Neighborhood Effects on Crime Clearance Rates

Nhan Le

University of Rochester

Abstract

Juvenile criminal activity remains an important public safety concern. We already know that neighborhood effects influence crime incidence rates but not much is known about crime clearance rates. We wish to examine whether law enforcement treated all crimes fairly across race, gender, status, and various neighborhood effects. We found strong patterns between how often crimes are closed and all tested neighborhood effects but one.

Neighborhood Effects on Crime Clearance Rates

Despite recent downward trends in incidents of both property crimes and violent crimes (Figure 1), presence of juvenile criminal activity remains an important public safety concern. Due to the limitless amount of possible factors that could contribute to the proliferation of crimes, we must attempt to understand structural foundations of crimes before we are able to formulate possible solutions.

MacDonald et al. (2009) demonstrated how neighborhood effects affect crime incidence rates. From education, wage, and unemployment to age and gender, the community's socioeconomic conditions are strongly associated with and frequently predictive of the community's crime incidence rates. MacDonald et al.'s study heavily implied one-way causation as opposed to mere correlation because of experimental manipulations in which at-risk neighborhoods received socioeconomic boosts via business improvement districts.

While the influence of neighborhood effects on crime incidence may be common knowledge, knowledge on crime clearance is relatively obscure. Studying how crimes are resolved is just as important as studying why crimes are committed. Does law enforcement provide equal justice to each crime incidence?

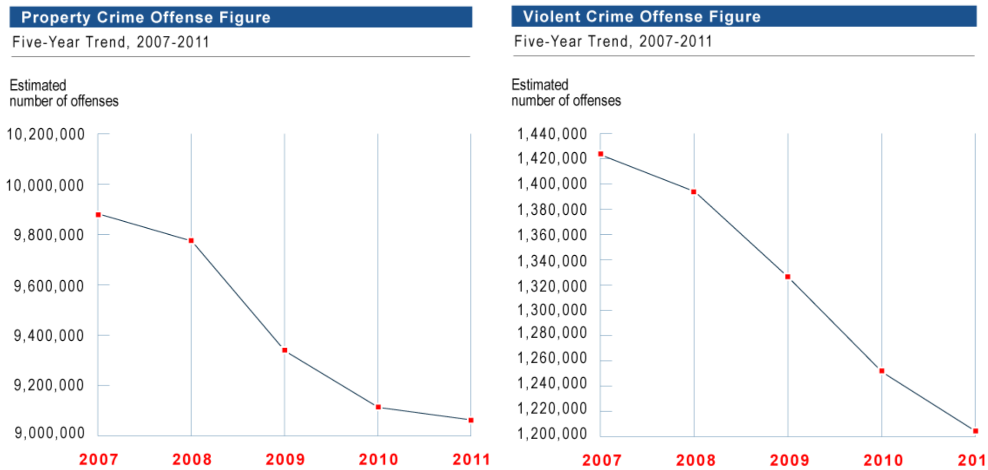


Figure 1. Recent trends in incidence rates of property crimes (FBI, 2012a) and violent crimes (FBI, 2012b). The incidence rates of both property crimes and violent crimes consistently declined between 2007 and 2011.

Insights into law enforcement practices can be gained by examining the crime clearance rate, which is defined as the percentage of reported crimes that result in lawful arrests. Lee (2005), using data from Los Angeles County, exposed systemic biases in crime clearance rates, suggesting that factors such as gender, race, ethnicity, and various neighborhood effects affect how law enforcement agencies allocate resources to solve crimes. The research demonstrated that certain victims “receive more law than others” based on systemic biases. However, the research failed to account for whether neighborhood effects were social versus economic and positive versus negative, which this paper will attempt to elucidate.

