1)	(2)
Age	−0.003	−0.004
	(0.005)	(0.005)
Foreign Born	−0.326*	−0.246
	(0.197)	(0.188)
Sex	−0.519***	−0.376***
	(0.149)	(0.144)
Left-right	0.314***	0.276***
	(0.055)	(0.052)
Secular	−0.425*	−0.353
	(0.256)	(0.246)
Traditional	0.236	0.287
	(0.285)	(0.272)
Ultra Orthodox	0.650***	0.495**
	(0.241)	(0.232)
High-school	0.368	0.250
	(0.337)	(0.321)
Primary-school	1.013**	0.634
	(0.466)	(0.458)
Undergrad	0.570*	0.430
	(0.340)	(0.325)
High income	−0.474*	−0.384
	(0.270)	(0.260)
Low income	−0.096	−0.180
	(0.216)	(0.208)
Very high income	0.151	−0.108
	(0.402)	(0.387)
Very low income	−0.002	0.048
	(0.202)	(0.195)
Mixed	−0.358	−0.478*
	(0.283)	(0.269)
Other	0.232	−0.186
	(0.456)	(0.438)
Sephardic	0.132	0.137
	(0.163)	(0.157)
Month		0.931***
		(0.296)
Never		1.486***
		(0.259)
Week		0.469
		(0.311)
Year		1.166***
		(0.312)
Constant	3.846***	3.046***
	(0.556)	(0.560)
Observations	375	372
R^2	0.289	0.370
Adjusted R^2	0.255	0.332
Residual Std. Error	1.345 (df = 357)	1.277 (df = 350)
F Statistic	8.537*** (df = 17; 357)	9.785*** (df = 21; 350)
Note:		*p<0.1; **p<0.05; ***p<0.01

Table 2

	Dependent variable:			
	Cooperation with PCI (=1)			
	(1)	(2)	(3)	(4)
Social Distance (binary)	−0.820***	−0.631**	−0.639**	−0.579*
	(0.243)	(0.306)	(0.264)	(0.319)
Age		−0.002		−0.002
		(0.008)		(0.008)
Foreign Born		0.137		0.053
		(0.318)		(0.325)
Sex		0.318		0.286
		(0.245)		(0.251)
Left-right		−0.208**		−0.198**
		(0.093)		(0.095)
Secular		−0.916**		−1.097***
		(0.411)		(0.424)
Traditional		−0.728		−0.919*
		(0.459)		(0.470)
Ultra Orthodox		−0.357		−0.501
		(0.386)		(0.394)
High-school		0.826		0.974
		(0.583)		(0.604)
Primary school		0.934		1.060
		(0.773)		(0.792)
Undergrad		0.469		0.604
		(0.590)		(0.613)
High		−0.085		−0.036
		(0.435)		(0.441)
Low		−0.409		−0.430
		(0.348)		(0.353)
Very high		0.158		0.086
		(0.625)		(0.632)
Very low		−0.356		−0.326
		(0.323)		(0.326)
Mixed		−0.135		−0.137
		(0.456)		(0.460)
Other		0.593		0.998
		(0.715)		(0.752)
Sephardic		−0.312		−0.262
		(0.267)		(0.270)
Trust in Arabs			0.536*	0.419
			(0.291)	(0.341)
Constant	−0.045	1.035	−0.271	0.867
	(0.213)	(0.905)	(0.249)	(0.942)
Observations	439	375	432	371
Log Likelihood	−274.275	−226.808	−268.588	−222.142
Akaike Inf. Crit.	552.549	491.617	543.176	484.285
Note:				*p<0.1; **p<0.05; ***p<0.01

3. Insert some twist based on your gained knowledge from Stats I and II

- For my twist, I first checked the model-fit of the ordered multinomial regression that the authors included in their replication code. Then, I ran an unordered model as well.

0.0.3 Ordered Model Checks

  - **NOTE:** The authors did not include this table in their main manuscript or in their appendix materials. They do include a footnote though that states "Ordered logit regression provides substantively similar results." (pg. 749).

