Israel Diego (UMID: israeldi)
Stats 504
October 24th, 2019

Testing for Non-Addtivity between Temporal Variation in Chicago Crime Reports

By working with Chicago crime data we can gain insight into the time trends of number incidents reported per crime. For example, we note that for the most common crimes in Chicago, the number of incidents reported have been decreasing over the long-term, although recently, some reported crimes have begun to increase. Our initial approach for modeling the frequency of crime reports incidents over the long-term and within each year is to use quasi-Poisson regression that is additive in its predictors, which include predictors that capture the long-term trend, the seasonal trend, and day-of-week effect. We aim to study whether temporal variation in the frequency of incidents reported should instead be non-additive. In other words, a non-additive model would imply that the day-of-week effect would be different depending on the time of year.

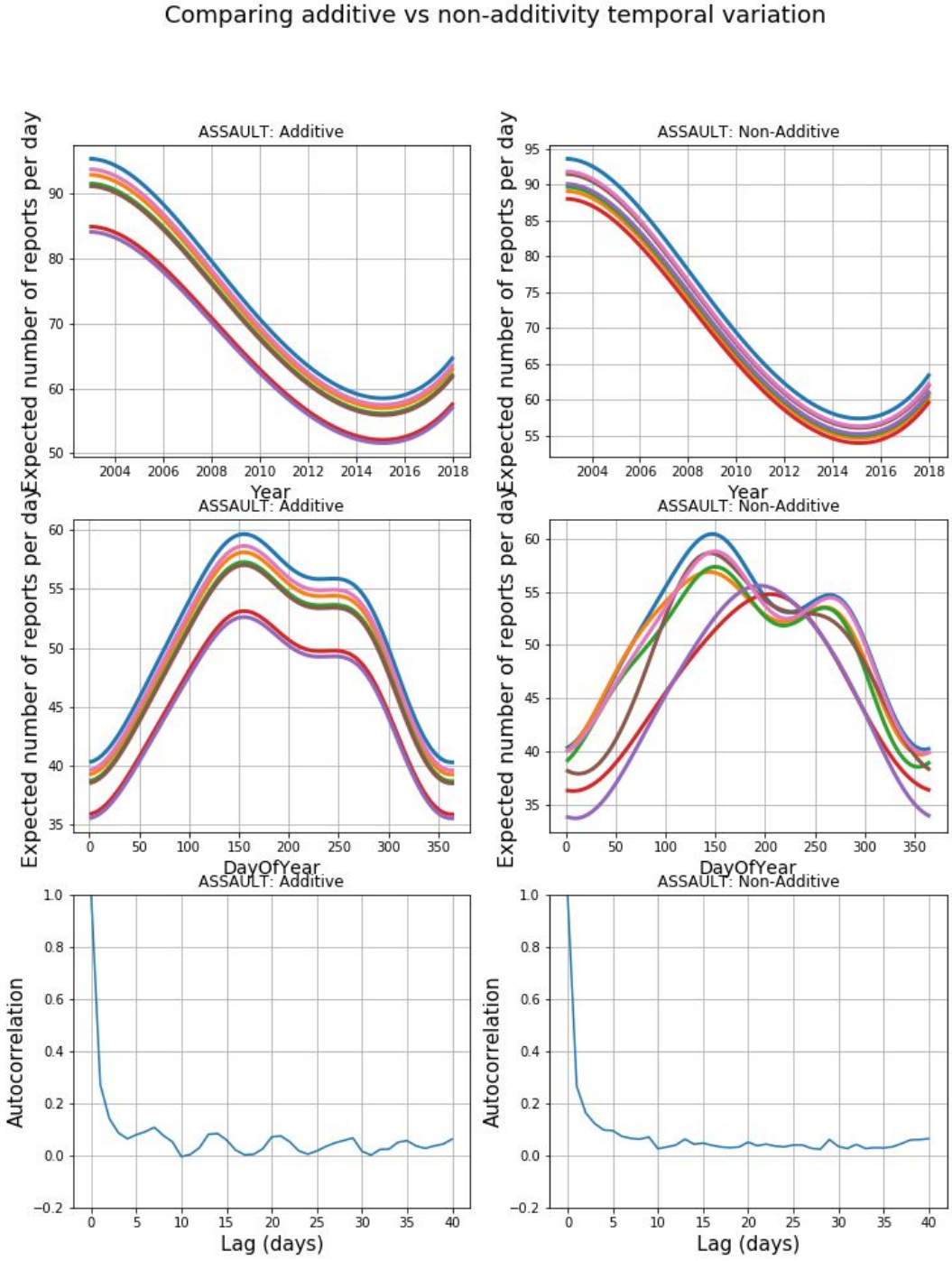
Chicago crime data comes from the city of Chicago's online data portal. The data contains information on crime reports across 79 community areas in Chicago. For our analysis, we aggregate crime reports for all community areas and focus on the 10 most commonly reported crimes in Chicago. As a result, we are working with daily data on the number of crimes reported for each of the 10 crimes from 2003-2019. The crimes include: assault, battery, burglary, criminal damage, deceptive practice, motor vehicle theft, narcotics, other offense, robbery, and theft.