For this paper, using data from Chicago, we will focus on crime prosecution rather than crime commitment, as the former is less understood than the latter. Our main regression will attempt to explain the crime clearance rate in response to crime incidence rate, unemployment rate, black composition of population, high school degree acquirement rate, bachelor’s degree acquirement rate, percent of population belonging to various age groups (under 15, 15 to 24, 25 to 34, 35 to 54, and over 54), and male composition of population. Because Jacob and Lefgren (2003) suggested that being in school affects how crimes are incapacitated and coordinated, school may additionally affect crime clearance rates in unknown ways, and thus we will account for the effects of different types of school days before and during our main regression.

age of arrest	% distribution
under 10	0.1
10-12	0.8
13-14	2.7
15	2.5
16	3.2
17	3.7
18	4.6
19	4.9
20	4.8
21	4.4
22	4
23	3.7
24	3.4
25-29	14.6
30-34	11
35-39	8
40-44	7.4
45-49	6.7
50-54	4.9
55-59	2.5
60-64	1.1
over 64	0.7

age of arrest	% distribution
under 15	0.1
under 15	0.8
under 15	2.7
15-19 (15-24)	18.9
20-24 (15-24)	20.3
25-29 (25-34)	14.6
29-34 (25-34)	11
35-39 (35-54)	8
40-44 (35-54)	7.4
45-49 (35-54)	6.7
50-54 (35-54)	4.9
over 54	2.5
over 54	1.1
over 54	0.7

Figure 2. Percent distributions of arrests by age (FBI, 2012c). The left panel displays the percent distributions of arrests grouped by FBI's reported age ranges. The right panel displays the percent distributions of arrests regrouped by Chicago's age ranges. Because FBI's age ranges are inconsistently divided, we must have subdivisions (not in parentheses) within divisions (in parentheses) so that we can compare equal amounts of years between subdivisions.

Our results show that incidence rates for property crimes and violent crimes are not affected by school; that crime clearance rates tend to be positively correlated with bad social neighborhood effects (such as blackness, low education, at-risk age, and maleness) possibly due to subconscious discrimination and negatively correlated with good social neighborhood effects (such as high education and being either younger or older than at-risk age) and bad economic neighborhood effects (such as unemployment rate and crime incidence rates) possibly due to low law enforcement resource allocation; and that weekend crimes have higher crime clearance rates than non-weekend crimes despite having statistically insignificant differences in crime incidence rates. The data support the conclusion that law enforcement is biased against certain peoples and communities.

Methods

In this paper, we seek to confirm Jacob and Lefgren's finding that crime incidence rate varies as a function of types of school days and to explore crime clearance rates as a function of economic neighborhood effects (unemployment rates and crime incidence rates), social neighborhood effects (race, education, age, etc.), and types of school days. Based on prior information from MacDonald et al., Figure 2, and Figure 6's data visualization and Tukey's test of Figure 2, we can designate each neighborhood effect as either "bad" or "good" (Figure 3). Listing the neighborhood effects that we already know will help us develop our model.

neighborhood effect	effect type	effect influence
violent crime incidence rate	economic	bad
property crime incidence rate	economic	bad
% unemployment	economic	bad
% black	social	bad
% high school degree	social	bad
% bachelor's degree	social	good
at-risk age (15-24, 25-34) (Figure 6)	social	bad
not-at-risk age (under 15, 35-54, over 54) (Figure 6)	social	good
% male	social	bad
population density	social	bad

Figure 3. Known types of neighborhood effects. The figure shows the impacts of the neighborhood effects we already know from MacDonald et al., Figure 2, and Figure 6 (in which we identified which age ranges are at risk for high crime incidence rates).

Now that we have basic ideas of what we want to regress, we must consider our model. The common OLS regression model will not suffice in this instance since OLS assumes that dependent variables are continuous, normally distributed, and linearly related to independent variables. Because crime rates are positively skewed and low (Figure 7), oftentimes containing nontrivial zeroes, OLS assumptions are violated. We will rely on the negative binomial regression model like Jacob and Lefgren. The negative binomial regression model has been the standard for the regression of crime-related research since Osgood (2000) first introduced the model.

