QMSS-GR5069 Data Challenge 2

Team 2

```
stopifnot(require(randomForest))
stopifnot(require(stats))
stopifnot(require(ggplot2))
stopifnot(require(dplyr))
stopifnot(require(stargazer))
stopifnot(require(gridExtra))
stopifnot(require(grid))
```

```
d <- read.csv("ConfrontationsData_170209.csv")
```

Task 1

Queston 1

Is there a dfference in the most important factors that predict civilian dead or wounded for states with more organized crime activity compared to states that are not so?

With this question we wanted to understand the particular dynamics of civilian collateral damage in the interventions. The first thing we are assuming here is that the incidence of injury or death of civilians is collateral damage. The analysis will stand even if this is not the case; understanding under what circumstances there are civilians affected is important for we assume, also, that the goal of peace-keeping forces is to reduce direct civilian involvement as much as possible. They should be defending the civilian population.

Understanding what factors go into civilian involvement, in states where organized crime activity, is necessary to identify situations in which peace-keeping forces should be more careful. We hypothesize that the variables that predict civilian involvement in both groups of states (the ones with high organized crime activity and the ones with low organized crime activity) will be different, and, hence, that the variables that predict it in low activity states indicate particular cases where the activity went far enough to involve civilians. We assume that civilian involvement in high activity states is related to a more systemic presence of organized crime, so civilian involvement in low activity states should be related to factors that are more controllable by peace-keeping forces. Deployment of force should take these factors into consideration.

One of our main assumptions here is that states can be characterized as having different amounts of organized crime activity. For this particular task we have used a document released by the International Crisis Group ("El Desafío de Peña Nieto: los cárteles y el Estado de Derecho en México." International Crisis Group. Informe sobre América Latina N°48. 2013), where they classify Mexican states by the amount of Cartel-related murders. We are assuming that states that had more than 1,000 such murders between 2006 and 2012 have, generally, higher organized crime activity. This was confirmed in exploratory analysis when most of the confrontations took place in these states.

For the analysis we chose to use Random Forests to determine the most important variables to predict if a case had or did not have civilian involvement. In this sense our dependent variable is a newly created dummy that records if there were any civilians killed or wounded in the confrontation. At the same time, we have created a dummy variable that records if the state where the confrontation took place is of high or low organized crime activity, in accordance to our source. The Random Forest algorithm was run on both subsets of states separately for comparative results.

Question 1 data wrangling code

```
dq1 <- d

#dummy variable for states with high cartel activity

dq1$high_cartel_activity <- ifelse(dq1$state.abbr == "BC" |
                                   dq1$state.abbr == "SON" |
                                   dq1$state.abbr == "CHIH" |
                                   dq1$state.abbr == "COAH" |
                                   dq1$state.abbr == "NL" |
                                   dq1$state.abbr == "TAMPS" |
                                   dq1$state.abbr == "SIN" |
                                   dq1$state.abbr == "DGO" |
                                   dq1$state.abbr == "JAL" |
                                   dq1$state.abbr == "MICH" |
                                   dq1$state.abbr == "GRO", 1, 0)


#dummy for cases where civilians wehere either killed or wounded
dq1$civilian_affected <- ifelse((dq1$civilian.dead + dq1$civilian.wounded) > 0, 1, 0)
```

Question 1 Random Forests code

```
#Random Forest for high cartel activity states

dq1 <- dq1[complete.cases(dq1), ]

set.seed(666)
rf1 <- randomForest(x = dq1[dq1$high_cartel_activity == 1, c(9, 11:19, 22:30, 32:36)],
                    y = as.factor(dq1[dq1$high_cartel_activity == 1, 38]),
                    importance = TRUE, proximity = TRUE)



#Random Forest for low cartel activity states

set.seed(666)
rf2 <- randomForest(x = dq1[dq1$high_cartel_activity == 0, c(9, 11:19, 22:30, 32:36)],
                    y = as.factor(dq1[dq1$high_cartel_activity == 0, 38]),
                    importance = TRUE, proximity = TRUE)
```

