

Safer at Home? Domestic Violence in Chicago During the Pandemic

Samuel, Natalie, & Stacy

INTRODUCTION AND DATA

(BLAH BLAH from doc)

METHODOLOGY

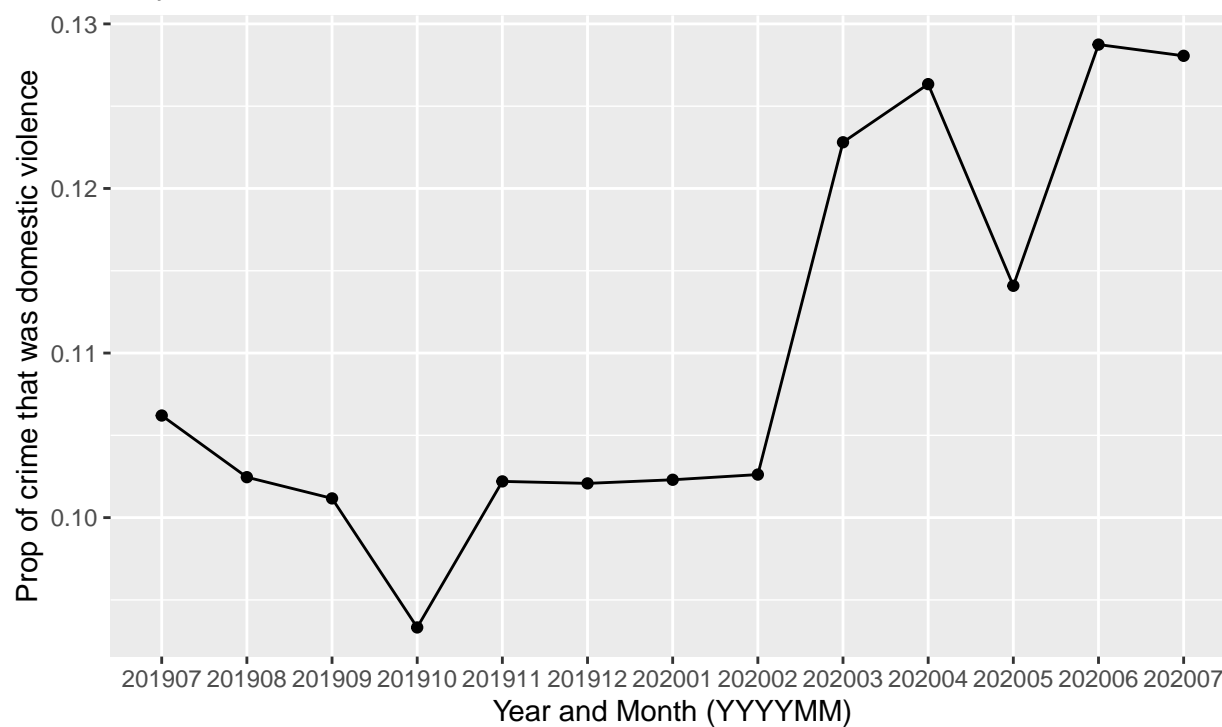
(BLAH BLAH from doc)

```
## Rows: 231,002
## Columns: 23
## $ CASE. <chr> "JD163753", "JD212847", "JC497784", "JC459410...
## $ DATE..OF.OCCURRENCE <chr> "02/24/2020 08:15:00 PM", "04/10/2020 10:56:0...
## $ BLOCK <chr> "031XX W LEXINGTON ST", "005XX W 103RD ST", "...
## $ IUCR <chr> "1153", "0560", "0860", "0560", "0810", "0820...
## $ PRIMARY.DESCRPTION <chr> "DECEPTIVE PRACTICE", "ASSAULT", "THEFT", "AS...
## $ SECONDARY.DESCRPTION <chr> "FINANCIAL IDENTITY THEFT OVER $ 300", "SIMPL...
## $ LOCATION.DESCRPTION <chr> "", "RESIDENCE", "DEPARTMENT STORE", "SIDEWAL...
## $ ARREST <chr> "N", "N", "N", "N", "N", "N", "N", "N", "N", "N", ...
## $ DOMESTIC <chr> "N", "N", "N", "N", "N", "N", "N", "N", "N", "N", ...
## $ BEAT <int> 1134, 2232, 1924, 122, 123, 2433, 312, 914, 3...
## $ WARD <int> 24, 9, 44, 4, 25, 48, 20, 11, 5, 26, 27, 37, ...
## $ FBI.CD <chr> "11", "08A", "06", "08A", "06", "06", "08A", ...
## $ X.COORDINATE <int> NA, 1174583, NA, NA, NA, NA, 1180030, 1171590...
## $ Y.COORDINATE <int> NA, 1836593, NA, NA, NA, NA, 1862317, 1887793...
## $ LATITUDE <dbl> NA, 41.70700, NA, NA, NA, NA, 41.77747, 41.84...
## $ LONGITUDE <dbl> NA, -87.63629, NA, NA, NA, NA, -87.61556, -87...
## $ LOCATION <chr> "", "(41.707000821, -87.636288063)", "", "", ...
## $ MONTH <int> 2, 4, 11, 10, 5, 12, 5, 5, 4, 5, 4, 5, 5, 5, ...
## $ DAY <int> 24, 10, 3, 4, 24, 5, 7, 3, 28, 7, 25, 7, 7, 5...
## $ YEAR <int> 2020, 2020, 2019, 2019, 2020, 2019, 2020, 202...
## $ DATEINT <int> 202002, 202004, 201911, 201910, 202005, 20191...
## $ isPM <dbl> 1, 1, 0, 0, 1, 1, 0, 0, 1, 1, 1, 0, 0, 0, 1, ...
## $ HOUR <int> 8, 10, 11, 6, 9, 2, 12, 2, 6, 5, 3, 8, 6, 12,...
```

