

Food Insecurity During the COVID-19 Pandemic

BST260 Final Project

Sara O'Brien

December 9, 2023

Harvard T.H. Chan School of Public Health

Introduction

The COVID-19 pandemic had a drastic impact on agricultural production and global supply chains, resulting in widespread food shortages and inflated food prices. Supply issues coupled with high unemployment rates, school closures, and a lack of access to grocery stores during quarantine led to a food insecurity crisis in the United States.¹ Food insecurity is defined by the U.S. Department of Agriculture (USDA) as a household-level economic and social condition of limited or uncertain access to adequate food.² This socioeconomic condition is associated with an increased risk of diabetes, obesity, heart disease, mental health disorders, and other chronic diseases, making it a high-priority issue in public health research and policy.³

The Supplemental Nutrition Assistance Program (SNAP) is a federal program that aims to reduce poverty and food insecurity while stimulating economic growth. Following the declaration of a public health emergency during the COVID-19 pandemic, the USDA granted flexibility to states to ease SNAP program administration.⁴ The aim of this project was to assess the association between COVID-19 cases and food insecurity on a county-level in the U.S. A secondary aim was to consider how SNAP participation, poverty, and income influence the relationship between COVID-19 cases and food insecurity. These aims were motivated by initial exploratory analyses that revealed a positive linear trend between COVID-19 case rates and food insecurity rates (Figure 1). Leveraging county-level data from the Centers for Disease Control and Prevention (CDC),⁵ USDA,⁶ Feeding America,⁷ and U.S. Census Bureau,⁸ this project visualized the associations of interest in a series of time trends, summary plots, and regression models.

Methods

The project was carried out in three main steps: (1) data wrangling to create an analysis data set, (2) exploratory data analyses, and (3) regression analyses.

1. Data Wrangling

The primary data wrangling required for this project was the cleaning and merging of county-level data from four different sources to pull in COVID-19 cases, food insecurity, SNAP participation, poverty and income, and population estimates.

COVID-19 Cases

The CDC continues to maintain a COVID-19 case surveillance database even after the end of the U.S. public health emergency declaration on May 11, 2023. The case reports are patient-level, de-identified, and reported by U.S. states and autonomous reporting entities, including New York City and the District of Columbia.⁹ All cases reported in 2021 were reported using the standardized case definitions for COVID-19 or lab-confirmed/probable cases, but the data is considered dynamic and provisional as health departments and entities receive new information. However, for the year being used in this project, 2021, the CDC has already reconciled and finalized case reporting.

SNAP Participation

During the pandemic, there was a government response called Pandemic Electronic Benefits Transfer (P-EBT) that provided benefits similar to SNAP benefits, allowing recipients to purchase foods eligible for SNAP.¹⁰ However, this was a temporary measure that could be received in addition to and separate from SNAP, with the majority of households eligible for both or neither. This project focuses on participation in SNAP during the pandemic rather than P-EBT since this measure could be captured longitudinally and was more reflective of individuals needing long-term assistance due to food insecurity and unemployment during the pandemic.

SNAP serves two groups of low-income households: Public Assistance (PA) and Non-Public Assistance (NPA). PA households include only members who receive or are eligible to receive income from Temporary Assistance for Needy Families, General Relief, or Supplemental Security Income. NPA households are those with at least one member not eligible for one of the PA programs.¹¹ The Bi-Annual State Project Area/County Level Participation and Issuance Data from the USDA Food and Nutrition Service contains counts of PA and NPA participation on the person and household level. Since COVID-19 case data is on the individual level, this project leveraged the summed SNAP total PA and NPA persons participation (as opposed to households) as a measure of SNAP participation. Since the other data sources were reported at the end of 2021 and SNAP data was only available at the county level bi-annually in January or July, SNAP data was pulled from January 2022 to reflect benefits dispersed in a similar time frame as the 2021 COVID-19 cases and counts of food insecurity.

Of note, not all states report SNAP participation at the county level. The following states were removed for only reporting state-level SNAP data: AK, CT, ID, MA, ME, MO, MT, NE, NH, NY, OR, RI, UT, VT, WA, and WY. The SNAP data set contained some government assistance programs other than SNAP, which were excluded.