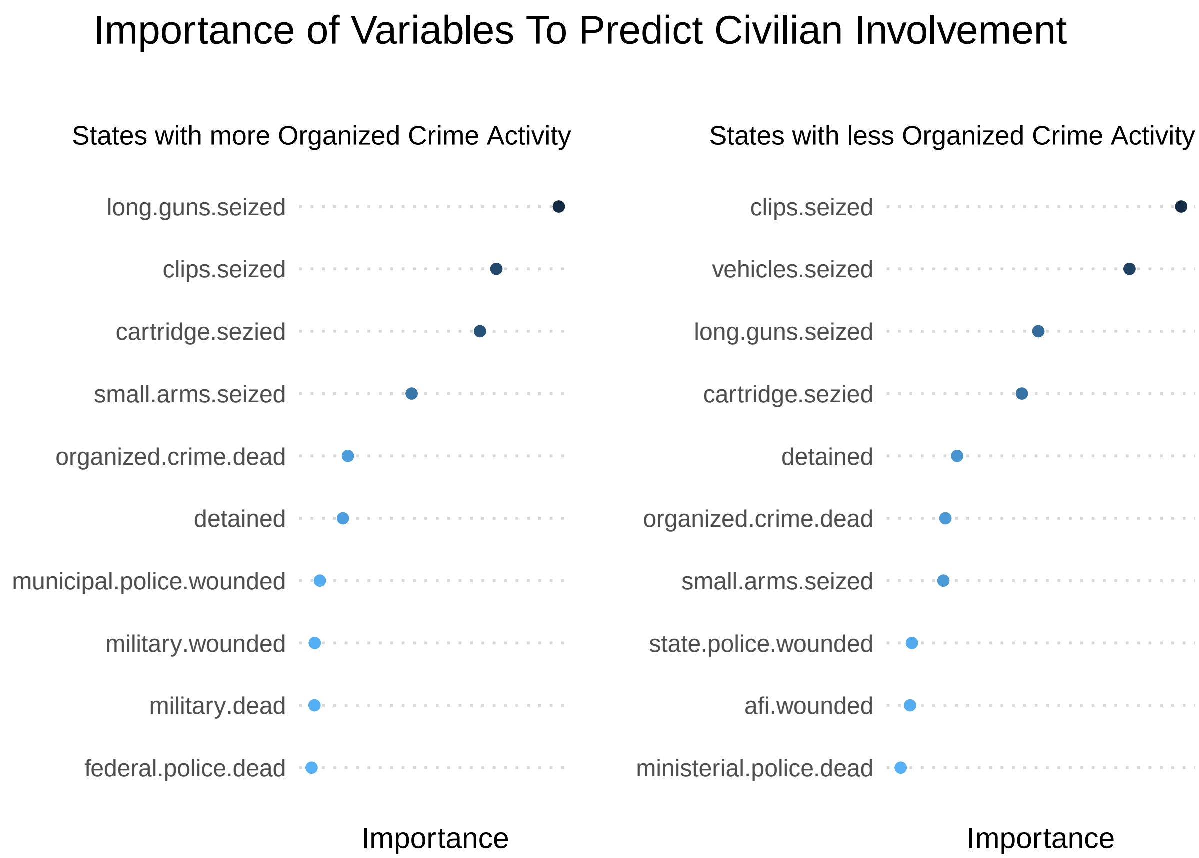
Question 1 Random Forests Plot code

```
#Data arranging for plots
rf1_plot <- as.data.frame(rf1$importance)
rf1_plot$variables <- rownames(rf1_plot)
rf1_plot <- arrange(rf1_plot, desc(MeanDecreaseAccuracy))

rf2_plot <- as.data.frame(rf2$importance)
rf2_plot$variables <- rownames(rf2_plot)
rf2_plot <- arrange(rf2_plot, desc(MeanDecreaseAccuracy))

#Plot code
```

```r
g1 <- ggplot(rf2_plot[(1:10),]) +
  geom_point(aes(x = reorder(factor(variables), MeanDecreaseAccuracy),
                 y = MeanDecreaseAccuracy, color = -MeanDecreaseAccuracy),
                 stat = "identity") +
  ggtitle("States with less Organized Crime Activity") +
  labs(x = "", y = "Importance") +
  theme(panel.background = element_blank(),
        panel.grid.major = element_line(color = "#d8d8d8", linetype = "dotted"),
        panel.grid.major.x = element_blank(),
        plot.title = element_text(size = 10, hjust = 1),
        axis.line.y = element_blank(),
        axis.ticks.y = element_blank(),
        axis.text.y = element_text(hjust = 1),
        axis.line.x = element_blank(),
        axis.text.x = element_blank(),
        axis.ticks.x = element_blank(),
        legend.title = element_blank(),
        legend.position = "none") +
  coord_flip()


g2 <- ggplot(rf1_plot[(1:10),]) +
  geom_point(aes(x = reorder(factor(variables), MeanDecreaseAccuracy),
                 y = MeanDecreaseAccuracy, color = -MeanDecreaseAccuracy),
                 stat = "identity") +
  ggtitle("States with more Organized Crime Activity") +
  labs(x = "", y = "Importance") +
  theme(panel.background = element_blank(),
        panel.grid.major = element_line(color = "#d8d8d8", linetype = "dotted"),
        panel.grid.major.x = element_blank(),
        plot.title = element_text(size = 10, hjust = 1),
        axis.line.y = element_blank(),
        axis.ticks.y = element_blank(),
        axis.text.y = element_text(hjust = 1),
        axis.line.x = element_blank(),
        axis.text.x = element_blank(),
        axis.ticks.x = element_blank(),
        legend.title = element_blank(),
        legend.position = "none") +
  coord_flip()
```

