Syjane Eang

Christian Krestel

Econ 687: Economic Applications of R Programming

17 December 2023

The effect of pollution on crime: Evidence from data on particulate matter and ozone

Introduction:

Connections between exposure to pollution and health outcomes have been identified across various fields of study. Both epidemiological and toxicological investigations have established a robust foundation of evidence supporting a broad spectrum of short- and long-term health impacts linked to the exposure to airborne particulate matter. Extensive research has confirmed the detrimental effects of air pollution on physical health and cognitive function, resulting in significant costs for individuals and society as a whole. Although the impact on mental health is an aspect that has received limited exploration, the potential consequences emphasize the necessity for a comprehensive understanding of the societal expenses associated with pollution.

The paper aligns with the wider body of literature exploring factors that influence criminal behavior and the associated societal expenses, as discussed by Bishop and Murphy (2011). Some scholars have suggested a direct connection between the frequency of criminal activity and the likelihood of arrest. Meanwhile, other studies have emphasized the influence of unexpected losses in Sunday night football games on domestic violence. Additionally, research has established connections between diminishing daylight and rising crime rates, as well as linking crime to oil and gas development.

The only existing study by Lu et al. (2018) combines yearly crime data with annual pollution metrics, uncovering a connection between elevated air pollution levels and heightened criminal activity. Moreover, three unpublished papers in the field of economics investigate the correlation between pollution and crime. Research also suggests that temporary rises in pollution can contribute to increased crime rates in urban settings, as demonstrated by Bondy et al.'s (2018) examination of daily pollution exposure and crime in London. Exploiting observations on EPA policy reveals a reduction in pollution on monitored days, which is linked to lower crime rates.

This study distinguishes itself in three key ways. Firstly, it creates a daily, county-level dataset covering the entire continental United States. Secondly, it utilizes unique data properties to estimate dose-response functions, explore avoidance behavior, and evaluate the impact of external factors such as wildfire smoke and temperature. Lastly, despite these methodological distinctions, the study reveals remarkably similar effects to those observed in previous research, collectively presenting compelling evidence of a new consequence of pollution on aggressive behavior.

Data and Descriptive Statistics:

The authors used an extensive dataset that integrated multiple types of data to examine the relationship between air pollution and crime. This included detailed crime data, which provided information on the types and frequencies of crimes across different geographic locations and time periods. Air pollution data, specifically measurements of particulate matter and ozone levels, were central to the study, enabling the authors to assess the correlation between pollution levels and crime rates. Additionally, weather data was incorporated to account for the influence of weather conditions on both pollution levels and crime. The integration of these diverse datasets was very important to provide a robust analysis of the potential impact of air pollution on crime rates, ensuring that the study accounted for various factors that could influence or bias this relationship.

The crime data is central to the study, providing detailed insights into the types, frequencies, and locations of crimes. This information is crucial for identifying patterns and understanding whether there's a notable correlation between air pollution levels and crime rates. On the other hand, the air pollution data, focusing on particulate matter and ozone, offers a direct measure of environmental quality. This allows the researchers to quantitatively assess the extent to which air quality impacts criminal behavior. Lastly, incorporating weather data adds another layer to the analysis. Weather conditions can influence both the dispersion of pollutants and human behavior, making it a necessary factor to consider for a more accurate and comprehensive understanding. By integrating these diverse datasets, the study aims to provide a nuanced view of how environmental factors might influence social outcomes, bridging the gap between environmental science and social economics.

The inclusion of the demographic and socioeconomic data in the study is crucial. It allows the researchers to control for factors that might influence both crime rates and the distribution of pollution, ensuring that any observed relationships are not confounded by these underlying variables. By accounting for these demographic and socioeconomic factors, the analysis becomes more robust, and the researchers can be more confident that they are observing a true effect of pollution on crime.

Methodology:

 eq (1)

Equation (1) displays the empirical specification used by the authors to measure the effect of pollution on crime. Crimect, represents the count of crime of type j in county c on day t; PM25ct is a measure of air pollution in county c on day t; Ozonect, is the daily mean ozone level in ppm in county c on day t; Xct is a vector encompassing control variables such as minimum and maximum temperature and precipitation; 𝜙ct accounts for county-by-year-by-month-by-day fixed effects and the error term captures any unobserved factors that may influence crime counts but are not explicitly accounted for in the specified variable.

