Impact of Public Transit Use on COVID-19 Spread

Miranda Reid

Advanced Data Analysis

April 28, 2020

**INTRODUCTION**

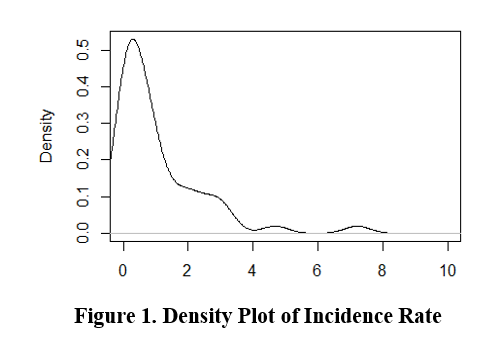
The novel coronavirus and its resulting disease, COVID-19, are a significant threat to the health and wellbeing of the United States with recent models suggesting that the healthcare system would be overwhelmed and millions could die without significant action from federal, state, and local governments.1,2 Given the very real possibility that the U.S. will face multiple waves of the disease before a vaccine is available,3 it is essential to determine which factors increased the incidence and prevalence rates of the disease, so that state and local governments can take appropriate preventative action. If high public transit use is significantly associated with increased COVID-19 incidence, then transit could be a high impact target for future prevention interventions. However, if public transit use is not associated with COVID-19 incidence or prevalence rate, then current measures, such as increasing disinfection of buses and stations and encouraging mask wearing,4 may be sufficient. This could potentially indicate that cities could start bringing back routes and increasing service, as well as making efforts to attract riders again, as current ridership is significantly decreased, placing a financial strain on many public transit systems.4 This is especially relevant now that states are starting to re-open a wide range of businesses,5 the increased travel associated with these openings mean that many public transit systems will soon need to make key decisions on what preventive measures to put in place and how broadly they themselves should reopen services.

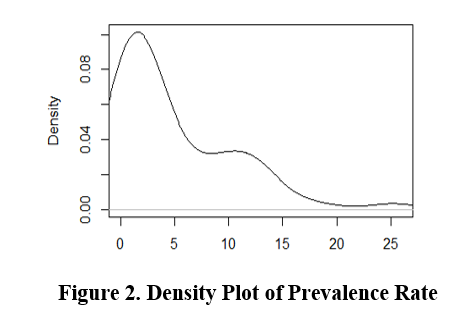
In order to determine the impact of public transit use on the spread of COVID-19, this paper will assess the relationship between state-level public transit use and both the incidence rate and prevalence rate 2 weeks after the first infection in the state. If there is a significant association between public transit use and either incidence rate or prevalence rate at that point in time, then public transit use may be associated with the rate of spread of COVID-19.

**METHODS**

Impact of public transit was modeled at the state level. This included all fifty states and the District of Columbia. All states and the District of Columbia were included as data on the dependent variable, independent variable, and included covariates were easily available for them. Data on public transit use were collected from the U.S. Department of Transportation.6 Public transit use was defined as the percentage of workers who use public transportation (buses, metro, subway etc.) in their commute to work. This definition excludes taxis or active transportation, like walking or biking.

Incidence rate and prevalence rate were calculated as the incidence rate or prevalence rate 2 weeks after the first confirmed case of COVID-19 in that state from data collected by the New York Times based on reports from state and local health agencies.7 Two weeks was selected in order to compare the states at a similar amount of time after the first case and a time early enough in the pandemic that incidence rates had not already peaked. However, since the incidence rate was often zero this early on in the pandemic, both for states that had successfully limited community spread at this point and for states who may have had trouble with setting up daily reporting at that time prevalence rate was also modeled. This was in the hope that the prevalence rate model would differentiate between states that had very low incidence and prevalence at 2 weeks and states that just coincidentally had artifically low incidence rates at 2 weeks, e.g. the 2 week mark fell on a Saturday in a state that tends to do less testing and reporting on weekends. There will always be inherent uncertainty when modeling as the event occurs given variability in states’s testing capabilities and reporting systems, but by separately modeling incidence rate and prevalence rate, this paper hopes to more thoroughly investigate the impact of public transit use on COVID-19 spread.