Results

```r
grid.arrange(g2, g1, ncol=2,
             top = textGrob("Importance of Variables To Predict Civilian Involvement \n",
                            gp = gpar(fontsize = 15, font = 1)))
```



Brief

The main difference between states with high and low organized criminal activity is the presence of seized vehicles in low states, and peace-keeping bodies involved in the confrontations for each state.

Expanded

First of all, the most striking result is that, in the variables with more importance to predict if civilians were hurt or killed, both types of states have roughly the same variables as most important; long guns, cartridges, and small arms seized are indicators of the size of the confrontation. It is fair to say that civilian involvement becomes more likely when the size of the confrontation is bigger. This could be an understatement, but it is something worth considering.

On the side of the differences, the biggest one is the importance of vehicles seized in predicting civilian involvement for low organized crime activity states. Though this is also an indicator of the possible size of the confrontation, it is notable that this variable is not within the ten most important for predicting in high criminal activity states. This could indicate that civilians are more likely to be hurt in "compound raids" of sorts, where peace-keeping forces enter a compound or headquarters of the criminal organization. These compounds are probably located in residential areas, which can explain the phenomenon.

The last point comes together with the difference in the type of peace-keeping bodies that are present in each group of states. Here we are assuming that the incidence of wounded or dead of a particular body is an indicator of their presence and vice-versa. Where in high activity states the predominant forces are municipal police, military, and federal police, in low activity states they are state police, afi, and ministerial police. Though we don't know much about the differences between the bodies, the presence of the military for high activity states tells us that these may be cases where a bigger display of force was needed; a pressing situation that required not police activity (which will probably look to build a case and bring the actors to justice), but military activity (which acts in a different way, objective-centered and less focused on posterior legal process). This could be related to the vehicle seizure variable: in states with less organized crime presence, civilians are more likely to be killed in confrontations that happen more under the umbrella of a police operation, where the objective is to build a case (hence the vehicle seizure, these are central locations for the cartels). In states with high organized crime activity, civilians are more likely to be hurt in the confrontations that happen comparatively more: force-centered confrontations that may be more pressing and less planned than police raids.

Our conclusions are limited to the fact that these are only variables of importance for Random Forest prediction, which could change slightly on each iteration and don't lend themselves to straight forward analytic needs. On top of this, our data does not depict all the cases but just the cases where a confrontation happened, which hampers the comparative potential of the analysis. Finally, we are bounded by the data to assume the peace-keeping bodies' presence in a confrontation through their number of killed and wounded, which is not a strong assumption to have.

Queston 2

What peace-keeping body's participation is more related to there being detentions in a given confrontation?