- The code that the authors included is as follows:

```
                                as.factor(as.character(religiosity)) + sex +
                                as.factor(education) + as.factor(income) +
                                left_right +
                                as.factor(ethnicity),
                              data=data.analysis)

sumOrderedModTHEM <- summary(arab.accept.polr) #substantively similar
    results to lm

sumOrderedModTHEM
```

  - Note that it is their comment in the script that states that the results are substantively similar to lm
  - There was no further justification of this from the authors

- First, I made sure that this model was run properly by creating my own ordered factor variable and running it again

```
data.analysis$arab_acceptTEST <- factor(data.analysis$arab_accept,
    ordered = TRUE)

class(data.analysis$arab_acceptTEST)

data.analysis$arab_acceptTEST<- factor(data.analysis$arab_accept_
    character,
                                              levels = c("relative", "friend"
    , "neighbor",
                                                          "coworker", "citizen
    ", "visitor", "none"),
                                              ordered = T)

data.analysis$arab_acceptTEST

```

```
arab.accept.polrTEST <- polr(arab_acceptTEST ~ age + foreign_born +
                                as.factor(as.character(religiosity)) + sex
    +
                                as.factor(education) + as.factor(income) +
                                left_right +
                                as.factor(ethnicity),
                              data=data.analysis)
summary(arab.accept.polrTEST)
```

- I found that the results are the same, shown in Table 3

Table 3

	Dependent variable:	
	arab_acceptTEST	arab_accept
	(1)	(2)
age	−0.007	−0.007
	(0.007)	(0.007)
foreign_born1	−0.426	−0.426
	(0.275)	(0.275)
as.factor(as.character(religiosity))secular	−0.617*	−0.617*
	(0.360)	(0.360)
as.factor(as.character(religiosity))trad	0.225	0.225
	(0.409)	(0.409)
as.factor(as.character(religiosity))u_orthodox	0.849**	0.849**
	(0.354)	(0.354)
sex1	−0.778***	−0.778***
	(0.219)	(0.219)
as.factor(education)high	0.264	0.264
	(0.485)	(0.485)
as.factor(education)primary	1.211*	1.211*
	(0.702)	(0.702)
as.factor(education)undergrad	0.445	0.445
	(0.490)	(0.490)
as.factor(income)high	−0.606	−0.606
	(0.372)	(0.372)
as.factor(income)low	−0.018	−0.018
	(0.307)	(0.307)
as.factor(income)v_high	0.010	0.010
	(0.550)	(0.550)
as.factor(income)v_low	0.059	0.059
	(0.290)	(0.290)
left_right	0.406***	0.406***
	(0.082)	(0.082)
as.factor(ethnicity)mixed	−0.403	−0.403
	(0.404)	(0.404)
as.factor(ethnicity)other	0.250	0.250
	(0.612)	(0.612)
as.factor(ethnicity)sep	0.213	0.213
	(0.240)	(0.240)
Observations	375	375
Note:	*p<0.1; **p<0.05; ***p<0.01	

- Then I checked with categories corresponded to each number in the arab_accept factor and made sure that this matched the arab_accept_character factor

```
table(data.analysis$arab_accept, data.analysis$arab_accept_character)
```

- I found that relative is coded as 1, friend as 2, neighbor as 3, coworker as 4, citizen as 5, visitor as 6, and none as 7
- So, the further an individual is along the latent variable, the more prejudiced they are towards PCR.
- Now that I confirmed that their model was run correctly, I ran several other tests to check whether this model would have been more appropriate than both a linear regression model and an unordered multinomial regression.
- First, I checked the intercept values and plotted them

```
sumOrderedModTHEM$zeta
```

