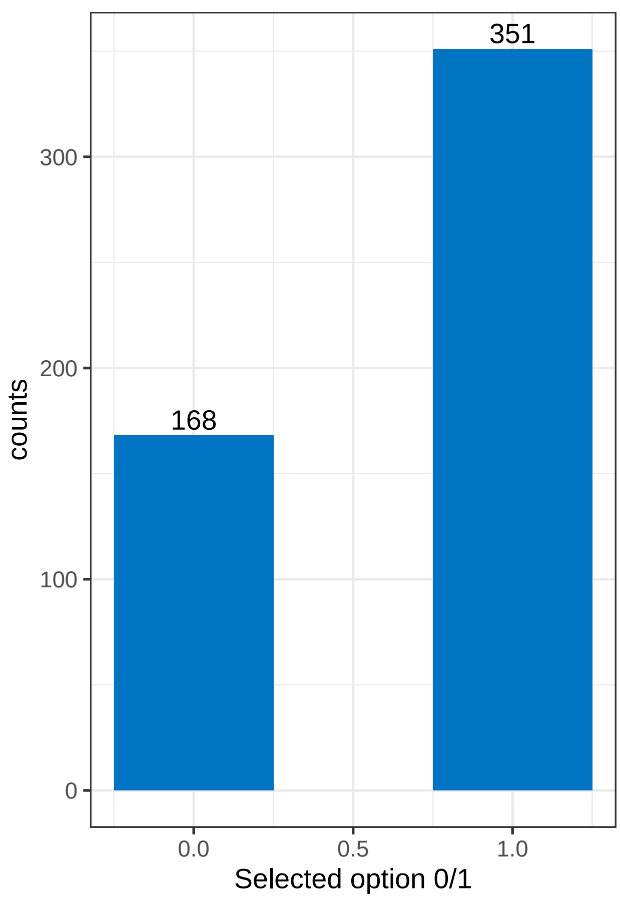
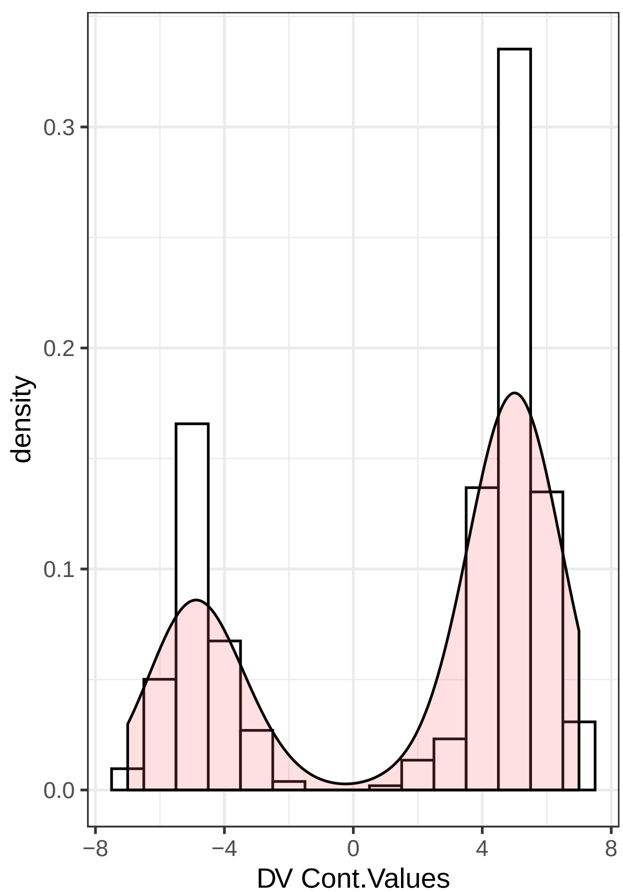
Experimental Data: Description and Analysis

This document details multiple procedures for an exploration and analysis of experimental data. In this version of the file, I use data collected in a survey experiment conducted on June 2019. The sample consists of 519 respondents. The material included a vignette describing an international conflict, then I described two potential policy options to address the conflict, and asked respondents for their preferred alternative.

Data description

The first set of procedures includes presenting the data (distribution of main variables). There are two dependent variables: binary and continuous one.