The authors employ the Poisson quasi-maximum likelihood (PQML) methodology to estimate models for the non-negative count variable crime\_jct. PQML is chosen due to over-dispersion in the crime variable, a situation where the variance of crime observations exceeds what is expected under a Poisson distribution. This method is robust for handling over-dispersion and provides consistent parameter estimates under the assumption of a correctly specified conditional mean function.

To further adjust for overdispersion, the authors also calculate fully cluster-robust standard errors, adjusted for correlations within counties, ensuring correct inference. Additionally, various specifications, including OLS, negative binomial, Bayesian Information Criterion, and PQML without robust standard errors are employed to assess the robustness of PQML. The findings suggest that PQML is the most suitable specification for the dataset, since estimates for other specifications were not meaningfully different from PQML estimates. The PQML robust standard error model is chosen over the standard PQML model because in the standard model, standard errors are smaller due to overdispersion being present, which distort our understanding of the estimates and overall relationship, since the estimates are inconsistent and biased.

To mitigate endogeneity and control for unobservables related to the correlation between pollution and crime rates with location and time-varying factors, the authors incorporate high-dimensional fixed effects (county-by-year-by-month-by-day of the week) in order to achieve unbiased estimates for mean\_Ozone and pm25. These fixed effects account for constant and time-variant factors, such as changes in policies, demographics, population density, and seasonal variations. Temperature and precipitation splines are also included to control for weather effects, acknowledging the correlation between crime and temperature. This comprehensive approach helps minimize the risk of biased estimates and enhances the robustness of the findings.

Replication & Results:

Table 3 (figure 2 in the appendix section) presents the outcomes of equation (1) with a focus on violent crimes, employing various fixed effects to examine the impact of unobservable factors on the relationship between air pollution and crime rates.

In Column 1, Year-by-month, state, and day of the week fixed effects are included. Column 2 incorporates County fixed effects, while Column 3 introduces County-by-year-by-month and day of the week fixed effects. The most comprehensive model is found in Column 4, encompassing County-by-year-by-month-by-day of the week fixed effects.

The model estimates represent average effects. In Column 1, ceteris paribus, a one-unit increase in pm2.5 is associated with a 1.1% rise in the likelihood of violent crime. Similarly, in column 2, ceteris paribus, a one unit increase in mean\_Ozone is associated with a 0.9576 rise in the likelihood of violent crime.

To account for potential positive correlations between air pollution (PM2.5 or ozone) and crime rates with unobservable factors, the authors expect coefficient estimates to decrease from Column 1 to Column 4. Indeed, the estimate on PM2.5 decreases notably between Columns 1 and 2 due to the inclusion of a county fixed effect, and continues a modest decline from Columns 2 to 4. Crucially, both PM2.5 and ozone coefficients remain relatively stable after including some form of county fixed effect (Columns 2–4), suggesting effective control for endogeneity in air pollution through the county fixed effect.

Columns 5 and 6 examine individual pollutants, with Column 5 focusing on pm2.5 and Column 6 on mean ozone. The pm2.5 effect remains consistent even without ozone, but excluding pm2.5 slightly amplifies the ozone effect. Despite a low correlation of 0.15 between pm2.5 and ozone, both pollutants appear to impact violent crime. The authors conclude that the optimal model is in Column 4, incorporating both pm2.5 and mean ozone, along with the full set of fixed effects.

In order to replicate table 3, we started by defining new variable names for various categories, such as `stmax\_vars`, `stmin\_vars`, `sppt\_vars`, and `dowd\_vars`. We proceeded to build six Poisson regression models using the `fepois` function from the `fixest` package, dedicating each model to a specification outlined on in the empirical methodology section, with the primary predictors being `pm25\_wt` (particulate matter) and `mean\_Ozone` (ozone levels). We ensured the inclusion of high-dimensional fixed effects tailored to each model's specifications and clustered standard errors by `fipsid` to address potential correlations within clusters. For the final step, we generated a comprehensive table summarizing the models. We compiled a list of models, labeled `model\_list` and spanning Models 1 to 6. We utilized the `etable` function from the `fixest` package to create a tabular summary, displaying coefficient estimates, standard errors, statistical significance, number of observations and the model’s R-squared.