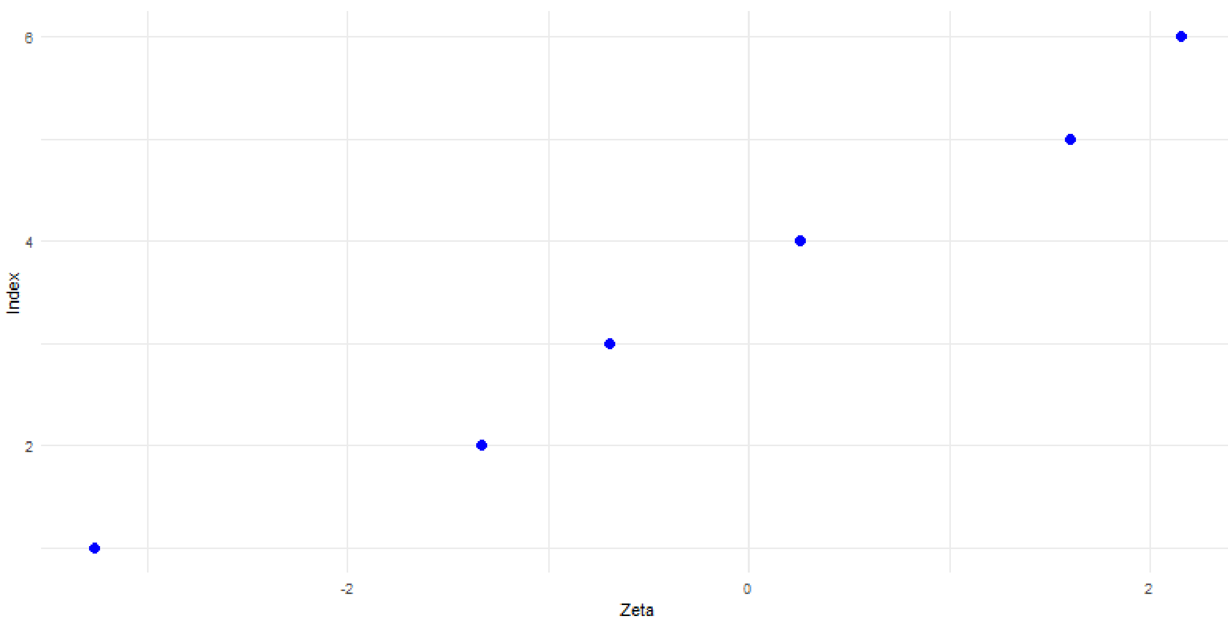


Figure 3: Intercepts

- The reason I did this was to check the distance between each of the intercept values
- Long (1997) states that "in general, the results of the LRM only correspond to those of the ORM *if* the thresholds are all about the same distance apart. When this is not the case, the LRM can give very misleading results.

- Aside from theoretical reasons about the outcome variable not being continuous and therefore not being a good fit for a linear regression model, this also serves as a first indication that the linear regression model might be giving skewed results.

- Next, I checked the author's assumptions about whether the ordered logit model produces similar results to the linear model.

- Referring back to Table 3, using either of the models since they produce the same results, we can see that the following variables are significant at the $p < 0.001$ level: sex/gender, and political ideology (left/right). The following are significant at the $p < 0.05$ level: ultra-Orthodox religious identity (compared to orthodox as the reference category). The following are significant at the $p < 0.01$ level: secular religious identity (compared to orthodox as the reference category), and primary education level (compared to grad).

- Comparing this to the results from the linear model (Table 1, left column), we see that the following variables are significant at the $p < 0.001$ level: sex/gender, political ideology (left/right), and ultra-Orthodox (compared to Orthodox). The following are significant at the $p < 0.05$ level: primary educational level (compared to grad as reference category). The following are significant at the $p < 0.01$ level: foreign born (compared to not foreign born as the reference category), secular religious identity (compared to orthodox as reference category), undergraduate educational level (compared to grad as reference category), and high income (compared to average as the reference category) and primary education level (compared to grad).

- We can see from these results that, although the sex, political ideology, ultra-orthodox, secular, and primary education variables are signficant in both models (although at varying levels), the linear model contains significance for the variables foreign born, undergraduate, and high income as well.

- Although the authors did not state how they were defining substantively similar, I am going to assume that they consider this result to be similar enough to use OLS as the primary model.

- I also checked the p-values by converting the results to odd-ratio form and constructing confidence intervals at the 95% level for both the log odds and odd ratios in the ordered multinomial model

```r
#now lets change everything to odds ratios - with CIs
ord_OR <- exp(cbind(OR = coef(arab.accept.polr), confint(arab.accept.polr
    )))

#print these
print(ord_OR)

#export to latex
stargazer(ord_OR)
```

```r


#confidence intervals for log odds ----
#https://stats.oarc.ucla.edu/r/dae/ordinal-logistic-regression/

#save summary of model
coefs <- (summary(arab.accept.polr))

#these are the coefficients of the omodel
coef(summary(arab.accept.polr))

#make table of values and std errors
ctable <- coef(summary(arab.accept.polr))
ctable


## calculate and store p values
p <- pnorm(abs(ctable[, "t value"]), lower.tail = FALSE) * 2
print(p)


## combined table
ctable <- cbind(ctable, "p value" = sprintf("%.20f", p))

#adding p values to the table
ctable

# Obtain confidence intervals for coefficients
ci <- confint(arab.accept.polr)
ci
# Extract coefficients and their confidence intervals
coefficients <- coef(arab.accept.polr)
ci_lower <- ci[, 1]
ci_upper <- ci[, 2]

# Combine coefficients and confidence intervals into a data frame
coef_ci <- data.frame(coefficients, Lower_CI = ci_lower, Upper_CI = ci_upper)
coef_ci

print(coef_ci)
stargazer(coef_ci)


#can now look at this in latex
```

- The results, shown in Tables 4 and 5, indicate that at the $p < 0.05$ level are: ultra-Orthodox, sex, and political ideology.
- This is used as an added check to make sure our results are in line with the p values that are calculated.