There are two treatments in this design, below are the sizes of groups for each treatment.

```
table(MyData$trt1)
```

```
##
##   0   1
## 243 276
```

```r
table(MyData$trt2)
```

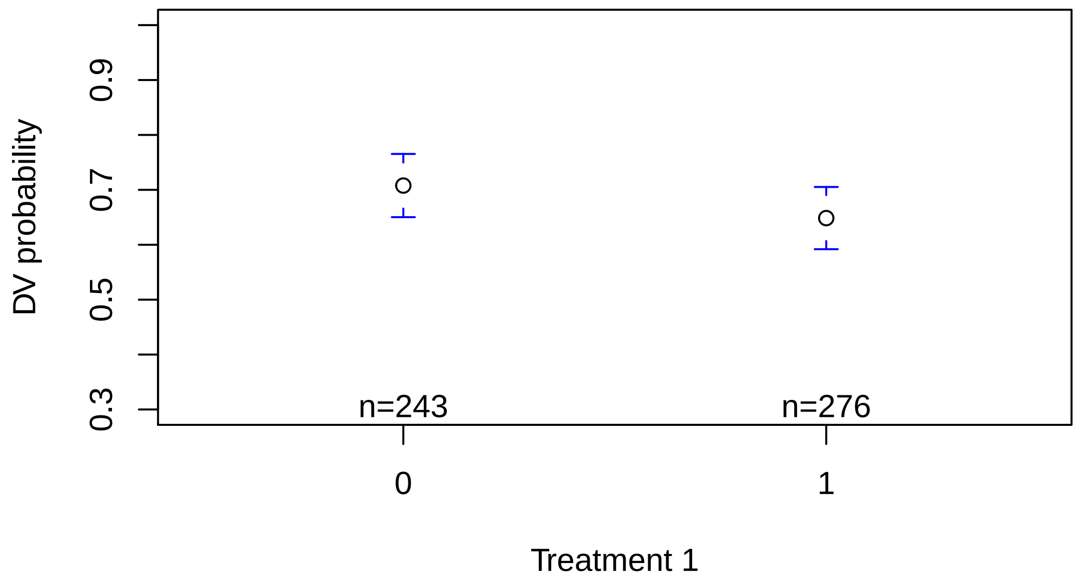
```
##
##   0   1
## 253 266
```

Means

In the following, I present various options to compute and present the mean response for the dependent variable, separated by both conditions for each treatment.

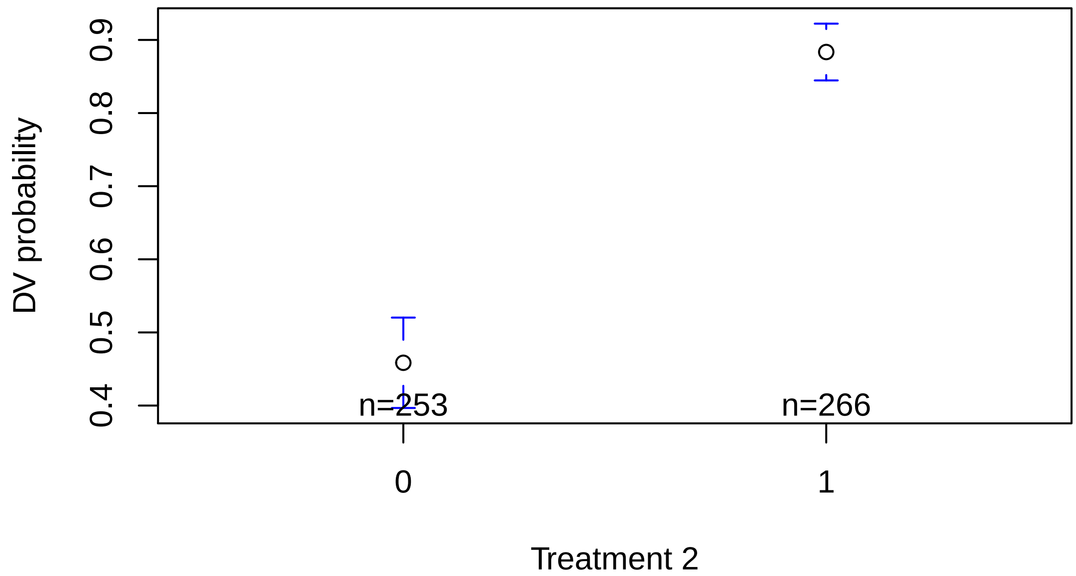
```
## [1] "Mean values of DV: treatment 1, two conditions"

##         0         1
## 0.7078189 0.6485507
```




```
## [1] "Mean values of DV: treatment 2, two conditions"

##         0         1
## 0.4584980 0.8834586
```