We believe that due process should be conducing to solving societal problems such as organized crime. In this context, detentions are an opportunity to gather important information, and formalize a conflict tainted by violence. Understanding the conditions under which detentions are made can help planning confrontations, and clarify what factors go into process-oriented confrontations. Different peace-keeping bodies have different ways of proceeding, and depicting which bodies are more likely to favor detentions can be useful if the confrontations with organized crime are planned.

For this, we performed a logistic regression with a dependent dummy variable that codes if there where any detentions made. Our independent variables are a series of dummies that code if a particular peace-keeping body (and others) were involved in the conflict; this was done under our previous assumption that a member of a body killed or wounded is an indicator of that body's participation in the confrontation, and vice-versa. As controls, we have indicators of confrontation size and civilian involvement.

Question 2 data wrangling code

```r
#Variable Creation

dq2 <- d

#Dummy for detention incidence
dq2$detention_incidence <- ifelse(dq2$detained > 0, 1, 0)

#Dummies for participation of different groups
for(i in c("military", "navy", "federal.police",
           "afi", "state.police", "ministerial.police",
           "municipal.police", "public.prosecutor",
           "organized.crime", "civilian")) {

  dq2[[i]] <- ifelse(dq2[[paste(i, "dead", sep = ".")]] > 0 |
                         dq2[[paste(i, "wounded", sep = ".")]] > 0, 1, 0)

}
```

Question 2 Logistic Regression Code

```r
logitq2 <- glm(detention_incidence ~ military + navy + federal.police +
               afi + state.police + ministerial.police + municipal.police +
```

```
               long.guns.seized + small.arms.seized + organized.crime + civilian,
               data = dq2, family = binomial(link = "logit"))

stargazer(logitq2, title = "Logistic Regression Results")
```

Question 2 Logistic Regression Plot Code

```
#Data for plot
g3d <- summary(logitq2) %>% coefficients %>% as.data.frame() %>% subset(, c(1:2, 4))
g3d$ub <- g3d$Estimate + g3d$`Std. Error`
g3d$lb <- g3d$Estimate - g3d$`Std. Error`

g3d$`Pr(>|z|)` <- ifelse(g3d$`Pr(>|z|)` <= 0.001, "***",
                  ifelse(g3d$`Pr(>|z|)` <= 0.01, "**",
                  ifelse(g3d$`Pr(>|z|)` <= 0.1, "*", "Not Significant")))

g3d$`Pr(>|z|)` <- ordered(g3d$`Pr(>|z|)`, levels = c("***", "**", "*", "Not Significant"))

g3d$variables <- rownames(g3d)
g3d$variables <- as.factor(as.character(g3d$variables))
g3d <- arrange(g3d, desc(`Pr(>|z|)`))
g3d$variables <- ordered(g3d$variables, levels = c(unique(as.character(g3d$variables))))


#Plot code
g3 <- ggplot() +
  geom_point(aes(x = variables, y = Estimate, color = `Pr(>|z|)`), data = g3d) +
  scale_colour_discrete(h=c(170,230), l = seq(0, 80, length.out = 5)) +
  geom_errorbar(aes(x = variables, y = Estimate,
                    color = `Pr(>|z|)`, ymin = lb, ymax = ub),
                    data = g3d, width = 0.1) +
  ggtitle("Coefficients of Logistic Regression") +
  labs(x = "", y = "Estimate") +
  theme(panel.background = element_blank(),
        panel.grid.major = element_line(color = "#d8d8d8", linetype = "dotted"),
        panel.grid.major.x = element_blank(),
        plot.title = element_text(hjust = 0.5),
        axis.line.y = element_blank(),
        axis.ticks.y = element_blank(),
        axis.text.y = element_text(hjust = 1),
        legend.key = element_blank()) +
  geom_hline(aes(yintercept = 0), color = "#d8d8d8", linetype = "longdash") +
  coord_flip()
```