Table 4

	OR	2.5 %	97.5 %
age	0.993	0.979	1.006
foreign_born1	0.653	0.381	1.122
as.factor(as.character(religiosity))secular	0.540	0.264	1.086
as.factor(as.character(religiosity))trad	1.252	0.560	2.792
as.factor(as.character(religiosity))u_orthodox	2.337	1.159	4.661
sex1	0.459	0.298	0.704
as.factor(education)high	1.302	0.493	3.335
as.factor(education)primary	3.358	0.861	13.795
as.factor(education)undergrad	1.560	0.586	4.047
as.factor(income)high	0.545	0.263	1.131
as.factor(income)low	0.982	0.537	1.792
as.factor(income)v_high	1.010	0.348	3.053
as.factor(income)v_low	1.061	0.598	1.870
left_right	1.501	1.279	1.767
as.factor(ethnicity)mixed	0.668	0.303	1.487
as.factor(ethnicity)other	1.284	0.394	4.492
as.factor(ethnicity)sep	1.237	0.774	1.983

Table 5: Regression Coefficients and Confidence Intervals

Variable	Coefficient	Lower CI	Upper CI
age	-0.007	-0.021	0.006
foreign_born1	-0.426	-0.965	0.115
religiosity (secular)	-0.617	-1.332	0.083
religiosity (traditional)	0.225	-0.579	1.027
religiosity (ultra-orthodox)	0.849	0.147	1.539
sex1	-0.778	-1.211	-0.351
education (high)	0.264	-0.707	1.205
education (primary)	1.211	-0.150	2.624
education (undergrad)	0.445	-0.534	1.398
income (high)	-0.606	-1.337	0.123
income (low)	-0.018	-0.622	0.583
income (very high)	0.010	-1.057	1.116
income (very low)	0.059	-0.515	0.626
left_right	0.406	0.246	0.569
ethnicity (mixed)	-0.403	-1.194	0.397
ethnicity (other)	0.250	-0.930	1.502
ethnicity (separate)	0.213	-0.257	0.685

- We can now run logit models for each category in order to see if the parallel line assumption holds

```
# Iterate over each unique category of the outcome variable
for (i in 1:length(levels(data.analysis$arab_acceptTEST))) {
  # Create a model for each category
  assign(paste("logit_model_", i, sep=""),
         glm(arab_acceptTEST == levels(arab_acceptTEST)[i] ~ age +
    foreign_born +
                   as.factor(as.character(religiosity)) + sex +
                   as.factor(education) + as.factor(income) +
                   left_right +
                   as.factor(ethnicity),
               data = data.analysis, family = binomial),
         envir = globalenv())
}
```

- Results are shown in Table 6

Table 6

	Dependent variable:						
	arab_acceptTEST == levels(arab_acceptTEST)[i]						
	(Relative)	(Friend)	(Neighbor)	(Coworker)	(Citizen)	(Visitor)	(None)
age	−0.031	−0.031*	0.013	−0.008	0.033***	0.010	−0.021**
	(0.039)	(0.019)	(0.020)	(0.012)	(0.009)	(0.013)	(0.008)
foreign_born1	−18.790	1.953***	0.184	−0.500	0.088	−0.018	−0.356
	(7,232.651)	(0.736)	(0.713)	(0.512)	(0.362)	(0.479)	(0.346)
as.factor(as.character(religiosity))secular	19.305	0.179	−0.756	−0.351	1.246**	0.616	−1.145***
	(9,884.922)	(0.980)	(0.865)	(0.536)	(0.574)	(0.638)	(0.431)
as.factor(as.character(religiosity))trad	18.834	−1.224	−0.871	−0.881	0.851	−0.072	0.185
	(9,884.922)	(1.226)	(1.036)	(0.712)	(0.619)	(0.739)	(0.450)
as.factor(as.character(religiosity))u_orthodox	−0.118	−0.842	−2.310**	−1.820***	0.734	0.085	0.649*
	(11,153.780)	(1.102)	(1.051)	(0.578)	(0.556)	(0.625)	(0.381)
sex1	0.349	1.135*	−0.029	0.102	0.663**	0.016	−0.865***
	(1.365)	(0.630)	(0.613)	(0.393)	(0.291)	(0.365)	(0.255)
as.factor(education)high	18.275	−1.529	1.073	−0.881	0.810	15.536	−0.305
	(13,139.300)	(0.940)	(1.332)	(0.721)	(0.743)	(861.706)	(0.581)
as.factor(education)primary	1.839	−17.209	−15.837	−1.165	0.800	14.892	0.618
	(19,616.320)	(1,415.656)	(3,812.976)	(1.257)	(0.945)	(861.706)	(0.779)
as.factor(education)undergrad	−0.556	−3.270**	1.078	−0.556	0.775	15.683	−0.207
	(14,362.630)	(1.318)	(1.322)	(0.725)	(0.747)	(861.706)	(0.587)
as.factor(income)high	0.180	0.838	1.891	−0.385	−0.307	−0.046	−0.452
	(1.742)	(1.056)	(1.175)	(0.643)	(0.491)	(0.688)	(0.490)
as.factor(income)low	−0.375	1.481	−15.752	−0.161	−0.148	−0.731	0.193
	(1.586)	(0.962)	(1,652.911)	(0.531)	(0.375)	(0.597)	(0.364)
as.factor(income)v_high	−17.820	−0.230	−16.210	−1.130	0.497	0.499	−0.276
	(16,584.300)	(1.706)	(4,497.914)	(1.152)	(0.682)	(0.899)	(0.695)
as.factor(income)v_low	0.230	0.740	1.808	−0.076	−0.928**	0.345	0.180
	(1.548)	(0.977)	(1.137)	(0.506)	(0.383)	(0.483)	(0.339)
left_right	−0.651	−0.759***	−0.580**	−0.163	−0.039	0.334**	0.279***
	(0.449)	(0.216)	(0.240)	(0.140)	(0.106)	(0.152)	(0.095)
as.factor(ethnicity)mixed	1.456	1.301	−0.163	−1.118	0.701	−0.196	−0.495
	(1.788)	(0.946)	(0.983)	(0.815)	(0.533)	(0.820)	(0.501)
as.factor(ethnicity)other	−17.986	0.024	0.944	−16.232	−14.705	1.234	0.225
	(20,241.190)	(1.304)	(1.049)	(1,205.275)	(726.996)	(0.842)	(0.794)
as.factor(ethnicity)sep	0.500	0.650	−0.606	−0.837*	0.032	0.172	0.217
	(1.574)	(0.709)	(0.773)	(0.432)	(0.316)	(0.412)	(0.278)
Constant	−37.035	1.336	−1.921	1.242	−4.108***	−20.199	−0.216
	(16,442.410)	(1.849)	(2.076)	(1.292)	(1.161)	(861.707)	(0.945)
Observations	375	375	375	375	375	375	375
Log Likelihood	−13.329	−51.361	−43.917	−107.725	−174.212	−116.887	−212.732
Akaike Inf. Crit.	62.658	138.721	123.835	251.451	384.425	269.775	461.465

