

Investigating the causal relationship between lockdowns and SARS-CoV-2 infection rates in US counties

Abuzar Royesh, Christina Saviddes, Eric Weiner

Table of Content:

Introduction	2
Method Overview and Assumptions	4
Data	4
Experiment Setup	5
Implementation	6
Defining treatment and control	6
Measurement of Outcome Variable	8
Pairwise Matching	8
Results	11
Appendix	13

Introduction

In an effort to curtail the spread of SARS-CoV-2 transmission, many counties across the United States have implemented lockdowns ¹. The focus of this paper is to explore the causal effect of these lockdowns on rates of new cases of SARS-CoV-2 infection. Although many states chose to implement lockdown procedures, tallies of confirmed infections as well as the implementation, execution, and enforcement of lockdowns were often left to county governments ². Thus, we focus our analysis on the data from the 3,141 counties in the US.

The first case of COVID-19 was observed in the US on January 21, 2020 ³. As cases rose in the US, California became the first state to issue a statewide lockdown on March 19, 2020 ⁴. A lockdown (also known as a “stay-at-home” order), is the common term to refer to a large scale physical distancing measure that restricts movement of a population. The World Health Organization (WHO) has advocated for

lockdowns as a means to slow and curb SARS-CoV-2 transmission by limiting contact between people as ample evidence points to SARS-CoV-2 being primarily transmitted through person-to-person contact ⁵. If lockdowns are effective, areas imposing stay-at-home orders should observe a decrease in the rate of infection. This notion is the crux of our analysis.

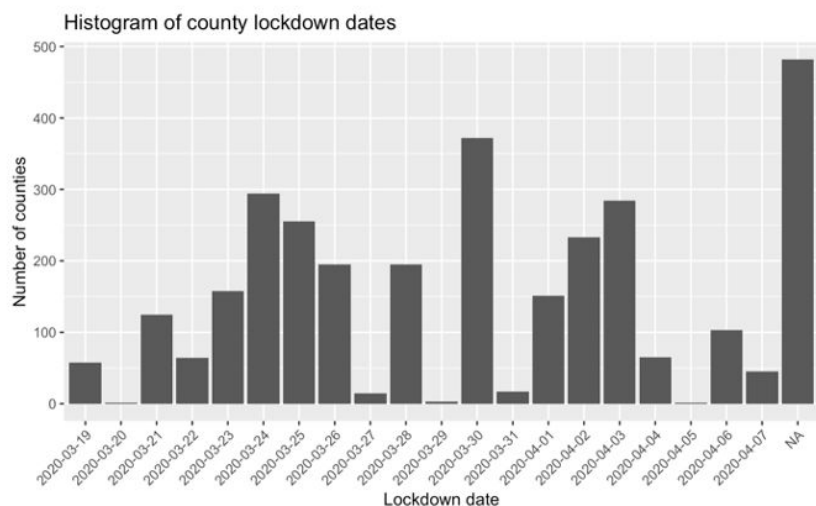


Figure 1: Histogram of Lockdown dates of US Counties

¹ "Three out of four Americans under some form of ... - BBC.com." 31 Mar. 2020, <https://www.bbc.com/news/world-us-canada-52103066>. Accessed 19 Nov. 2020.

² "Lockdowns, closures: How is each US state handling" 14 Apr. 2020, <https://www.aljazeera.com/news/2020/4/14/lockdowns-closures-how-is-each-us-state-handling-coronavirus>. Accessed 19 Nov. 2020.

³ "First Travel-related Case of 2019 Novel Coronavirus Detected" 21 Jan. 2020, <https://www.cdc.gov/media/releases/2020/p0121-novel-coronavirus-travel-case.html>. Accessed 19 Nov. 2020.

⁴ "California becomes first state to order lockdown - KSLA.com." 20 Mar. 2020, <https://www.ksla.com/2020/03/20/california-becomes-first-state-order-lockdown/>. Accessed 19 Nov. 2020.

⁵ "Physical distancing, face masks, and eye ... - The Lancet." [https://www.thelancet.com/journals/lancet/article/PIIS0140-6736\(20\)31142-9/fulltext](https://www.thelancet.com/journals/lancet/article/PIIS0140-6736(20)31142-9/fulltext). Accessed 19 Nov. 2020.

As infections surge nationwide, evaluating the efficacy of lockdowns and their ability to curb SARS-CoV-2 infection has important policy implications. Preliminary analyses have shown the effect of restrictions, ranging from lockdowns to mask mandates, on curbing the spread of the virus ⁶. However, some critics have pointed to the adverse effects of lockdowns on mental health and the economy ^{7 8}. Furthermore, some have criticized lockdowns as a blunt tool to control the virus and point to other regions and nations (e.g. New Zealand, Taiwan, and South Korea) that have been able to curb infection without widespread lockdowns ⁹. Unfortunately for the US, lack of widespread contact tracing and lax enforcement of stay at home orders makes solutions that worked in other regions infeasible domestically.

We build upon previous literature that investigates correlations and causal relationships between lockdowns and infection rates. Recent academic studies investigating the effectiveness of lockdowns have indicated that they indeed are effective in curbing the spread of the virus. For example, Lau et al. found that the doubling time of the virus (time for the number of cases to double) increased from two to four days after a lockdown measure implemented in the Hubei province of China ¹⁰. A secondary analysis conducted by Fang et al. found, using counterfactual simulations, that without the late-January lockdown in Wuhan, cases would have been 64.81% higher in the 347 cities outside of the Hubei province and 52.64% higher in the non-Wuhan cities within the Hubei province ¹¹.

Given the numerous factors affecting virus infections, it is important to be cognizant of the assumptions and limitations of causal studies. For example, Gupta et al. have found that lockdowns are often introduced as a response to rising infection levels ¹². Thus, even small differences in

⁶ "See How Coronavirus Restrictions Compare to Case Counts" 18 Nov. 2020, <https://www.nytimes.com/interactive/2020/11/18/us/covid-state-restrictions.html>. Accessed 19 Nov. 2020.

⁷ "Impact of COVID-19 and lockdown on mental ... - NCBI - NIH." 24 Aug. 2020, <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7444649/>. Accessed 19 Nov. 2020.

⁸ "The Great Lockdown: Worst Economic Downturn ... - IMF Blog." <https://blogs.imf.org/2020/04/14/the-great-lockdown-worst-economic-downturn-since-the-great-depression/>. Accessed 19 Nov. 2020.

⁹ "Emerging COVID-19 success story: South Korea learned the" 30 Jun. 2020, <https://ourworldindata.org/covid-exemplar-south-korea>. Accessed 19 Nov. 2020.

¹⁰ "positive impact of lockdown in Wuhan on containing the" 17 Mar. 2020, <https://academic.oup.com/jtm/article/27/3/taaa037/5808003>. Accessed 19 Nov. 2020.

¹¹ "Human Mobility Restrictions and the Spread of the Novel" 27 Mar. 2020, <https://www.nber.org/papers/w26906>. Accessed 19 Nov. 2020.

¹² "Tracking Public and Private Responses to" 23 Apr. 2020, <https://www.nber.org/papers/w27027>. Accessed 19 Nov. 2020.

pre-lockdown infection rates can result in considerable differences post-lockdown resulting in reverse causality. We detail later in the methods section how we account for this by forcing pre-treatment matching on infection rates.

In addition to the factor above, there are numerous other confounders that make causal investigation nuanced and complex. For example, packaged policies (e.g. social distancing *and* mask mandates) make attributing effectiveness to anyone policy difficult through lack of identifiability. Another complexity is that citizens voluntarily may take precautions before an order is imposed. This may introduce a larger bias and invalidate some assumptions around the lockdown being a well-defined treatment. There may also be spillovers from region to region. For example, lockdowns in New York City may have affected virus prevalence in surrounding regions or major air travel hubs that connect frequently to NYC resulting in the doubt of the validity of SUTVA¹³. Finally, variation in the timing of the policy may result in a bias as well. When treatments are staggered and effects change over time (as is often the case with regional lockdowns), it can often result in bias. In the following section, we detail our methodology and how we get around some of these issues¹⁴.

Method Overview and Assumptions

Data

In order to elucidate the causal effects of lockdown on growth rate of SARS-CoV-2 infections, we have assembled a dataset that contains key demographic information about all counties in the United States, number of new SARS-CoV-2 cases per day, and date of lockdown (if the county was subject to lockdown). The following table lists the variables and sources from which we collected the data.

¹³ "Using Difference-in-Differences to Identify Causal Effects of" https://cdn.vanderbilt.edu/vu-my/wp-content/uploads/sites/2318/2020/05/11154933/Covid-DD_v2.pdf. Accessed 19 Nov. 2020.

¹⁴ "Difference-in-Differences with Variation in Treatment Timing" 6 Sep. 2018, <https://www.nber.org/papers/w25018>. Accessed 19 Nov. 2020.

