

# HW2-INTL 601

Halil İbrahim Ergül

27 March 2022

## 1 Model Specification

Negative binomial is by far the best GLM model for the purpose of this paper. For one thing, the discrete outcome variable (nwound-n-p22-number of wounded) consists of counts that can take only values between 0 and infinity. Poisson distribution is not suitable for this case as it relies on huge assumption that expected value and variance are equal. The outcome variable in this model, however, contains a lot of zeros that indicate no attack or existence of wounded people in some districts. This means that the underlying distribution of my outcome variable carries the risk of over-dispersion owing to lots of zero points. Besides, the incidence of being wounded due to PKK attack can be considered as rare event that does not typically occur that much. Therefore, over-dispersed distribution of my outcome variable harboring lots of 0s and the nature of being a rare event, negative binomial stands as the best GLM model to proceed.

## 2 Control Variables

The main explanatory variable "margin" is numeric variable scaled from 0 to 1. As it approaches to 0, this means that closeness of vote share between Kurdish parties and the government increases as well. As control variables, I determined to include border (dummy variable with 1 as TRUE and 0 as FALSE), number of valid votes (validvotes2), urbanization rate of the

provinces (urbanization-rate), unemployment (unemployment), and infant mortality per thousand (infant-mort-perthousand).

Several reasons I can mention on why I have chosen them as controls. First, I assume that the number of valid votes that a municipality has shows as well the proportion or density of eligible voters that have right to vote in that particular municipality. Thus, I suspected that PKK might be targeting where they attack by considering both the vote margin (between government and Kurdish parties) and the significant number of eligible voters to create a terrorized atmosphere that may possibly go in favor of Kurdish parties. Also, I added the border variable, representing the closeness of particular province to the border of Iraq and Syria, because understanding PKK terrorist attacks without considering the geographical context would be meaningless. Indeed, this particular border serves as critical 'gateway' for PKK and various terrorist groups for years.

Besides, I assumed that urbanization rate, rate of unemployment, household income, and infant mortality per thousand are potential indicators or measures for socio-economic development that a particular municipality enjoys. Setting the issue of vote margin aside, I determined them as control variables because of the question that is it simply the case that because PKK is more likely to be well organized and historically rooted in less developed provinces due to relative authority gap, they are more likely to organize attacks in these peculiar municipalities.

After involving in exploratory data analysis with this data, I have found in bivariate analysis that there seems to be a meaningful pattern and a positive correlation between unemployment and number of wounded people in provinces like Hakkari where PKK is historically well-rooted. Similarly, I included infant mortality per thousand since literature on social welfare suggest less numbers of infant mortality is a potential indicator of developed health services which is reasonably related with economic development. Therefore, because these factors have high possibility of influencing the outcome variable in addition to main explanatory variable, I preferred to include them in the model to avoid the danger of lurking variable; thereby, arriving at healthier causal inferences and insights.

### 3 Results and Discussion

Table 1:

		<i>Dependent variable:</i>	
		nwound_n_p22	
margin		11.159***	(4.046)
border		2.998***	(1.081)
validvotes2		−0.00001**	(0.00001)
urbanization_rate		0.018	(0.016)
unemployment		0.197***	(0.065)
HHI		−25.287***	(6.462)
infant_mort_perthousand		0.063***	(0.013)
Constant		6.364**	(2.574)
Observations		904	
Log Likelihood		−357.050	
$\theta$	3	0.049***	(0.008)
Akaike Inf. Crit.		730.100	