Before the main regression, we must first regress for property crime incidence rate and violent crime incidence rate, both of which are absent from our dataset. We can derive the formulas for the incidence rates of property crimes and violent crimes as functions of the types of school days (Figure 4), both of which are adapted from Jacob and Lefgren.

$$\begin{aligned} &property\ crime\ incidence\ rate \\ &\quad = B_0 + B_1(school\ day\ with\ no\ attendance) + B_2(weekend) \\ &\quad + B_3(summer\ break) + B_4(holiday) + \varepsilon \end{aligned}$$

$$\begin{aligned} &violent\ crime\ incidence\ rate \\ &\quad = B_0 + B_1(school\ day\ with\ no\ attendance) + B_2(weekend) \\ &\quad + B_3(summer\ break) + B_4(holiday) + \varepsilon \end{aligned}$$

Figure 4. Equations for incidence rates of property crimes and violent crimes. The panels show the relationships between the independent variables (types of school days) and the dependent variables (incidence rates for property crimes and violent crimes).

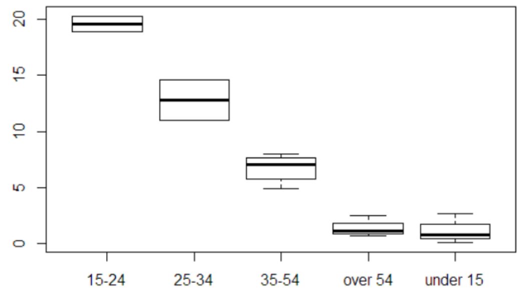
For the main regression, we will use Figure 5's formula. The formula will predict crime clearance rate as a function of known neighborhood effects and the types of school days. We will not be regressing crime clearance rate against population since knowing the population without knowing the population density is not useful since each police beat might be of unequal size.

$$crime\ clearance\ rate = \frac{\#\ of\ arrests}{\#\ of\ crimes}$$

$$= B_0 + B_1(violent\ crime\ incidence\ rate)$$
$$+ B_2(property\ crime\ incidence\ rate) + B_3(\%\ unemployment)$$
$$+ B_4(\%\ black) + B_5(\%\ high\ school\ degree) + B_6(\%\ bachelor's\ degree)$$
$$+ B_7(\%\ under\ 15) + B_8(\%\ 15\ to\ 24) + B_9(\%\ 25\ to\ 34) + B_{10}(\%\ 35\ to\ 54)$$
$$+ B_{11}(\%\ male) + B_{12}(school\ day\ with\ no\ attendance) + B_{13}(weekend)$$
$$+ B_{14}(summer\ break) + B_{15}(holiday) + \varepsilon$$

Figure 5. Equation for crime clearance rate. The figure shows the relationship between the independent variables (violent crime incidence rate; property crime incidence rate; unemployment rate; black composition of population; high school degree acquirement rate; bachelor's degree acquirement rate; percents of population under 15, 15 to 24, 25 to 34, 35 to 54, and, over 54; male composition of population; and types of school days) and the dependent variable (crime clearance rate).

Descriptive Statistics




```
  Tukey multiple comparisons of means
    99.97% family-wise confidence level

Fit: aov(formula = x1$X..distribution ~ x1$age.of.arrest)

$`x1$age.of.arrest`
                        diff        lwr       upr     p adj
25-34-15-24       -6.8000000 -17.327520  3.727520 0.0067929
35-54-15-24      -12.8500000 -21.967099 -3.732901 0.0000189
over 54-15-24    -18.1666667 -27.776933 -8.556400 0.0000016
under 15-15-24   -18.4000000 -28.010267 -8.789733 0.0000014
35-54-25-34       -6.0500000 -15.167099  3.067099 0.0057034
over 54-25-34    -11.3666667 -20.976933 -1.756400 0.0000796
under 15-25-34   -11.6000000 -21.210267 -1.989733 0.0000676
over 54-35-54     -5.3166667 -13.357193  2.723859 0.0058372
under 15-35-54    -5.5500000 -13.590526  2.490526 0.0043972
under 15-over 54  -0.2333333  -8.829017  8.362350 0.9995641
```