Variable category	Metrics	Source	Date
COVID-19	Confirmed cases, deaths	https://github.com/CSSEGISandData/COVID-19/tree/master/csse_covid_19_data/csse_covid_19_daily_reports	Oct 20, 2020
Demographics	Total_population, race, gender, age	CDC Wonder	2019
Income & urban	Urban/rural, overall poverty, poverty for specific age groups, median household income	U.S. Census Bureau	2018
Education	Level of education	U.S. Census Bureau	2014-18
Employment	Unemployment, labor force	U.S. Census Bureau	2019
Health	Life expectancy, Physical activity, obesity	http://www.healthdata.org/us-health/data-download	2010-2011
Health	Diabetes	Institute for Health Metrics and Evaluation (IHME)	2012
Lockdown data	County level lockdown dates	Kaggle	2020

Table 2: Data sources used for experiment

Experiment Setup

In the absence of any pre-treatment randomization mechanism, we designed a study combining pairwise matching and difference-in-differences analyses in an attempt to isolate the causal impact of lockdowns on the number of new cases of SARS-CoV-2 infection. Through matching similar counties that did and did not undergo lockdown at specific time intervals and using difference-in-differences, we hoped to reduce any inherent differences between the counties in treatment and in control. For each matched pair, our estimator was as follow:

For each pair:

$$\hat{\tau}^{DiD} = \Delta \bar{y}_{treatment} - \Delta \bar{y}_{control}$$

Where

$$\Delta \bar{y}_{treatment} = \bar{y}_{treatment}^{post} - \bar{y}_{treatment}^{pre}$$

$$\Delta \bar{y}_{control} = \bar{y}_{control}^{post} - \bar{y}_{control}^{pre}$$

$z \in \{\text{treatment, control}\}$

$$y_z^{pre} = \frac{y_{t-4} + \dots + y_t}{population}$$

$$y_z^{post} = \frac{y_{t+10} + \dots + y_{t+14}}{population}$$

Where y_t is the number of new cases on March 23, 2020, and *population* is the population of the county.

We proceeded with a Fisher's test to test for the null hypothesis that our diff-in-diff estimator was zero. For robustness check, we also ran a Wilcox test. Each of these steps will be explained in further detail in the next sections.

Implementation

Defining treatment and control

Throughout the course of 2020, 45 states issued stay-at-home orders. Of the states that had partial or no lockdown, the states that did not undergo lockdown had disproportionately less dense populations and other key demographic differences.

For instance, as can be seen in Figure 1 below, the states that had lockdowns and those that did not had big differences in urban versus rural, education, unemployment, diabetes, racial makeup, and levels of poverty.

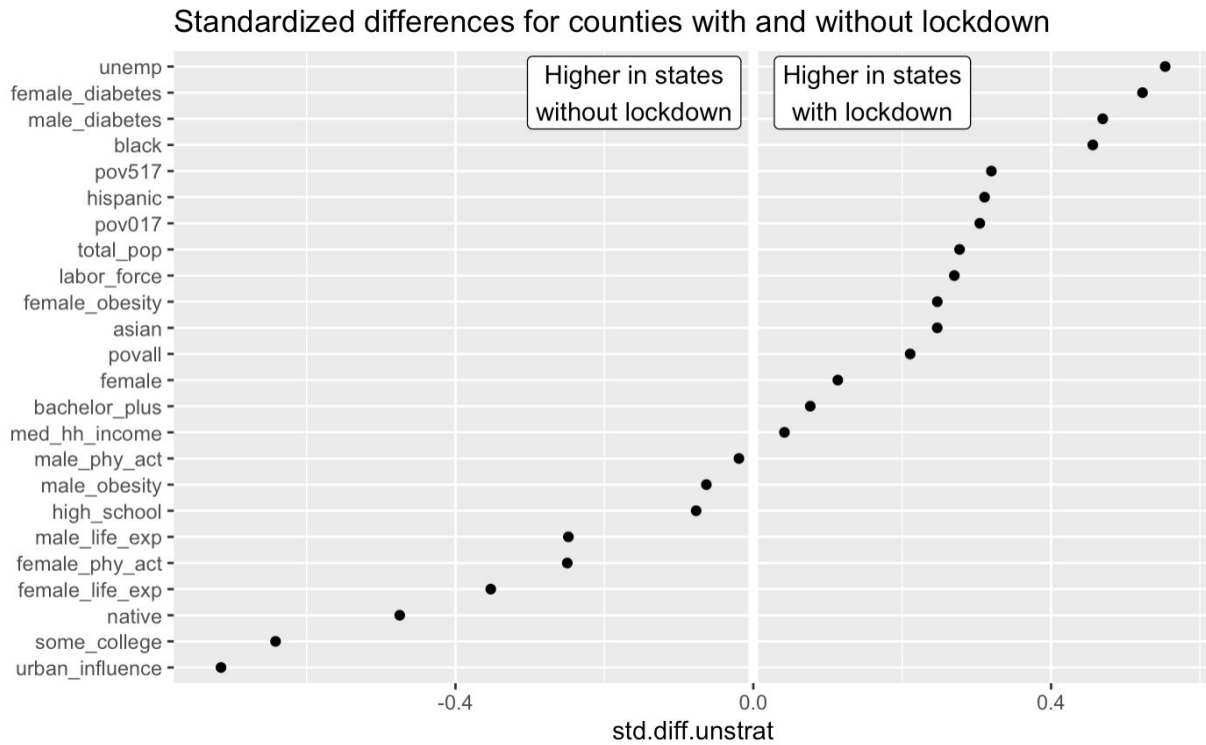


Figure 2: Standardized differences for counties with and without lockdowns

Given that most counties in the United States had some kind of lockdown procedure, in order to have a reasonable “no treatment” arm to compare, we chose to call the “treatment group” counties that had lockdown before March 24th, and the “no treatment group” counties that did not have lockdowns by April 3rd, 2020. We selected the cutoff of March 24th for counties to be assigned to treatment because a diverse set of counties underwent lockdown by this date. Between March 19th and 24th, 406 counties were in lockdown in the US.

It is well understood that the impacts of lockdown on transmission are not immediately obvious due to the incubation period of SARS-CoV-2. Infected individuals would not immediately test positive for the virus given the limitations of current testing methods; therefore decreases in transmission would not manifest for a few days. The incubation period of SARS-CoV-2 is on average 4.54 days, with a range of 2 to 14 days¹⁵. Therefore the effects of lockdown on transmission and number of new cases would not be observed until 5+ days after lockdown began. It is also worth noting that even after beginning to exhibit symptoms, most individuals who have been infected with SARS-CoV-2 wait a number of days before seeking medical care. While studying COVID-19, Zhang

¹⁵ Incubation Period of Severe Acute Respiratory Syndrome Novel Coronavirus 2 that Causes Coronavirus Disease 2019: A Systematic Review and Meta-Analysis

and colleagues found that an average of 2.5 days elapsed between symptom onset and first healthcare consultation ¹⁶. In the absence of routine population wide screening and testing, the date which an individual seeks medical care is also the earliest date which they would be tested for the virus.

In order to allow for sufficient time to observe the impact of lockdown on rates of transmission, we chose to compare the five day average of new cases on March 20th to 24th to the five day average of new cases on April 3rd to April 7th. We propose that a 10 day separation between these two dates should allow for sufficient time to begin to observe the impacts of lockdown. Counties that underwent lockdown on or after April 3rd were assigned to the no-treatment group because we assume that any impact of lockdown on the average number of new cases would not be observed for at least 5 days after lockdown began. Counties that underwent lockdown between March 25th and April 2nd were not considered in this analysis. After running this, our dataset included 406 treatment and 722 control counties.

We chose to keep comparison dates consistent regardless of the exact date of lockdown for several reasons. Given news is often broadcast at a national level, it is easy to assume that in addition to county wide rulings, people's behavior and social distancing practices are also impacted by national messaging through entities like the CDC, White House Briefings, etc. By not time-shifting the data, we are controlling for this national messaging. Additionally, a large limitation of relying on caseload data is that the number of positive cases is related to the amount of testing being conducted. Day of the week, proximity to holidays, and other factors may influence when individuals choose to get tested. For example, more individuals may get tested on weekends to avoid taking time off from work. Alternatively, proximity to a holiday like Easter may spur more individuals to get tested if they are afraid of spreading their illness to family members, or perhaps labs disproportionately process tests on certain days of the week. Keeping the dates of comparison consistent helps mitigate these fluctuations in testing. An important assumption in our study is that access to testing, and propensity to seek out testing, is the same among comparison counties. In other words, we are assuming differences in rates of change of SARS-CoV-2 cases are not due to changes in testing capacity, accessibility, or frequency.

Measurement of Outcome Variable

In order to circumvent biases that arise due to difficulties in accessing PCR testing, entities like the New York Times have proposed that death is a better metric to understand the severity of an

¹⁶ Zhang J, Litvinova M, Wang W, Wang Y, Deng X, Chen X, et al. Evolving epidemiology of novel coronavirus diseases 2019 and possible interruption of local transmission outside Hubei Province in China: a descriptive and modeling study. medRxiv. 2020 Feb 23;

outbreak¹⁷. However, while SARS-CoV-2 is a particularly infectious virus, it is not a particularly lethal virus. With a mortality rate of less than 2%, the county wide mortality data would not offer enough variance to yield insight. Additionally, delays between infection, hospitalization, and eventually death can cause a significant delay before seeing the effects of an intervention on changes in mortality rate. To circumvent these issues, and obtain a more immediate proxy for the efficacy of the intervention, we chose to look at the number of new cases per capita.