In table 4 The findings indicate that the increase in violent crime is primarily attributed to assaults, While a slight, negative, and marginally significant correlation is observed between increased ozone exposure and robberies, there is no significant statistical impact of PM2.5 changes on property crimes or vehicle thefts.

Table 4 involved constructing 4 Poisson regression models to analyze the impact of air pollution (particulate matter and ozone levels) on various types of crimes assaults, robberies, property crimes, and vehicle thefts, using the primary model from Table 3. Using R's `fepois` function from the `fixest` package, models were developed for Assaults, Robberies, Property Crimes, and Vehicle Thefts. Key to this analysis was handling high-dimensional fixed effects and clustering, crucial for controlling unobserved heterogeneity. Then we built a function, `extract\_coef`, to extract estimates and standard errors for key pollutants from the complex `fepois` model summaries. This extraction was essential due to the unique structure of the output poison model, which standard methods couldn’t directly address.

One limitation that we encounter in the replication code has to be done with a `fepois` function's output, meaning changes in the function or package could necessitate code adjustments in R. Its specificity to the study's dataset and variables limits its direct applicability to other contexts without significant modification. Additionally, the reliability of the analysis hinges on the statistical assumptions of the Poisson model and the chosen model structure. Therefore, while the methodology is robust within the study's framework, it may not universally apply to other datasets or different types of regression analyses without careful consideration of underlying assumptions and contextual factors.

We also recreate Table 7 using the same PQML method. This table adds temperature and smoke interactions to test whether temperature or wildfire smoke plumes affect air pollution and violent crime. Literature indicates that temperature affects cognitive function and behavior. Smoke from wildfires does increase particulate matter PM2.5 and may increase ozone mean. Our first replication model in column 1 indicates that the effect of ozone on violent crime is largest in the 3rd quartile of ozone, which is the same deduction from the paper’s coefficients. However, our replication does not see that change in our 3rd quartile of ozone on assaults. The results indicate that violent crime rises approximately linearly with daily maximum temperatures. The results from the paper and our replication are consistent with prior literature on the effect of temperature on crime. This may signify the importance of sunlight in the ozone formation process, since more sunlight tends to mean more ozone. Sunlight, ozone, and daily maximum temperature go hand in hand when one increases.

There is no significant indication that the effect of wildfire smoke on crime is different from the effect of PM2.5. The smoke indicator, HMS, is highly correlated with PM2.5 and also the amount of smoke days is small relative to the number of observations. Only 1.5% of the dataset had smoke plumes because smoke plumes are not consistent on the daily so this is a caveat. The paper does mention potential measurement error because they applied the kriging process for estimating missing samples of data. But, they mention that it should not have introduced significant additional measurement error.

Figures 2 & 3:

Figure 2:

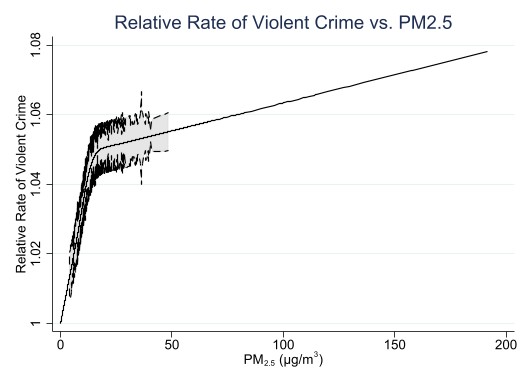
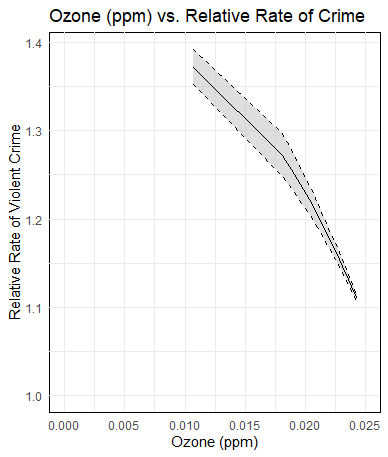