Other Data Sources

- **Food Insecurity:** The most recently available data (2021) from the Feeding America’s Map the Meal Gap initiative captures food insecurity estimates at the county level with an overall food insecurity rate (percentage of inhabitants who are food insecure) and a count of food insecure persons (number of inhabitants who are food insecure).
- **Poverty and Income:** The U.S. Census Bureau’s Small Area Income and Poverty Estimates program produces county-level estimates of income and poverty.¹² 2021 estimates were pulled into the analysis data set.
- **Population Counts:** The U.S. Census Bureau’s County Population Totals similarly estimate county-level population counts and changes each year.¹³ 2021 estimates were pulled into the analysis data set.

All data sources were merged using FIPS codes. The FIPS codes uniquely identify each county and are part of the Federal Information Processing Standards (FIPS) developed by the National Institute of Standards and Technology (NIST), U.S. Department of Commerce.

2. Exploratory data analysis

Exploratory data analysis aimed to uncover patterns and trends in SNAP participation rates, COVID-19 case rates, food insecurity rates, and poverty rates. Although count data was also explored given the intent to employ Poisson regression models, rates were the key metric utilized in EDA so as to not skew trends with inflated population sizes. All rates were derived using count data for the respective measure by dividing over the county population. Scatter plots displayed the majority of comparisons between variables at a single time point, 2021, whereas box-and-whisker plots were selected to capture longitudinal trends in county-level food insecurity rates.

3. Regression analysis

Since the outcome variable, food insecurity, was available as both a rate and a count measure, there was flexibility in the modeling approach for this project. Both multivariate linear models and Poisson regression models were run and compared to determine the model with the best fit and simplest interpretation to address the primary question of interest for this project. In the linear model, food insecurity was included as a continuous rate. Prior to running a linear model, the LINE assumptions were assessed to determine whether the regression estimates were likely to be biased. In the Poisson model, food insecurity was included as count data. Prior

to conducting a Poisson model regressing food insecurity counts on COVID-19 case counts, the mean-variance assumption was assessed to determine whether the model exhibited over-dispersion. A Negative Binomial model allowing for over-dispersion was also explored to model the outcome data. Both the linear and Poisson models considered as covariates county-level COVID-19 case rates, SNAP participation rates, poverty rates, and median household income. Stepwise model selection, likelihood ratio tests, model AICs, and overall model parsimony were used to select the final model.

Results

In 2021, there were 34,456,043 reported COVID-19 cases across all U.S. counties. Some cases were reported without a state or county and were excluded, leaving a total of 32,617,301 cases for analyses. Grouping across state and county, cumulative counts of the number of COVID-19 cases from January to December, 2021 were obtained, with a total of 1,787 counties in the initial analysis data set. There was a 1:1 match between the COVID-19 cases and food insecurity and poverty/income county-level data, but not between the COVID-19 and SNAP participation data, resulting in a final analytic data set of 1,467 observations/counties.

From 2019 to 2021, there was a general trend of food insecurity rates decreasing on the county level across all U.S. geographic regions. Figure 2 provides evidence contrary to existing literature¹ that there was not an overall increase in food insecurity during the time-frame of the COVID-19 pandemic according to Map the Meal Gap data. However, this box-and-whisker plot alone cannot capture the complexity of food insecurity and further exploration of trends between COVID-19 cases and food insecurity motivated the exploration of regression modeling techniques. Namely, plots of county-level COVID-19 case counts versus food insecurity counts and SNAP participants counts predictably followed a linear relationship, as seen in Figures 3 and 4. This trend was expected, given that counties with larger populations would experience more COVID-19 cases and higher SNAP participation compared to smaller counties. In Figures 1 and 5, the trend was still roughly linear and increasing when plotting COVID-19 case rates against food insecurity rates and SNAP participation rates. These exploratory plots suggested that even when controlling for population size, there was a relationship worth exploring between the predictors and outcome.

Linear Model

Model diagnostics for the multiple linear regression model were run to detect influential points and assess normality, homoscedasticity, and linearity. One key consideration for this data was that it was on the county level rather than the individual level, meaning that food insecurity rates are less clearly independent. However, proceeding with caution, the other LINE assumptions appeared to be satisfied based on the residual plots (Figures 6.1-6.4). The qq-plot was normal, there was equal variance of the residuals, and there was no indication of non-linear trends in the data. Furthermore, while there were a few influential observations based on

Cook's distance, none were of major concern. Proceeding with a linear model in light of these considerations produced the final linear model summarized in Table 1.