Figure 6. Boxplot and Tukey's test for Figure 2's right panel. The top panel shows the percent distribution for arrests by age groups. The bottom panel shows Tukey's test for multiple comparisons of means by age groups. Both panels show that 15-24 and 25-34 have statistically significantly high crime at the 99.97% confidence interval. Therefore, we can designate age ranges 15 to 24 and 25 to 34 as at risk for high crime incidence rates.

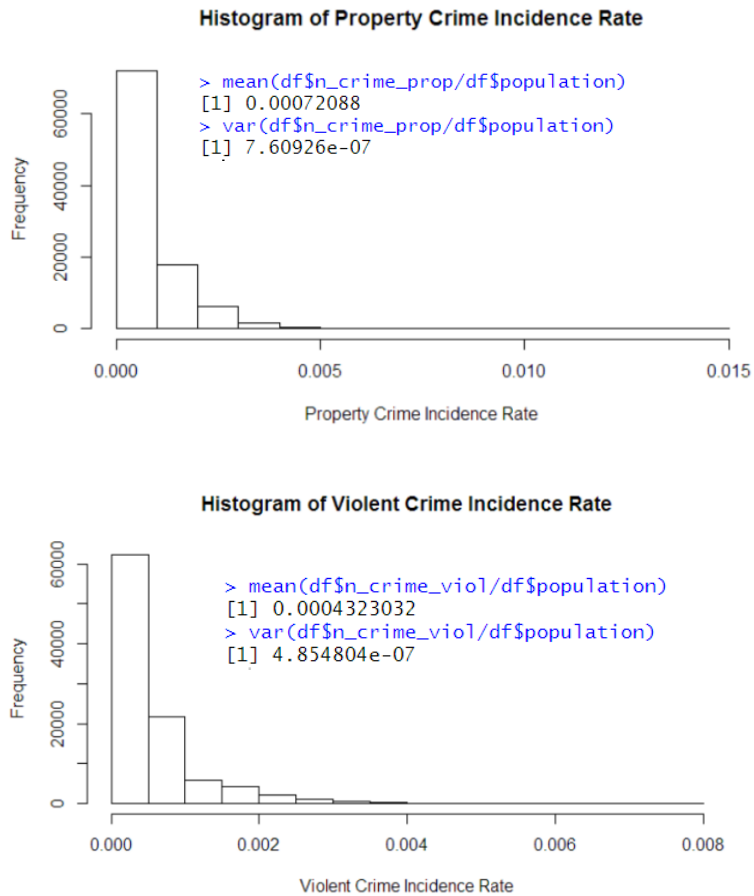


Figure 7. Histograms, means, and variances of incidence rates of property crimes and violent crimes. The panels display relevant statistics and charts for incidence rates for property crimes and violent crimes. For both incidence rates, the histograms display lack of normality, being positively skewed and containing nontrivial zeroes. The histograms indicate that we should regress using the negative binomial regression model.

Results

The regressions for incidence rates of property crimes and violent crimes (Figure 8) yield surprising results. For both regressions, the incidence rate is predicted by the intercept (indicating “school day with attendance”) and the error term.

$property\ crime\ incidence\ rate = -7.28559 + \varepsilon$
$violent\ crime\ incidence\ rate = -7.7240 + \varepsilon$

Figure 8. Regressions for incidence rates of property crimes and violent crimes. The summary statistics for the top panel are in Figure 9. The summary statistics for the bottom panel are in Figure 10.

Because having negative values of crime incidence rates makes no sense, we can assume that the true predictors for crime incidence rates lie somewhere within the error terms. The summary statistics for the regressions (Figure 9 and Figure 10) support this assumption. Both Figure 9 and Figure 10 show that only the intercept is statistically significant even at the 90% confidence interval.