Results

Table 1: Logistic Regression Results

	Dependent variable:
	detention_incidence
military	-0.509^{***}
	(0.130)
navy	0.055
	(0.369)
federal.police	0.302
	(0.192)
afi	0.557
	(0.548)
state.police	0.543^{***}
	(0.199)
ministerial.police	0.043
	(0.212)
municipal.police	0.110
	(0.150)
long.guns.seized	0.034^{***}
	(0.010)
small.arms.seized	0.340^{***}
	(0.032)
organized.crime	0.036
	(0.072)
civilian	-0.266^{**}
	(0.108)
Constant	-0.776^{***}
	(0.062)
Observations	3,835
Log Likelihood	−2,405.188
Akaike Inf. Crit.	4,834.376
Note:	*p<0.1; **p<0.05; ***p<0.01

Brief

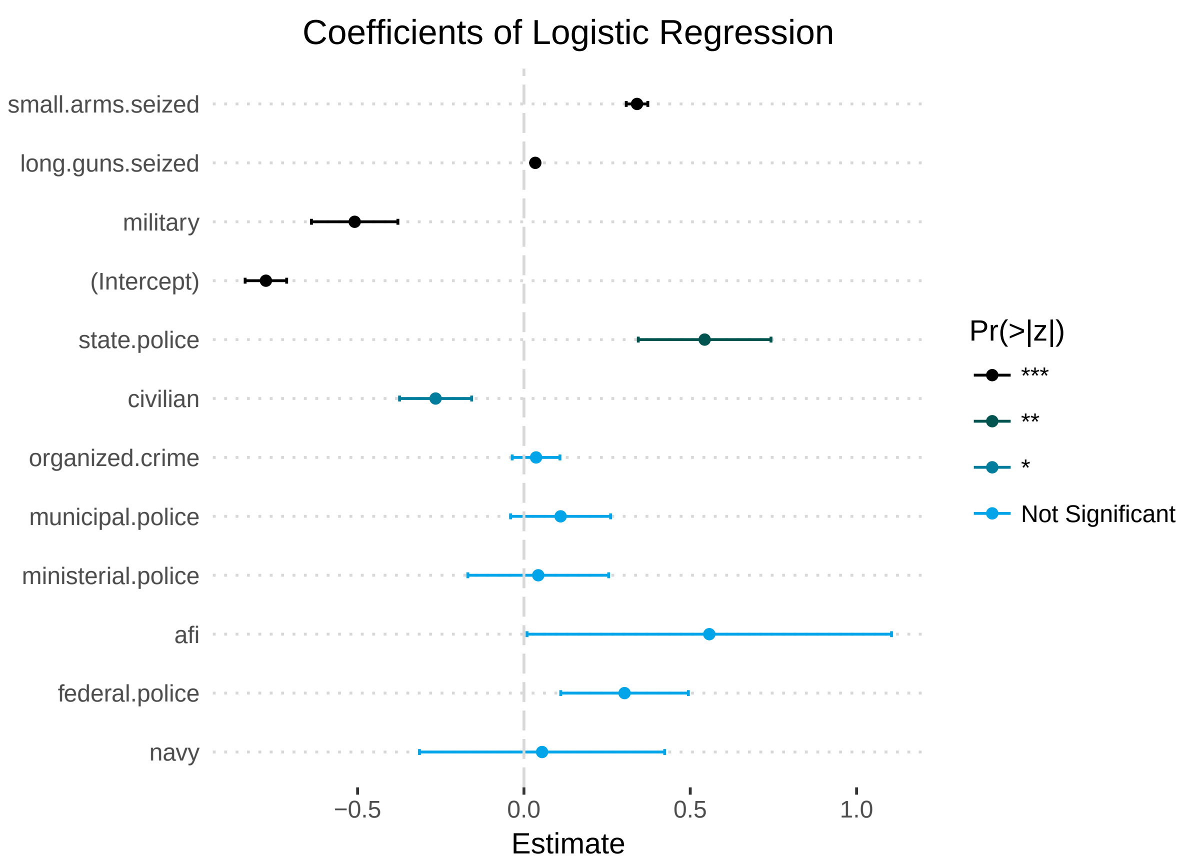
Military involvement negatively affects the incidence of there being detentions in a confrontation, almost with the same magnitude that state police involvement affects it positively.

