Has expansion of Medicaid under ACA affected statewide crime rates in the states that have implemented expansion?

Taylor Hertzler

Follow this and additional works at: https://scholarship.law.upenn.edu/prize_papers

Part of the Law Commons
Overview

This project examines the connection between crime rates and healthcare access. It assesses whether the expansion of Medicaid under the Affordable Care Act (ACA) has led to an increase or decrease in crime rates. Answering this question will involve finding crime-rate data throughout the country (at the county level). These data are available from the Department of Justice’s Uniform Crime Reporting Program. It will also require data on Medicaid expansion (which states have expanded and when). These data are available from many sources, but this interactive map (https://www.kff.org/medicaid/issue-brief/status-of-state-medicaid-expansion-decisions-interactive-map/) from KFF.org is one of the most intuitive.

I consulted three faculty members to help design this project: Dr. Jonathan Klick, Professor of Law, Penn Law; Dr. Douglas J. Wiebe, Professor of Epidemiology, Perelman, Department of Biostatistics, Epidemiology and Informatics; and Dr. Jesse Yenchih Hsu, Assistant Professor of Biostatistics and Epidemiology, Perelman, Department of Biostatistics, Epidemiology and Informatics.

Dr. Klick helped me to come up with the idea and pointed me to some data sources, including the many papers that have already studied the expansions in other ways. He also helped me to refine my methods, particularly the regressions at the end. Dr. Wiebe helped me to take my original question (“Have the state Medicaid expansions under ACA affected crime rates?”) and make it more specific and testable (see below). He said that the best way to do this would be a before-and-after analysis rather than a cross-sectional one that compares implementation states with non-implementation states at one point in time only. He also suggested that I analyze a specific year or an average of a few years, rather than a trend line across years; doing this for a before and an after for each group of states will allow me to complete a difference-in-differences analysis, which is a particularly strong study design. Dr. Hsu cautioned me to be conservative in any conclusions I draw, since the large possibility of confounding variables means that I
might not be able to assert a straight causal connection, and to discuss the most likely confounding variables.

Final Project Github Repository (https://github.com/hzler/BMIN503_Final_Project)

**Introduction**

My topic is "Has expansion of Medicaid under ACA affected statewide crime rates in the states that have implemented expansion?" As part of ACA, many states have expanded their Medicaid programs, a process that began in 2014. This expansion, however, has not been uniform, as many states did not expand right away, and many states have not expanded at all. In expansion states, Medicaid expansion affects both access to and costs of healthcare. Many researchers have accordingly studied these expansions to see how those access/costs changes have affected other social phenomena.

One aspect that has received little attention, however, is the effects expansion has had on crime rates. Medicaid expansion expands the group of eligible adults to include all adults with incomes within 138% of the poverty level. Some effects that have resulted include: greater increases in coverage among the near poor than the increases seen in non-expansion states (1); greater reductions in uninsured rates than the reductions seen in non-expansion states (1),(2); and greater reductions than non-expansion states in the number of individuals who reported skipping care because of costs (3),(4); *but also* greater difficulty in accessing physician care than that experienced in non-expansion states (1),(2),(5).

Effects like these suggest a common trend: Medicaid expansion improved healthcare access for poor folks. That trend on its own may suggest an effect on crime. Specifically, poverty may induce individuals to engage in crime (especially theft) to cover healthcare costs that they cannot afford; if those individuals now have healthcare access, they should commit fewer crimes. Indeed, poverty may induce individuals to engage in crime to cover any costs that they cannot afford, healthcare or otherwise; granting healthcare access to (and thus reducing healthcare costs for) these individuals should therefore lead them to commit fewer crimes. Thus, we might expect reductions in crime rates in expansion states. On the other hand, variables such as increased difficulty in accessing physicians may negate or at least diminish those effects, for increasing an individual’s healthcare coverage does not do much if his access to healthcare does not also improve. It is therefore worthwhile assessing Medicaid expansion’s effects on crime rates.

