Safer at Home? Domestic Violence in Chicago During the Pandemic

Samuel, Natalie, & Stacy

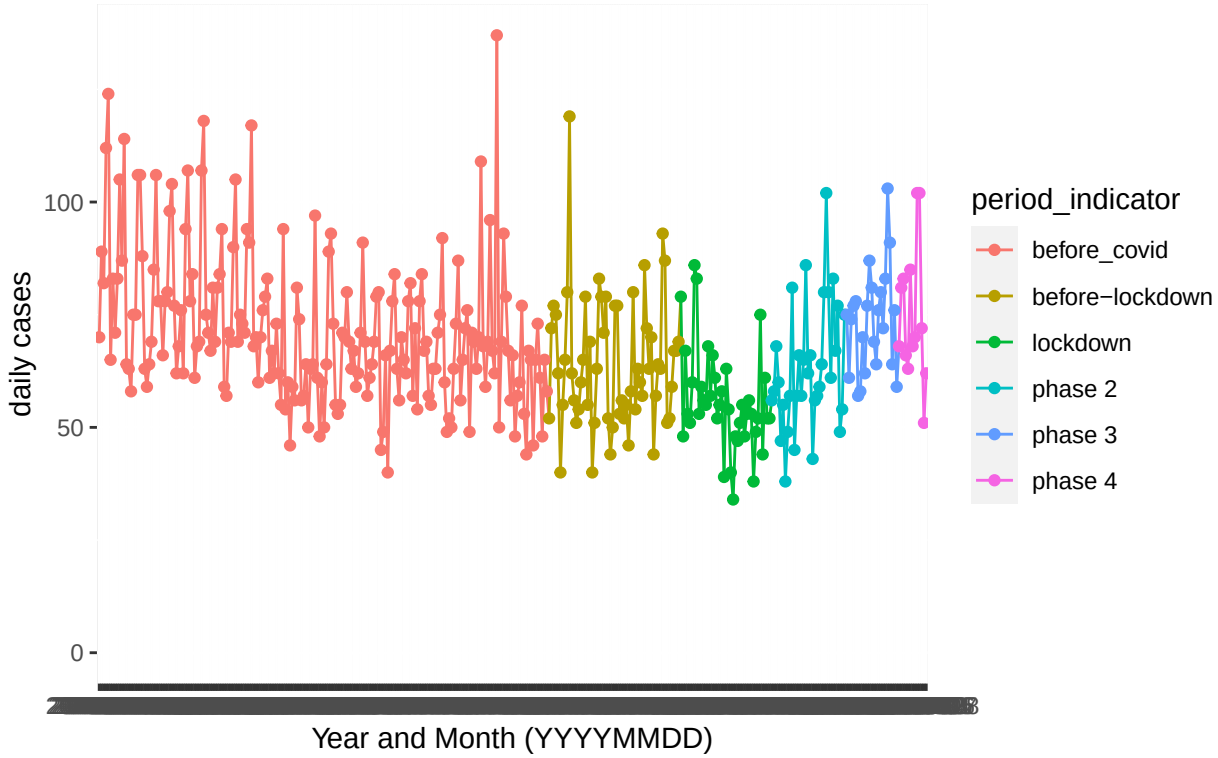
INTRODUCTION AND DATA

(BLAH BLAH from doc)