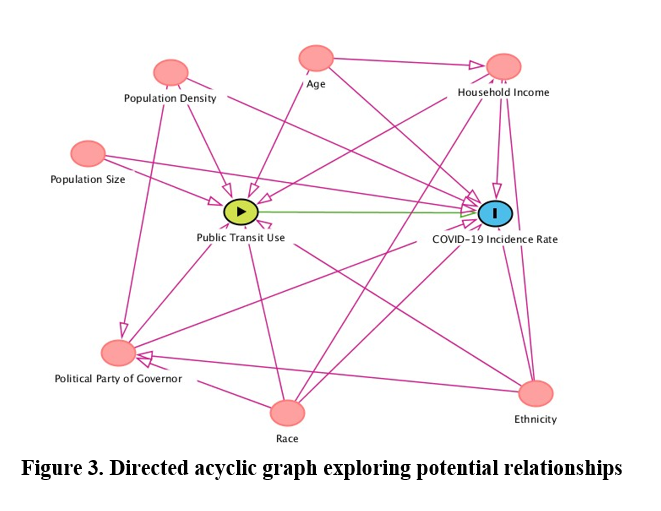
The model included the following covariates: overall population size, population density, race, ethnicity, age, sex, household income, and political party of the Governor. Data for covariates were collected from the 2018 American Community Survey, the 2010 Census, and the Kaiser Family Foundation.8–10 Population density was defined as the average population per square mile. Race was defined as the percentage of the population who identifies as white. Ethnicity was defined as the percentage of the population that identifies as hispanic. Age was defined as median age. Sex was defined as the percentage of the population that was male. Household income was defined as median household income. Analysis revealed that almost all continuous variables were skewed with the exception of sex, so descriptive statistics were analyzed using median and interquartile range for continuous variables and counts and proportions for categorical variables.

****Both incidence rate (Figure 1) and prevalence rate (Figure 2) were right skewed, so a poisson model was appropriate. However, due to overdispersion, a negative binomial model with robust standard errors was used. A likelihood ratio test confirmed that the negative binomial model was statistically significantly better than the poisson model (p<0.001). Likelihood ratio tests also confirmed that models including the covariates were statistically significantly better (p<0.001) than univariate models at predicting rates. Incidence rate ratios and prevalence rate ratios were calculated to help determine whether any results that were statistically signficant were also clinically meaningful.

**RESULTS**

|  |  |
| --- | --- |
| **Table 1. State Characteristics (n=51)** | |
| **Characteristic** | **Median (IQR)** |
| Incidence Rate (per 100,000) | 0.57 (0.15,1.46) |
| Prevalence Rate (per 100,00) | 2.45 (1.14, 8.75) |
|  |  |
| Population (in thousands) | 4,467(1,789, 7,446) |
| Population Density (per sq. mile) | 101.20 (45.80, 221.45) |
| Commute via Public Transit (%) | 1.60 (0.75,3.95) |
|  |  |
| Age | 38.65 (37.27, 39.73) |
| Male (%) | 49.00 (49.00, 50.00) |
| White (%) | 71.00 (58.00, 80.00) |
| Hispanic (%) | 10.00 (5.00, 13.50) |
| Household Income ($) | 58,570 (53,069, 68,373) |
|  |  |
| Political Party of Governor | N (%) |
| Democrat | 25 (49.0) |
| Republican | 26 (51.0) |

Public transit use was low at the state-level with a median of 1.60% of the population using public transportation to commute to work (Table 1). The median incidence rate (0.57 cases per 100,000) and prevalence rate at 2 weeks was fairly low (2.45 cases per 100,000) this early on in the pandemic, by April 7 the national prevalence rate was 119.6 cases per 100,000.11 There was little variation between states in terms of age (med.=38.65; IQR =37.27-39.73) or sex (med. % male = 49.00; IQR= 49.00-50.00). There was more substantial variation on race (med. % white= 71.00; IQR= 58.00-80.00), ethnicity (med. % Hispanic = 10.00; IQR =5.00-13.50), and household income (med.= 58,570; IQR=53,069-68,373). States were fairly evenly split between Democratic governors (n=25) and Republican governors (n=26).[[1]](#footnote-1) The median state population size was 4.5 million (IQR=1.8-7.5) and population density was 101.2 people per square mile (IQR= 45.80-221.45).