Most of the debate around ACA has concerned its efficacy: improving access to healthcare, decreasing healthcare costs, improving the efficiency of our healthcare system etc. Less of the debate concerns ACA’s secondary effects: how its policies affect social phenomena like employment, education, secondary governmental policies (e.g., drug laws or welfare programs) etc. One important secondary effect to address is crime rates and how, if at all, ACA’s expansion programs have affected them. Any causal connections that do exist between the two should play a part in the nation’s continued debate over our national healthcare scheme.

This problem is interdisciplinary because it involves informatics, healthcare and criminology. Healthcare
and crime continue to be pressing topics throughout our nation, at various governmental levels, making this an important topic. Both also tend to have lots of data, making them ripe for empirical analysis. The project begins with healthcare data, as the first steps are grouping all fifty states into expansion and non-expansion states and identifying expansion years for analysis. The criminology aspects concern not just crime data but which types of crime data to use. Most states compile total criminal offenses but then subdivide those into violent and non-violent (i.e., property) offenses. The informatics component informs which parts of these data to use and how to use them to answer the above question. This problem thus incorporates these three disciplines to answer a question that ultimately concerns public policy.

## Methods

This project will culminate in linear regressions and a difference-in-differences analysis. Crime data from expansion states will be examined to determine the difference in per capita crime rates across those states after Medicaid expansion was implemented. The same will be done using crime data from non-expansion states. Those differences will then be compared to see whether expansion of Medicaid likely contributed to those changes. Finally, the same processes will be performed on violent-crime data only and then property-crime data only to see if Medicaid expansion may have affected those subtypes of crimes.

The first step is to obtain criminal data from the Department of Justice. DOJ's Uniform Crime Reporting Program has county-level crime data going back several decades. The crime data are both comprehensive, covering most of the nation, as well as detailed, dividing crimes into violent and property and counting each individual crime type (e.g., murder, arson etc.) within those categories.

The analyses to be conducted will require data from two different years. The first year should be shortly before any states expanded Medicaid. Expansion began in 2014, so 2012 data will be used for the earlier year. The second year should be as far after expansion as possible. The most recent UCRP data are from 2016, so those data will be used for the later year.

```r
library(haven)
library(dplyr)

##
## Attaching package: 'dplyr'

## The following objects are masked from 'package:stats':
##
## filter, lag
## The following objects are masked from 'package:base':
## intersect, setdiff, setequal, union

```r
arrests_2012 <- read_dta("ICPSR_35019/DS0001/35019-0001-Data.dta") # nationwide crime data for all ages, 2012
arrests_2012 <- dplyr::rename(arrests_2012, CNTY POP = CPOPARST, TOT CRIMES = P1TOT, VLNT CRIMES = P1VLNT, PROP CRIMES = P1PRPTY) # changing a few names
arrests_2012 <- subset(arrests_2012, select =
c(FIPS ST, FIPS CTY, CNTY POP, TOT CRIMES, VLNT CRIMES, PROP CRIMES)) # removing unwanted columns
arrests_2012$CRIMES_PERCAP <- arrests_2012$TOT CRIMES/arrests_2012$CNTY POP # adding a column showing per capita crime rates
arrests_2012$VLNT_PERCAP <- arrests_2012$VLNT CRIMES/arrests_2012$CNTY POP # adding a column showing per capita violent crime rates
arrests_2012$PROP_PERCAP <- arrests_2012$PROP CRIMES/arrests_2012$CNTY POP # adding a column showing per capita property crime rates
arrests_2012 <- subset(arrests_2012, CNTY POP != 0) # removing zero-population counties