Note: *p<0.1; **p<0.05; ***p<0.01

- We can see that the parallel line assumption **does not appear to hold**, as we have sign changes as we move along the acceptance categories
- For example, in the age column, a one unit change in age is associated on average, with a 0.031 decrease in the the log odds of accepting a PCR as a relative, holding all other variables constant.
- Moving along to the neighbor logit regression, we can see that a one unit change in age is associated on average, with a 0.013 increase in the the log odds of accepting a PCR as a neighbor, holding all other variables constant.
- Next, we see that for the coworker logit regression, a one unit change in age is associated on average, with a 0.008 decrease in the the log odds of accepting a PCR as a coworker, holding all other variables constant.
- These results are confusing, because if this were a truly ordinal scale we would see that these coefficients would move in the same direction. Based on just the 'relative' logit regression, would imply that older participants are less accepting of PCR. We would expect to see this effect consistent through all the logit regressions. Instead, we see sign flips throughout implying that the variable may not be truly ordinal.
- This effect is repeated throughout the other variables as well.
- In this case, we can run an unordered model to see if this is a better fit

0.0.4 Unordered Multinomial Model

- First I converted the outcome variable to an unordered factor variable, set the reference category to 'none,' and ran the unordered multinomial regression. Table 7 shows the log odds regression results and Table 8 shows the coefficients converted to odds ratios.

```
data.analysis$arab_acc_fact_unord <- factor(data.analysis$arab_accept_character,
                                            levels = c("relative", "friend", "neighbor",
                                                       "coworker", "citizen", "visitor", "none"),
                                            ordered = F)

# checking to make sure i converted this correctly
arab_acc_fact_char <- as.character(data.analysis$arab_acc_fact_unor)
arab_accept_character_char <- as.character(data.analysis$arab_accept_character)

# Compare character vectors
mismatches <- which(arab_acc_fact_char != arab_accept_character_char)

mismatches

```

```
#no mismatches
table(data.analysis$arab_acc_fact_unord, data.analysis$arab_accept)

#now relevel with reference as none
data.analysis$arab_acc_fact_unord <- relevel(data.analysis$arab_acc_fact_unord, ref = "none")



# run model
arabAcceptUnord <- multinom(arab_acc_fact_unord~ age + foreign_born+ sex + left_right +
                                as.factor(as.character(religiosity))  + as.factor(education) + as.factor(income) +
                                 as.factor(ethnicity),
                               data=data.analysis)


summary(arabAcceptUnord)
```

