1. The reason we can’t simply take data from multiple cities and run the regression of “Crime” on “Police” is due to the reverse causality problem that surrounds this regression. If we are trying to estimate the causal effect of the number of police officers in a city has on its crime rate, there is an issue of whether some cities have more police officers due to having a high crime rate, or if a city has a low crime rate because it has a high number of police officers. There is also an issue of potential lead-lag effects in regards to these variables. An immediate increase in the amount of police officers may not cause an immediate decrease in crime, or vis versa. Therefore, without using some sort of exogenous instrument that affects the number of police officers in a city, but does not affect the crime rate, we will be unable to find a causal estimate between them.
2. Researchers at UPenn were able to isolate the effect that the number of police officers in a city has on its crime rate by using the “Terrorism Alert System” as an instrumental variable. They found that when the system was placed at a terror alert level of orange in Washington D.C., the city increased the number of police officers it had patrolling the streets and in public places, and yet it would be highly unlikely that this terror alert would affect the number of crimes committed during this time period. Because of this, the terror alert level satisfied both criteria for being an instrumental variable since it was found to be correlated with the endogenous explanatory variable “Police”, since it caused an increase in the number of officers on the street, but was not correlated with the crime rate (due to this increase not being associated with the crime rate), or other factors in the explanatory equation such as the number of tourists in the city during this alert. By using this instrument, they found that the crime rate did in fact decrease during the time periods when the terrorism alert level was at orange, and thus when more cops were patrolling the streets. With this method, they were able to find that high alert days were correlated with a 7.316 unit decrease in the number of crimes committed, and a 6.046 unit decrease in the number of crimes committed during high alert days, when accounting for METRO ridership. Both of these estimates were statistically significant at the 5% level of significance.
3. The reason that the researchers controlled for Metro ridership is due to the fact that, if the number of metro riders in the city also decreased during high alert days, it is likely that the number of crimes that occurred in the city would decrease as well. By controlling for the effects of ridership on the number of crimes, they can show that, although high alert days decrease the crime rate in Washington D.C, its effect is not as high as when accounting for the number of tourists or citizens using the metro system in the city.
4. This regression is estimating the effect that high alert days have on crime in different areas of D.C., with a focus on the National Mall (District 1). Separating D.C. into districts finds that, when the city is on high alert, there is a large decrease in the crime rate in District 1 of 13.679 units that is statistically significant at the 1% level, holding all else fixed, while there is only an 11.629 unit decrease in other districts during high alert days that is not statistically significant, holding all else fixed. This is likely due to the fact that large numbers of new police officers are stationed at the mall during high alert days due to it being a likely target for terroristic acts, and thus in need of more security. Due to this large increase in police officers, there is a strong, statistically significant, decrease in the crime rate during this time period that is not directly due to the crime rate of the area itself.