Pairwise Matching

Even after choosing a cutoff of March 24, there were still systematic differences between the treatment and control. Running a standardized differences for stratified comparisons test between the counties in the treatment and control groups showed a large imbalance in the covariates (Figure 3). This indicated that counties that had ordered lockdowns and those that did not were fundamentally different from one another along these variables. For the full table of results see Table 1 in Appendix. An analysis of the propensity scores also corroborated this finding. As shown on Figure 4 below, the treatment status of counties could be predicted very accurately using the remaining variables in the dataset.

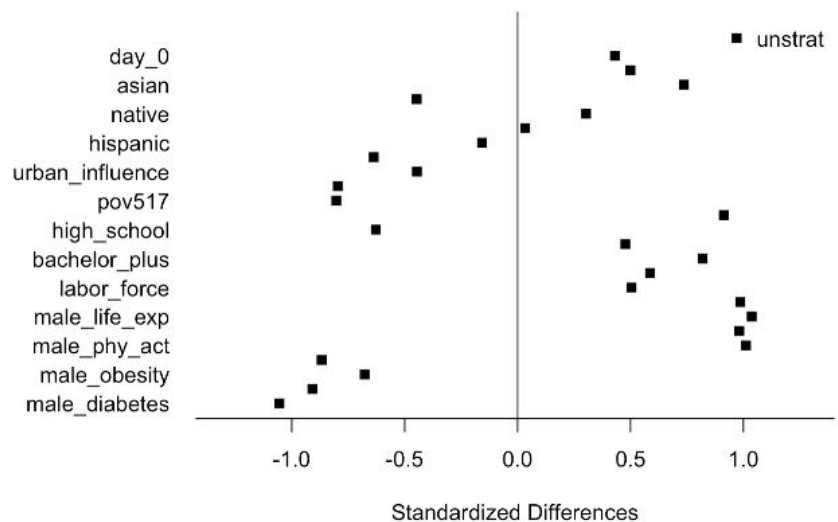


Figure 3: Standardized Differences for Unstratified Comparisons

¹⁷ "Monitoring the Coronavirus Outbreak in Metro Areas Across" 23 Apr. 2020, <https://www.nytimes.com/interactive/2020/04/23/upshot/five-ways-to-monitor-coronavirus-outbreak-us.html>. Accessed 20 Nov. 2020.

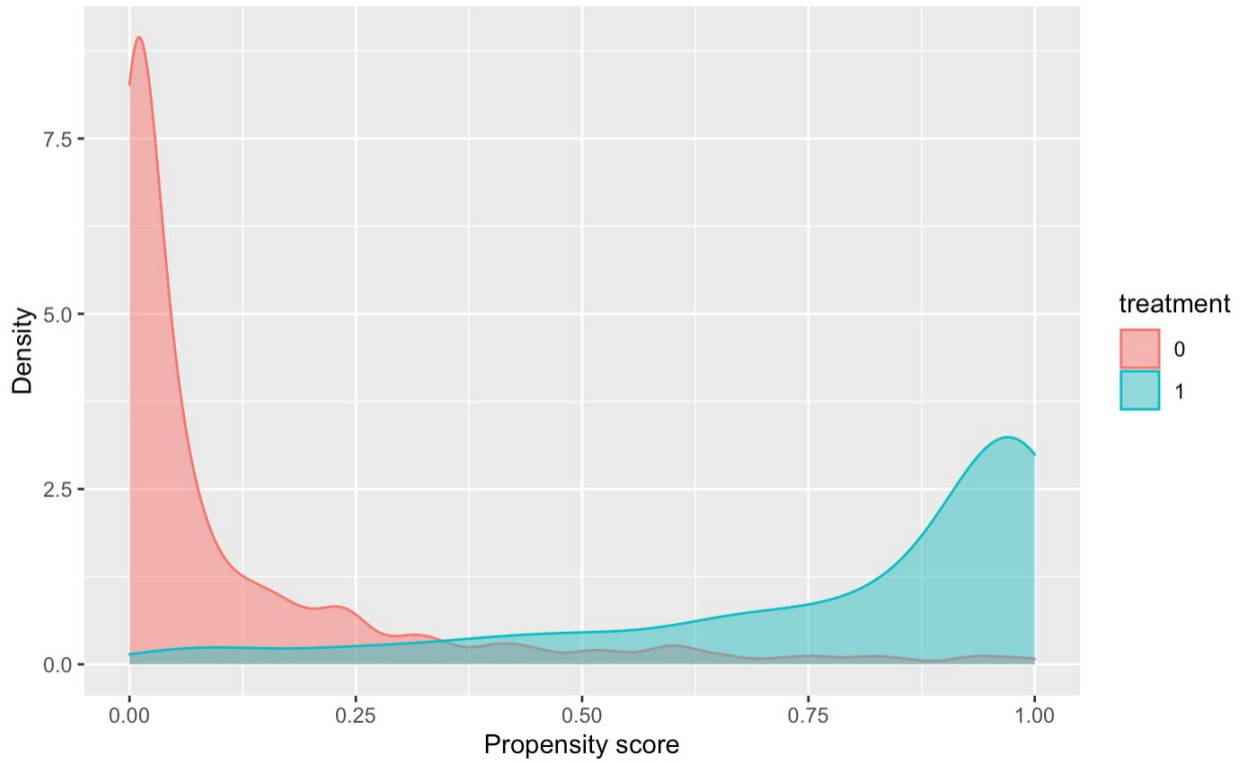


Figure 4: Propensity Scores for treatment and control pairs

To address this problem, we decided to run a pairwise matching to match each county in the treatment group with a similar county in the control group. We forced matching based on propensity scoring, using a caliper of 0.1. This process matched the 406 counties in the treatment to their closest counterpart in the control group.

Even though pairwise matching helped balance the two groups, the process of matching did not entirely fix the initial imbalance between the two groups, as can be seen in

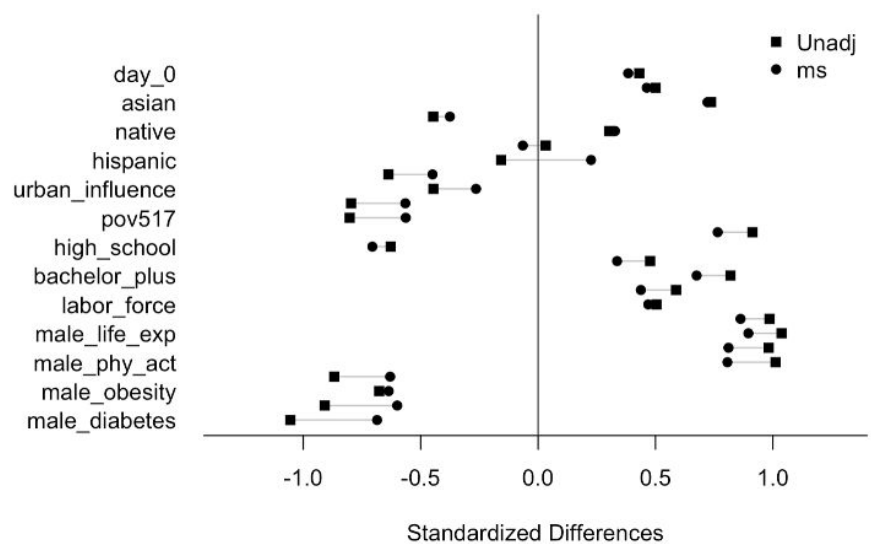


Figure 2: Standardized Differences of Matched Groups

Figure 5. In particular, it did not improve the variable day_0, which contained the number of cases per capita for all counties on March 24th, which we believed was a major problem. In particular, if a county had more per capita cases on March 24th, the number of cases was likely to grow much faster than a county that had few cases or had not been exposed to the virus at all. We therefore filtered out any treatment-control matches that had a large difference (equal to or larger than 0.00001) in the number of per capita cases on March 24th. This left us with 168 pairs, which are plotted below.