Analysis

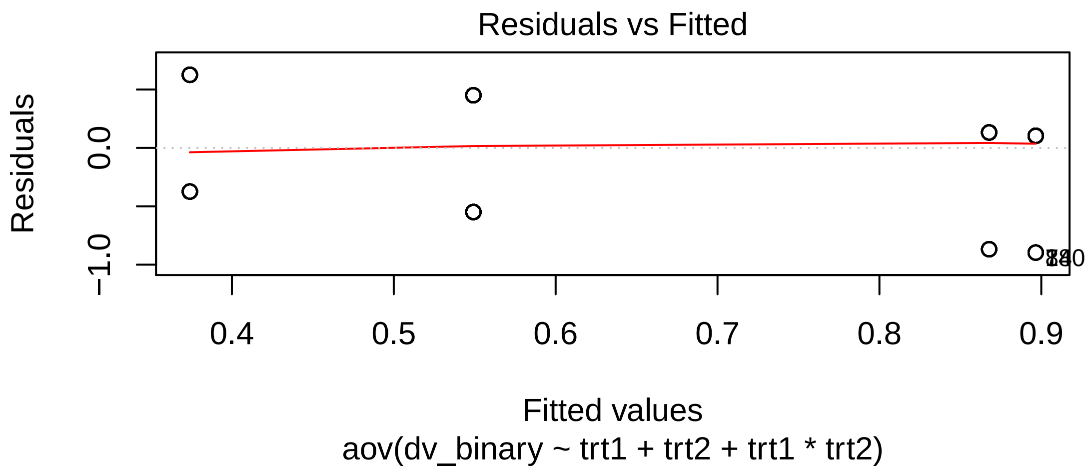
I run different models to analyze the data. First, I use an ANOVA model for the binary dependent variable. The code below includes both treatments and their interaction (no co-variates). Then, the 'model.tables' command presents various options for the mean values of the dependent variable: the grand mean of the entire sample; the means by treatment groups (including the size of each group); and finally, the means of the DV based on all four conditions of both treatments (including the size of each group).

```
##              Df Sum Sq Mean Sq F value  Pr(>F)
## trt1          1   0.45   0.454   2.650 0.10415
## trt2          1  23.61  23.614 137.865 < 2e-16 ***
## trt1:trt2     1   1.34   1.342   7.834 0.00532 **
## Residuals   515  88.21   0.171
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

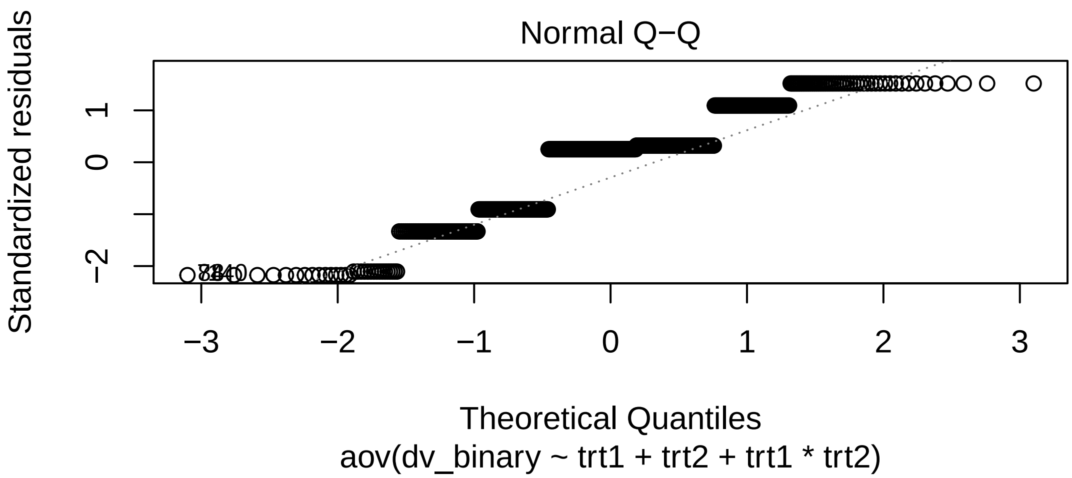
```
## Tables of means
## Grand mean
##
## 0.6763006
##
##  trt1
##            0        1
##       0.7078   0.6486
## rep 243.0000 276.0000
##
##  trt2
##            0        1
##       0.4577   0.8842
## rep 253.0000 266.0000
##
##  trt1:trt2
##      trt2
## trt1  0      1
##   0     0.55   0.87
```

```
##    rep 122.00 121.00
##    1      0.37   0.90
##    rep 131.00 145.00
```

When using an ANOVA model, it is useful to test its main assumptions: first, I plot the *residuals versus fits plot* to check the homogeneity of variances. In the plot below, it appears that there is no relationship between the residuals and fitted values. I also employ the Levene test. The results show that we can reject the null, meaning that the variance across groups is statistically significant. Second, we check the normality assumption of the residuals.