Figure 3:



In the construction of figures 2 and 3, the author adopted a bifurcated methodology to estimate the dose-response functions. The initial phase involved the representation of PM2.5 concentrations in equation (1) via a restricted cubic spline. This spline was anchored at knots located at the 25th, 50th, 75th, and 95th percentile with distributions of PM2.5, equating to concentrations of 5.62, 8.37, 12.26, and 28.47 g/, respectively. Subsequently, a parallel approach was utilized for modeling ozone levels, applying a restricted cubic spline with knots at the 25th, 50th, 75th, and 95th percentile values, which correspond to 0.023, 0.031, 0.038, and 0.049 ppm. This dual-spline methodology enabled a comprehensive exploration of the dose-response dynamics between PM2.5 and ozone concentrations.

Regarding Figure 2, the derived spline-based model elucidates that the relative incident of violent crime escalates concomitantly with PM2.5 concentrations up to approximately 20 g/. Beyond this concentration, the curve exhibits a plateau, suggesting a diminishing marginal impact of additional PM2.5 on violent crime rates.

As for Figure 3, it illustrates that the impact of ozone on violent crime amplifies up to about one standard deviation above the mean daily ozone level. Past this junction, the graph indicates an inflection point, where the rate of violent crime commences a downward trajectory. This pattern may signify a nonlinear correlation, wherein the influence of ozone on violent crime intensifies up to a certain ozone concentration but subsequently recedes beyond that threshold. The observed decline post one standard deviation above the mean could be indicative of various phenomena, such as a saturation effect, where heightened ozone levels cease to further augment the likelihood of violent crime.

Conclusion & Policy Implications:

The study conducted by the authors presents a comprehensive analysis of the relationship between air pollution and crime rates, drawing upon an extensive dataset that includes detailed crime data, air pollution measurements, and weather conditions. Employing the Poisson quasi-maximum likelihood (PQML) methodology to account for overdispersion in the crime variable, the study meticulously adjusts for potential biases and endogeneity issues. This is accomplished through the use of high-dimensional fixed effects, cluster-robust standard errors, and the inclusion of control variables like temperature and precipitation.

The key findings of the study, as presented in Tables 3 and 4, suggest a nuanced relationship between air pollution and crime. It is observed that increases in PM2.5 and ozone levels are associated with a rise in violent crimes, particularly assaults. However, the impact on other types of crimes, such as property crimes or vehicle thefts, appears to be insignificant. This differential effect underscores the complexity of the relationship between environmental factors and social outcomes.

Moreover, the study goes beyond mere statistical correlations to explore the dose-response relationship between air pollution and crime. Figures 2 and 3, using a bifurcated methodology with restricted cubic splines, reveal interesting patterns in the impact of PM2.5 and ozone on violent crime rates. These patterns suggest a nonlinear relationship, with the influence of pollutants on crime increasing up to a certain threshold and then plateauing or even declining.

The research also takes into account the potential effects of temperature and wildfire smoke on the relationship between air pollution and crime. This aspect is particularly relevant given the increasing frequency of extreme weather events and wildfires, which can significantly alter local air quality.

In terms of policy implications, the study sheds light on an often-overlooked social cost of pollution: its potential to increase violent crime rates. This finding is crucial for policymakers, as it suggests that efforts to improve air quality are not only beneficial for public health and the environment but could also contribute to reducing crime rates. The study thus calls for a more integrated approach in policy formulation, where environmental and social factors are considered in tandem.

Additionally, the study's methodology, particularly its approach to addressing over-dispersion and endogeneity, sets a precedent for future research in this area. It demonstrates the importance of rigorous statistical methods and the careful consideration of various confounding factors to arrive at more accurate and reliable conclusions.

In conclusion, the study offers significant insights into the complex interplay between environmental quality and social phenomena like crime. Its findings emphasize the multifaceted impacts of pollution and provide a compelling argument for more holistic environmental and social policies.