Figure 8's finding is in contradiction with Jacob and Lefgren's finding that being in school increases property crime incidence rates and decreases violent crime incidence rates. Perhaps Chicago, which was not in Jacob and Lefgren's sample, has stronger predictors for crime incidence rates than types of school days.

```
Deviance Residuals:
     Min        1Q    Median        3Q       Max
-0.10609  -0.03858  -0.03607   0.12027   0.43377

Coefficients:
                                        Estimate Std. Error z value Pr(>|z|)
(Intercept)                             -7.28559    0.38217 -19.063   <2e-16 ***
df$day_typeSchool day with no attendence  0.08228    0.41716   0.197    0.844
df$day_typeWeekend                      -0.05218    0.45751  -0.114    0.909
df$day_typeSummer break                  0.13910    0.49000   0.284    0.777
df$day_typeHoliday                       0.07438    0.80482   0.092    0.926
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for Negative Binomial(9899.158) family taken to be 1)

    Null deviance: 1028.0  on 98549  degrees of freedom
Residual deviance: 1027.7  on 98545  degrees of freedom
AIC: 1100

Number of Fisher Scoring iterations: 1


              Theta:  9899
          Std. Err.:  63639
Warning while fitting theta: iteration limit reached

 2 x log-likelihood:  -1087.971
```

Figure 9. Summary statistics for types of school days vs. property crime incidence rate. The figure shows the predictors (types of school days) along with standard errors and z-values for incidence rates of property crimes.

```
Deviance Residuals:
     Min        1Q    Median        3Q       Max
-0.08361  -0.02902  -0.02806   0.10752   0.32577

Coefficients:
                                        Estimate Std. Error z value Pr(>|z|)
(Intercept)                             -7.77240    0.48750 -15.943   <2e-16 ***
df$day_typeSchool day with no attendence -0.06792    0.53900  -0.126    0.900
df$day_typeWeekend                       0.13151    0.56861   0.231    0.817
df$day_typeSummer break                  0.15271    0.62339   0.245    0.806
df$day_typeHoliday                       0.04026    1.04033   0.039    0.969
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for Negative Binomial(5339.378) family taken to be 1)

    Null deviance: 660.05  on 98549  degrees of freedom
Residual deviance: 659.64  on 98545  degrees of freedom
AIC: 707.78

Number of Fisher Scoring iterations: 1


              Theta:  5339
          Std. Err.:  32554
Warning while fitting theta: iteration limit reached

 2 x log-likelihood:  -695.784
```

Figure 10. Summary statistics for types of school days vs. violent crime incidence rate. The figure shows the predictors (types of school days) along with standard errors and z-values for incidence rates of violent crimes.

$$property\ crime\ incidence\ rate = \frac{\#\ of\ property\ crimes}{\#\ of\ residents}$$

$$violent\ crime\ incidence\ rate = \frac{\#\ of\ violent\ crimes}{\#\ of\ residents}$$

Figure 11. Alternate equations for incidence rates of property crimes and violent crimes.

Since, in Chicago, types of school days are poor predictors for crime incidence rates, we can substitute “violent crime incidence rate” and “property crime incidence rate” in our main regression via Figure 11’s simplified alternate equations instead of Figure 4’s equations.

The regression for crime clearance rate (Figure 12) yields results similar to Figure 3’s predictions. Expectedly, all bad economic neighborhood effects and good social neighborhood effects decreased crime clearance rate while all bad social neighborhood effects (except statistically insignificant “% 15-24”) increased clearance rate. Unexpectedly, “weekend” decreased clearance rate.