In order to study the time trends on the frequency of crime reports in Chicago, we use a GLM approach. Our decision to use GLM rather than classical linear regression is motivated by modeling the mean-structure of the crime report data. We observe that for each crime type, the empirical conditional variance grows linearly with the empirical conditional mean. Thus, quasi-Poisson regression is a good candidate for our GLM. The long-term trend is modeled by basis splines with 4 degrees of freedom. The day-of-year seasonality effect is modeled by Fourier basis: $\sum_{n=1}^{3} a_n cos(\frac{2\pi nx}{365.25}) + b_n sin(\frac{2\pi nx}{365.25})$ and the day-of-week effect is modeled by indicator variables.

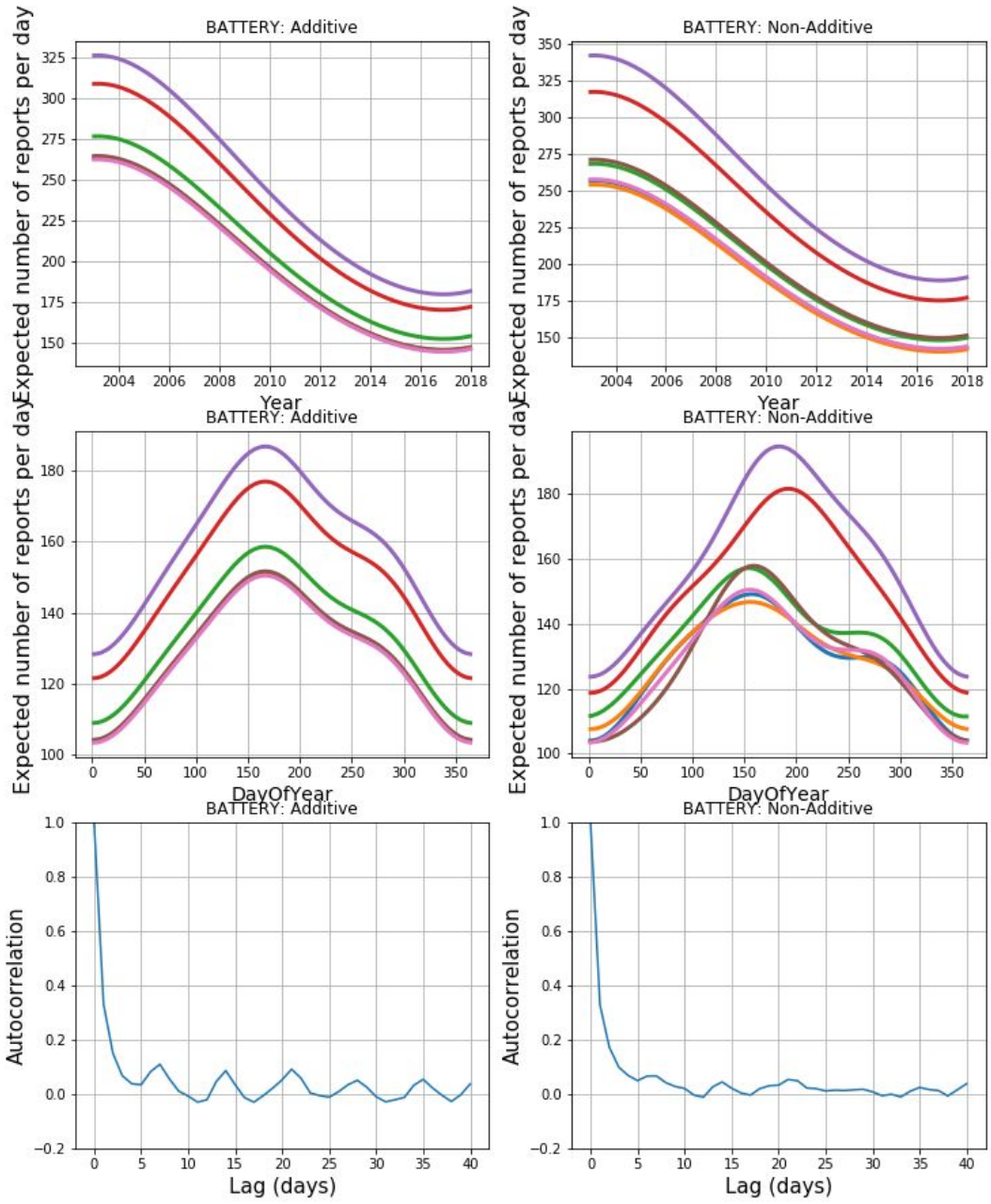
We assess for presence of non-additivity between temporal variation by plotting the autocorrelation of the residuals for each crime type and compare the residual plots of the additive model and the non-additive model with interactions between the temporal predictors. For the additive model, crimes such as assault and battery display a persistent sinusoidal pattern in the residuals that occurs weekly and oscillate between 0 and 0.1. On the other hand, the autocorrelation in the residuals for other crimes, such as deceptive practice and theft, spike upward roughly on the 31st day. Both situations situations suggest that there is temporal variation not captured by our current (additive) model.