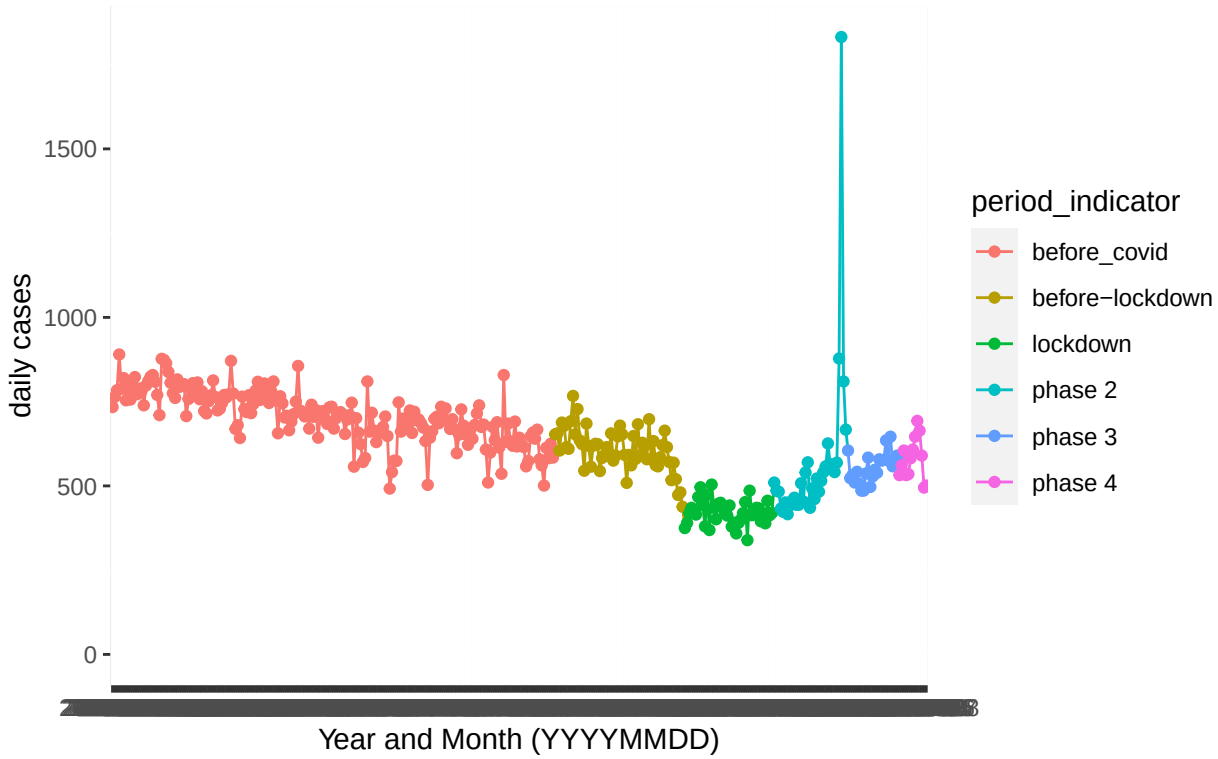
METHODOLOGY

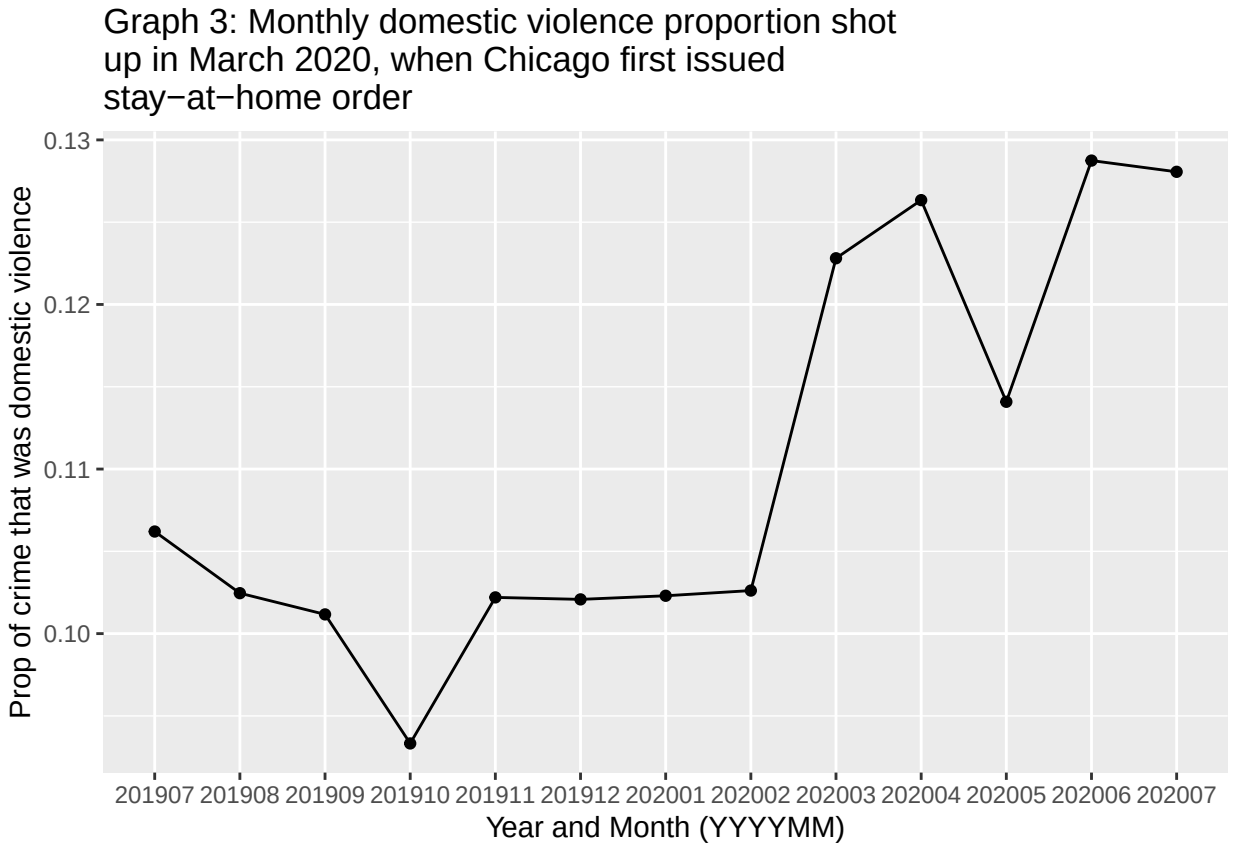
Visualizations To visualize the fluctuations in domestic violence cases during the time period of interest, we first plotted the daily number of cases of all crimes and those reported as domestic violence in the past year. We also plotted the proportion of crime cases that were related to domestic violence by month. Crimes were classified as cases of domestic violence if they were reported as domestic violence or aggravated domestic assault in the CLEAR dataset.