$$\begin{aligned}
&crime\ clearance\ rate \\
&\quad = -0.88811 - 322.03223(violent\ crime\ incidence\ rate) \\
&\quad - 71.17529(property\ crime\ incidence\ rate) - 0.62911(\%\ unemployment) \\
&\quad + 0.45035(\%\ black) + 1.17687(\%\ high\ school\ degree) \\
&\quad - 0.91756(\%\ bachelor's\ degree) - 0.95590(\%\ under\ 15) \\
&\quad + 0.74508(\%\ 25\ to\ 34) - 2.74930(\%\ 35\ to\ 54) + 0.78791(\%\ male) \\
&\quad - 0.07225(weekend) + \varepsilon
\end{aligned}$$

Figure 12. Regression crime clearance rate. The summary statistics are in Figure 13.

```
Deviance Residuals:
    Min       1Q   Median       3Q      Max
-1.1042  -0.7387  -0.5656   0.8773   2.0295

Coefficients:
                                          Estimate Std. Error z value Pr(>|z|)
(Intercept)                               -0.88811    0.13937  -6.372 1.86e-10 ***
prop_crime_percent                      -322.03223    9.06665 -35.518  < 2e-16 ***
viol_crime_percent                       -71.17529    9.66268  -7.366 1.76e-13 ***
df$p_unemployed                           -0.62911    0.14326  -4.391 1.13e-05 ***
df$p_black                                 0.45035    0.03079  14.629  < 2e-16 ***
df$p_hsdegree                              1.17687    0.16841   6.988 2.78e-12 ***
df$p_bsdegree                             -0.91756    0.18600  -4.933 8.09e-07 ***
df$p_popunder15                           -0.95590    0.18662  -5.122 3.02e-07 ***
df$p_pop15_24                              0.30935    0.19281   1.604  0.10862
df$p_pop25_34                              0.74508    0.18950   3.932 8.43e-05 ***
df$p_pop35_54                             -2.74930    0.28008  -9.816  < 2e-16 ***
df$p_male                                  0.78791    0.24792   3.178  0.00148 **
df$day_typeSchool day with no attendence  -0.02510    0.02218  -1.132  0.25773
df$day_typeWeekend                        -0.07225    0.02415  -2.992  0.00277 **
df$day_typeSummer break                   -0.01868    0.02664  -0.701  0.48324
df$day_typeHoliday                        -0.01058    0.04420  -0.239  0.81076
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for Negative Binomial(12457.75) family taken to be 1)

    Null deviance: 60530  on 85517  degrees of freedom
Residual deviance: 57733  on 85502  degrees of freedom
  (13032 observations deleted due to missingness)
AIC: 96359

Number of Fisher Scoring iterations: 1


              Theta:  12458
          Std. Err.:  6176
Warning while fitting theta: iteration limit reached

 2 x log-likelihood:  -96325.29
```

Figure 13. Summary statistics for neighborhood effects vs. crime clearance rate

neighborhood effect	effect type	effect influence	estimator sign
violent crime incidence rate	economic	bad	negative
property crime incidence rate	economic	bad	negative
% unemployment	economic	bad	negative
% black	social	bad	positive
% high school degree	social	bad	positive
% bachelor's degree	social	good	negative
at-risk age (15-24)	social	bad	insignificant
at-risk age (25-34)	social	bad	positive
not-at-risk age (under 15)	social	good	negative
not-at-risk age (35-54)	social	good	negative
% male	social	bad	positive
weekend	day		negative

Figure 14. Figure 3 revisited.