Elements Not Reproduced In This Report:

The tables within the appendix, include methodological details and robustness checks, we decided not replicate these to maintain focus on the core results of the paper. The appendices offer rigorous explanations and supplementary analyses that are too intricate for the main text. They are important for understanding the research methods but are not essential for the study's main conclusions.

Moreover, the appendix's complex statistical models, like the cubic distributed lag model, are used to refine the analysis of pollution's effects on crime over time, addressing multicollinearity issues and ensuring a smooth temporal relationship. While these models offer in-depth insights and validate the study's approach, reproducing them in the main body would unnecessarily complicate the presentation of the key findings, which are the study's primary focus. The replication of these models is facilitated through the online data repository, enabling interested readers to explore these nuances further.

Furthermore, we believed it was unnecessary to replicate tables 5 and 6, despite having the NIBRS data detailing information on crimes both at home and outside, for several reasons.

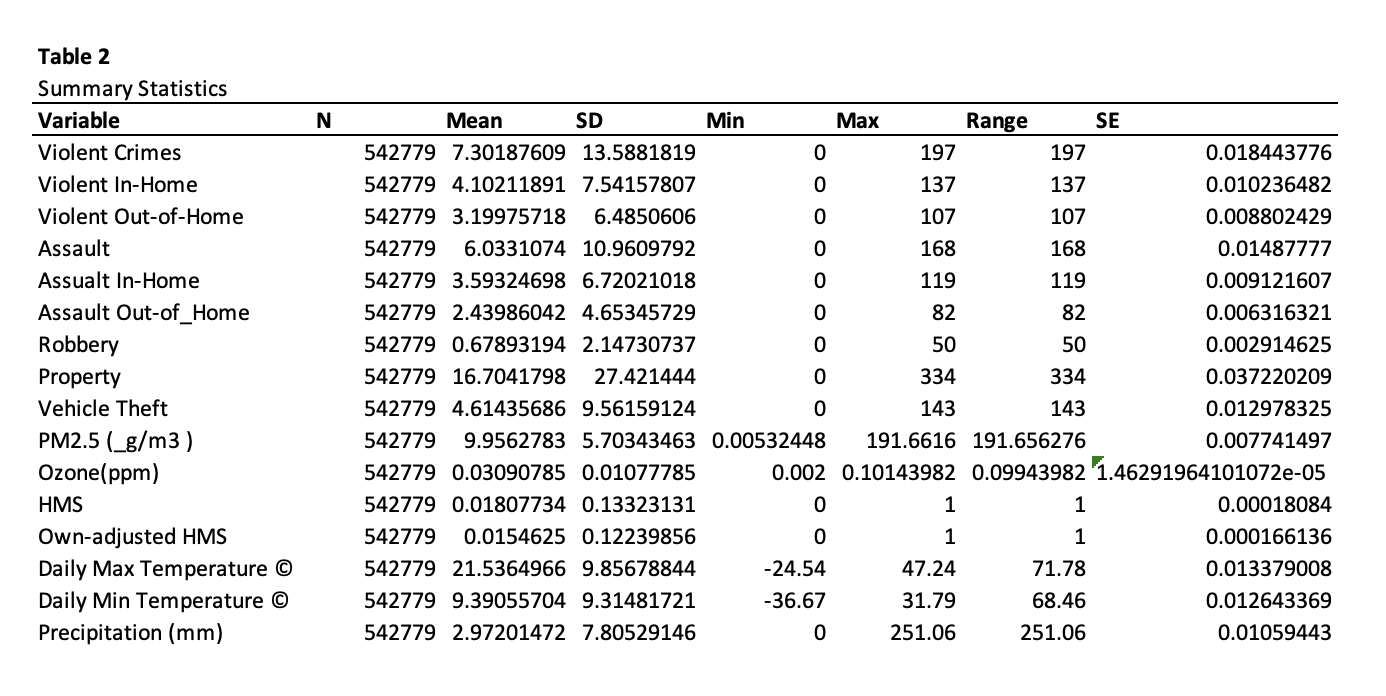
Firstly, there is a potential redundancy in reproducing results. The tables we replicated already indicate that the effects of changes in PM2.5 and ozone are present both indoors and outdoors. Replicating these findings might be redundant, particularly if it does not contribute to a deeper understanding of causal relationships.