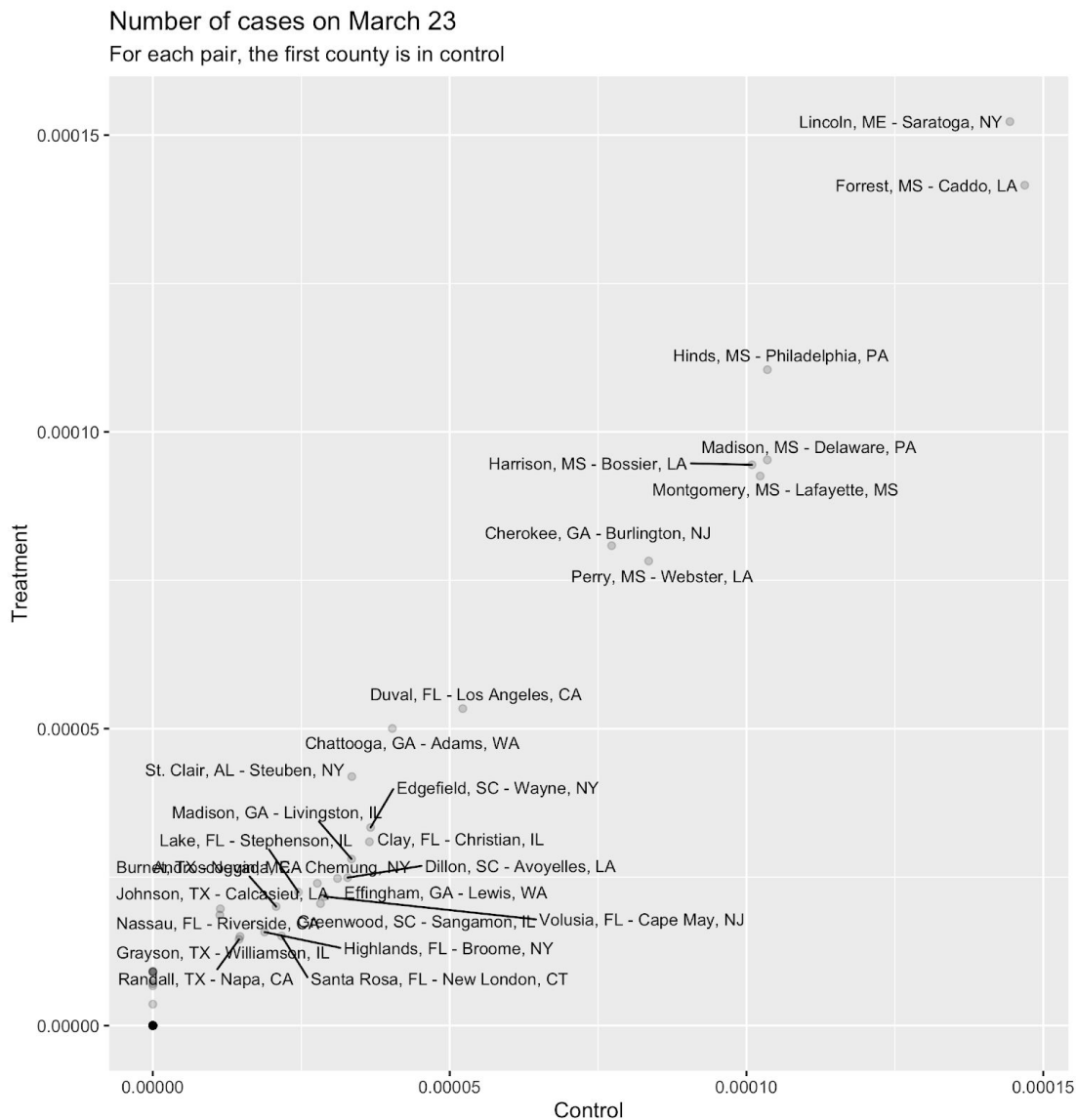


Figure 6: Filtered pairs used for analysis

Results

We performed a Fisher Randomization Test on the change in new cases per capita between treated and untreated matched pairs, and found that that lockdown was associated with an increased average number of new cases, with a p-value of 1.26×10^{-9} . This allows us to reject the sharp null hypothesis that there was no treatment effect.

Pvalue	Deviate	Statistic	Expectation	Variance
1.261061e-09	5.960017	1204.228	152.0425	31166.65

We decided to explore this interesting phenomenon further and investigate the differences in the distributions of the treated pairs and the lockdown pairs. Below is a histogram of the distributions of change of average number of new cases before and after lockdown. When analyzing the distributions, we see some right-skewness in the treatment, or lockdown, group. To further dive into the populations and pairs, we ran a Wilcoxon rank sum test with a continuity correction. Specifically, we wanted to test if the distribution of the control pairs were identical to the treatment pairs. With an extremely low p-value of 1.55×10^{-7} we reject the null and confirm that they are non-identical distributions.

Overall, while the Pre-lockdown average for the counties in control and treatment were slightly different ($1.2\text{e-}06$ and $1.4\text{e-}06$, respectively), the difference between the two groups increased for the post-lockdown dates ($2.07\text{e-}05$ and $3.41\text{e-}05$, respectively). This means that the pre- and post-lockdown difference in the average number of new cases per capita for the control group was $1.95\text{e-}05$, while the number was $3.27\text{e-}05$ for the treatment group. The above three findings, taken together, would suggest that compared to the control counties, the prevalence of the virus grew faster in counties that had lockdowns.

	N	Pre-lockdown average	Post-lockdown average	Delta
Counties in Control	168	$1.2\text{e-}06$	$2.07\text{e-}05$	$1.95\text{e-}05$
Counties in Treatment	168	$1.4\text{e-}06$	$3.41\text{e-}05$	$3.27\text{e-}05$

Table 3: Results

Discussion

While paradoxical to our intuition, the observed results that lockdown was associated with an increased average number of new cases, may stem from a multitude of reasons, two of which are based on our assumption of when we thought the effects of lockdown would become apparent and our assumption that rates of testing were consistent between counties. While we assumed the impact of lockdown would become apparent after 10 days, other studies suggest the peak of the infection curve occurs anywhere between 14-18 days after the lockdown ¹⁸. If we were to conduct this analysis again, it would be interesting to see whether measuring the average number of new cases on days 20-25 yielded different trends. Given that lockdowns are often responses to sharply rising caseloads, it follows to reason that if we measured the change in cases too soon after lockdown was implemented, we would observe higher caseloads rising in counties in lockdown. Perhaps in addition to matching on the number of cases per population we also should have matched on change in caseload as well to control for this possibility.

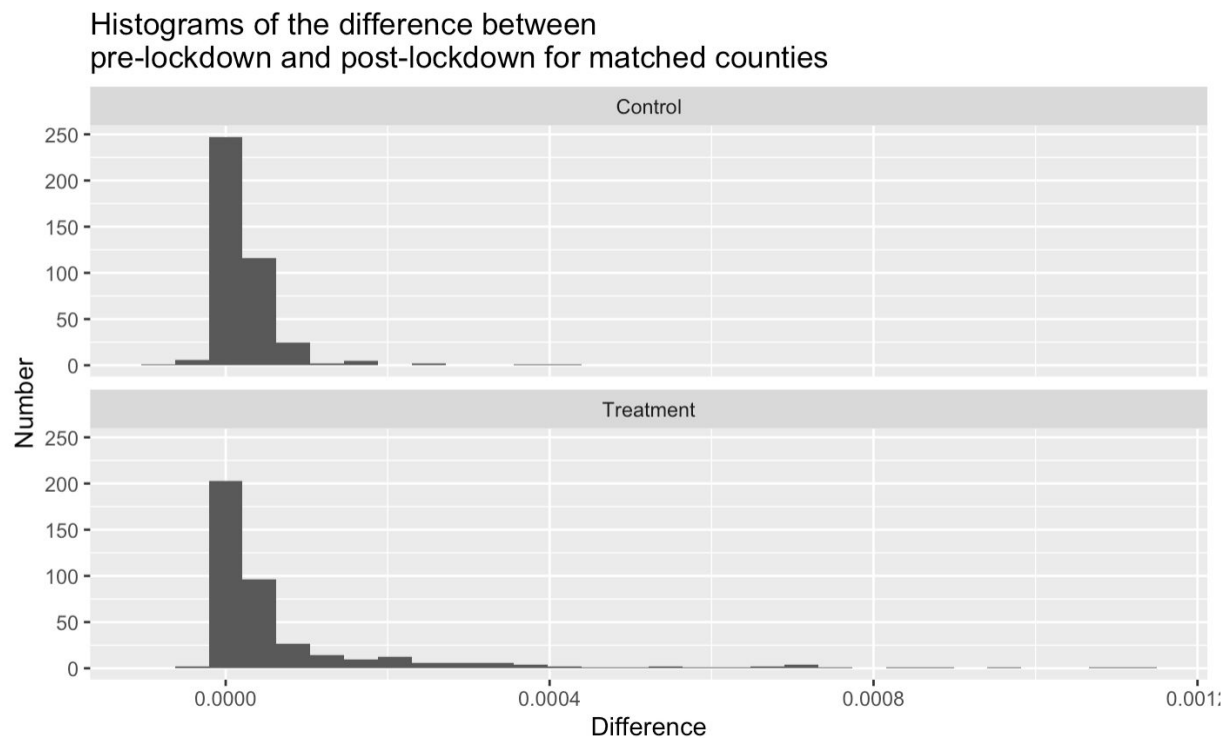


Figure 7: Treatment and Control distributions of matched counties

¹⁸ "Lockdown timing and efficacy in controlling ... - The Lancet." 13 Jul. 2020, [https://www.thelancet.com/journals/eclinm/article/PIIS2589-5370\(20\)30201-7/fulltext](https://www.thelancet.com/journals/eclinm/article/PIIS2589-5370(20)30201-7/fulltext). Accessed 19 Nov. 2020.