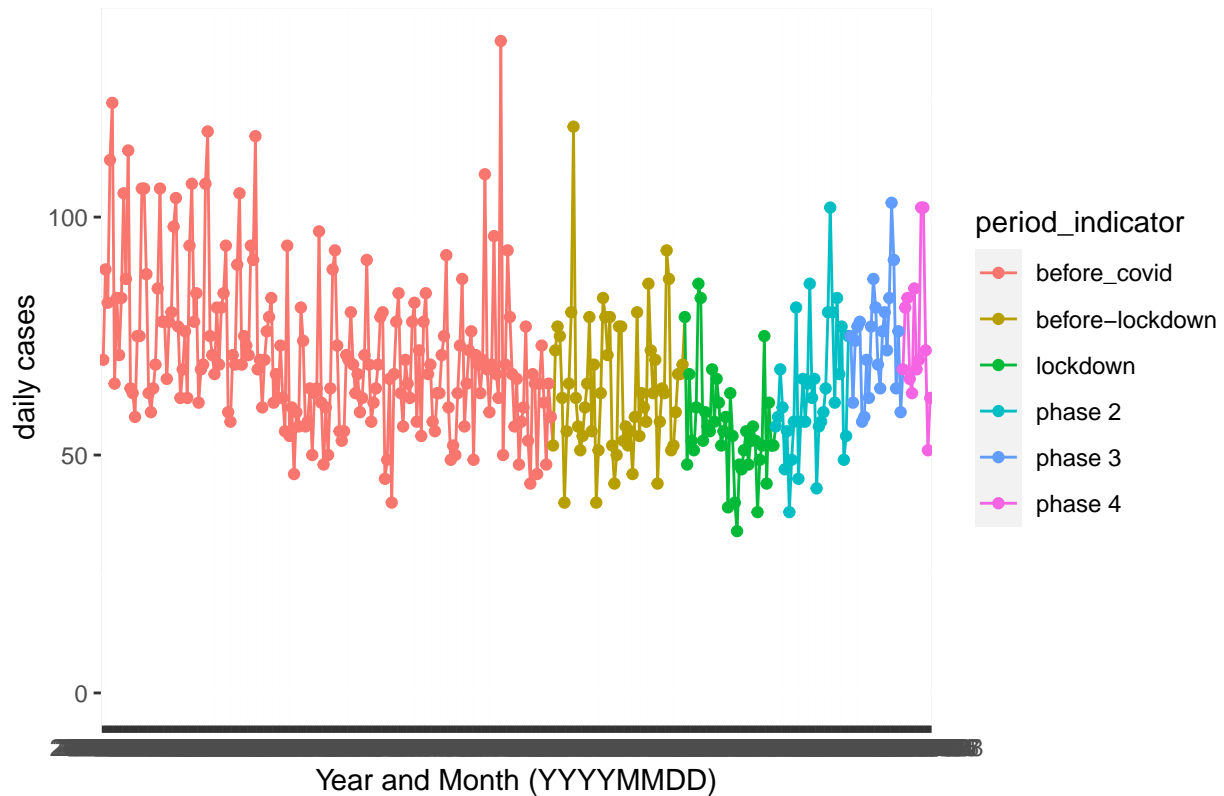
RESULTS

(BLAH BLAH from Doc)

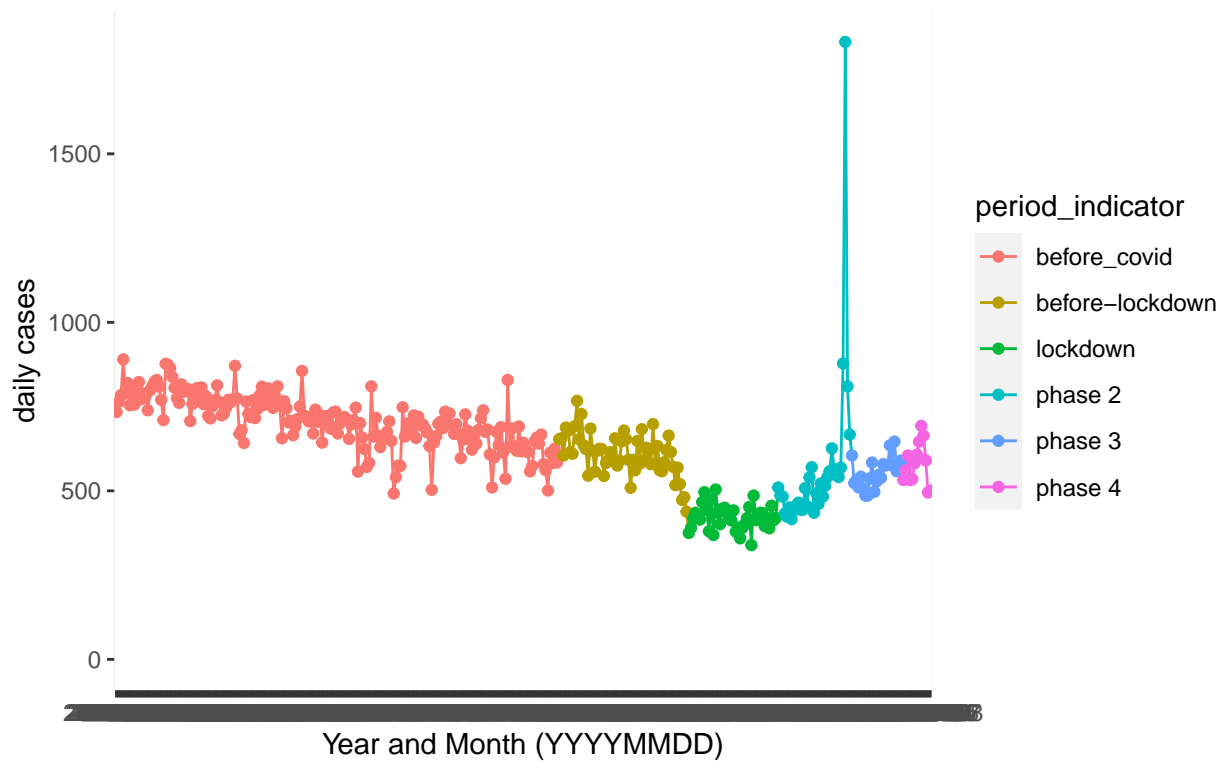
Monthly domestic violence proportion shot up in March 2020, when Chicago first issued stay-at-home order