Incorporating non-additivity by means of interacting season and day-of-week predictors, in the case of assault and battery, the autocorrelation between lagged residuals looks closer to zero and the interaction variables help remove the weekly jumps in the autocorrelation function. Although assault and battery have similar autocorrelation plots, they differ in that assault is predicted to be most frequent in the summer weekdays, whereas battery is most predicted to occur most frequently in the summer weekends. Burglary also displays similar improvement on its residuals. However, for all other crimes, the residual plots look the same as in the additive model. For example, for crimes with jumps at the 31st lag, the interaction terms did not contribute to capturing this monthly dependence. Thus, we can infer that there is no evidence of non-additivity in the temporal variation for all other crimes. In addition, introducing interaction variables was not effective in mitigating short-range and long-range dependence. Examples of models demonstrating short-range dependence were criminal damage, while models with long-range dependence included robbery, narcotics, motor vehicle theft, and burglary.

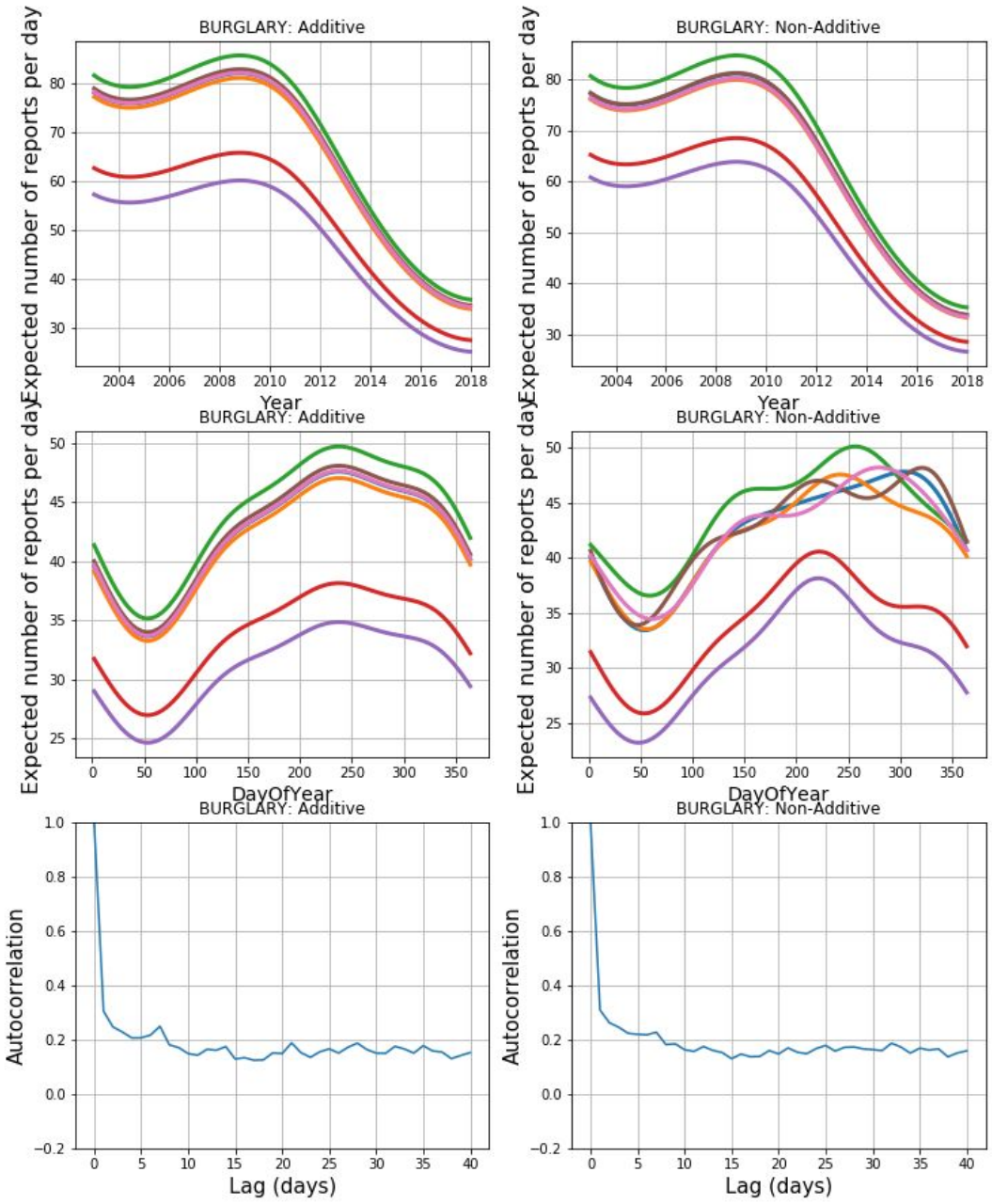
We found that only assault and battery exhibited non-additivity between season and day-of-week. All other crimes did not experience an improvement in the autocorrelation of the residuals, implying that we do not have evidence of non-additivity being present in these crimes.