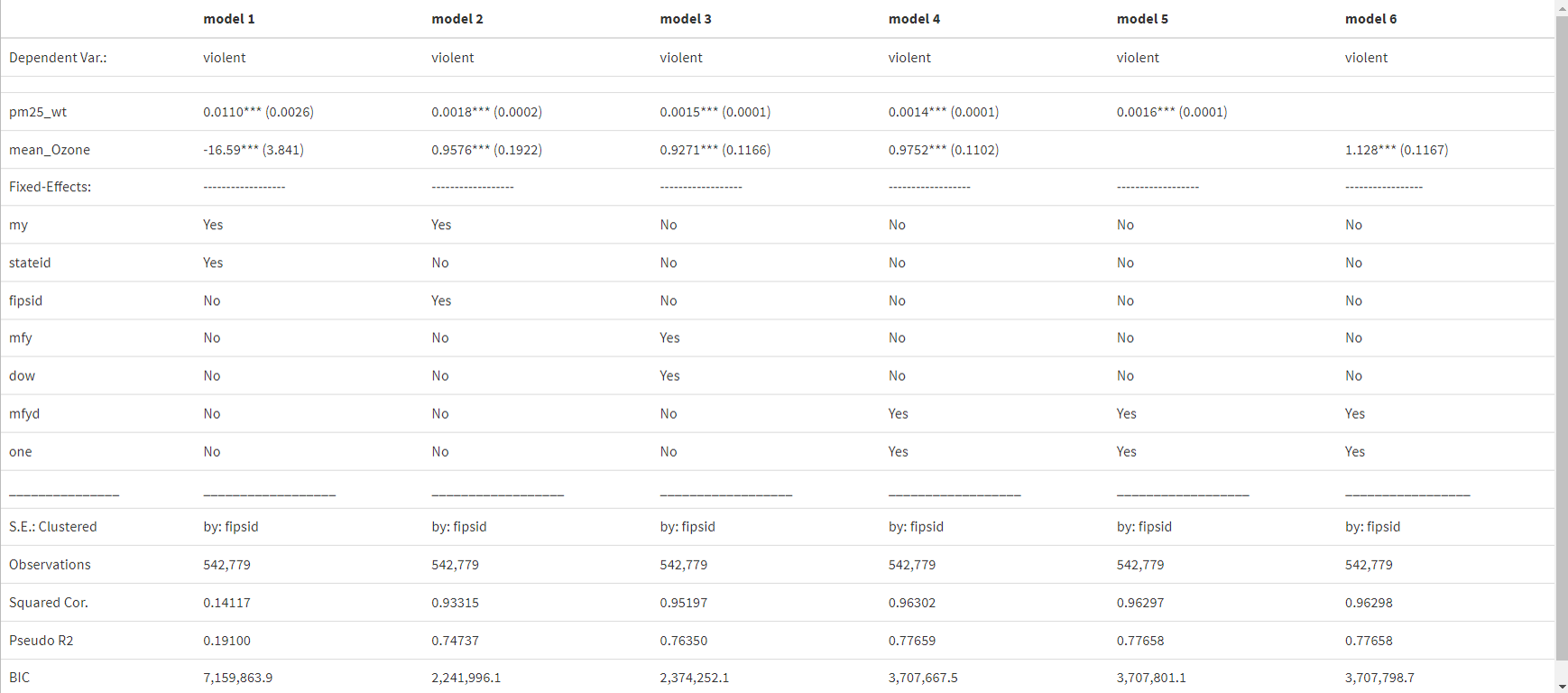
Secondly, the study's primary objective needs to be considered. The main goal was to explore the contemporaneous effect of PM2.5 and ozone changes on crime rates, thus replicating results with lags (table 6) does not align with the initial objective, which we fulfilled well with the tables we reproduced in this report.