Another possibility would be that the increased average caseload for the treatment group stems from the outliers in the data. As Figure 7 above shows, while most of the histogram for the two groups is similar, counties in treatment had more outliers, where the difference between the post- and pre-lockdown prevalence rates are higher.

Another important assumption that we made which may impact the validity of our results is our assumption that testing was consistent between counties. That is to say, we assumed that an individual in either of our paired counties would have equal access and incentive to undergo testing. However early on in the pandemic, while testing was limited, this was likely not the case. Early problems with the primers delayed the CDCs ability to distribute testing supplies, and counties were left scrambling to implement testing on their own. Counties had very different requirements to be eligible for testing, and this makes it difficult to compare caseload data from county to county.

Decisions on what units to match are incredibly important to the validity of causal inferences. This project was a rewarding opportunity to utilize the tools we have learned throughout the quarter to ask questions about our world during this global pandemic. Learning how to ask meaningful questions from real world data and grappling with the difficulties and challenges associated with these types of analyzes was a wonderful opportunity to test our comprehension.

Appendix

Table 1: Full results of stratified and unstratified differences

	std.diff.unstrat	z.unstrat	p.unstrat
day_0	0.4308689	6.8046368	0
total_pop	0.4998142	7.8380232	0
asian	0.736522	11.1980309	0
black	-0.4466092	-7.0424662	0
native	0.302653	4.8302785	0.0000014
female	0.032844	0.5296283	0.5963697
hispanic	-0.1573792	-2.5309261	0.0113762
povall	-0.6374335	-9.8297112	0
urban_influence	-0.4455404	-7.0263517	0
pov017	-0.7961103	-11.9922083	0
pov517	-0.8024781	-12.0757615	0
med_hh_income	0.9129399	13.4837129	0
high_school	-0.6273695	-9.6875356	0
some_college	0.4766015	7.4925559	0
bachelor_plus	0.8193937	12.2964635	0
unemp	0.5868711	9.1096915	0
labor_force	0.50405	7.900784	0
female_life_exp	0.9852269	14.3618938	0
male_life_exp	1.0366555	14.9655858	0
female_phy_act	0.9818781	14.321974	0
male_phy_act	1.0109469	14.6660047	0
female_obesity	-0.8683949	-12.925447	0

male_obesity	-0.6769418	-10.3822354	0
female_diabetes	-0.9084418	-13.4279287	0
male_diabetes	-1.0547767	-15.1741095	0

Appendix-Code

Nov 20, 2020

Reading in libraries

```
library(tidyverse)
library(lubridate)
library(RITools)
library(DOS2)
library(optmatch)
source("utility.R")

covid_cases_file <- "covid_cases_time_series.csv"
lockdown_file <- "lockdown_us.csv"
county_file <- "covid_cleaned.csv"
```

Cleaning case data

```
# cases <-
#   read_csv("time_series_covid19_confirmed_US.csv") %>%
#   mutate(fips = as.integer(FIPS)) %>%
#   filter(!str_detect(Admin2, "Unassigned/Out of")) %>%
#   filter(!is.na(fips)) %>%
#   filter(!Province_State == "Puerto Rico") %>%
#   select(
#     fips, everything(), -UID, -iso2, -iso3, -code3, -Province_State,
#     -Country_Region, -Lat, -Long_, -Combined_Key, -FIPS,
#   ) %>%
#   gather(key = "date", value = "cases", -Admin2, -fips) %>%
#   arrange(fips) %>%
#   # mutate(date = as.Date(date)) %>%
#   separate(date, into = c("month", "day", "year"), sep = "/") %>%
#   mutate(
#     year = str_c("20", year),
#     date = make_date(year, month, day)
#   ) %>%
#   select(fips, county_name = Admin2, date, cases) %>%
#   filter(
#     date >= as.Date("2020-03-10"),
#     date <= as.Date("2020-04-15"),
#   )
#
# cases %>%
#   write_csv("covid_cases_time_series.csv")
```

Reading in cleaned case data

```
cases <-
  read_csv(covid_cases_file) %>%
  group_by(fips, county_name) %>%
  mutate(new_cases = cases - lag(cases)) %>%
  ungroup()

##
## -- Column specification -----
## cols(
##   fips = col_double(),
##   county_name = col_character(),
##   date = col_date(format = ""),
##   cases = col_double()
## )
```

Getting number of cases on day zero and average pre and post cases

```
day_0 <-
  cases %>%
  filter(date == as.Date("2020-03-23")) %>%
  select(fips, day_0 = cases)

pre_post <-
  cases %>%
  mutate(
    days =
      case_when(
        date >= as.Date("2020-03-19") & date <= as.Date("2020-03-23") ~ "pre_lockdown",
        date >= as.Date("2020-04-02") & date <= as.Date("2020-04-06") ~ "post_lockdown",
      )
  ) %>%
  filter(!is.na(days)) %>%
  group_by(fips, days) %>%
  summarize(new_cases = mean(new_cases)) %>%
  spread(key = days, value = new_cases)

## `summarise()` regrouping output by 'fips' (override with `groups` argument)
```

Reading in lockdown data

```
state_lockdowns <-
  read_csv(lockdown_file) %>%
  filter(is.na(County)) %>%
  select(-Country, -County, -Type, -X6) %>%
  rename(
    state = State,
    lockdown_date_2 = Date
  )

## Warning: Missing column names filled in: 'X6' [6]

county_lockdowns <-
  read_csv(lockdown_file) %>%
```

```

filter(!is.na(County)) %>%
select(-Country, -Type, -X6) %>%
rename(
  county_name = County,
  state = State,
  lockdown_date_1 = Date
)

```

Warning: Missing column names filled in: 'X6' [6]

County level characteristics

```

data <-
  read_csv(county_file) %>%
  left_join(state_lockdowns, by = "state") %>%
  left_join(county_lockdowns, by = c("state", "county_name")) %>%
  mutate(
    lockdown =
      case_when(
        lockdown_date_1 <= lockdown_date_2 ~ lockdown_date_1,
        is.na(lockdown_date_1) & !is.na(lockdown_date_2) ~ lockdown_date_2,
        !is.na(lockdown_date_1) & is.na(lockdown_date_2) ~ lockdown_date_1
      )
  ) %>%
  separate(lockdown, into = c("month", "day", "year"), sep = "/") %>%
  mutate(
    year = str_c("20", year),
    lockdown = make_date(year, month, day),
    treatment =
      case_when(
        lockdown <= as.Date("2020-03-23") ~ 1,
        lockdown >= as.Date("2020-04-02") ~ 0
      )
  ) %>%
  select(
    treatment, lockdown, everything(), -lockdown_date_1, -lockdown_date_2,
    -confirmed, -deaths
  )

```

```

##
## -- Column specification -----
## cols(
##   .default = col_double(),
##   county_name = col_character(),
##   state = col_character()
## )
## i Use `spec()` for the full column specifications.

```

Histogram of date of lockdowns

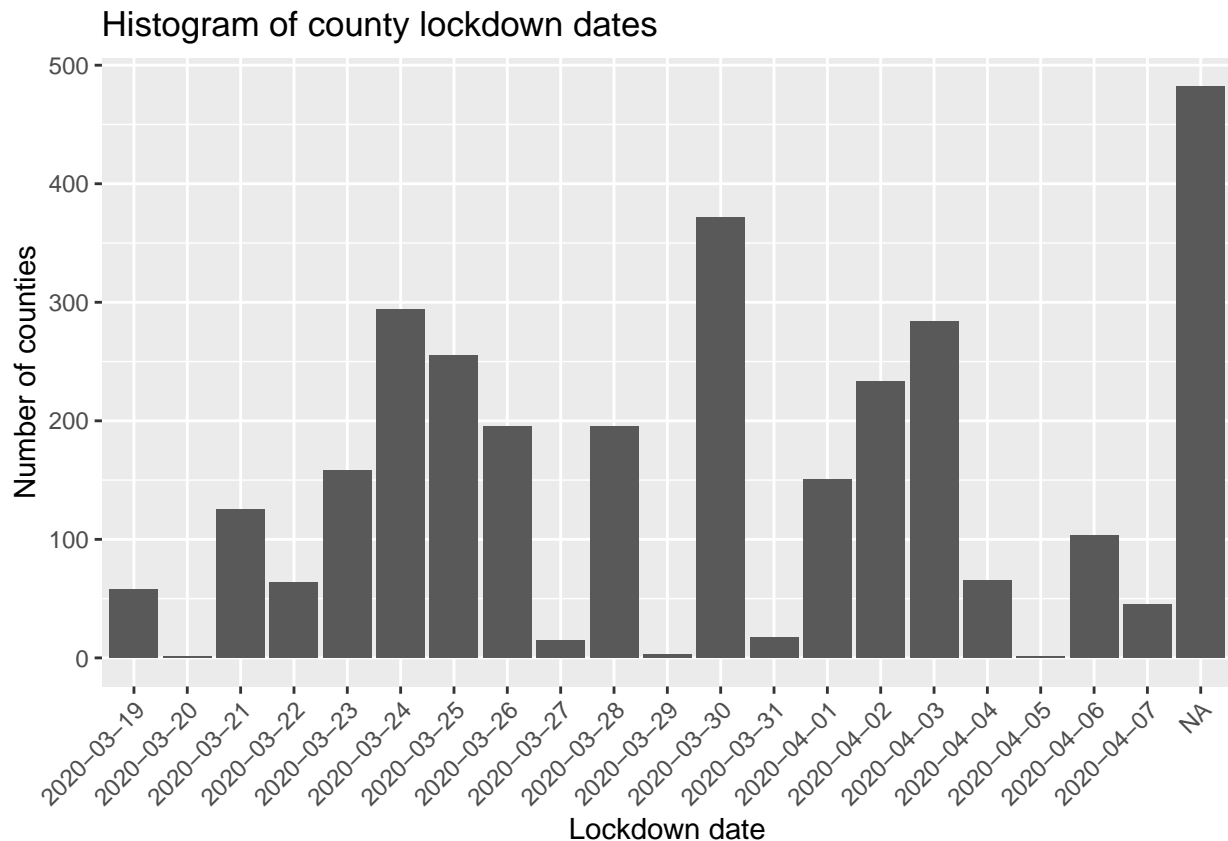
```

data %>%
  mutate(lockdown = as.character(lockdown)) %>%
  ggplot(aes(lockdown)) +