****

**Figure 3. Directed acylic graph of the relationship between public transit use, COVID-19 Incidence Rate, and covariates.**

All covariates included in the model were potential confounders as seen in Figure 3. Age, race, ethnicity, and income all impact public transit use on an individual level,12 while population size, population density, and political party of governor impact public transit use by impacting the likelihood that a well-developed public transit system will be built and maintained. Political party of governor is associated with the isolation measures ordered at the state level (e.g. stay-at-home orders) which can then impact COVID-19 incidence rate, and age, race, ethnicity, income, population density, and population size all have a demonstrated relationship with COVID-19 incidence rate.13–15

|  |  |  |
| --- | --- | --- |
| **Table 2. Negative binomial distribution of prevalence rate.** | | |
| ***Variable*** | ***Prev. Rate Ratio*** | ***95% CI*** |
| Public Transit Use | 1.054 | [0.972-1.143] |
| Population Density | 1.000 | [1.000-1.000] |
| Population Size | 1.000\* | [1.000-1.000] |
| Age | 0.985 | [0.905-1.072] |
| Male | 1.116 | [0.752-1.656] |
| Non-Hispanic white | 0.985 | [0.964-1.007] |
| Hispanic | 0.956 \* | [0.920-0.994] |
| Household Income | 1.000 \*\* | [1.000-1.000] |
| Republican Governor | 0.522 \* | [0.275-0.988] |
| *Note: Punctuation indicates significance at the following levels: '\*\*' 0.01 ;'\*' 0.05 ; '.' 0.1* | | |

A negative binomial model of the impact of public transit use on prevalence rate adjusted for population density, population size, age, sex, race, ethnicity, income, and political party of governor did not show a statistically significant relationship (p=0.203) (Table 2). A state with a 1 percentage point higher public transit use would have a 5.4% increase in the prevalence rate (PRR=1.054; 95% CI: 0.972-1.143). Given that the 95% confidence interval crosses 1 and the p-value is greater than 0.05, it is not certain that higher public transit use is associated with an increased prevalence rate. In the model, only population size, household income, Hispanic-ethnicity, and the political party of the governor had a statistically significant relationship with prevalence rate. Of these only the political party of the governor and Hispanic ethnicity had meaningful prevalence rate ratios. States with Republican governors had a 47.8% decrease in the prevalence rate compared to states with Democratic governors (PRR=0.522; 95% CI: 0.275-0.988; p<0.05). A state with a 1 percentage point higher proportion of Hispanic residents would have a 4.4% decrease in the prevalence rate (PRR=0.956; 95% CI:0.920-0.994; p<0.05).

|  |  |  |
| --- | --- | --- |
| **Table 3. Negative binomial distribution of incidence rate.** | | |
| ***Variable*** | ***Incid. Rate Ratio*** | ***95% CI*** |
| Public Transit Use | 1.002 | [0.899-1.117] |
| Population Density | 1.000 | [1.000-1.000] |
| Population Size | 1.000 | [1.000-1.000] |
| Age | 0.890 \* | [0.812-0.975] |
| Male | 0.947 | [0.602-1.492] |
| Non-Hispanic white | 0.989 | [0.969-1.009] |
| Hispanic | 0.949 \* | [0.907-0.992] |
| Household Income | 1.000 | [1.000-1.000] |
| Republican Governor | 0.418 \* | [0.209-0.837] |
| *Note: Punctuation indicates significance at the following levels: '\*\*' 0.01 ;'\*' 0.05 ; '.' 0.1* | | |