Lastly, in table 6, the initial lag of PM2.5 shows statistical significance; however, the coefficient is 40% smaller than the primary coefficient, possibly indicating the influence of serial correlation in pollution. Furthermore, none of the ozone lags exhibit statistical significance when violent crime is the outcome. Thus, the results of table 6 do not add significantly to the main question at hand, hence our decision to not replicate the findings.

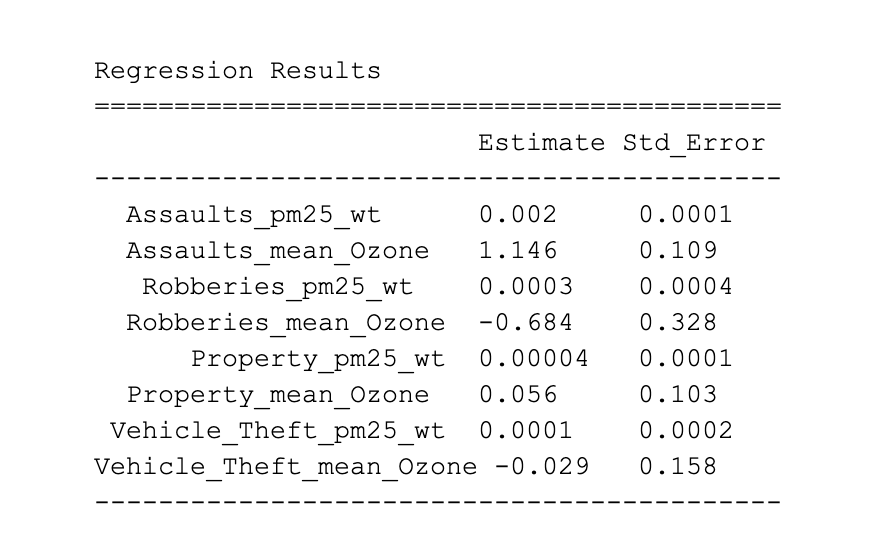
Appendix:

*Figure 1 - Table Three Replication:*

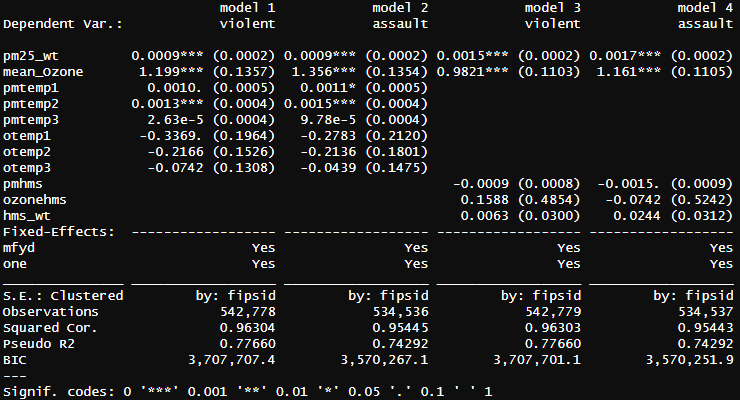


*Figure 2 - Table Three Replication:*  


*Figure 3 - Table Four Replication:*

**

*Figure 4 - Table Seven Replication:*



References

Bishop, K.C., Murphy, A.D., 2011. Estimating the willingness to pay to avoid violent crime: a dynamic approach. 101, 625–629.

Lu, J.G., Lee, J.J., Gino, F., Galinsky, A.D., 2018. Polluted morality: air pollution predicts criminal activity and unethical behavior. Psychol. Sci. 29, 340–355. Mapou, A.E., Shendell, D., Ohman-Strickland, P., Madrigano, J., Meng, Q., Whytlaw, J., Miller, J., 2017. Environmental factors and fluctuations in daily crime rates.

J. Environ. Health 80, 8–23.

Bondy, M., Roth, S., Sager, L., 2018. Crime Is in the Air: the Contemporaneous Relationship between Air Pollution and Crime.

Burkhardt, Jesse, et al. “The effect of pollution on crime: Evidence from data on Particulate Matter and ozone.” *Journal of Environmental Economics and Management*, vol. 98, 2019, p. 102267, https://doi.org/10.1016/j.jeem.2019.102267.