```

```
geom_histogram(stat = "count") +
  theme(
    axis.text.x = element_text(angle = 45, hjust = 1)
  ) +
  labs(
    x = "Lockdown date",
    y = "Number of counties",
    title = "Histogram of county lockdown dates"
  )
```

Warning: Ignoring unknown parameters: binwidth, bins, pad



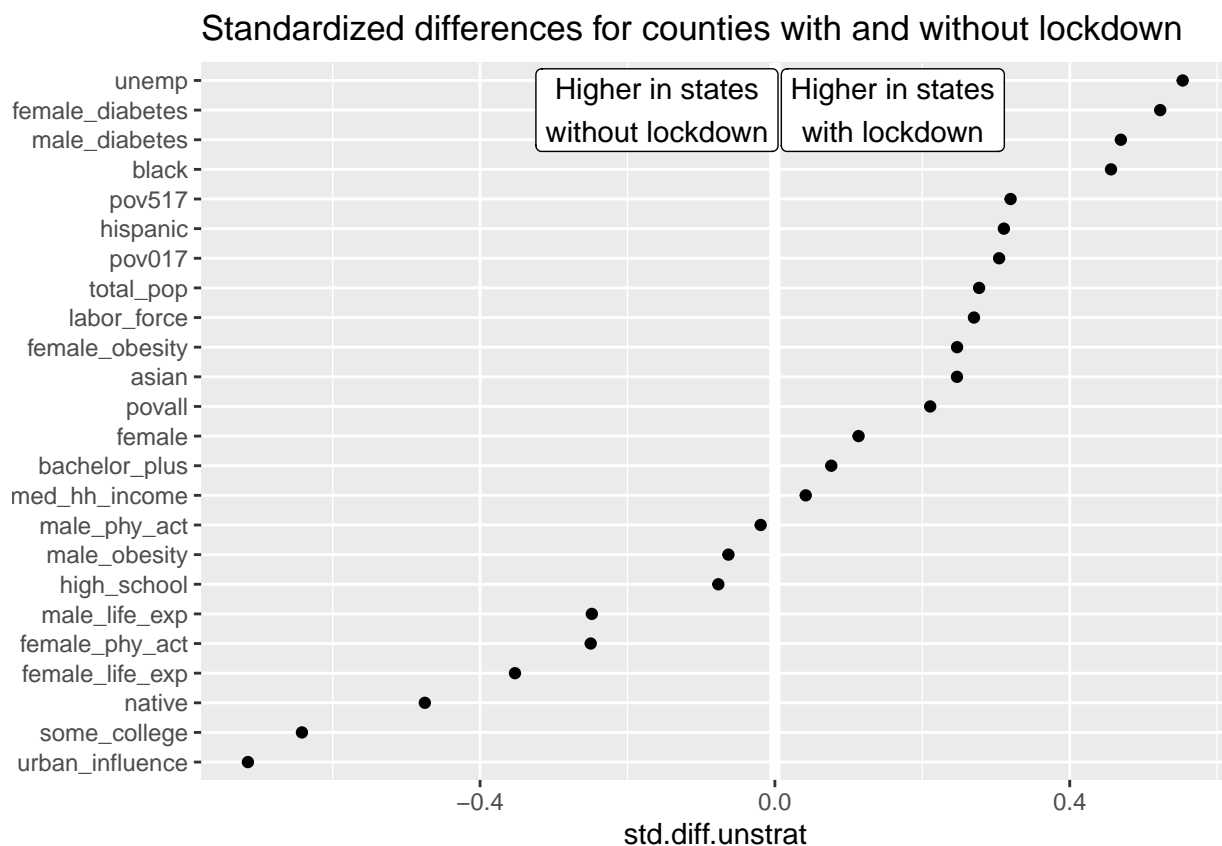
Plotting difference between states with lockdown and those without

```
test_df <-
  data %>%
  mutate(treatment = if_else(is.na(lockdown), 0, 1)) %>%
  select(
    -c(
      lockdown, fips, county_name, state, month, day, year,
      starts_with(c("age", "mort"))
    )
  )

test_res <- xBalance(treatment ~ ., data = test_df)

as_tibble(test_res$results) %>%
  cbind(variable = names(test_df %>% select(-treatment))) %>%
```

```
mutate(variable = fct_reorder(variable, std.diff.unstrat)) %>%
ggplot(aes(std.diff.unstrat, variable)) +
geom_vline(xintercept = 0, color = "white", size = 2) +
geom_point() +
annotate(
  "label",
  label = c("Higher in states\nwith lockdown", "Higher in states\nwithout lockdown"),
  x = c(0.16, -0.16),
  y = 23,
) +
labs(
  y = NULL,
  title = "Standardized differences for counties with and without lockdown"
)
```



Merging and final cleaning of the data

```
merged <-
data %>%
filter(!is.na(treatment)) %>%
left_join(pre_post, by = "fips") %>%
left_join(day_0, by = "fips") %>%
filter(post_lockdown >= 0 & pre_lockdown >= 0) %>% # dropping counties with negative averages
mutate_at(
  vars(day_0, pre_lockdown, post_lockdown),
  ~ (. / total_pop)
```

```

    ## 100000 #normalizing by population and scaling up for readability
  ) %>%
  select(
    day_0, pre_lockdown, post_lockdown, everything(),
    -month, -day, -year, -starts_with(c("age", "mort"))
  ) %>%
  filter_all(~ !is.na(.)) %>%
  mutate(y = post_lockdown - pre_lockdown) %>%
  rename(z = treatment)

merged %>%
  count(lockdown)

```

```

## # A tibble: 11 x 2
##   lockdown      n
##   <date>    <int>
## 1 2020-03-19    58
## 2 2020-03-20     1
## 3 2020-03-21   125
## 4 2020-03-22    64
## 5 2020-03-23   158
## 6 2020-04-02   231
## 7 2020-04-03   282
## 8 2020-04-04    64
## 9 2020-04-05     1
## 10 2020-04-06    99
## 11 2020-04-07    45

```

Checking balance between treatment and control groups

```

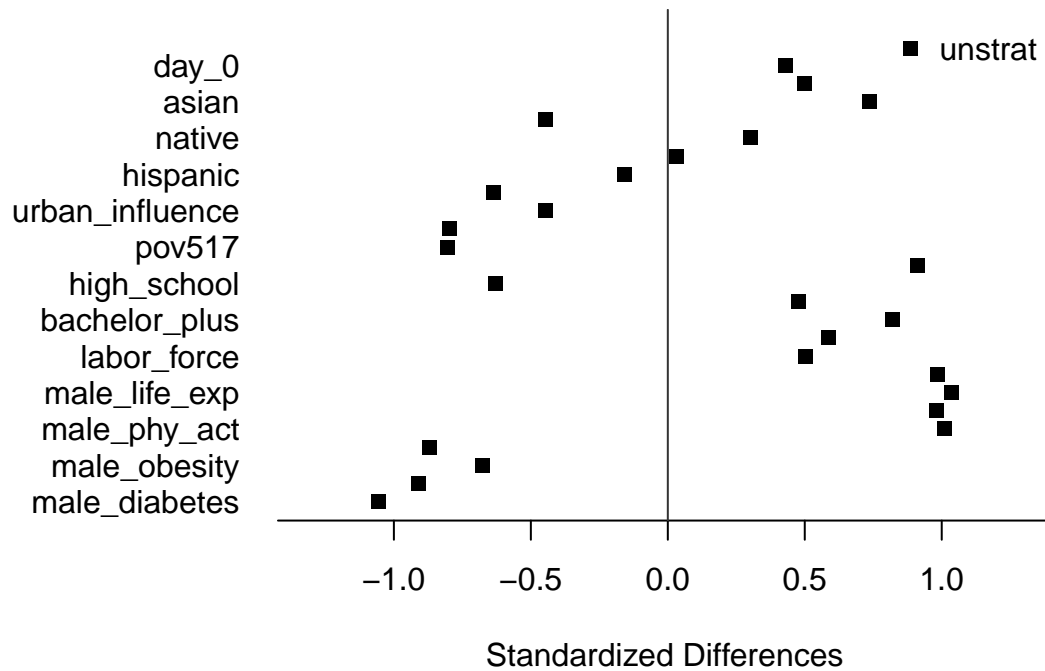
temp <-
  merged %>%
  select(-c(pre_lockdown, post_lockdown, y, fips, county_name, lockdown, state))

xBalance(z ~ ., data = temp) %>%
  knitr::kable()

plot(xBalance(z ~ ., data = temp))