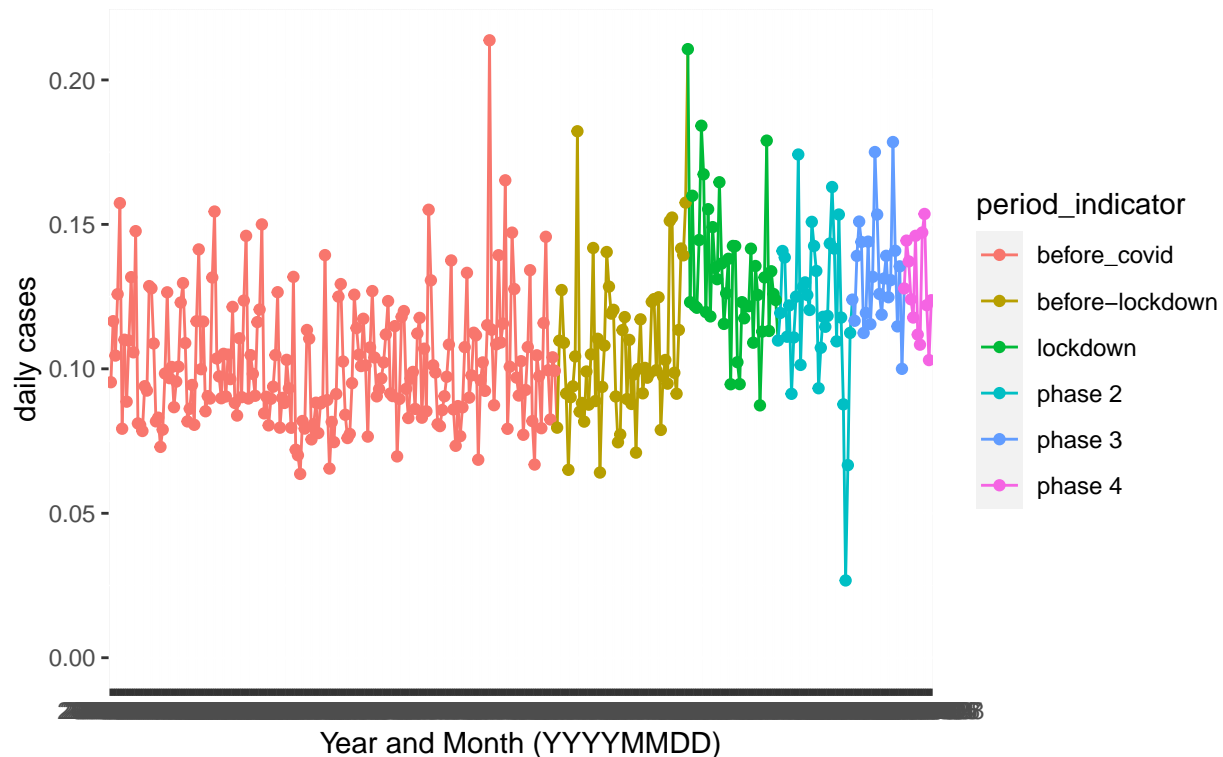
daily domvio cases over time shows domestic crime rates over different phases



daily crime cases over time shows crime rates over the different phases as
note the increased crimes around the time of George Floyd's death



prop crime rates over the different phases aren't constant /n
note the increased crimes around the time of George Floyd's death



Chi-Square Test Since the data (Table 1) satisfies the independent sampling assumption and is large enough (i.e. each cell > 10), we will be performing a chi-square test at the $\alpha = 0.05$ significance level. We test the two hypotheses below:

H_0 : Domestic violence cases in Chicago is unrelated to the 3 periods of the pandemic. H_1 : Domestic violence cases in Chicago is related to the 3 periods of the pandemic.