# same processes for 2016 data
arrests_2016 <- read_dta("ICPSR_37059/DS0001/37059-0001-Data.dta")
arrests_2016 <- dplyr::rename(arrests_2016, CNTY POP = CPOPARST, TOT CRIMES = P1TOT, VLNT CRIMES = P1VLNT, PROP CRIMES = P1PRPTY)
arrests_2016 <- subset(arrests_2016, select =
c(FIPS ST, FIPS CTY, CNTY POP, TOT CRIMES, VLNT CRIMES, PROP CRIMES))
arrests_2016$CRIMES_PERCAP <- (arrests_2016$TOT CRIMES/arrests_2016$CNTY POP)
arrests_2016$VLNT_PERCAP <- arrests_2016$VLNT CRIMES/arrests_2016$CNTY POP
arrests_2016$PROP_PERCAP <- arrests_2016$PROP CRIMES/arrests_2016$CNTY POP
arrests_2016 <- subset(arrests_2016, CNTY POP != 0)
```

Next, the data for each county in 2016 must be labeled based on whether that county lies in a state that has expanded Medicaid. The following states are those that (a) never expanded Medicaid, (b) adopted expansion but have yet to implement it or (c) expanded Medicaid after December 31, 2016: AL, FL, GA, ID, KS, ME, MS, MO, NE, NC, OK, SC, SD, TN, TX, VT, VA, WI, WY. All other fifty states (minus LA) plus DC had expanded Medicaid by January 1, 2016. LA will be excluded from the calculations because it expanded Medicaid on July 1, 2016.
library(stringr)

# removing Louisiana from both sets
arrests_2016 <- subset(arrests_2016, FIPS_ST != 22)
arrests_2012 <- subset(arrests_2012, FIPS_ST != 22)

# new column for year status (2012 [before exp] = 1, 2016 [after exp] = 2)
arrests_2016$AFTER_EXP <- 2
arrests_2012$AFTER_EXP <- 1

# creating new column (for setting expansion status)
arrests_2016$EXPANSION_STATUS <- arrests_2016$FIPS_ST
arrests_2012$EXPANSION_STATUS <- arrests_2012$FIPS_ST

# setting status as 1 for non-expansion states, 2 for expansion states in both sets
arrests_2016 <- arreststs_2016 %>%
mutate(EXPANSION_STATUS = if_else(str_detect(FIPS_ST, "(1|12|13|16|20|23|28|29|31|37|40|45|46|47|48|50|51|55|56)"), 1, 2))
arrests_2012 <- arreststs_2012 %>%
mutate(EXPANSION_STATUS = if_else(str_detect(FIPS_ST, "(1|12|13|16|20|23|28|29|31|37|40|45|46|47|48|50|51|55|56)"), 1, 2))

# combining state and county fips codes
colnames(arrests_2012)[colnames(arrests_2012) == "FIPS_ST"] <- "FIPS"
arrests_2012$FIPS <- sprintf("%02d", arrests_2012$FIPS)
arrests_2012$FIPS_CTY <- sprintf("%03d", arrests_2012$FIPS_CTY)
arrests_2012$FIPS <- paste0(arrests_2012$FIPS, arrests_2012$FIPS_CTY)
arrests_2012 <- subset(arrests_2012, select = -c(FIPS_CTY))

# same for 2016
colnames(arrests_2016)[colnames(arrests_2016) == "FIPS_ST"] <- "FIPS"
arrests_2016$FIPS <- sprintf("%02d", arrests_2016$FIPS)
arrests_2016$FIPS_CTY <- sprintf("%03d", arrests_2016$FIPS_CTY)
arrests_2016$FIPS <- paste0(arrests_2016$FIPS, arrests_2016$FIPS_CTY)
arrests_2016 <- subset(arrests_2016, select = -c(FIPS_CTY))

# ensuring that all data are numeric (minus AFTER_EXP and EXPANSION_STATUS)
arrests_2012$FIPS <- as.numeric(arrests_2012$FIPS)
arrests_2016$FIPS <- as.numeric(arrests_2016$FIPS)
arrests_2016$AFTER_EXP <- as.factor(arrests_2016$AFTER_EXP)
arrests_2012$AFTER_EXP <- as.factor(arrests_2012$AFTER_EXP)
arrests_2016$EXPANSION_STATUS <- as.factor(arrests_2016$EXPANSION_STATUS)
arrests_2012$EXPANSION_STATUS <- as.factor(arrests_2012$EXPANSION_STATUS)
```r
sapply(arrests_2012, class)