Graph 1: # Daily domestic violence cases remained relatively constant this past year



Graph 2: Total # daily crime cases dipped at the onset of lockdown





As seen in graphs 1 and 2, daily numbers of domestic violence cases remained fairly constant during the pandemic, while the total number of crime cases dipped markedly, starting at the beginning of lockdown. These observations are supported by graph 3, where the largest spike in the proportion of domestic violence cases occurred in March 2020 (at the beginning of Chicago's stay-at-home order), in accordance with our initial hypothesis.

Note: the dramatic spike in graph 2 (May 31st) and dip in graph 3 (May) can be explained by the extraordinary number of non-domestic violence crimes (i.e. looting, riots, violence) related to the Black Lives Matter protests over the murder of George Floyd in Chicago on that day, which overshadowed domestic violence cases [8].

Chi-Square Test Next, we assessed whether the relationship between these fluctuations in the proportion of domestic violence cases and quarantine status was statistically significant using a chi-square test of independence.

In order to simplify the analysis, we generalized the 6 phases of the pandemic into 3 periods: pre-lockdown (BEFORE COVID-19 and BEFORE LOCKDOWN), during lockdown (LOCKDOWN and PHASE 2), and post-lockdown (PHASE 3 and PHASE 4). These 3 periods served as our time categories for our chi-square test. Since there were enough independent observations in each period (n>10 for each cell) to satisfy the model assumptions, we conducted a chi-square test at the $\alpha = 0.05$ significance level, which tests the two hypotheses below:

H~0 : The proportion of crime cases that were domestic violence in Chicago is unrelated to the 3 periods of the pandemic. H~1 : The proportion of crime cases that were domestic violence in Chicago is related to the 3 periods of the pandemic.

```
##
##     during   post    pre
##   0  31109  17599 157397
```

```
##   1   4311   2667  17919
```

```
##
##  Pearson's Chi-squared test
##
## data:  table(domvio_mut$isdomviolence, domvio_mut$PERIOD)
## X-squared = 247.63, df = 2, p-value < 2.2e-16
```

Step Down 2 Proportion Z-Tests Given the significant p-value of our chi-square test statistic, we conducted three step-down 2 sample tests of proportion to identify any pairwise differences between periods.

For each test, our null hypothesis was that the proportion of crimes reported as domestic violence of the periods of interest was the same, while the alternative hypothesis was that the proportion of crimes reported as domestic violence of the periods of interest was different.

During-vs-post

```
##
##  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
```

Post-vs-pre

```
##
##  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
```

During-vs-pre

```
##
##  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
```

Regression Analysis Knowing that there is a general difference between the proportion of crimes reported that were related to domestic violence in relation to the lockdown status of the city, we moved to quantify the relationship with a logistic regression model. Here, we aimed to answer the question: how much more likely was a reported crime to be related to domestic violence, given the phase in which it occurred?