Comparing additive vs non-additivity temporal variation
ASSAULT: Additive
ASSAULT: Non-Additive
Expected number of reports per day
Year
DayOfYear
Autocorrelation
Lag (days)

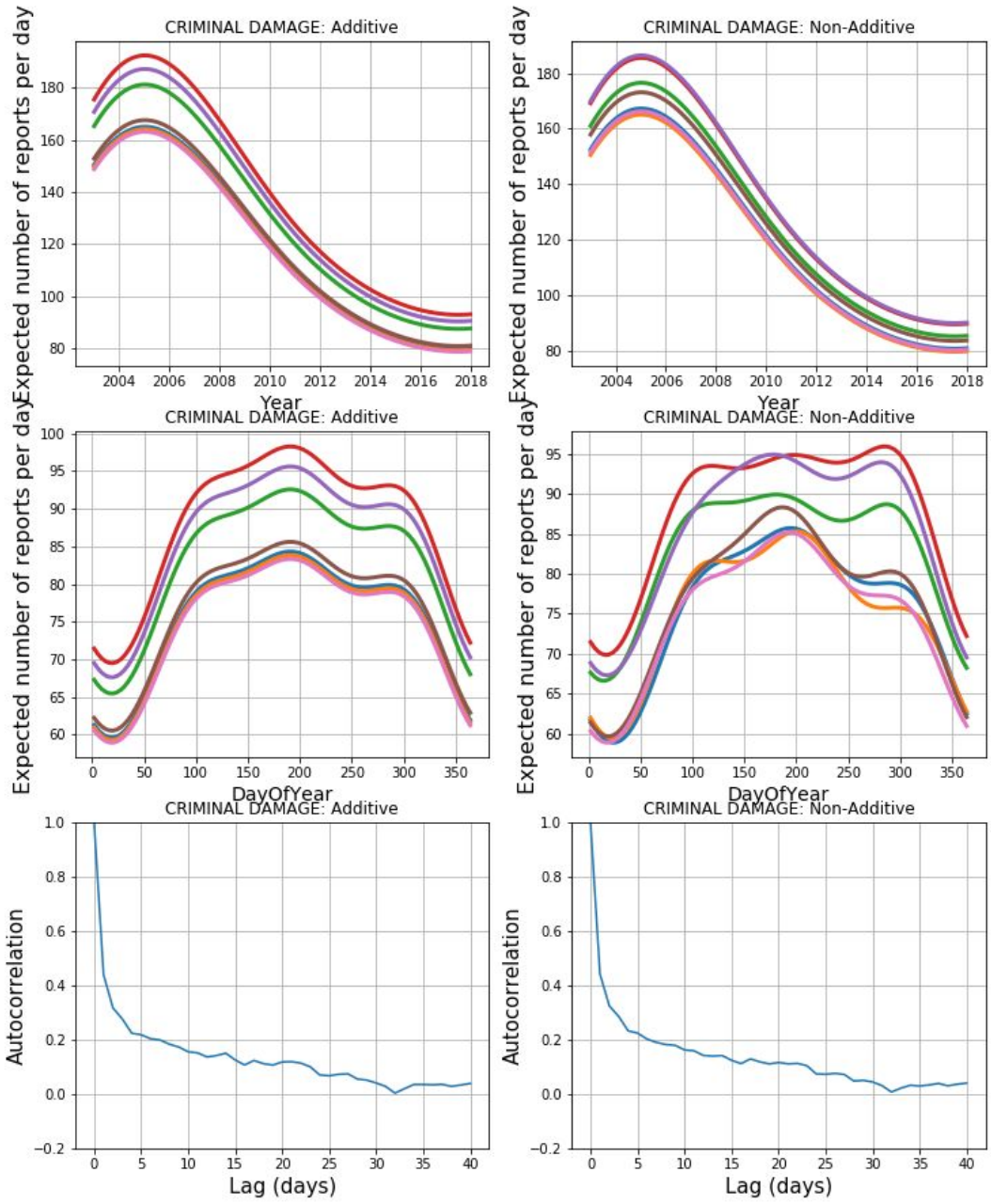
Comparing additive vs non-additivity temporal variation