# A tibble: 6 x 10
#  FIPS CNTY_POP TOT_CRIMES VLNT_CRIMES PROP_CRIMES CRIMES_PERCAP VLNT_PERCAP
#       <dbl>    <dbl>      <dbl>       <dbl>       <dbl>         <dbl>       <dbl>
# 1  1001   57161        220          10         210       0.00385   0.000175
# 2  1003   18747        586          26         560       0.00313   0.000139
# 3  1005    27228       113           5          108       0.00415   0.000184
# 4  1007    22907        64           3          62       0.00279   0.000131
# 5  1009    57909        68           3          65       0.00117   0.0000518
# 6  1011    10584        24           1          23       0.00227   0.0000945
```

Showing summaries of 2012 data as an example

```r
sapply(arrests_2016, class)

# A tibble: 6 x 10
#  FIPS CNTY_POP TOT_CRIMES VLNT_CRIMES PROP_CRIMES CRIMES_PERCAP VLNT_PERCAP
#       <dbl>    <dbl>      <dbl>       <dbl>       <dbl>         <dbl>       <dbl>
# 1  1001   57161        220          10         210       0.00385   0.000175
# 2  1003   18747        586          26         560       0.00313   0.000139
# 3  1005    27228       113           5          108       0.00415   0.000184
# 4  1007    22907        64           3          62       0.00279   0.000131
# 5  1009    57909        68           3          65       0.00117   0.0000518
# 6  1011    10584        24           1          23       0.00227   0.0000945
```

Showing summaries of 2012 data as an example
Summary (arrests_2012)

#creating four separate datasets, distinguished by year and expansion status
arrests_2012_expansion <- subset(arrests_2012, EXPANSION_STATUS == 2)
arrests_2012_nonexpansion <- subset(arrests_2012, EXPANSION_STATUS == 1)
arrests_2016_expansion <- subset(arrests_2016, EXPANSION_STATUS == 2)
arrests_2016_nonexpansion <- subset(arrests_2016, EXPANSION_STATUS == 1)

Four separate datasets were then created, based on year and dividing between states that had expanded Medicaid by 2016: 2012 expansion states, 2012 non-expansion states, 2016 expansion states and 2016 non-expansion states.

Results

Total Crimes

Since all the data for the entire nation are available, and the number of instances (counties) is on the order of magnitude of thousands (making it feasible to conduct analyses on the whole set), all data will be used for the analyses, rather than using a random sample from each dataset.

The analyses will have to control for population. Not only will population levels have changed from 2012 to 2016, but population totals will differ between expansion and non-expansion states. Below is an
Population will therefore be controlled for by using per capita crime rates. To conduct statistical analyses, per capita crime rate will be treated as a continuous variable beginning at 0: i.e., for every county, how many crimes per person were committed in the year in question? That will allow for linear regressions, with year serving as a binary variable: i.e., before or after expansion (2012 = 1, 2016 = 2).

First, below is a visualization of how per capita crime rates changed from 2012 to 2016 between expansion v. non-expansion states.
#creating a MATRIX holding the average per capita crime rate for each year