Expanded

It seems that military and state police involvement are inversely related to the likelihood of there being detentions in a given confrontation. Specifically, the presence of the military affects the log-odds of detentions made negatively by -0.51, statistically significant, whereas the presence of state police affects them positively by 0.54, statistically significant. Regarding our controls, the size of the conflict seems to positively affect the log-odds of detentions for a given confrontation (0.03, and 0.34 for long guns and small guns seized, respectively. Statistically significant). This could sound weird, but we believe it is related to the weakness of using these variables as indicators of size of confrontation; in reality, what we believe is happening is that confrontations where there was seizure of guns are confrontations that were supposed to make seizures. That is, these confrontations are supposed to build a case, and, hence, seizure is more important in these cases. This goes to explain why they predict higher log-odds of detentions; detentions are probably made more where there is due process in mind. This explains why military presence negatively predicts the log-odds of detentions: their function is different from the police, and they are looking to neutralize an enemy, not collect evidence. Military intervention is probably more violent and less process-focused. This idea relates to the fact that civilian deaths and wounded also negatively affects the log-odds of detentions (-2.7, statistically significant). This opens the question, does the decision of involving the military implies taking a risk of involving civilians?

Our results are limited by the assumptions we have made: We can't completely assure that wounded and killed is an indicator of a particular actor's involvement in a conflict. At the same time, the analysis does not take into account the joint action of actors, which could add nuance to the analysis. Once more, the data is incomplete when it comes to involvement and it only refers to the cases that did happen, we do not have a clear view of different types of confrontations or what happened at the 0 level of our variables.

The next plot is a representation of our regression coefficients with confidence intervals



Task 2

Hypothesis 1

Going forward with what was observed in the previous section, we want see if we can distinguish differing effects of the presence of peace-keeping bodies and the type of confrontation observed. If the armed forces interventions are less process-focused, and this is observable in how armed forces presence affects the likelihood of detentions, are armed forces interventions related to a higher count of civilian deaths?

Our first hypothesis is that the presence of armed forces, when acting alone (not jointly with police forces), in cases where there are no seizures, has a higher incidence on civilian deaths than other possible scenarios.

We are assuming that such cases (because peace-keeping forces are not trying to build a case) are less measured in use of force and, hence, present a higher risk of civilian casualties. We are also assuming that the navy and military work in a similar way that is different from all police forces, who, in turn, work in a similar way themselves, i.e. trying to build a case and, thus, more focused on evidence collection. This could have an incidence in civilian casualties. Our main assumption is, again, that we can measure presence of a given group by sheer accounting of their wounded or dead. This is considering the fact that the data only refers to cases where there was a confrontation, leaving us with no reference for other cases.

Linear regression and data wrangling code

```
dq3 <- dq2

#dummy for cases where either military or navy involved:
dq3$armed_forces <- ifelse((dq3$military + dq3$navy) > 0, 1, 0)


#dummy for cases where state, ministerial, municipal or federal police involved;
dq3$police_forces <- ifelse((dq3$state.police + dq3$ministerial.police +
                                dq3$municipal.police + dq3$federal.police) > 0, 1, 0)

#Dummy for incidence of seizure
dq3$seizures <- ifelse(dq3$long.guns.seized + dq3$small.arms.seized +
                         dq3$cartridge.sezied + dq3$clips.seized +
                         dq3$vehicles.seized > 0, 1, 0)


ols <- lm(civilian.dead ~ armed_forces + police_forces +
            seizures + armed_forces * police_forces +
            armed_forces * seizures + police_forces * seizures +
            armed_forces * police_forces * seizures, data = dq3)


stargazer(ols, title = "Linear Regression Results")
```

Results

Table 2: Linear Regression Results

	Dependent variable:
	civilian.dead
armed_forces	−0.075
	(0.053)
police_forces	0.041
	(0.035)
seizures	−0.076***
	(0.019)
armed_forces:police_forces	0.100
	(0.238)
armed_forces:seizures	0.125**
	(0.063)
police_forces:seizures	0.036
	(0.050)
armed_forces:police_forces:seizures	−0.160
	(0.275)
Constant	0.134***
	(0.015)
Observations	3,835
R^2	0.006
Adjusted R^2	0.005
Residual Std. Error	0.515 (df = 3827)
F Statistic	3.499*** (df = 7; 3827)
Note:	*p<0.1; **p<0.05; ***p<0.01

Brief

There are no significant effects on civilian killed for armed forces involvement (with no collaboration of police forces) in confrontations where there were no seizures.