*Note:*

\*p<0.1; \*\*p<0.05; \*\*\*p<0.01

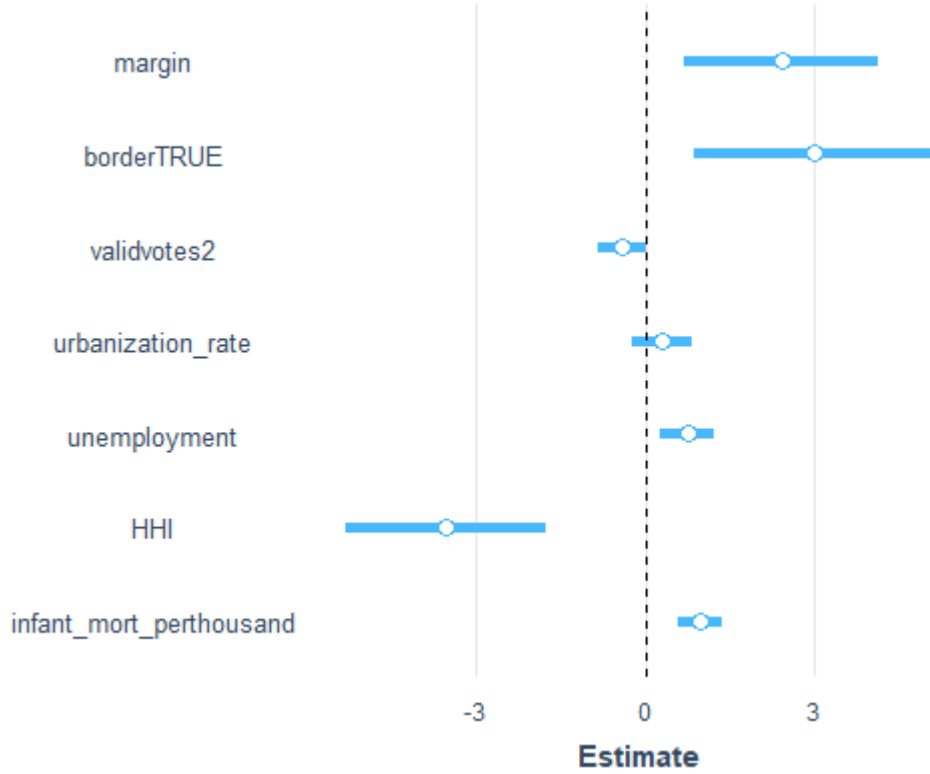


Figure 1: Coefficient Plot of the Model

The basic question this model tries to answer is whether or not the PKK is more or less likely to attack where the vote margin from the previous election is close. The results of negative binomial GLM model in which number of wounded people is dependent variable are given in the Table 1. The results of my model suggest that PKK is less likely to attack to provinces in where vote margin from the previous election was close. To put it differently, as the vote difference increases on behalf of one party, PKK is more likely to organize attacks resulting in high number of people to get wounded.

Table 1 shows that the main explanatory variable (margin) is one of the most important estimator of the number of wounded people owing to PKK attacks and it has a statistically significant and positive relationship with

the number of wounded people. Thus, it can be interpreted that holding other control variables at constant, as the closeness of the previous electoral results between the government and the Kurdish parties decrease, the number of wounded people increase accordingly by 11.15 with every one unit increase in margin. This is significant in the sense that PKK is presumably increasing its attacks in provinces where Kurdish or government parties enjoyed wide majority and ascendancy in the previous election as opposed to each other. Thus, the number of wounded people due to terrorist attacks is high in municipalities where there is a significant vote difference between Kurdish parties and the government. Moreover, contrary to my hypothesis, the control variable I added as to represent the number of eligible voters (`validvotes2`) has significant and negative relationship with the outcome variable. Hence, as the number of eligible voters in a particular municipality increases, the number of wounded people decreases accordingly.

Arguably, there might be two potential rationale why PKK is more likely to target places where vote differences are high. First, because PKK is organizing various attacks in places where government enjoys electoral ascendancy despite majority of Kurdish people in the east, one can argue that by targeting these electoral context PKK endeavours to create a fearful and forcing atmosphere for voters to prefer for Kurdish parties in the election. Second, it might be the case that PKK is more likely to organize attack in provinces where Kurdish parties already enjoy electoral ascendancy as to consolidate and reinforce their social support and suppress those voting for government so that Kurdish parties can manage to take all of the municipalities in a particular province.

Control variable of border also seems statistically significant and positively related in this model; thereby, showing that there is a significant difference between close and far provinces to this border in terms of the number of wounded people. Holding other variables at constant, the results suggest that the number of wounded people are high in those provinces which are closer to the border of Iraq and Syria as opposed to those provinces being far from it. Thus, it is important to bear in mind that geographical context stands as an important factor and estimator to predict the attacks of PKK

to particular provinces.

The control variables, which I added to make sure their influence on number of wounded people due to PKK attacks, that I included as indicators of socio-economic development in local context were unemployment, household income and infant mortality are significant in this model. For one thing, holding other variables at constant, unemployment seems to have a positive relationship with my dependent variable. Thus, as unemployment goes up, the number of wounded people in provinces owing to PKK attacks increase as well. Besides, there is a negative relationship between household income and number of wounded people. This demonstrates that as household income increases at municipality level, the number of wounded people decreases significantly. Thus, it can be inferred that PKK is actively organizing attacks in provinces where household income is relatively low. Another important indicator of economic development that has a significantly positive relationship with outcome variable is infant mortality. Results show that as infant mortality per thousand increases by one unit, the number of wounded people increase with a decent magnitude. However, urbanization rate that I included as control variable to represent economic development does not have a meaningful relationship with the number of wounded people.

Taking these indicators of economic development into consideration, it seems that PKK concentrate its activities and involve in more violent attacks in provinces where economic development is relatively low suffering from relatively high unemployment, high infant mortality, and low household income. Two potential reasons can be mentioned why PKK attacks are high in less developed provinces. First, PKK might have well-placed and organized in these municipalities in where government spending on economic infrastructure is very low. This relative authority gap or power vacuum may have given PKK the chance to take root and amplified its power. Second, it could be simply the case that because less developed municipalities do exist in the eastern region and they are inhabited mainly by Kurds and since PKK organize attacks mostly in the eastern cities, economic indicators seem significant in the model. But these questions cannot be answered by solely looking into this particular regression model.