```r
a <- sum(arrests_2012_expansion$TOT_CRIMES)/sum(arrests_2012_expansion$CNTY_POP)
b <- sum(arrests_2016_expansion$TOT_CRIMES)/sum(arrests_2016_expansion$CNTY_POP)
c <- sum(arrests_2012_nonexpansion$TOT_CRIMES)/sum(arrests_2012_nonexpansion$CNTY_POP)
d <- sum(arrests_2016_nonexpansion$TOT_CRIMES)/sum(arrests_2016_nonexpansion$CNTY_POP)

avgs_matrix <- matrix(c(a, b, c, d), 1, 4)
```

The barplot suggests that time was a significant predictor of crime rates. If this effect were limited to expansion states, that might suggest that Medicaid expansion on its own is a significant predictor of crime rates. But the effect exists for both expansion and nonexpansion states, so crime rates went down nationwide from 2012 to 2016. This makes the difference-in-differences analysis necessary.

Here are the actual percentage decreases in per capita crime rates between 2012 and 2016, for expansion and non-expansion states, and the ratio between those two decreases.
Across expansion states, the per capita crime rate decreased by 0.175% from 2012 to 2016.

Across non-expansion states, the per capita crime rate decreased by 0.182% from 2012 to 2016.

Non-expansion states thus saw a greater decrease than expansion states in per capita crime rates from 2012 to 2016. The decrease in non-expansion states was 1.043 times greater than the decrease seen in expansion states.

Linear regressions will be performed to assess whether that 1.043x difference is significant. Below are boxplots showing the distributions of per capita crime rates across the four samples (1 = before expansion, 2 = after).
ggplot(nonexp_df, aes(x = AFTER_EXP, y = CRIMES_PERCAP)) + geom_boxplot() + ggtitle("Non-expansion states: per capita crime rates")
Aside from some outliers toward the top, these plots fail to show any major differences before and after Medicaid expansion for each group. Below are linear regressions assessing whether time was a significant predictor of per capita crime rates, for both expansion and non-expansion states.

```r
# expansion states linear regression
exp_lin.reg <- lm(CRIMES_PERCAP ~ AFTER_EXP, data = exp_df)
summary(exp_lin.reg)
```
## Call:
## `lm(formula = CRIMES_PERCAP ~ AFTER_EXP, data = exp_df)`
##
## ## Residuals:
##          Min        1Q    Median        3Q       Max
## -0.004645 -0.002422 -0.000385  0.001824  0.036531
## ##
## ## Coefficients:
##                Estimate Std. Error t value Pr(>|t|)
## (Intercept)  4.645e-03  9.413e-05  49.343  < 2e-16 ***
## AFTER_EXP2  -4.467e-04  1.331e-04  -3.357 0.000801 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.003376 on 2572 degrees of freedom
## Multiple R-squared:  0.004361,   Adjusted R-squared:  0.003974
## F-statistic: 11.27 on 1 and 2572 DF,  p-value: 0.0008008

#nonexpansion states linear regression
nonexp_lin.reg <- lm(CRIMES_PERCAP ~ AFTER_EXP, data = nonexp_df)
summary(nonexp_lin.reg)

## Call:
## `lm(formula = CRIMES_PERCAP ~ AFTER_EXP, data = nonexp_df)`
##
## ## Residuals:
##          Min        1Q    Median        3Q       Max
## -0.005057 -0.002580 -0.000528  0.001905  0.042868
## ##
## ## Coefficients:
##                Estimate Std. Error t value Pr(>|t|)
## (Intercept)  5.057e-03  8.516e-05  59.38  < 2e-16 ***
## AFTER_EXP2  -6.684e-04  1.204e-04  -5.55 3.06e-08 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.003599 on 3570 degrees of freedom
## Multiple R-squared:  0.008555,   Adjusted R-squared:  0.008277
## F-statistic: 30.81 on 1 and 3570 DF,  p-value: 3.061e-08
Using p < 0.05 as a significance level, time was a significant predictor of crime rates for both expansion states (p = 0.0008) and non-expansion states (p = 3.06e-08). This fits with what the barplot suggested. Further, time was a more significant predictor for the non-expansion states than for the expansion states (3.06e-08 < 0.0008), as predicted by the greater decrease in crime rates seen in non-expansion states.

To assess whether that difference is statistically significant, below is a difference-in-differences analysis.