A negative binomial model of the impact of public transit use on incidence rate adjusted for population density, population size, age, sex, race, ethnicity, income, and political party of governor did not show a statistically significant relationship (p=0.975) (Table 3). A state with a 1 percentage point higher public transit use would have a 0.2% increase in the incidence rate (IRR=1.001; 95% CI: 0.899-1.117). Given that the 95% confidence interval crosses 1 and the p-value is greater than 0.05, it is not certain that higher public transit use is associated with an increased incidence rate. In the model, only age, Hispanic ethnicity, and the political party of the governor had a statistically significant relationship with incidence rate. States with Republican governors had a 59.2% decrease in the incidence rate compared to states with Democratic governors (IRR=0.418; 95% CI: 0.209-0.837; p<0.05). A state with a 1 percentage point higher proportion of Hispanic residents would have a 5.1% decrease in the incidence rate (IRR=0.956; 95% CI:0.907-0.992; p<0.05). A state with a one-year increase in median age would have 11% decrease in incidence rate (IRR=0.890; 95% CI: 0.812-0.975; p<0.05).

While the direction of the relationship between public transit use and incidence rate or prevalence rate was consistent in both models, the lack of statistical significance or a meaningful rate ratio makes it difficult to draw conclusions on the nature of the relationship based on these models. The only variables with both statistical significance and meaningful rate ratios in both models were Hispanic ethnicity and the political party of the governor.

**DISCUSSION**

These results indicate that public transit use may not impact the spread of COVID-19 to a significant extent. While there was a consistent positive relationship between public transit use and rate ratios, both the prevalence rate ratio and incidence rate ratio were small and the 95% confidence interval for both included 1, indicating uncertainty in both the direction and even the existence of a relationship between public transit use and the spread of COVID-19. As public transit use was overall fairly low at the state-level, it may just not be common enough to have a significant impact on COVID-19 spread. It is still important to use appropriate social distancing and sanitation in public transportation, as in any public space. Additionally, special protections should be included for transit workers, as early reports indicate that they are dying at higher rates and research has shown that they are more susceptible to infection even under non-pandemic conditions.16,17 Overall, these results may indicate that with appropriate precautions, public transit may not significantly contribute to the spread of COVID-19.

While interesting that both an increased proportion of Hispanic residents and having a Republican governor decreased the incidence and prevalence rates of COVID-19, it is difficult to draw actionable conclusions from this information. Additionally, as prevalence rate and incidence rate were measured at only 2 weeks after the first case in each state, most governors had not taken significant action to limit travel or movement at this point, indicating that the impact of the governor’s political party may have more to do with which states have Republican governors than any specific actions by those governors to address the crisis.

Due to the nature of trying to assess the spread of a pandemic while it is still happening, this paper has significant limitations. States may have had significantly different levels of testing and reporting capabilities early on in the pandemic, which would have impacted the reported incidence and prevalence of COVID-19 in each states. Additionally, there was no ability to parse out different types of public transit, so some types may have had more or less of an impact than others. Additionally, some states had extremely low prevalence and incidence rates 2 weeks after their first case, as they were not experiencing community transmission at that time. These states generally had higher public transit usage than the median potentially biasing the results.

Despite these limitations, the results indicate that public transit use may not have as significant an impact on COVID-19 transmission as many had feared. Future analyses could include more recent data and could use panel data to look at the impact of public transport use on incidence and prevalence rate trends. With appropriate precautions, especially for transit workers, it may be appropriate to gradually increase hours and services to reach the increasing need for transit as states reopen.

**BIBLIOGRAPHY**

1. Glanz J, Leatherby L, Bloch M, et al. Coronavirus Could Overwhelm U.S. Without Urgent Action, Estimates Say. The New York Times. https://www.nytimes.com/interactive/2020/03/20/us/coronavirus-model-us-outbreak.html. Published March 20, 2020. Accessed March 24, 2020.

2. Fink S. White House Takes New Line After Dire Report on Death Toll . The New York Times. https://www.nytimes.com/2020/03/16/us/coronavirus-fatality-rate-white-house.html. Published March 16, 2020. Accessed March 24, 2020.