```

	std.diff.unstrat	z.unstrat	p.unstrat
day_0	0.4308689	6.8046368	0.0000000
total_pop	0.4998142	7.8380232	0.0000000
asian	0.7365220	11.1980309	0.0000000
black	-0.4466092	-7.0424662	0.0000000
native	0.3026530	4.8302785	0.0000014
female	0.0328440	0.5296283	0.5963697
hispanic	-0.1573792	-2.5309261	0.0113762
povall	-0.6374335	-9.8297112	0.0000000
urban_influence	-0.4455404	-7.0263517	0.0000000
pov017	-0.7961103	-11.9922083	0.0000000
pov517	-0.8024781	-12.0757615	0.0000000
med_hh_income	0.9129399	13.4837129	0.0000000
high_school	-0.6273695	-9.6875356	0.0000000
some_college	0.4766015	7.4925559	0.0000000
bachelor_plus	0.8193937	12.2964635	0.0000000
unemp	0.5868711	9.1096915	0.0000000
labor_force	0.5040500	7.9007840	0.0000000
female_life_exp	0.9852269	14.3618938	0.0000000
male_life_exp	1.0366555	14.9655858	0.0000000
female_phy_act	0.9818781	14.3219740	0.0000000
male_phy_act	1.0109469	14.6660047	0.0000000
female_obesity	-0.8683949	-12.9254470	0.0000000
male_obesity	-0.6769418	-10.3822354	0.0000000
female_diabetes	-0.9084418	-13.4279287	0.0000000
male_diabetes	-1.0547767	-15.1741095	0.0000000

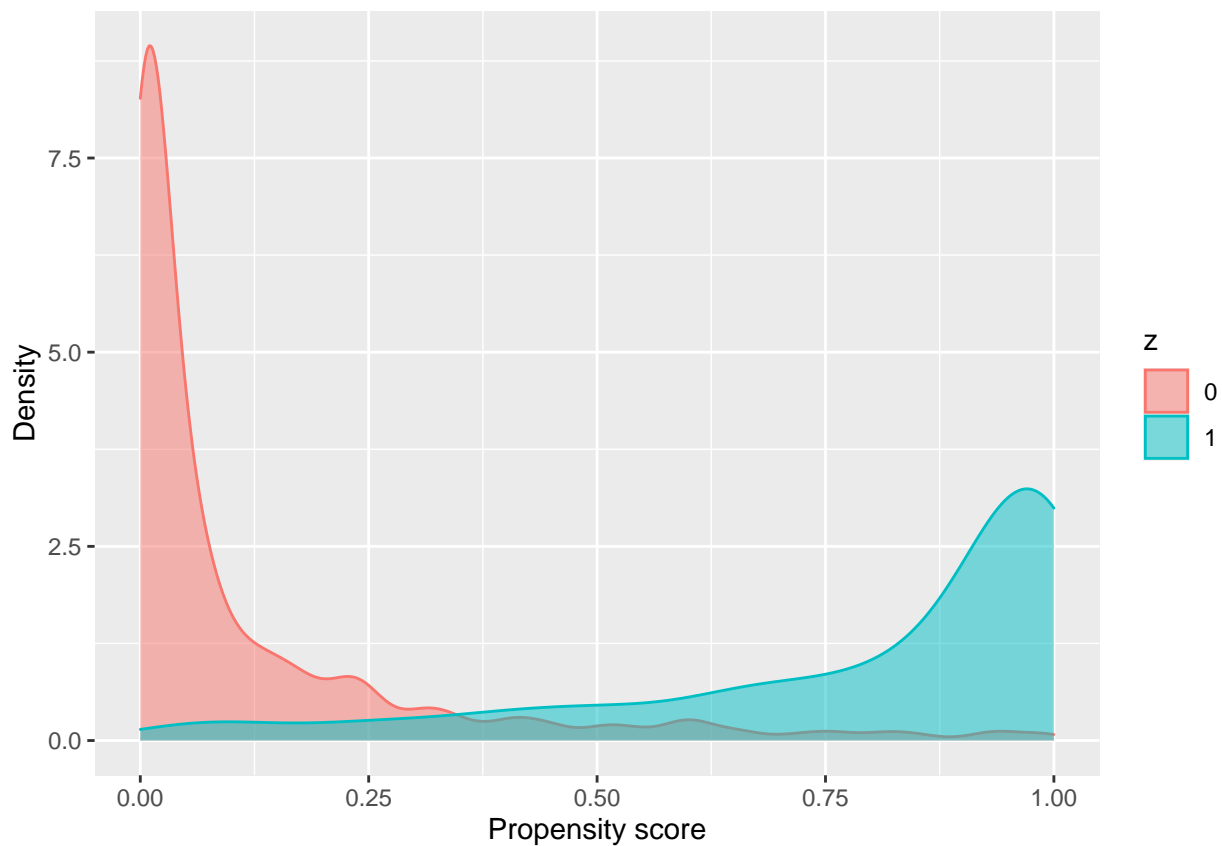


Propensity score

```
propensity_fit <- glm(z ~ ., family = "binomial", data = temp)
```

```
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
```

```
temp %>%  
  mutate(z = as.factor(z)) %>%  
  mutate(p_score = propensity_fit$fitted.values) %>%  
  ggplot(aes(p_score, fill = z, color = z)) +  
  geom_density(alpha = 0.5) +  
  labs(  
    x = "Propensity score",  
    y = "Density"  
  )
```



```
## Matching control and treat  
matrix_1 <-  
  smahal(z = temp$z, X = as.matrix(temp %>% select(-z)))  
# matrix_1 <- smahal(z = temp$treatment, X = as.matrix(temp %>% select(day_0, asian)))  
  
## Forcing matching on day 0  
# matrix_2 <-  
#   addalmostexact(matrix_1, z = temp$treatment, f = temp$day_0, mult = 10)  
  
## Adding caliper to force match on day_0  
matrix_3 <- addcaliper(matrix_1, z = temp$z, p = propensity_fit$fitted.values, caliper = 0.1)
```

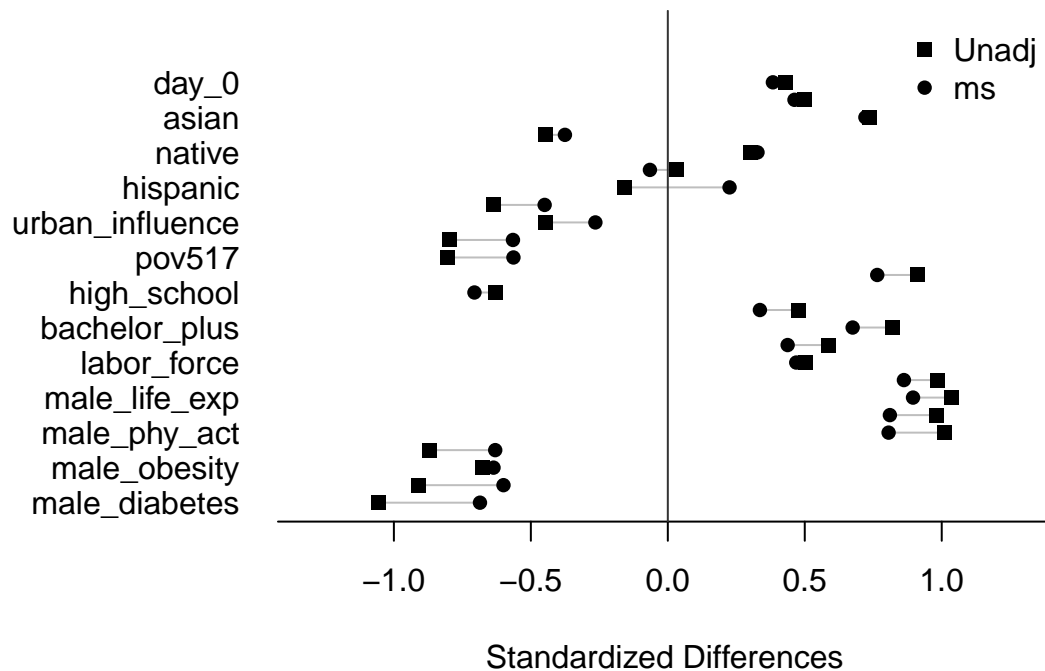


```
ms <- pairmatch(matrix_3, data = temp)

matched_1 <- cast.senm(merged, ms)
```