The logistic regression below models the relationship between the logit of the probability that a case was an incidence of domestic violence (a "success") given the phase of the pandemic in which it occurred. In order to control for an expected increase in incidences of domestic violence outside of the 9-to-5 time period (when individuals are either more likely to be at work or outside the home) before strict quarantine, our model included dummy variables for the time of the day as well as the phase of the pandemic (BEFORE COVID-19, BEFORE LOCKDOWN, LOCKDOWN, PHASE 2, PHASE 3, and PHASE 4) in which the crime was reported. BEFORE COVID-19 was used as the reference baseline category for the phase, while occurrence of a crime outside of the 9-to-5 period was used as the baseline category for time of day. We hypothesized that a crime was more likely to be related to domestic violence if it occurred outside of the 9am to 5pm period and/or during the lockdown phases. Additionally, we expected to see a decrease of the probability of success (i.e. that a case reported related to domestic violence) as the city of Chicago began to gradually reopen in recovery phases 2, 3, and 4.

```
## # 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
```

We obtained the following 95% confidence intervals for the odd ratios corresponding to each predictor, conditional on all the other predictors in our model:

BEFORE LOCKDOWN: (1.014, 1.097) LOCKDOWN: (1.292, 1.419) PHASE 2: (1.058, 1.166) PHASE 3: (1.279, 1.427) PHASE 4: (1.210, 1.393) 9-to-5: (0.711, 0.751)

RESULTS

Chi-square test Under the null hypothesis, our test statistic has a chi-square distribution with 2 degrees of freedom and 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 the proportion of crime cases that were domestic violence in Chicago is related to the periods of the pandemic.

Step Down 2 Proportion Z-Tests To account for multiple comparisons, we performed the Bonferroni correction and thus assessed our results relative to the adjusted $\alpha = 0.05/3$ level. At the adjusted $\alpha = 0.05/3$ level, we found all three pairwise differences in the proportion of crimes reported as domestic violence to be statistically significant. That is, there is sufficient evidence to suggest that the proportion of crime cases that were related to domestic violence in Chicago differs between each period of the pandemic.

We chose to focus in on the last two pairwise comparisons (post lockdown vs. pre lockdown and during lockdown vs. pre lockdown), where the respective 95% confidence intervals of (0.0245, 0.0343) and (0.0158, 0.0232) were positive. This suggests that the proportion of crime cases that were domestic violence has increased since issuance of stay-at-home orders.

Regression Analysis Our fitted logistic regression model, where each predictor corresponds to the predicted logit of the probability of success (i.e. a reported crime being related to domestic violence) of our outcome, conditionally on all other variables, is as follows:

$\hat{\beta}_0$ + $\hat{\beta}_1$ * (PHASE == before-lockdown) + $\hat{\beta}_2$ * (PHASE==lockdown) + $\hat{\beta}_3$ * (PHASE==phase2) + $\hat{\beta}_4$ * (PHASE==phase3) + $\hat{\beta}_5$ * (PHASE == phase4) + $\hat{\beta}_6$ * (is9_5)

At the $\alpha = 0.05$ significance level, the fitted β coefficients for each of our dummy variables relative to the baseline category of BEFORE COVID-19 were statistically significant, as was the fitted slope coefficient corresponding to the logit of the probability of success during the 9-to-5 workday, relative to the baseline category of not 9-to-5 workday (i.e. outside this window of time). Therefore, there is sufficient evidence to suggest that the true slope corresponding to each predictor is not equal to 0. There is some relationship between the nature of the crime (i.e. domestic violence or not domestic violence) and the phase in which the crime occurred, as well as time of day, conditional on the other variables. We were specifically interested in β_2, which corresponds to the logit of the probability of success that a reported crime was related to domestic violence during lockdown, as compared to before the epidemic hit Chicago:

H~0: $\beta_2 = 0$ H~1: $\beta_2 \neq 0$

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 < 0.001. Thus, at the $\alpha = 0.05$, we reject our null hypothesis; we have sufficient evidence to suggest that the true value of $\beta_2 \neq 0$, and that there is a relationship between the lockdown phase and the logit of the probability that a reported case is related to domestic violence, while holding time of day constant.