Table 7

	Dependent variable:					
	relative	friend	neighbor	coworker	citizen	visitor
	(1)	(2)	(3)	(4)	(5)	(6)
age	-0.024	-0.019	0.017	0.003	0.036^{***}	0.021
	(0.041)	(0.020)	(0.021)	(0.014)	(0.011)	(0.014)
foreign_born1	-18.421^{***}	2.039^{**}	0.558	-0.178	0.367	0.188
	(0.00000)	(0.792)	(0.769)	(0.573)	(0.425)	(0.522)
sex1	1.153	1.802^{***}	0.734	0.715^{*}	1.063^{***}	0.455
	(1.412)	(0.668)	(0.650)	(0.429)	(0.327)	(0.389)
left_right	-1.078^{**}	-1.063^{***}	-0.881^{***}	-0.422^{***}	-0.242^{*}	0.130
	(0.489)	(0.243)	(0.270)	(0.162)	(0.124)	(0.162)
as.factor(as.character(religiosity))secular	14.080^{***}	0.718	-0.120	0.353	1.831^{***}	1.310^{*}
	(0.923)	(1.022)	(0.938)	(0.604)	(0.635)	(0.687)
as.factor(as.character(religiosity))trad	12.588^{***}	-1.617	-1.271	-1.056	0.609	-0.126
	(0.957)	(1.255)	(1.110)	(0.763)	(0.660)	(0.767)
as.factor(as.character(religiosity))u_orthodox	-7.714^{***}	-1.521	-2.830^{***}	-2.047^{***}	0.330	-0.154
	(0.000)	(1.140)	(1.098)	(0.619)	(0.591)	(0.652)
as.factor(education)high	10.577^{***}	-1.171	1.032	-0.711	0.620	27.344^{***}
	(1.182)	(1.053)	(1.465)	(0.801)	(0.825)	(0.476)
as.factor(education)primary	-1.296^{***}	-46.280	-23.232^{***}	-1.662	0.129	26.348^{***}
	(0.00000)	(NA)	(0.000)	(1.340)	(1.017)	(0.894)
as.factor(education)undergrad	-8.259^{***}	-3.053^{**}	0.742	-0.581	0.518	27.443^{***}
	(0.00000)	(1.407)	(1.460)	(0.812)	(0.827)	(0.482)
as.factor(income)high	1.071	1.510	2.335^{*}	0.296	0.207	0.291
	(1.797)	(1.147)	(1.248)	(0.733)	(0.589)	(0.756)
as.factor(income)low	-0.103	1.492	-31.733^{***}	-0.120	-0.163	-0.766
	(1.619)	(1.015)	(0.000)	(0.584)	(0.426)	(0.629)
as.factor(income)v_high	-11.807^{***}	-0.230	-17.987^{***}	-0.893	0.594	0.655
	(0.00000)	(1.822)	(0.00000)	(1.263)	(0.800)	(0.987)
as.factor(income)v_low	0.376	0.793	1.757	-0.062	-0.799^{*}	0.190
	(1.592)	(1.024)	(1.176)	(0.551)	(0.426)	(0.514)
as.factor(ethnicity)mixed	1.622	1.688	0.118	-0.596	0.965	0.167
	(1.839)	(1.036)	(1.071)	(0.890)	(0.624)	(0.871)
as.factor(ethnicity)other	-16.002^{***}	-0.723	0.054	-30.247^{***}	-27.855^{***}	0.822
	(0.00000)	(1.498)	(1.232)	(0.000)	(0.000)	(0.937)
as.factor(ethnicity)sep	0.132	0.337	-0.809	-0.887^{*}	-0.112	0.016
	(1.529)	(0.737)	(0.804)	(0.470)	(0.353)	(0.435)
Constant	-21.131^{***}	2.811	0.161	2.361	-2.519^{*}	-30.769^{***}
	(1.182)	(2.054)	(2.326)	(1.477)	(1.320)	(1.046)
Akaike Inf. Crit.	1,090.955	1,090.955	1,090.955	1,090.955	1,090.955	1,090.955
Note:						*p<0.1; **p<0.05; ***p<0.01