Expanded

As it stands, none of our interaction terms of interest have a significant coefficient in the regression for civilian wounded or dead. The effects of all combinations of our dummy variables hold no statistical significance. In this sense, we can say that our hypothesis was not right; confrontations led by the armed forces with no backup from police forces where there were no seizures of any kind do not significantly predict, in average, a higher number of civilian deaths. However, we can see that our significant coefficients do add nuance to the story: the incidence of seizures where there is no presence of armed or police forces does predict, in average, a 0.08 fewer civilian deaths (statistically significant). At the same time, it is interesting that the presence of armed forces wit no police forces, in cases where there are seizures does significantly predict 0.13 more civilian deaths. How are we to interpret these findings? We believe that the dichotomy comes from the duality in which we can interpret seizures: they are, perhaps, both an indicator of due-process oriented confrontations (as reflected by the coefficient of *seizures*), and of the size of the confrontation (as reflected by the coefficient of the interaction term *armed_forces:seizures*). The cases where armed forces are solely involved and there are seizures are more likely to be cases where the magnitude of the confrontation is bigger, given the nature of the variables that go into *seizures*.

The main limitations of our results are that there is no way to establish causality of civilian deaths in this case. The fact that we are assuming presence of bodies by killed or wounded does not assure us that there are no other factors involved (a fact reflected by the extremely small R2 of our regression). Related to this, our data is not big enough to make sure that there are enough cases where the different conditions of our interactions take place separately, so the regression can be applied to a less-than-desirable size and, thus, the results will not be completely accurate. Finally, the use of a linear regression assumes a linear relation between dependent and independent variables, and we have not run tests to confirm this is so.

Marginal Effects

Now, if we were still interested in the marginal effects of there being an armed forces presence, without police presence and no seizures, on civilian deaths, we would go as follows

Marginal effects and standard error code

```
beta1 <- coef(ols)                # vector of betas
varcov1 <- as.matrix(vcov(ols))   # estimated covariance matrix
se1 <- sqrt(diag(vcov(ols)))      # vector of standard errors

#we input 0 (zero) for the value of the interaction terms because they are dummies
#and we are interested in the marginal effects of armed_forces alone

mfx.1 <- as.numeric(beta1["armed_forces"]) +
  as.numeric(beta1["armed_forces:seizures"])*0 +
  as.numeric(beta1["armed_forces:police_forces"])*0  #which is the same as the beta for
                                                     #armed_forces in the model
                                                     #with interactions

mfx.1.se <- se1["armed_forces"] +
  0^2*se1["armed_forces:seizures"] +
  0^2*se1["armed_forces:police_forces"] +
  2*0*varcov1["armed_forces",
              "armed_forces:seizures"] +
  2*0*varcov1["armed_forces",
              "armed_forces:police_forces"] +
  2*0*varcov1["armed_forces:seizures",
              "armed_forces:police_forces"]     #which is the same as the se for
                                                #armed_forces in the model with
                                                #interactions

(cbind(mfx.1, mfx.1.se))
```

```
##                    mfx.1   mfx.1.se
## armed_forces -0.07543977 0.05320184
```

Hypothesis 2

Now that our more global concerns with the relation of armed forces, police forces, seizures, and civilian deaths have been clarified, we want to go back to check on something that our first logit model suggested. In order to substantiate our claim that police forces may be more process-oriented, and that this orientation is indicated by the incidence of seizures after the confrontation, we want to compare the effects of the presence of each body on seizures. Because our theoretical focus is on seizures in general, we propose a logit model to check on the likelihood of there being seizures at all.

Our hypothesis is, then, that the presence of police forces, without the presence of armed forces, will predict a higher likelihood of seizures than the joint presence of police forces and armed forces or armed forces alone.

We are assuming that seizures, in any form, are an indicator of the process-orientation of a certain action that ends in a confrontation. At the same time we are assuming that the presence of a body is indicated by there being members of this body killed or wounded.