*The 6 phases of the pandemic is divided into 3 periods: before-lockdown (before COVID and before lockdown), during lockdown (lockdown and phase 2), and post-lockdown (phases 3 and 4)

```
##
##      before lockdown recovery
## 0 157397      31109      17599
## 1  17919       4311       2667
##
## Pearson's Chi-squared test
##
## data:  table(domvio_mut$isdomviolence, domvio_mut$PHASE)
## X-squared = 247.63, df = 2, p-value < 2.2e-16
```

Under the null hypothesis, our test statistic has a chi-square distribution with 2 degrees of freedom. We performed the test and obtained a chi-square value of 247.63, which corresponds to a p-value of < 0.001 . Thus, at an $\alpha = 0.05$ significance level, we reject the null hypothesis; there is sufficient evidence to suggest that domestic violence cases in Chicago is related to the 3 periods of the pandemic.

Step Down 2 Proportion Z-Tests Since the overall Chi-square test was significant, we stepped down to identify where the differences are. We conducted three 2 proportion z-tests. To account for multiple

comparisons, we will perform the Bonferroni correction and thus assess our results relative to the adjusted $\alpha = 0.05/3$ level.

```
##
## 2-sample test for equality of proportions with continuity correction
##
## data:  c(4311, 2667) out of c(35420, 20266)
## X-squared = 11.411, df = 1, p-value = 0.0007303
## alternative hypothesis: two.sided
## 95 percent confidence interval:
##  -0.01569439 -0.00408326
## sample estimates:
##      prop 1      prop 2
## 0.1217109 0.1315997

##
## 2-sample test for equality of proportions with continuity correction
##
## data:  c(2667, 17919) out of c(2667 + 17599, 17919 + 157397)
## X-squared = 166.3, df = 1, p-value < 2.2e-16
## alternative hypothesis: two.sided
## 95 percent confidence interval:
##  0.024497 0.034283
## sample estimates:
##      prop 1      prop 2
## 0.1315997 0.1022097

##
## 2-sample test for equality of proportions with continuity correction
##
## data:  c(4311, 17919) out of c(35420, 17919 + 157397)
## X-squared = 118.55, df = 1, p-value < 2.2e-16
## alternative hypothesis: two.sided
## 95 percent confidence interval:
##  0.01579582 0.02320653
## sample estimates:
##      prop 1      prop 2
## 0.1217109 0.1022097
```

We found that all three pairwise difference in proportions are significant at the adjusted $\alpha = 0.05/3$ significance level. There is sufficient evidence to suggest that the proportion of crime in Chicago related to domestic violence is different for each period of the pandemic.

Regression Analysis (BLAH BLAH copy from doc)

```
## # A tibble: 7 x 5
##   term                estimate std.error statistic  p.value
##   <chr>              <dbl>    <dbl>    <dbl>    <dbl>
## 1 (Intercept)        -2.07      0.0101   -205.    0.
## 2 PHASEbefore-lockdown 0.0528     0.0196    2.69 7.11e- 3
## 3 PHASElockdown        0.303      0.0244   12.4 1.87e- 35
## 4 PHASEphase 2         0.105      0.0249    4.23 2.37e- 5
## 5 PHASEphase 3         0.301      0.0276   10.9 8.43e- 28
## 6 PHASEphase 4         0.262      0.0355    7.37 1.77e- 13
## 7 as.factor(is9_5)1    -0.314     0.0142   -22.2 1.00e-108
```

Our logistic regression model corresponding to the probability of success that a reported crime for a given time period will be one of domestic violence is as follows:

$$\hat{\beta}_0 + \hat{\beta}_1 * (\text{PHASE} == \text{lockdown}) + \hat{\beta}_2 * (\text{PHASE} == \text{phase 2}) + \hat{\beta}_3 * (\text{PHASE} == \text{phase 3}) + \hat{\beta}_4 * (\text{PHASE} == \text{phase 4}) + \hat{\beta}_5 * (\text{PHASE} == \text{before-lockdown}) + \hat{\beta}_6 * (\text{is9_5})$$

At the $\alpha = 0.05$ significance level, the true β coefficients corresponding to the logit of the probability that a case reported during that time period was related to domestic violence for each of our dummy variables (all relative to the baseline/reference category of before the COVID-19 pandemic) were statistically significant, as was the beta coefficient corresponding to the logit of the probability of success (see above) during the 9-5 work day, as compared to outside this timeframe. We were specifically interested in β_1 , which corresponds to the logit of the probability of success (that a reported case was related to domestic violence) during lockdown, as compared to before the pandemic began, and whether there was a relationship between the probability of a reported crime being related to domestic violence during lockdown, when compared to before the pandemic.