Table 8

	Dependent variable:					
	relative	friend	neighbor	coworker	citizen	visitor
	(1)	(2)	(3)	(4)	(5)	(6)
(Intercept)	6.651812e-10	16.62011	1.174264	10.6033	0.08055344	4.337311e-14
age	0.9766339	0.9813665	1.0170968	1.0027272	1.0362798	1.0211608
foreign_born1	9.993412e-09	7.680720	1.746663	0.8367114	1.442817	1.206234
sex1	3.167965	6.061605	2.082368	2.044595	2.895577	1.576944
left_right	0.3402059	0.3454747	0.4144849	0.6555768	0.7852415	1.1388934
as.factor(as.character(religiosity))secular	1.302773e+06	2.051171	0.8867322	1.42354	6.238936	3.706946
as.factor(as.character(religiosity))trad	2.929564e+05	0.1985388	0.2804708	0.3477121	1.837836	0.8815897
as.factor(as.character(religiosity))u_orthodox	0.000446421	0.2185649	0.05903657	0.1291617	1.390973	0.8576777
as.factor(education)high	39234.80	0.3101697	2.807297	0.4909893	1.858677	750693100000
as.factor(education)primary	0.2736024	7.959215e-21	8.13683e-11	0.1896932	1.137379	2.771624e+11
as.factor(education)undergrad	0.0002588371	0.04723659	2.100926	0.5594632	1.678288	8.285699e+11
as.factor(income)high	2.919292	4.528083	10.32993	1.34446	1.229912	1.337493
as.factor(income)low	0.9019111	4.446884	1.654761e-14	0.886939	0.8492521	0.4649023
as.factor(income)v_high	7.451886e-06	0.7944283	1.542646e-08	0.4094867	1.81096	1.92551
as.factor(income)v_low	1.4566613	2.2092725	5.7967432	0.9400315	0.4495974	1.2089442
as.factor(ethnicity)mixed	5.06502	5.408902	1.125156	0.550868	2.624276	1.181947
as.factor(ethnicity)other	1.123243e-07	0.4855062	1.055101	7.309264e-14	7.992025e-13	2.275707
as.factor(ethnicity)sep	1.1416274	1.4000588	0.4451049	0.4117924	0.8942012	1.0161241
Note:						*p<0.1; **p<0.05; ***p<0.01

- Next we can look at the prediction accuracy

```
#make predictions
predictions <- predict(arabAcceptUnord, newdata = subset_data, na.rm =
   TRUE)
length(predictions)
table(predictions)
data.analysis$arab_accept_character

length(data.analysis$arab_accept_character)


#create a table of actual vs. predicted classes
pred_table <- table(data.analysis$arab_accept_character, predictions)
pred_table


#add margins
pred_tab_marg <- addmargins(prediction_table)
print(pred_tab_marg)
```

- The results are shown in Table 9

Table 9: Prediction Accuracy

Actual Outcomes	**Predictions**							**Sum**
	none	relative	friend	neighbor	coworker	citizen	visitor	
citizen	40.00	0.00	2.00	1.00	2.00	38.00	0.00	83.00
coworker	19.00	0.00	2.00	1.00	7.00	8.00	0.00	37.00
friend	5.00	0.00	8.00	1.00	1.00	3.00	1.00	19.00
neighbor	5.00	0.00	0.00	4.00	1.00	5.00	0.00	15.00
none	157.00	0.00	2.00	1.00	3.00	13.00	1.00	177.00
relative	1.00	1.00	1.00	0.00	0.00	1.00	0.00	4.00
visitor	26.00	0.00	0.00	1.00	2.00	10.00	1.00	40.00
Sum	253.00	1.00	15.00	9.00	16.00	78.00	3.00	375.00

- We can see in this table that the model tends to overpredict the "none" outcome and underpredicts the visitor outcome.
- Furthermore, we can see that the results of the unordered multinomial regression are very messy and hard to interpret due to the number of outcomes possible and the amount of predictors used in the model.
- To make this slightly easier to look at, I also decided to run this as a binary variable, which the authors do in their second model of cooperation. This counts everyone who said they would accept a PCI as a coworker or closer as low exclusionary preference and all other options as high exclusionary preference. Results are shown in Table 9 (log odds) and Table 10 (odds ratios)

```
arabAcceptUnordBin <- multinom(arab_reject_binary~ age + foreign_born+ sex + left_right +
                              as.factor(as.character(religiosity)) + as.factor(education) + as.factor(income) +
                              as.factor(ethnicity),
                            data=data.analysis)
summary(arabAcceptUnordBin)
ORUnordBin <- exp(coef(arabAcceptUnordBin))
```

Table 10

	Dependent variable:
	arab_reject_binary
age	0.014
	(0.010)
foreign_born1	−0.227
	(0.392)
sex1	−0.565*
	(0.320)
left_right	0.588***
	(0.124)
as.factor(as.character(religiosity))secular	0.500
	(0.468)
as.factor(as.character(religiosity))trad	1.169**
	(0.577)
as.factor(as.character(religiosity))u_orthodox	2.173***
	(0.512)
as.factor(education)high	1.082*
	(0.617)
as.factor(education)primary	2.606**
	(1.261)
as.factor(education)undergrad	1.383**
	(0.639)
as.factor(income)high	−0.778
	(0.526)
as.factor(income)low	−0.308
	(0.466)
as.factor(income)v_high	1.227
	(0.988)
as.factor(income)v_low	−0.535
	(0.437)
as.factor(ethnicity)mixed	0.191
	(0.545)
as.factor(ethnicity)other	0.676
	(0.857)
as.factor(ethnicity)sep	0.544
	(0.362)
Constant	−3.922***
	(1.152)
Akaike Inf. Crit.	329.296
Note:	*p<0.1; **p<0.05; ***p<0.01