Logit regression and data wrangling code

```r
dq4 <- dq3

logitq4 <- glm(seizures ~ armed_forces + police_forces +
            civilian + organized.crime + detention_incidence +
            armed_forces * police_forces,
            data = dq4, family = binomial(link = "logit"))

summary(logitq4)

stargazer(logitq4, title = "Lgistic Regression Results. Hypothesis 2")
```

Results

Table 3: Lgistic Regression Results. Hypothesis 2

	Dependent variable:
	seizures
armed_forces	0.627^{***}
	(0.128)
police_forces	-0.578^{***}
	(0.103)
civilian	-0.323^{***}
	(0.104)
organized.crime	0.536^{***}
	(0.071)
detention_incidence	1.286^{***}
	(0.077)
armed_forces:police_forces	0.590
	(0.564)
Constant	-0.257^{***}
	(0.065)
Observations	3,835
Log Likelihood	−2,352.281
Akaike Inf. Crit.	4,718.563
Note:	*p<0.1; **p<0.05; ***p<0.01

Brief

Opposite to our expectations, the presence of police forces by themselves actually decreases the likelihood of, whereas the presence of armed forces increases it.

Expanded

Our results are the opposite of what we expected. The presence of police forces alone decreases the log-odds of seizures, in average, by 0.58, statistically significant. At the same time, the presence of armed forces alone increases the log-odds of seizures, in average, by 0.62. At the same time, the presence of both these bodies does not significantly predict the likelihood of seizures. This could completely contradict our hypothesis, if it was not for the fact that detention incidence does strongly predicts an increase in the log-odds of seizures by 1.29, the larges coefficient in the regression; this could go with the idea that case-building logic is related to seizure of evidence. Together with this, the fact that civilian presence (killed or wounded) decreases, in average, the log-odds of seizures by 0.32 seems to point to the same direction, for messier confrontations that affect civilians may not be particularly centered evidence-gathering. Against the hypothesis, the presence of organized crime (killed or wounded) increases the log-odds of seizures, in average, by 0.54; if there are more killed or wounded members of organized crime (and given the data set we assume most, if not all, the confrontations involve organized crime members), that should indicate a bloodier confrontation and not one centered on evidence gathering, if we follow our previous logic.

In definitive, what we are looking at is the same problem of seizures as an indicator as before. The dichotomy of the regression results can be explained, perhaps, by the fact that seizures could indicate both process-orientation but also size of confrontation. A next step would be to confirm this with the data available. Our hypothesis remains false.

The main limitations of the results have to do with the inadequacy of seizures to indicate what we want to measure. Together with this, because of the flimsiness of our variables we can not really say that these are actual effects; we do not know if our assumptions are correct and, hence, we can not confirm that we are measuring sole presence of the bodies. We do not know what people were actually involved and we do not know what other types of incidents are there; this heavily limits our ability to establish causality or even a sense of representativity in the research. Finally, and this applies to much of our analysis, we are using a large amount of dummy variables as predictors, which can create problems of interpretation and it makes assumptions that are not completely correct for the analytic methods we are using.

Marginal Effects

Now, if we were still interested in the marginal effects of there being an police forces presence, without armed forces presence, on seizure, we would go as follows

Marginal effects and standard error code

```
beta2 <- coef(logitq4)                   # vector of betas
varcov2 <- as.matrix(vcov(logitq4))   # estimated covariance matrix
se2 <- sqrt(diag(vcov(logitq4)))      # vector of standard errors

#we input 0 (zero) for the value of the interaction terms because they are dummies
#and we are interested in the marginal effects of police_forces alone

mfx.2 <- as.numeric(beta2["police_forces"]) +
  as.numeric(beta2["armed_forces:police_forces"])*0  #which is the same as the beta for
                                                      #police_forces in the model
                                                      #with interactions

mfx.2.se <- se2["police_forces"] +
  0^2*se2["armed_forces:police_forces"] +
  2*0*varcov2["police_forces",
              "armed_forces:police_forces"] #which is the same as the se for police_forces
                                            #in the model with interactions


(cbind(mfx.2, mfx.2.se))
```

```
##                    mfx.2  mfx.2.se
## police_forces -0.5776325 0.1033327
```