The estimated $\hat{\beta}_2$ coefficient was 0.303. Therefore, we would expect crimes reported during the lockdown phase to have approximately 1.354 (exp(0.303), with 95% CI (1.292, 1.419)), times the odds of being related to domestic violence compared to cases that were reported before the COVID-19 pandemic hit Chicago, while adjusting for time of day. This suggests that a reported crime was more likely to be related to domestic violence if it occurred during the lockdown.

The 95% confidence interval for the odds ratio corresponding to lockdown (exp(β_2)) is greater than (and does not overlap with) that of before lockdown (exp(β_1)), which suggests that compared to cases before COVID-19 hit Chicago, the odds of a crime committed during lockdown has a greater times the odds of being related to domestic violence than one committed before lockdown, while adjusting for time of day. The same case can be made for phases 3 and 4. Interestingly, the 95% confidence interval for the odds ratio corresponding to phase 2 (exp(β_4)) overlaps with that of before lockdown (exp(β_1)), suggesting that compared to cases before COVID-19 hit Chicago, the odds of a crime committed during phase 2 has comparable times the odds of being related to domestic violence than one committed before lockdown, while adjusting for time of day. However, it is important to note that this dip in the odds ratio corresponding to phase 2 may be inaccurate because it includes the extraordinary spike of non-domestic violences cases on May 31st due to the Black Lives Matter protests over the murder of George Floyd on that date. If we ignore phase 2 as an outlier, we see that contrary to our hypothesis, the city's gradual reopening did not decrease the likelihood of a crime being related to domestic violence to a statistically significant degree.

Lastly, the estimated β_6 coefficient corresponding to our binary indicator "is9_5" which determined whether a given crime was reported during the 9-to-5 workday was -0.314, which corresponds to an odds ratio of approximately 0.731. Therefore, we would expect crimes that were reported during the 9-to-5 timeframe to have 0.731 (0.711, 0.751) times the odds of being related to domestic violence as compared to cases that occurred outside of work hours, adjusting for the phase in which the crime was committed. As expected, a crime is less likely to be related to domestic violence if it were reported within the hours of 9-to-5.

DISCUSSION

Through our analysis, we observed that there was indeed a trend towards increased proportions of domestic violence cases over the phases of lockdown and reopening due to COVID-19, as we had originally hypothesized.

In particular, our analyses found statistically significant differences between the proportion of domestic violence cases reported during the three different periods of pre-lockdown, during lockdown, and post lockdown (corresponding to a p-value of < 0.001 for both the chi-square and subsequent step-down tests). We also found that when adjusting for time of day, the odds of a crime reported during lockdown being related to domestic violence was greater than the odds of a crime reported before the pandemic began being related to domestic violence, as evidenced by the statistically significant slope coefficient β_2 from our hypothesis test in the fitted logistic model. However, trends in crimes did not appear to exactly coincide with the phases of recovery, as we had predicted. On the contrary, we do not see a decrease in the odds of a crime being related to domestic violence during the later recovery periods, relative to before the pandemic (except for phase 2, which we classified as an outlier for including the day of the George Floyd protest), as evidenced by overlapping confidence intervals for our respective odds ratios. Perhaps the reopening plan issued by Chicago authorities has not been impactful enough to bring domestic violence rates back down to "normal" relative to before the stay-at-home order was implemented. This could be due in part to the fact that there have been many challenges with reopening the country - so much so that many states have even had to shut down once again as the number of cases continue to rise.

Though our analyses show increases in the proportion and likelihood of reported domestic violence cases starting at the beginning of lockdown in Chicago, we cannot conclude with certainty that this is a direct result of an actual rise in cases of domestic violence; it could be that domestic violence cases happened to be reported more frequently during lockdown, or even that non-domestic violence cases decreased in occurrence or reporting, but the underlying cases of domestic violence themselves stayed the same. Most likely, the decrease in total crime cases during the lockdown (as reported by some sources [8]) was more significant than the decrease in domestic violence cases, leading to an increase in the proportion of domestic violence crimes as a result. In order to address this, in the future it might be worthwhile to conduct an ANOVA in order to determine if the mean number of cases of domestic violence reported per day actually changed throughout the phases of the pandemic.