Table 11

	ORUnordBin
(Intercept)	0.02
age	1.01
foreign_born1	0.80
sex1	0.57
left_right	1.80
as.factor(as.character(religiosity))secular	1.65
as.factor(as.character(religiosity))trad	3.22
as.factor(as.character(religiosity))u_orthodox	8.78
as.factor(education)high	2.95
as.factor(education)primary	13.54
as.factor(education)undergrad	3.99
as.factor(income)high	0.46
as.factor(income)low	0.73
as.factor(income)v_high	3.41
as.factor(income)v_low	0.59
as.factor(ethnicity)mixed	1.21
as.factor(ethnicity)other	1.97
as.factor(ethnicity)sep	1.72

- These results make significantly more sense and are easier to interpret.
- We can also check the predictions, shown in Table 12

```
predictionsBin <- predict(arabAcceptUnordBin, newdata = subset_dataBIN,
    na.rm = TRUE)
length(predictionsBin)
table(predictionsBin)
data.analysis$arab_reject_binary

length(data.analysis$arab_reject_binary)


#create a table of actual vs. predicted classes
pred_tableBIN <- table(data.analysis$arab_reject_binary, predictions)
pred_tableBIN


#add margins
pred_tabBIN_marg <- addmargins(pred_tableBIN)
print(pred_tabBIN_marg)
```

Table 12

	0	1	Sum
0	27.00	48.00	75.00
1	14.00	286.00	300.00
Sum	41.00	334.00	375.00

- This prediction table implies that the model correctly predicts exclusionary attitudes relatively well (286/300) but not non-exclusionary attitudes (27/75).

0.1 Overall Findings of My Twist

- The authors found that "In this article, we have shown that social distance, a measure of exclusionary preferences, is strongly predictive of cooperation in a public goods game" (p. 753).

- The authors also state "Political ideology, education, and religiosity all appear to be strongly related to social distance, with more right-wing, more religious, and less-educated subjects expressing more exclusionary preferences" (p. 749).

- My results from Table 3 indicate that at the $p < 0.05$ level, right-wing and ultra-Orthodox participants were associated with a positive increase in the odds of holding higher exclusionary attitudes on an ordinal scale. At the $p < 0.01$ level, primary education, compared to graduate education, was associated with an increase in exclusionary attitudes. Therefore, less educated participants were associated with a positive increase in the odds of holding higher exclusionary attitudes on an ordinal scale. Furthermore, female participants were associated, on average, with a decrease in the log odds of holding exclusionary attitudes.

- All of this together, I would say that the conclusion that the 'results are substantively similar' is partially justified. Based on the conclusions drawn by the authors about which factors are significant in their overall analysis, the results are justified. However, this is based on the assumption that the social desirability scale is truly ordinal and that this simplification is not impactful of the overall results.

- Based on the results of the parallel line assumption test (Table 6), we can see that this assumption does not appear to hold.

- Within the results of the unordered multinomial model, we can see that many of the predictors appear to have abnormally strong associations with social closeness preferences.

- However, the predictive power of this model seems quite good for predicting the 'none' outcome for not accepting PCI at all. The other categories do not seem as strong though.

- Overall, it seems as though all of the models I tested seem to be relatively good at predicting a lack of acceptance towards PCI. However, each model has its own setback: the OLS model such that the outcome variable is not continuous in nature and the cut points of the ordinal model imply that this model might be skewed; the ordinal model does not pass the parallel line assumption test; and the unordered model produces strangely high odds ratios and has quite poor predictive power.

- The conclusion that I have drawn from this is that the scale from which the major conclusions were drawn from might be skewed, leading to abnormal results. This is supported by Weinfurt and Moghaddam (2001), who state the difficulties of using the social closeness scale as an ordinal measure in various cultural contexts

0.2 References

Enos, R. D., & Gidron, N. (2018). Exclusion and Cooperation in Diverse Societies: Experimental Evidence from Israel. *American Political Science Review*, *112*(4), 742–757. https://doi.org/10.1017/S0003055418000266

Weinfurt, K. P., & Moghaddam, F. M. (2001). Culture and Social Distance: A Case Study of Methodological Cautions. *The Journal of Social Psychology*, *141*(1), 101–110. https://doi.org/10.1080/00224540109600526