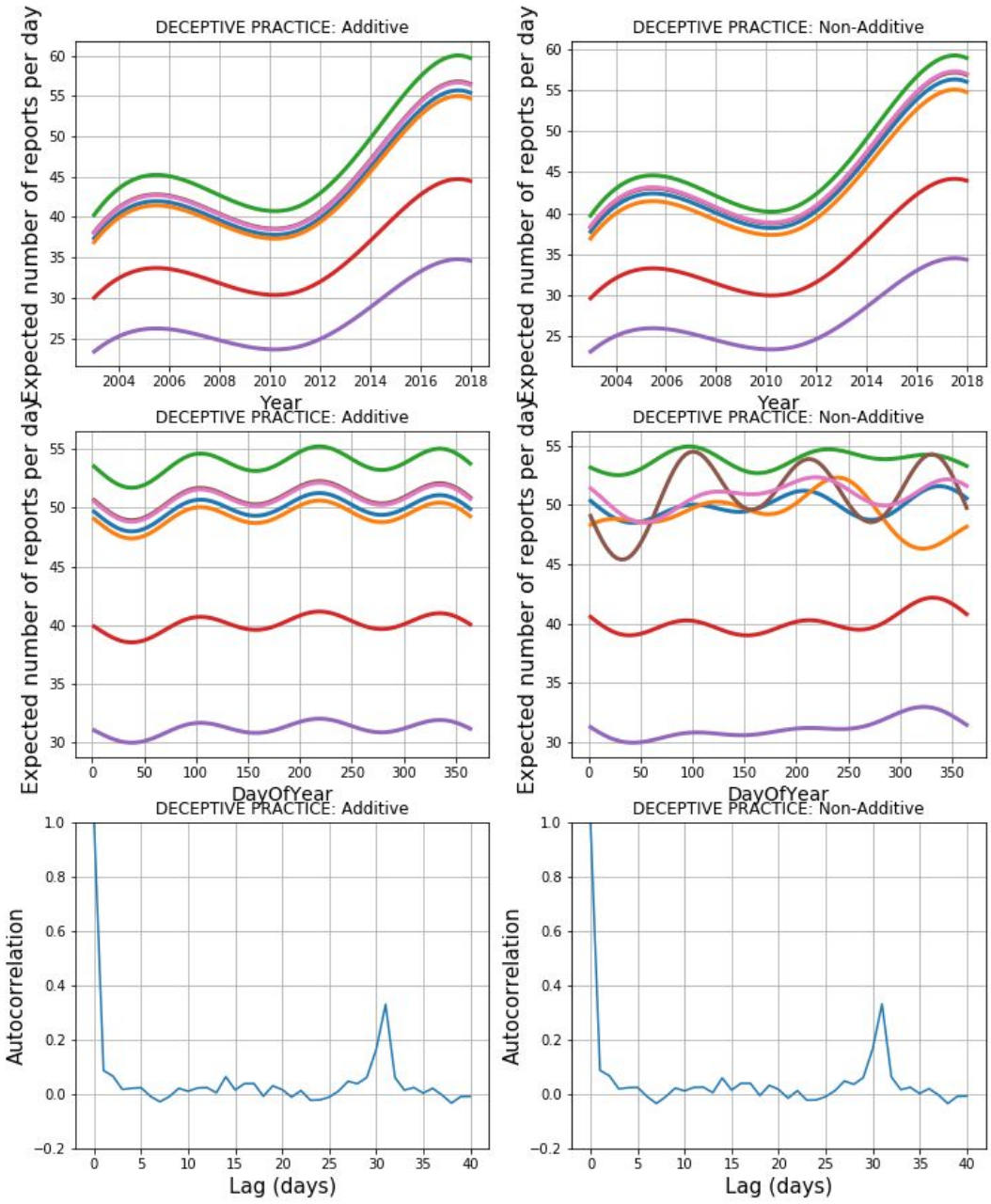
Comparing additive vs non-additivity temporal variation