```r
#creating a dataframe for all data from all four samples
agg_df <- rbind(exp_df, nonexp_df)

#creating an interaction term between expansion status and time for DiD analysis
agg_df$DiD <- (as.numeric(agg_df$AFTER_EXP)) * (as.numeric(agg_df$EXPANSION_STATUS))

#DiD analysis
fin_analysis <- lm(CRIMES_PERCAP ~ AFTER_EXP + EXPANSION_STATUS + DiD, data = agg_df)
summary(fin_analysis)
```

Using p < 0.05 as a significance level, the interaction term (DiD) was not a significant predictor of crime rates (p = 0.2215). This indicates that the difference between expansion states and non-expansion states
in their crime decreases is not statistically significant. These data, therefore, do not suggest that Medicaid expansion is a significant predictor of changes in crime rates.

Next, the same analyses will be conducted using violent-crime data only and then property-crime data only, beginning with violent crimes.

**Violent Crimes**

```r
#creating a MATRIX holding the average per capita violent crime rate for each year
e <- sum(arrests_2012_expansion$VLNT_CRIMES)/sum(arrests_2012_expansion$CNTY_POP)
f <- sum(arrests_2016_expansion$VLNT_CRIMES)/sum(arrests_2016_expansion$CNTY_POP)
g <- sum(arrests_2012_nonexpansion$VLNT_CRIMES)/sum(arrests_2012_nonexpansion$CNTY_POP)
h <- sum(arrests_2016_nonexpansion$VLNT_CRIMES)/sum(arrests_2016_nonexpansion$CNTY_POP)
avgs_matrix.vlnt <- matrix(c(e,f,g,h), 1, 4)

#barplot showing per capita violent crime rates in each of the four samples
barplot(avgs_matrix.vlnt, main = "All samples: per capita violent crime rates",
```
This barplot suggests that the 2012-to-2016 differences in both the expansion and non-expansion states were minimal.

\[
\begin{align*}
\text{r} & \leftarrow (\text{avgs_matrix.vlnt}[1, 1] - \text{avgs_matrix.vlnt}[1, 2]) / (\text{avgs_matrix.vlnt}[1, 1]) \\
\text{s} & \leftarrow (\text{avgs_matrix.vlnt}[1, 3] - \text{avgs_matrix.vlnt}[1, 4]) / (\text{avgs_matrix.vlnt}[1, 3]) \\
\text{paste} & ("Across expansion states, the per capita violent crime rate decreased by", \ \text{round}(r, \ \text{digits = 3}), \ "\% from 2012 to 2016.") \\
\text{paste} & ("Across non-expansion states, the per capita violent crime rate decreased by", \ \text{round}(s, \ \text{digits = 3}), \ "\% from 2012 to 2016.") \\
\text{paste} & ("Expansion states thus saw a greater decrease than expansion states in per capita violent crime rates from 2012 to 2016. The decrease in Expansion states was", \ \text{round}(r/s, \ \text{digits = 3}), \ "times greater than the decrease seen in expansion states.")
\end{align*}
\]

Linear regressions will be performed to assess whether that 2.087x difference is significant. Below are boxplots showing the distributions of per capita violent crime rates across the four samples.