H_0 : $\beta_1 = 0$ (There is no relationship between our predictor and the probability of success that a reported case is related to domestic violence, while holding time of day (either during the 9-5 workday or outside this timeframe) constant. The true population parameter β_1 is equal to 0).

H_1 : (There is a relationship between our predictor and the probability of success that a reported case is related to domestic violence, while holding time of day (either during the 9-5 workday or outside this timeframe) constant. The true population parameter β_1 is not equal to 0).

Under the null hypothesis, our test statistic follows a standard normal distribution. The value of our test statistic is equal to approximately 12.426, which corresponds to a p-value of less than 0.001. Thus, at the $\alpha = 0.05$, we reject our null hypothesis; we have sufficient evidence to suggest that the true value of β_1 is not equal to 0, such that this predictor tells us something about the probability of success of our outcome, while holding time of day (either during the 9-5 workday or outside this timeframe) constant.

We see from the model output above that the estimated $\hat{\beta}_1$ coefficient, corresponding to the difference of being in the “lockdown” phase of time in quarantine - relative to the time period before COVID-19 - on the logit of the probability of success that a reported case will be related to domestic violence, is approximately 0.303. This corresponds to an odds ratio of $\exp(0.303)$, which is approximately 1.354. Therefore, we would expect crimes reported during the lockdown phase of the stay-at-home order in the city of Chicago to have 1.354 times the odds of being related to domestic violence, relative to cases that were reported before the COVID-19 pandemic began, while adjusting for time of day, relative to the 9-5 work period. Moreover, we are 95% confident that the true population parameter, $\exp(\beta_1)$ lies within the interval (3.694, 4.059).

```
## [1] 1.05443
## [1] 1.05443
## [1] 1.353914
## [1] 1.110711
## [1] 1.351209
## [1] 1.298228
## [1] 0.730519
## [1] 1.013896
## [1] 1.096583
## [1] 1.291702
## [1] 1.419123
## [1] 1.057599
## [1] 1.16649
```

```
## [1] 1.279054
## [1] 1.427435
## [1] 1.209783
## [1] 1.393138
## [1] 0.7107465
## [1] 0.7508416
```

We can interpret the odds ratios and 95% CIs in the context of our model and research question as follows:

Adjusting for time of day, we would expect crimes reported prior to the issuing of the first stay at home order in the city of Chicago to have 1.054 times the odds of being related to domestic violence, compared to cases that were reported before the COVID-19 pandemic began, with 95% CI (1.014, 1.097).

Adjusting for time of day, we would expect crimes reported during the “lockdown” phase of the stay at home order in the city of Chicago to have 1.354 times the odds of being related to domestic violence, compared to cases that were reported before the COVID-19 pandemic began, with 95% CI (1.292, 1.419).

Adjusting for time of day, we would expect crimes reported during the phase 2 of reopening (modified stay at home order) in the city of Chicago to have 1.111 times the odds of being related to domestic violence, compared to cases that were reported before the COVID-19 pandemic began, with 95% CI (1.058, 1.166).

Adjusting for time of day, we would expect crimes reported during the phase 3 of reopening (modified stay at home order) in the city of Chicago to have 1.351 times the odds of being related to domestic violence, compared to cases that were reported before the COVID-19 pandemic began, with 95% CI (1.279, 1.427).

Adjusting for time of day, we would expect crimes reported during the phase 4 of reopening (modified stay at home order) in the city of Chicago to have 1.298 times the odds of being related to domestic violence, compared to cases that were reported before the COVID-19 pandemic began, with 95% CI (1.210, 1.393).

Adjusting for lockdown phase, we would expect crimes reported during the 9 to 5 workday in the city of Chicago to have 0.731 times the odds of being related to domestic violence, compared to crimes reported outside this timeframe, with 95% CI (0.711, 0.751).

DISCUSSION

(BLAH BLAH from doc)