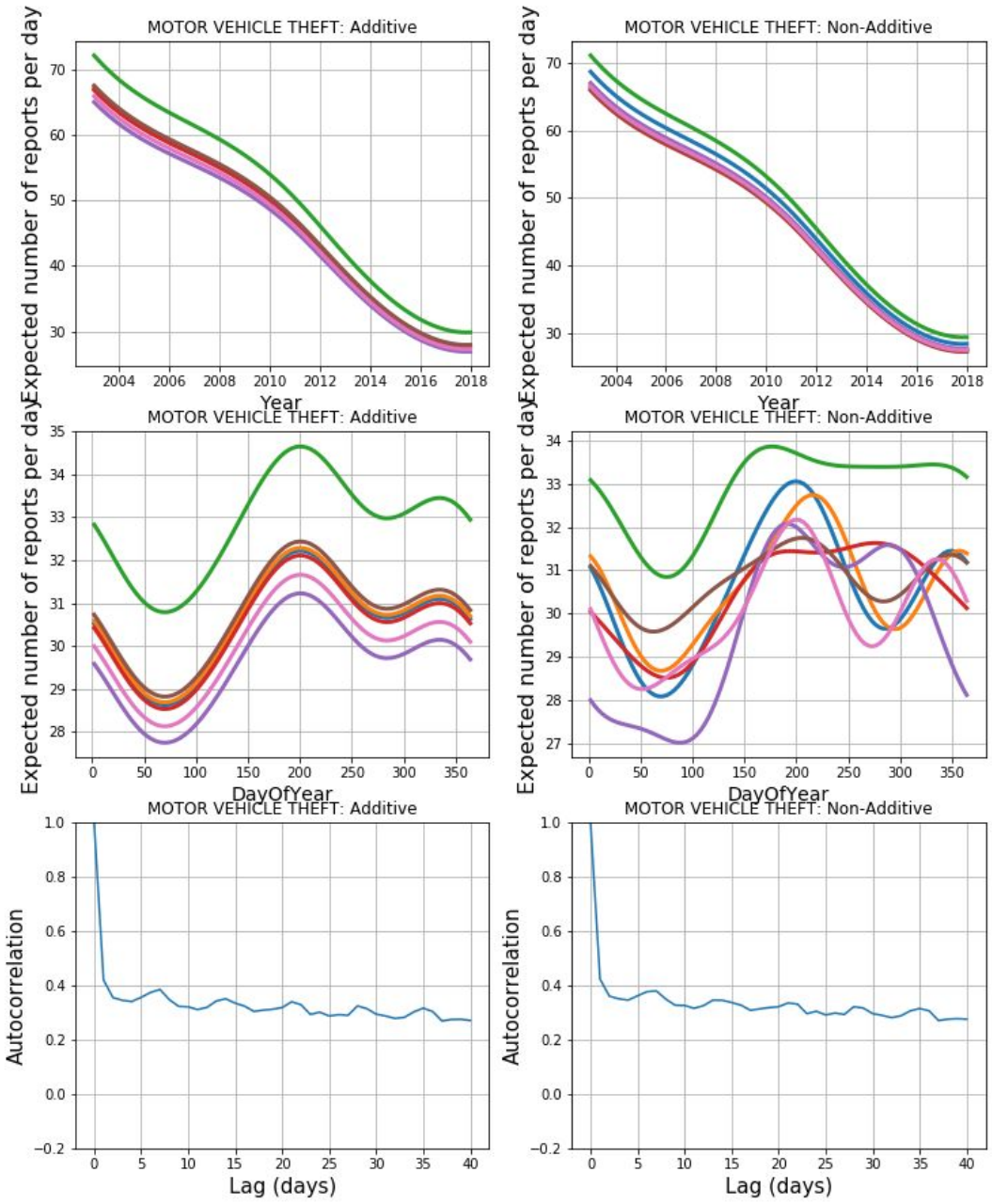
Comparing additive vs non-additivity temporal variation



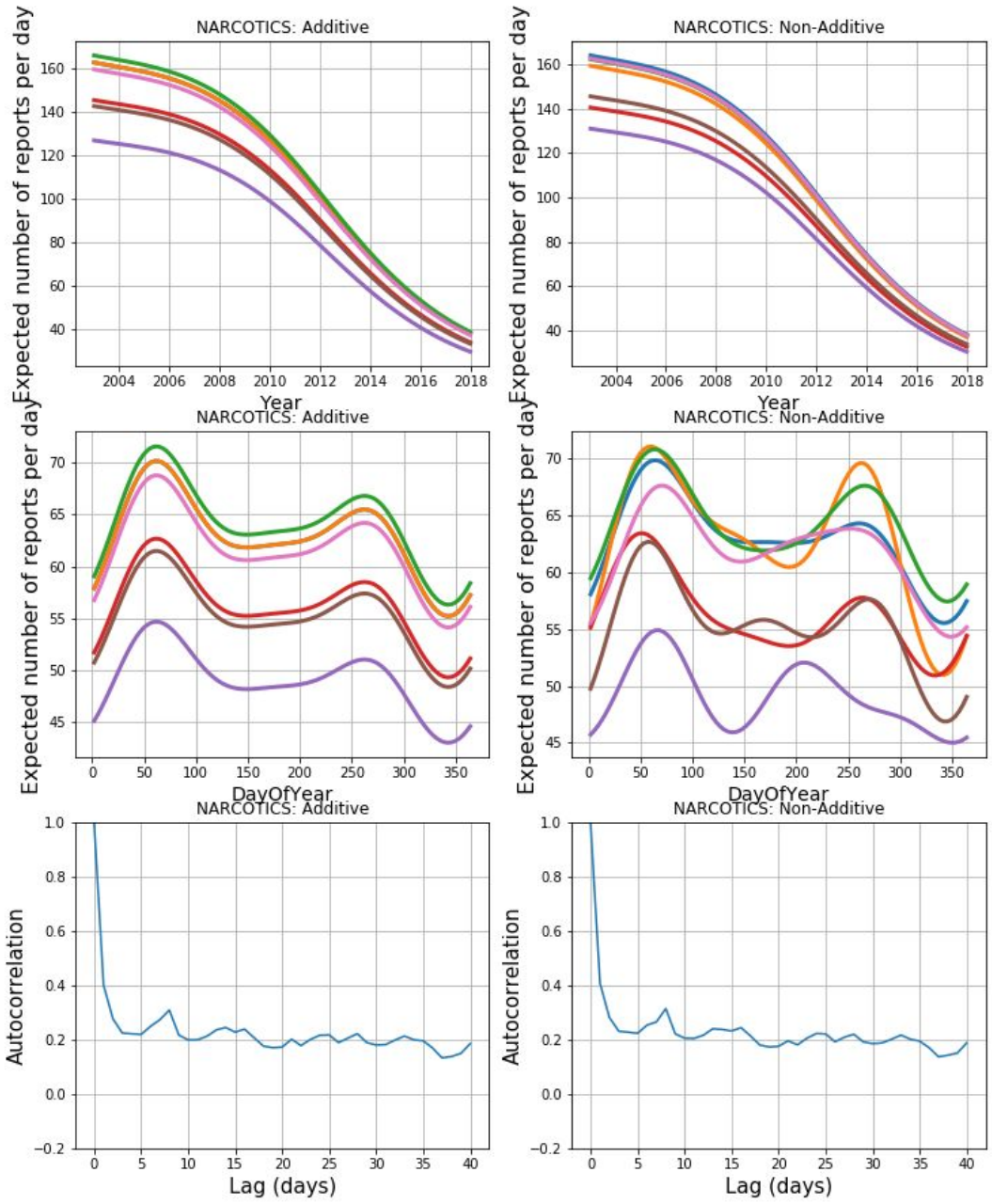
Comparing additive vs non-additivity temporal variation