3. Department of Health and Human Services. *PanCAP Adapted U.S. Government COVID-19 Response Plan*.; 2020. https://int.nyt.com/data/documenthelper/6819-covid-19-response-plan/d367f758bec47cad361f/optimized/full.pdf#page=1. Accessed March 24, 2020.

4. Levy A, Goldwyn E. How U.S. Public Transit Can Survive Coronavirus. City Lab. https://www.citylab.com/perspective/2020/04/coronavirus-public-transit-ridership-federal-funding-bus-subways/610453/. Published April 24, 2020. Accessed April 29, 2020.

5. Mervosh S, Lee JC. See Which States Are Reopening and Which Are Still Shut Down. The New York Times. https://www.nytimes.com/interactive/2020/us/states-reopen-map-coronavirus.html. Published April 27, 2020. Accessed April 28, 2020.

6. United States Department of Transportation Statistics. Commuting to Work. Bureau of Transportation Statistics. https://www.bts.gov/content/commuting-work. Published 2013. Accessed April 28, 2020.

7. The New York Times. Coronavirus (Covid-19) Data in the United States. https://github.com/nytimes/covid-19-data/blob/master/README.md. Accessed April 28, 2020.

8. National Governors Association (NGA). Current Governors 2020. Kaiser Family Foundation. https://www.kff.org/other/state-indicator/governors/?currentTimeframe=0&sortModel=%7B%22colId%22:%22Location%22,%22sort%22:%22asc%22%7D. Published 2020. Accessed April 28, 2020.

9. United States Census Bureau. American Community Survey. https://data.census.gov/cedsci/all?q=median age&hidePreview=false&tid=ACSST1Y2018.S0101&t=Age and Sex. Published 2018. Accessed April 28, 2020.

10. United States Census Bureau. 2010 Census: Population Density Data. 2010. https://www.census.gov/data/tables/2010/dec/density-data-text.html. Accessed April 28, 2020.

11. Bialek S, Bowen V, Chow N, et al. Geographic Differences in COVID-19 Cases, Deaths, and Incidence — United States, February 12–April 7, 2020. *Morb Mortal Wkly Rep*. 2020;69(15):465-471. doi:10.15585/mmwr.mm6915e4

12. Andersen M. Who relies on public transit in the U.S. . Pew Research Center. https://www.pewresearch.org/fact-tank/2016/04/07/who-relies-on-public-transit-in-the-u-s/. Published April 17, 2016. Accessed April 29, 2020.

13. Guan WJ, Ni ZY, Hu Y, et al. Clinical Characteristics of Coronavirus Disease 2019 in China. *N Engl J Med*. February 2020. doi:10.1056/NEJMoa2002032

14. Garg S, Kim L, Whitaker M, et al. Hospitalization Rates and Characteristics of Patients Hospitalized with Laboratory-Confirmed Coronavirus Disease 2019 — COVID-NET, 14 States, March 1–30, 2020. *Morb Mortal Wkly Rep*. 2020;69(15):458-464. doi:10.15585/mmwr.mm6915e3

15. Grinstein-Weiss M, Gupta B. We don’t need a map to tell us who COVID-19 hits the hardest in St. Louis. Brookings. https://www.brookings.edu/blog/up-front/2020/04/06/we-dont-need-a-map-to-tell-us-who-covid-19-hits-the-hardest-in-st-louis/. Published April 8, 2020. Accessed April 29, 2020.

16. Barkan R. New York’s Transit Workers Keep Getting Sick. The Nation. https://www.thenation.com/article/politics/mta-transit-driver-covid/. Published April 9, 2020. Accessed April 29, 2020.

17. Winkleby M, Ragland D, Fisher J, Syme S. Excess risk of sickness and disease in bus drivers: a review and synthesis of epidemiological studies. *Int J Epidemiol*. 1988;17(2):255-262. https://www.ncbi.nlm.nih.gov/pubmed/3042649. Accessed April 29, 2020.

1. Note: For the District of Columbia, Mayor Muriel Bowser (D) was counted as the governor as she is the leader of the district level executive branch of government. [↑](#footnote-ref-1)