```
## Levene's Test for Homogeneity of Variance (center = median)
##         Df F value    Pr(>F)
## group   3  22.721 7.903e-14 ***
##       515
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```



Regression Analysis

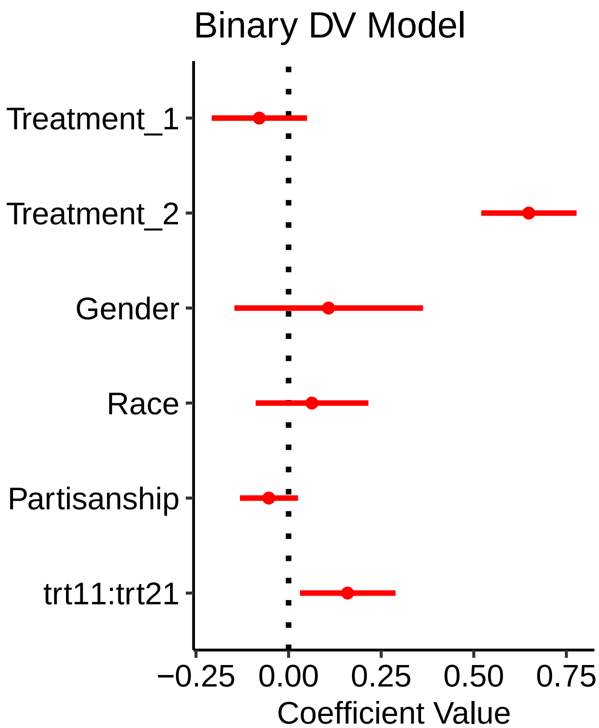
Next, I run regression models for both DVs. A probit model for the binary variable and OLS for the continuous dependent variable. Below, I present the results of all models in a single table. Models 1 & 3 include only the experimental treatments (baseline models); models 2 & 4 also include all the co-variates.

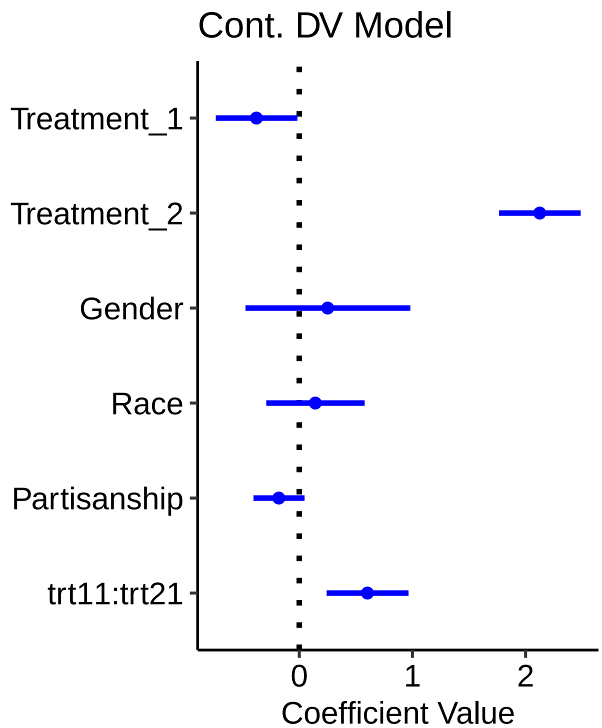
Table 1: Regression Results

	Dependent variable:			
	Binary DV		Cont. DV	
	probit		*OLS*	
	(1)	(2)	(3)	(4)
Treatment1	−0.075	−0.079	−0.364**	−0.379**
	(0.064)	(0.065)	(0.180)	(0.180)
Treatment2	0.644***	0.648***	2.127***	2.125***
	(0.064)	(0.065)	(0.180)	(0.180)
Gender		0.108		0.251
		(0.128)		(0.366)
Race		0.063		0.142
		(0.076)		(0.218)
Partisanship		−0.053		−0.180
		(0.039)		(0.113)
Treat1 x Treat2	0.148**	0.159**	0.575***	0.602***
	(0.064)	(0.065)	(0.180)	(0.181)
Constant	0.545***	0.519*	1.745***	1.937**
	(0.064)	(0.281)	(0.180)	(0.813)
Observations	519	519	519	519
R^2			0.234	0.239
Adjusted R^2			0.229	0.230
Akaike Inf. Crit.	540.126	542.948		
Note:	*p<0.1; **p<0.05; ***p<0.01			