Plotting the adjustment

```
plot(
  xBalance(
    z ~ day_0 + . + strata(ms) - 1,
    data = temp %>% select(-starts_with(c("age", "mort")))
  )
)
```



```
matched <-
  cbind(merged, enframe(ms, name = NULL), p_score = propensity_fit$fitted.values) %>%
  filter(!is.na(value))

matched_treat <- matched %>% filter(z == 1)
matched_control <- matched %>% filter(z == 0)

pairs <-
  matched_treat %>%
  left_join(matched_control, by = "value", suffix = c(".1", ".0")) %>%
  filter(abs(day_0.1 - day_0.0) < 0.00001)
```

Getting average treatment effect

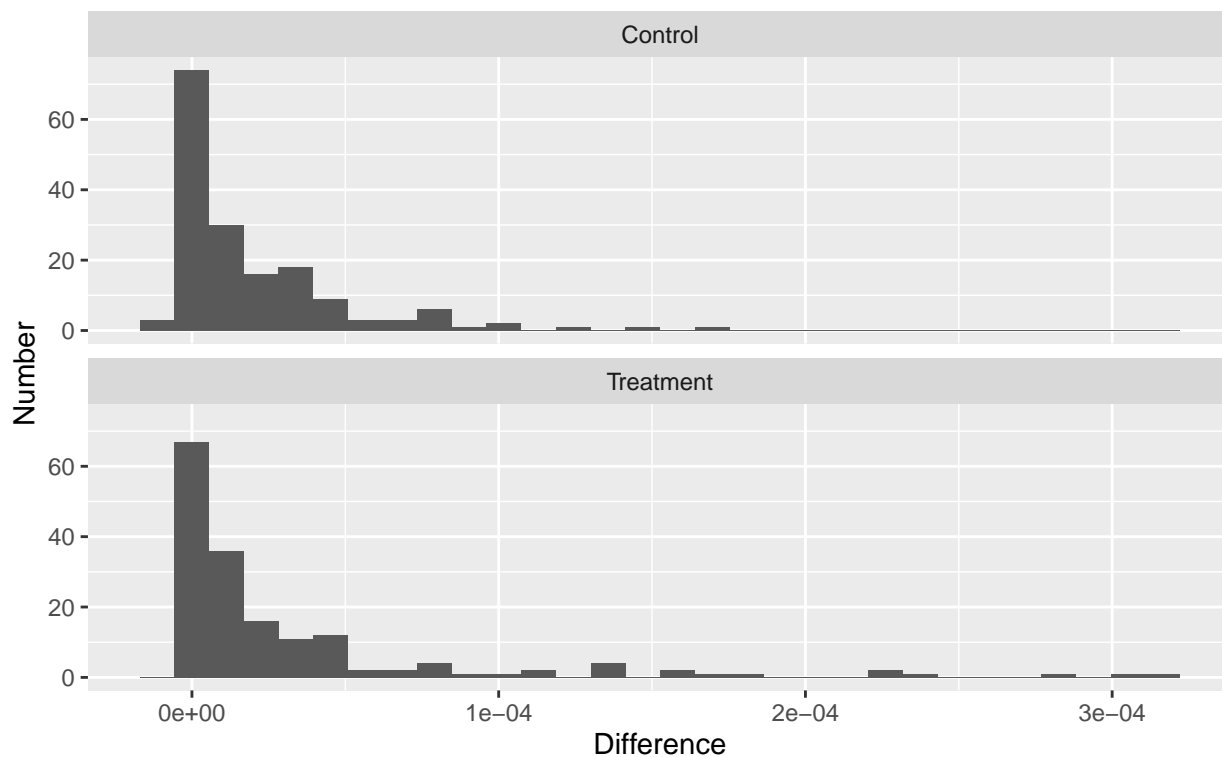
```
pairs %>%
  summarize_at(vars(starts_with(c("pre", "post", "y"))), mean) %>%
  knitr::kable()
```

pre_lockdown.1	pre_lockdown.0	post_lockdown.1	post_lockdown.0	y.1	y.0
1.4e-06	1.2e-06	3.41e-05	2.07e-05	3.27e-05	1.95e-05

```
## Comparing treatment to control
pairs %>%
  select(y.0, y.1) %>%
  gather(key = variable, value = value) %>%
  mutate(Group = if_else(variable == "y.0", "Control", "Treatment")) %>%
  ggplot(aes(value)) +
  geom_histogram(alpha = 1) +
  labs(
    x = "Difference",
    y = "Number",
    title = "Histograms of the difference between\npre-lockdown and post-lockdown for matched counties"
  ) +
  facet_wrap(vars(Group), ncol = 1)
```

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```

Histograms of the difference between pre-lockdown and post-lockdown for matched counties



Fischer's test

```
library(sensitivitymult)

##
## Attaching package: 'sensitivitymult'
```

```
## The following object is masked from 'package:DOS2':
##
##   planScheffe
```

```
senm(
  y = matched$y,
  z = matched$z,
  mset = matched_1$mset,
  gamma = 1.2,
  inner = 0,
  trim = Inf
)
```

```
## $pval
## [1] 1.261061e-09
##
## $deviate
## [1] 5.960017
##
## $statistic
## [1] 0.01204228
##
## $expectation
## [1] 0.001520425
##
## $variance
## [1] 3.116665e-06
```

Wilcox test

```
wilcox.test(pairs$y.0, pairs$y.1)
```

```
##
## Wilcoxon rank sum test with continuity correction
##
## data: pairs$y.0 and pairs$y.1
## W = 12697, p-value = 0.1037
## alternative hypothesis: true location shift is not equal to 0
```

Visualizing pairs

```
state_abbr <-
  read_csv("state_names.csv") %>%
  select(-Abbrev)
```

```
##
## -- Column specification -----
## cols(
##   State = col_character(),
##   Abbrev = col_character(),
##   Code = col_character()
## )
```

```
# png("pairs.png", width = 8, height = 8, units = "in", res = 300)
```

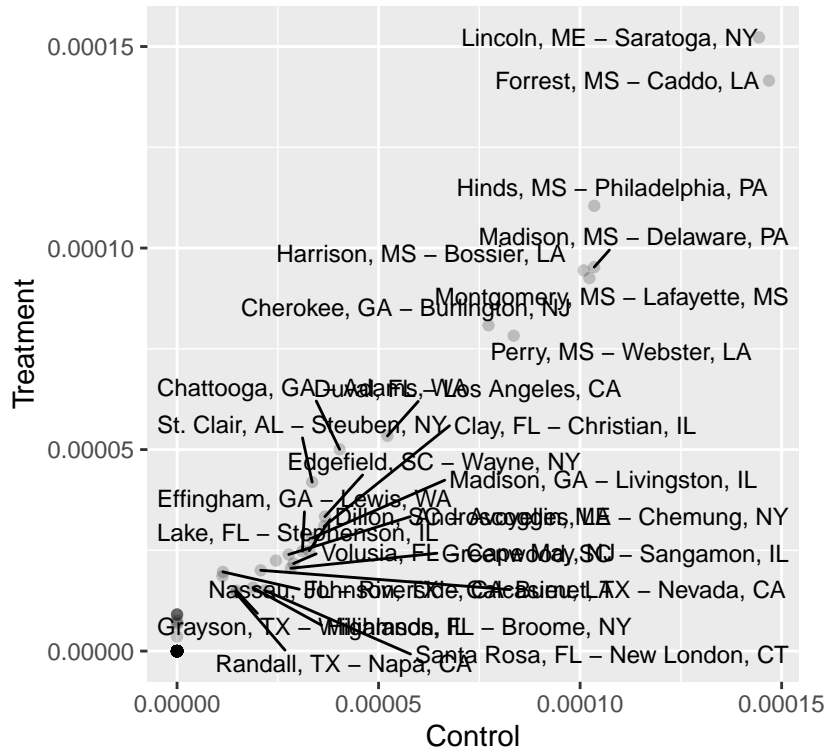
```

pairs %>%
  left_join(state_abbr, by = c("state.0" = "State")) %>%
  left_join(state_abbr, by = c("state.1" = "State"), suffix = c(".0", ".1")) %>%
  mutate(
    combined_name =
      str_c(county_name.0, ", ", Code.0, " - ", county_name.1, ", ", Code.1)
  ) %>%
  ggplot(aes(day_0.0, day_0.1)) +
  geom_point(alpha = 0.2) +
  ggrepel::geom_text_repel(
    aes(label = combined_name),
    data = . %>% filter(day_0.1 >= 0.00001 | day_0.0 >= 0.00001),
    size = 3
  ) +
  labs(
    x = "Control",
    y = "Treatment",
    title = "Number of cases on March 23",
    subtitle = "For each pair, the first county is in control"
  ) +
  coord_fixed()

```

Number of cases on March 23

For each pair, the first county is in control



```
# dev.off()
```