```r
#boxplots to show the per capita change
ggplot(exp_df, aes(x = AFTER_EXP, y = VLNT_PERCAP)) +
  geom_boxplot() +
  ggtitle("Expansion states: per capita violent crime rates")
```
ggplot(nonexp_df, aes(x = AFTER_EXP, y = VLNT_PERCAP)) +
  geom_boxplot() +
  ggtitle("Non-expansion states: per capita violent crime rates")
Aside from some outliers toward the top, these plots fail to show any major differences before and after Medicaid expansion for each group.
## Call:
## lm(formula = VLNT_PERCAP ~ AFTER_EXP, data = exp_df)
##
## Residuals:
##       Min        1Q   Median        3Q       Max
## -0.0011459 -0.0007155 -0.0002436  0.0003364  0.0199372
##
## Coefficients:
##             Estimate Std. Error  t value  Pr(>|t|)
## (Intercept) 1.094e-03   3.386e-05   32.317   <2e-16 ***
## AFTER_EXP   5.146e-05   4.787e-05    1.075    0.283
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.001214 on 2572 degrees of freedom
## Multiple R-squared:  0.000449,   Adjusted R-squared:  6.039e-05
## F-statistic: 1.155 on 1 and 2572 DF,  p-value: 0.2825

## Call:
## lm(formula = VLNT_PERCAP ~ AFTER_EXP, data = nonexp_df)
##
## Residuals:
##       Min        1Q   Median        3Q       Max
## -0.0010615 -0.0006039 -0.0001700  0.0004097  0.0105492
##
## Coefficients:
##             Estimate Std. Error  t value Pr(>|t|)
## (Intercept) 1.029e-03   2.136e-05   48.205   <2e-16 ***
## AFTER_EXP   3.207e-05   3.020e-05    1.062    0.288
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.0009025 on 3570 degrees of freedom
## Multiple R-squared:  0.0003157,   Adjusted R-squared:  3.566e-05
## F-statistic: 1.127 on 1 and 3570 DF,  p-value: 0.2884

#nonexpansion states violent crime linear regression
nonexp_lin.reg.v <- lm(VLNT_PERCAP ~ AFTER_EXP, data = nonexp_df)
summary(nonexp_lin.reg.v)
Using \( p < 0.05 \) as a significance level, time was not a significant predictor of violent crime rates for expansion states (\( p = 0.283 \)) or non-expansion states (\( p = 0.288 \)). This fits with the percentage decreases calculated above, for the percentage decreases in violent crime rates experienced by expansion states (0.068) and non-expansion states (0.032) are miniscule. True, time was a more significant predictor for expansion states than for non-expansion states (0.283 < 0.288). But since (a) that difference is miniscule, and (b) time is not a significant predictor for either, a difference-in-differences analysis is unnecessary.

### Property Crimes

```r
# creating a MATRIX holding the average per capita crime rate for each year
i <- sum(arrests_2012_expansion$PROP_CRIMES)/sum(arrests_2012_expansion$CNTY_POP)
j <- sum(arrests_2016_expansion$PROP_CRIMES)/sum(arrests_2016_expansion$CNTY_POP)
k <- sum(arrests_2012_nonexpansion$PROP_CRIMES)/sum(arrests_2012_nonexpansion$CNTY_POP)
l <- sum(arrests_2016_nonexpansion$PROP_CRIMES)/sum(arrests_2016_nonexpansion$CNTY_POP)
avgs_matrix.p <- matrix(c(i,j,k,l), 1, 4)

# barplot showing per capita crime rates in each of the four samples
```

![All samples: per capita property crime rates](image)
Across expansion states, the per capita property crime rate decreased by 0.214 % from 2012 to 2016.

Across non-expansion states, the per capita property crime rate decreased by 0.218 % from 2012 to 2016.

Non-expansion states thus saw a greater decrease than expansion states in per capita property crime rates from 2012 to 2016. The decrease in non-expansion states was 1.017 times greater than the decrease seen in expansion states.

Linear regressions will be performed to assess whether that 1.017x difference is significant. Below are boxplots showing the distributions of per capita property crime rates across the four samples.
ggplot(nonexp_df, aes(x = AFTER_EXP, y = PROP_PERCAP)) +
  geom_boxplot() +
  ggtitle("Non-expansion states: per capita property crime rates")
Aside from some outliers toward the top, these plots fail to show any major differences before and after Medicaid expansion for each group.