After controlling for both poverty rates and median income, SNAP participation rates were not a significant predictor of food insecurity in this model. For a 1-unit increase in COVID-19 cases/population, a county's food insecurity rate increased by an estimated 0.037 individuals/population, on average, holding all other covariates fixed. That is, the linear model suggested that there was a statistically significant harmful effect of COVID-19 on food insecurity at the county level. Moreover, the linear regression models revealed that poverty rates and median household income were highly significant predictors of food insecurity. For a 1-unit increase in the number of individuals in poverty/population and 1-dollar increase in median household income, a county's food insecurity rate increased by an estimated 0.388 individuals/population and decreased by an estimated 5.757×10^{-7} individuals/population, respectively.

Poisson Model

Based on a dispersion test, the original Poisson model did not satisfy the mean-variance assumption. An extension of the Poisson model, the Negative Binomial model, was selected to help with overdispersion and lead to more valid inference by relaxing the mean-variance assumption. Since there was not an issue of excess zeroes, this model extension was selected over a Zero-inflated Poisson model. The Negative Binomial model had a lower AIC and a likelihood ratio test comparing the Poisson model to the Negative Binomial model returned a significant p-value $< 2 \times 10^{-16}$, suggesting that the Negative Binomial model was preferred. However, stepwise model selection indicated that a final model without COVID-19 case rates was the best fit, and the Negative Binomial model including case rates did not indicate that this was a significant predictor of food insecurity counts. A summary of the Negative Binomial model is available in Table 2. Ultimately, this model type was less suited to the available analysis data than the linear model.

Conclusion

This project investigated the relationship between COVID-19 cases and food insecurity at the U.S. county level. A multiple linear regression model provided evidence of a significant harmful effect of COVID-19 on food insecurity. While SNAP participation rates were not a significant predictor of food insecurity in the final adjusted model, poverty rates and median household income emerged as strong predictors of food insecurity. One main limitation of this project was that the data was pulled from a wide range of sources that collected data at different time points via differing methods. Additionally, these findings are at the county level, and future research could more systematically observe trends between COVID-19 and food insecurity on an individual level by sampling on individuals' food security status, income, geographic location, and other demographics. These findings highlight an interplay of risk factors for food

insecurity and underscore the importance of socioeconomic policy interventions in addressing rising food insecurity concerns during public health emergencies.

References

1. <https://www.ncbi-nlm-nih-gov.ezp-prod1.hul.harvard.edu/pmc/articles/PMC9335023/>
2. <https://www.ers.usda.gov/topics/food-nutrition-assistance/food-security-in-the-u-s/definitions-of-food-security/#:~:text=Food%20insecurity%E2%80%94the%20condition%20assessed,may>
3. <https://www.nimhd.nih.gov/resources/understanding-health-disparities/food-accessibility-insecurity-and-health-outcomes.html>
4. <https://www.cbpp.org/research/food-assistance/most-states-are-using-new-flexibility-in-snap-to-respond-to-covid-19>
5. https://data.cdc.gov/Case-Surveillance/COVID-19-Case-Surveillance-Public-Use-Data-with-Ge/n8mc-b4w4/data_preview
6. <https://www.fns.usda.gov/pd/supplemental-nutrition-assistance-program-snap>
7. https://map.feedingamerica.org/?__ga=2.121274824.1141638753.1696475787-2079434502.1696475787
8. <https://www.census.gov/programs-surveys/saipe.html>
9. https://data.cdc.gov/Case-Surveillance/COVID-19-Case-Surveillance-Public-Use-Data-with-Ge/n8mc-b4w4/about_data
10. <https://www.fns.usda.gov/snap/pandemic-ebt-questions-and-answers-retailers#:~:text=For%20some%20>
11. https://www.dss.virginia.gov/geninfo/annual_report/benefits/index.html
12. <https://www.census.gov/data/datasets/2021/demo/saipe/2021-state-and-county.html>
13. <https://www.census.gov/data/tables/time-series/demo/popest/2020s-counties-total.html>

Appendix

Figure 1

County-Level COVID-19 Case Rate vs. Food Insecurity Rate in the U.S.