To find out why “% 15-24” was statistically insignificant, we regressed “% 15-24” against the other independent variables from our main regression (Figure 15). “% 15-24” is strongly inversely correlated to high school degree acquirement rate, which makes sense as increased high school graduation rate would mean decreased percent of population that are high school-aged, and is strongly correlated to black population rate, which makes sense since Chicago’s black population skews young (Black Demographics, 2014). “% 15-24” appears to be statistically insignificant due to redundancy.

```
Deviance Residuals:
    Min       1Q   Median       3Q      Max
-0.8558  -0.7334  -0.5836   0.7671   0.8634

Coefficients:
                                           Estimate Std. Error z value Pr(>|z|)
(Intercept)                              -7.867e-01  1.475e-01  -5.334 9.61e-08 ***
df$day_typeSchool day with no attendence  7.402e-05  2.822e-02   0.003  0.99791
df$day_typeWeekend                       -3.782e-04  3.050e-02  -0.012  0.99011
df$day_typeSummer break                  -8.744e-04  3.388e-02  -0.026  0.97941
df$day_typeHoliday                       -2.609e-04  5.564e-02  -0.005  0.99626
prop_crime_percent                        2.745e+00  9.568e+00   0.287  0.77420
viol_crime_percent                        8.150e+00  1.234e+01   0.660  0.50894
df$p_unemployed                           2.850e-01  1.867e-01   1.526  0.12704
df$p_black                                2.659e-01  4.067e-02   6.539 6.19e-11 ***
df$p_hsdegree                            -6.444e-01  2.093e-01  -3.079  0.00208 **
df$p_bsdegree                             9.693e-02  2.156e-01   0.450  0.65305
df$p_popunder15                          -2.383e+00  2.349e-01 -10.144  < 2e-16 ***
df$p_pop25_34                            -3.684e+00  2.331e-01 -15.802  < 2e-16 ***
df$p_pop35_54                            -4.930e+00  2.557e-01 -19.281  < 2e-16 ***
df$p_pop55more                           -3.264e+00  1.897e-01 -17.210  < 2e-16 ***
df$p_male                                 4.020e+00  2.377e-01  16.912  < 2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for Negative Binomial(759039.8) family taken to be 1)

    Null deviance: 56434  on 98549  degrees of freedom
Residual deviance: 55306  on 98534  degrees of freedom
AIC: 71820

Number of Fisher Scoring iterations: 1
```

Figure 15. Summary statistics for “% 15-24” vs. other neighborhood effects.

Conclusions

Based on our analysis in Figure 14, we can conclude that law enforcement tends to be biased against certain neighborhood effects in regards to crime clearance. Bad economic neighborhood effects likely have lower crime clearance rates because law enforcement is unwilling or unable to allocate adequae resources to poor neighborhoods, and bad social neighborhood effects likely have higher crime clearance rates due to subconscious discrimination against certain communities. However, because our dataset failed to distinguish race, gender, and status between victims and perpetrators, nothing is conclusive. (For example, bad social neighborhood effects may have different influences on victims and perpetrators.) We can only point out correlations and make assumptions.

References

Black Demographics. (2014). Chicago Metro Black Demographic Profile.

FBI. (2012). Property Crime. Retrieved from https://www.fbi.gov/about-us/cjis/ucr/crime-in-the-u.s/2011/crime-in-the-u.s.-2011/property-crime/property-crime

FBI. (2012). Violent Crime. Retrieved from https://www.fbi.gov/about-us/cjis/ucr/crime-in-the-u.s/2011/crime-in-the-u.s.-2011/violent-crime/violent-crime

FBI. (2012). Table 46. Retrieved from https://www.fbi.gov/about-us/cjis/ucr/crime-in-the-u.s/2011/crime-in-the-u.s.-2011/tables/table_46_arrests_cities_by_age_2011.xls

Jacob, B. A., Lefgren, L. (2003). Are Idle Hands the Devil's Workshop? Incapacitation, Concentration, and Juvenile Crime. *The American Economic Review*.

Lee, C. (2005). The value of life in death: Multiple regression and event history analyses of homicide clearance in Los Angeles County. *Journal of Criminal Justice*.

MacDonald, J., Bluthenthal, R. N., Golinelli, D., Kofner, A., Stokes, R. J., Sehgal, A., . . . Beletsky, L. (2009). Neighborhood Effects on Crime and Youth Violence. *RAND Infrastructure, Safety, and Environment*.

Osgood, D. W. (2000). Poisson-Based Regression Analysis of Aggregate Crime Rates. *Journal of Quantitative Criminology*.