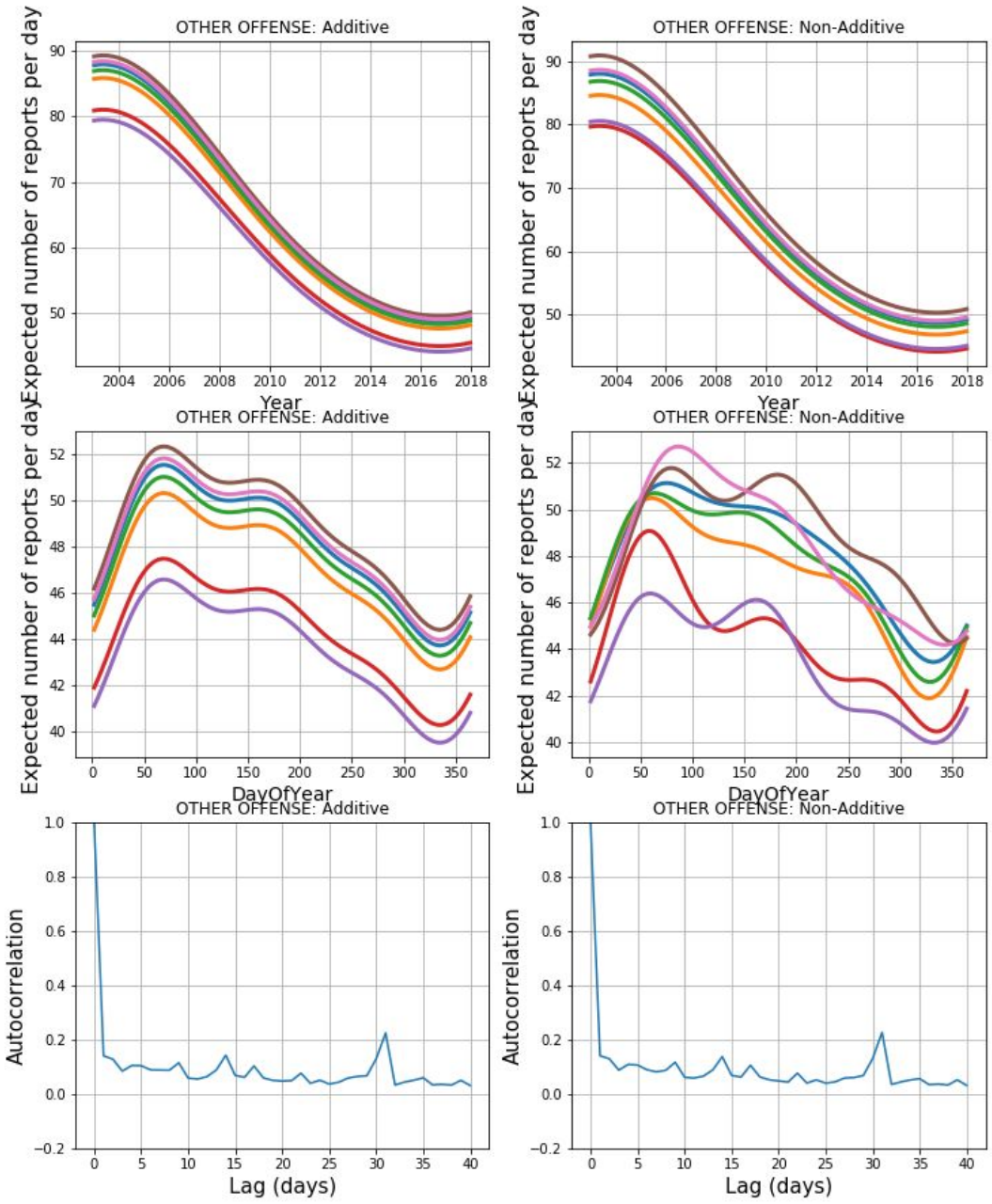
Comparing additive vs non-additivity temporal variation