```r
# expansion states property crime linear regression
exp_lin.reg.p <- lm(PROP_PERCAP ~ AFTER_EXP, data = exp_df)
summary(exp_lin.reg.p)
```
```r
#nonexpansion states property crime linear regression
nonexp_lin.reg.p <- lm(PROP_PERCAP ~ AFTER_EXP, data = nonexp_df)
summary(nonexp_lin.reg.p)
```

```r
## Call:
## lm(formula = PROP_PERCAP ~ AFTER_EXP, data = exp_df)
##
## Residuals:
##          Min         1Q     Median         3Q        Max
## -0.0035492 -0.0019634 -0.0003494  0.0014171  0.0206549
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  3.549e-03  7.236e-05  49.050  < 2e-16 ***
## AFTER_EXP2  -4.959e-04  1.023e-04  -4.848 1.32e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.002595 on 2572 degrees of freedom
## Multiple R-squared:  0.009055,  Adjusted R-squared:  0.00867
## F-statistic:  23.5 on 1 and 2572 DF,  p-value: 1.321e-06
```
Using p < 0.05 as a significance level, time was a significant predictor of property crime rates for both expansion states (p = 1.32e-06) and non-expansion states (p = 5.68e-12). This fits with what the barplot suggested. Further, time was a more significant predictor for the non-expansion states than for the expansion states (5.68e-12 < 1.32e-06), as predicted by the greater decrease in property crime rates seen in non-expansion states.

To assess whether that difference is statistically significant, below is a difference-in-differences analysis.

```
# DiD analysis
fin_analysis.p <- lm(PROP_PERCAP ~ AFTER_EXP + EXPANSION_STATUS + DiD, data = agg_df)
summary(fin_analysis.p)
```

Using p < 0.05 as a significance level, the interaction term (DiD) was not a significant predictor of property crime rates (p = 0.16751). This indicates that the difference between expansion states and non-expansion states in their property crime decreases is not statistically significant. These data, therefore, do not suggest that Medicaid expansion is a significant predictor of changes in property crime rates.

**Conclusions and Limitations**
The difference-in-differences analyses answer this project’s question in the negative: no, expansion of Medicaid under ACA has probably not affected statewide crime rates in the states that have implemented expansion. Between 2012 and 2016 (i.e., before and after expansion), crime rates did go down by a significant degree within expansion states. This is evidenced by the linear regression that showed time (specifically, before or after expansion) to be a significant predictor for crime rates. But during the same period, crime rates also went down in non-expansion states. The decrease in non-expansion states was greater than that seen in expansion states. But the difference between those two decreases is not significant enough to conclude that Medicaid expansion affected statewide crime rates.

The same answer is true even when subdividing into violent and property crimes. Both expansion and non-expansion states saw a decrease in both violent and property crimes following Medicaid expansion. The decreases in violent crimes were not significant. But in both groups of states, the decreases in property crimes were significant, and that decrease was greater in non-expansion states. However, the difference between those two decreases is not significant enough to conclude that Medicaid expansion affected statewide property crime rates, either. Based on the data used above, therefore, there is insufficient evidence to suggest that Medicaid expansion has significantly affected crime rates, be they total crime, violent crimes only or property crimes only.

Several factors limit this study, including the following examples. The primary limitation of this study is the difficulty of identifying causal factors for crime. The study assumes that poverty is a significant cause of crime—significant enough, at least, for a decrease in healthcare costs among the poor to decrease crime. Many factors cause crime, poverty being at best one of them; it is unclear what all factors those are and what proportionate part each plays in causing statewide crime. The study could much more accurately assess any causal effects of Medicaid expansion on crime if it could control for all other variables that cause crime. The portion of the study that assesses property crimes only is an attempt at doing that (i.e., by assuming that poverty might cause property crimes more than it causes violent crimes) but an incomplete one.

A second limitation is time. The study assumes that any causal effects of healthcare access on crime rates will be visible within two years. That may not be the case.

And a third limitation is the inability to control for confounding variables. E.g., the expansion states may all share non-healthcare-related policies that affected crime rates within the time period studied, policies not shared by non-expansion states. Further assessment of this question would therefore seek to control for as many of such confounding variables as possible.

References


2. Miller S, Wherry LR. Health and Access to Care during the First 2 Years of the ACA Medicaid