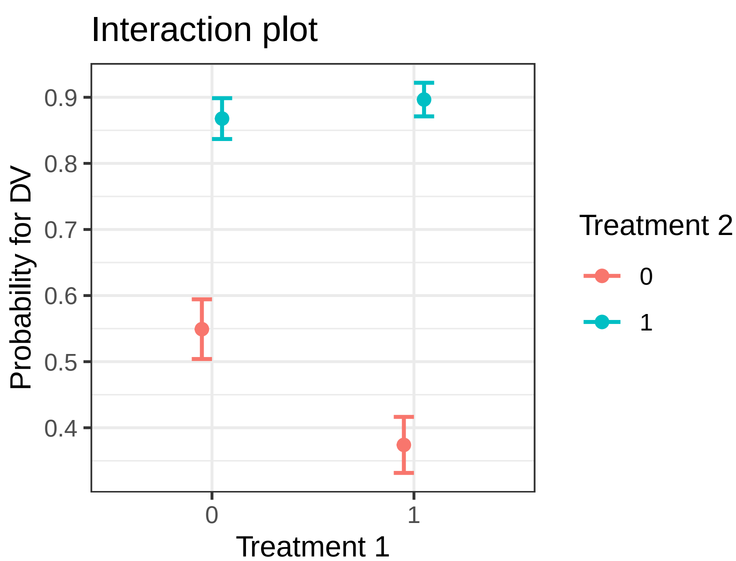
Visualizing the results

There are many ways to visualize the results of the analysis. First, I depict coefficient plots of the main models (binary and continuous DV, including all co-vars, models 2 and 4 in the table above).

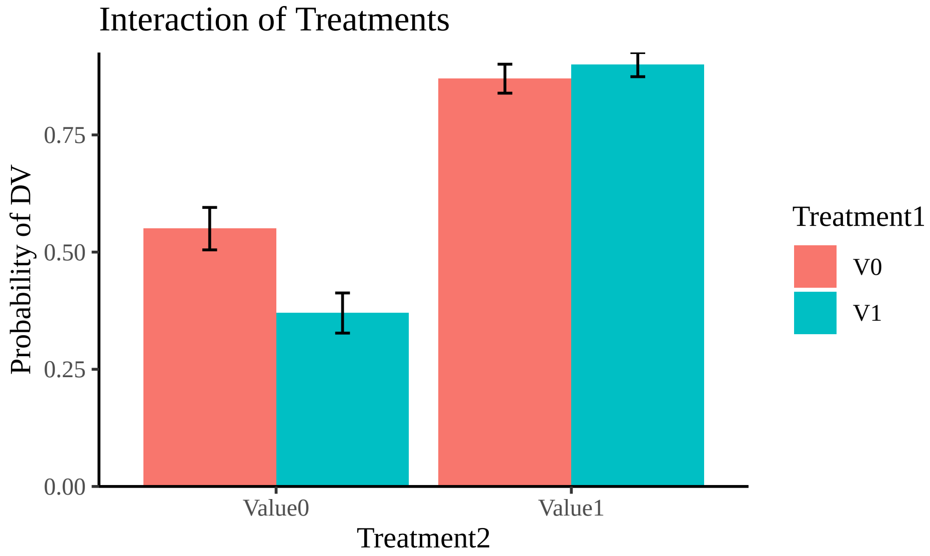




Presenting the model results using coefficient plots is useful for showing the effects of separate factors/variables (main effects in ANOVA models). We can use interaction plots to show more clearly how the conditional effect of both treatments affect the dependent variable. Below is the code for a plot that show the values of a binary DV when considering the conditional relations of both (binary) independent variables. The plot clearly shows the interaction effect between both treatments.



There are additional ways to display the same results. The code below uses grouped barplot to distinguish among the treatments and show the interaction.



The final plot offers a different way to display the mean values of the DV based on the different treatments. I create density plots that are based on bootstrapped sample in order to improve the accuracy of the results.

```
## No id variables; using all as measure variables
## No id variables; using all as measure variables
```