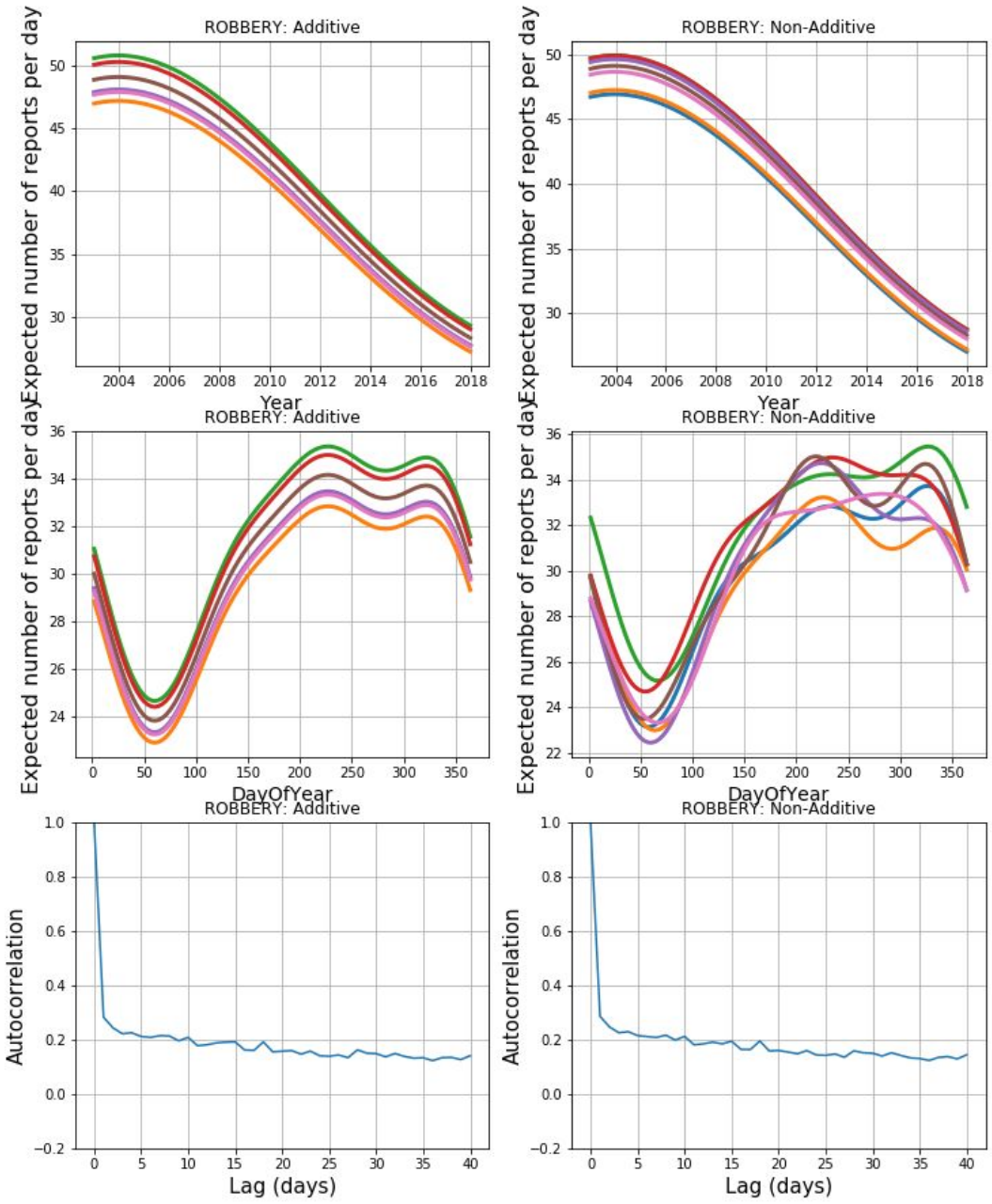
Comparing additive vs non-additivity temporal variation



Comparing additive vs non-additivity temporal variation



Comparing additive vs non-additivity temporal variation



Comparing additive vs non-additivity temporal variation