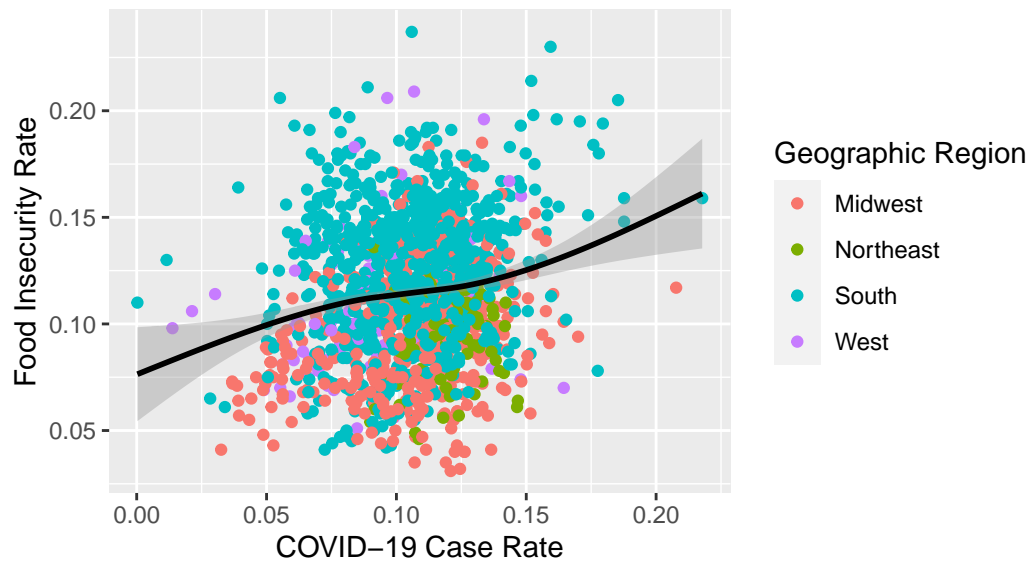
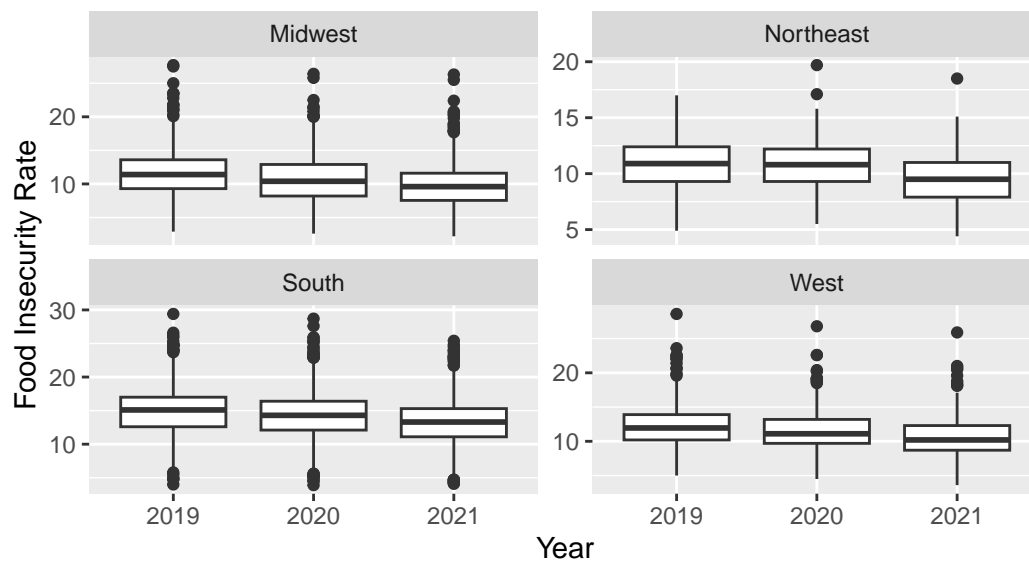