In our regression model that evaluated the relationship between the logit of the probability of a given crime being related to domestic violence and the phase during which the crime was reported, we only adjusted for the time of day of the crime (either within the hours of 9-5 or outside this time period). However, other confounding variables present in our dataset, such as the location of the crime (neighborhood, district, ward, etc.), may have had an impact on our predicted outcome, but were not accounted for. Additionally, though our 9-to-5 indicator was designed to account for an expected lower number of domestic violence cases during the day, given that individuals would not be at home during this time (prior to lockdown), and therefore subject to potential abuse, we did not distinguish between weekdays and weekends. Additionally, it might also have been useful to include an interaction term between our 9-to-5 indicator and the phases of the pandemic, given that we would expect this indicator to be a much more useful predictor of whether a reported case was related to domestic violence before lockdown than during lockdown, as there would likely be decreased separation between normal workday hours and time spent at home in quarantine.

Furthemore, variables not included in our dataset, such as race and socioeconomic status, could potentially be associated with incidence of domestic violence. Adjusting for these factors in our logistic regression model might improve our predictions and better showcase the true relationship between domestic violence cases and the timeline of the pandemic. For example, individuals from different backgrounds may tend to be affected more than others by domestic violence, which could have public policy implications for targeted interventions.

Were we to repeat our analyses, we would include and adjust for more potential confounders to improve the prediction accuracy of our model. Secondly, expanding the population of interest to citizens in multiple U.S. cities or cities in other countries would give us a better idea of how quarantine affected domestic violence cases differently by region/state, country, or population density (urban vs. rural). This could also give us insight into how successful or unsuccessful reopening and recovery measures in other countries have impacted cases of domestic violence, especially when compared to how the US has handled the pandemic thus far. For example, we might observe a more marked decline in terms of domestic violence cases throughout reopening, as we had originally hypothesized, in other countries that have had greater success returning to some sense of "normal" life after strict lockdown. Thirdly, instead of discarding outliers, we could take measures to control for the abnormal spikes and dips in our data resulting from protests. In order to assess how well our sample

represents our population of interest, we could also conduct research on whether police report data is actually reflective of the true number of domestic violence cases occurring during lockdown. It may be possible that domestic violence cases are underreported because victims are afraid to report their abusers, especially when they may no longer have a safe place to flee to.

Finally, it might also be interesting to continue to follow rates of domestic violence for an extended time period after the immediate threat of the virus has passed, and we are left to deal with its long-lasting economic, social, and political consequences. Sobering studies show that abusers are "more likely to murder their partners in the wake of personal crises, including lost jobs or major financial setbacks," meaning we could potentially see even more drastic effects of domestic violence as the nation begins to recover [1].

REFERENCES

[1] Taub, A. (2020, April 06). A New Covid-19 Crisis: Domestic Abuse Rises Worldwide. Retrieved from https://www.nytimes.com/2020/04/06/world/coronavirus-domestic-violence.html

[2] Crimes - One year prior to present: City of Chicago: Data Portal. (2020, July 29). Retrieved from https://data.cityofchicago.org/Public-Safety/Crimes-One-year-prior-to-present/x2n5-8w5q

[3] Gracia, E. (2004). Unreported cases of domestic violence against women: Towards an epidemiology of social silence, tolerance, and inhibition. Journal of Epidemiology & Community Health, 58(7), 536-537. doi:10.1136/jech.2003.019604

[4] About COVID-19. (n.d.). Retrieved from https://dph.illinois.gov/topics-services/diseases-and-conditions/diseases-a-z-list/coronavirus/symptoms-treatment

[5] COVID-19 in Illinois, the U.S. and the world: Timeline of the outbreak. (2020, May 15). Retrieved from https://www.chicagotribune.com/coronavirus/ct-viz-coronavirus-timeline-20200507-uvrzs32nljabrpn6vkzq7m2fpq-story.html

[6] NBC Chicago. (2020, June 09). Here's What to Know About Illinois' 5 Phases of Reopening. Retrieved from https://www.nbcchicago.com/news/coronavirus/heres-what-to-know-about-illinois-5-phases-of-reopening/2286825/

[7] NBC Chicago. (2020, June 26). Illinois Enters Phase 4 of Reopening Plan: Here's What's Changing. Retrieved from https://www.nbcchicago.com/news/local/illinois-enters-phase-4-of-reopening-plan-heres-whats-changing/2295816/

[8] Chicago Tribune staff. (2020, June 01). George Floyd fallout: Here's what happened May 31 in the Chicago area. Retrieved from https://www.chicagotribune.com/news/breaking/ct-george-floyd-chicago-protests-20200531-qghf4l7ysjgl3etxqu3jv6oq6a-story.html