Figure 2

County-Level Food Insecurity Rates by U.S. Geographic Region, 2019–20



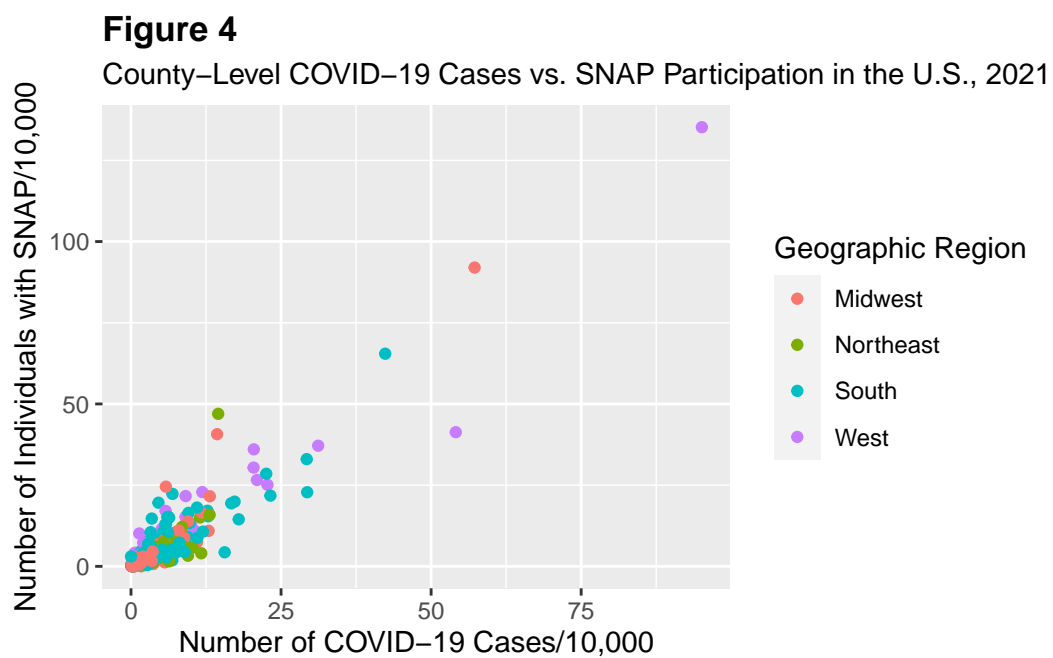
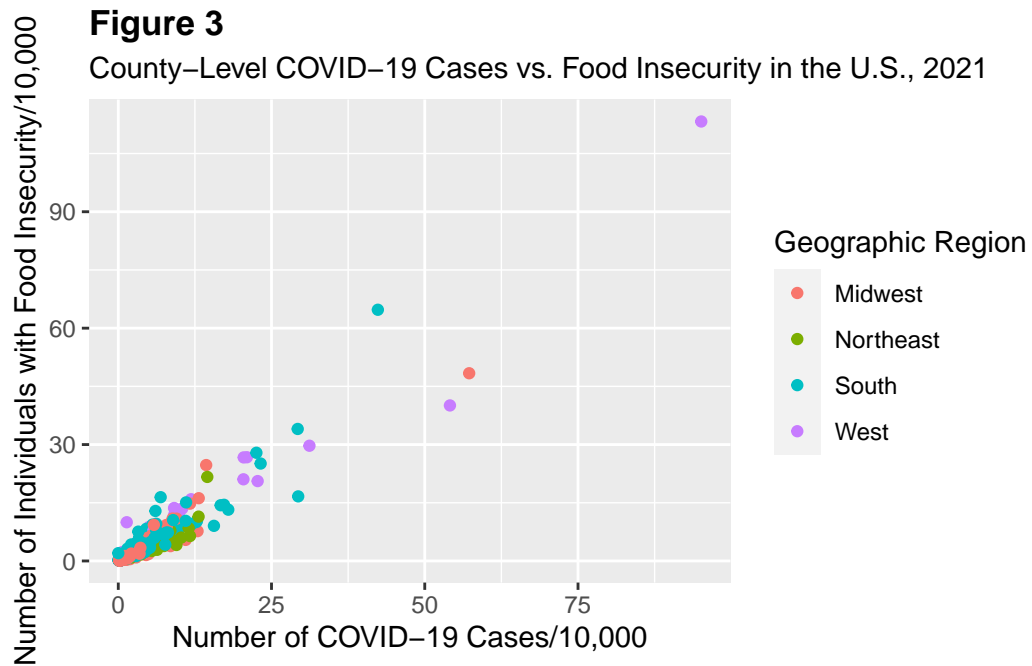


Figure 5

County-Level COVID-19 Cases vs. SNAP Participation in the U.S., 2021

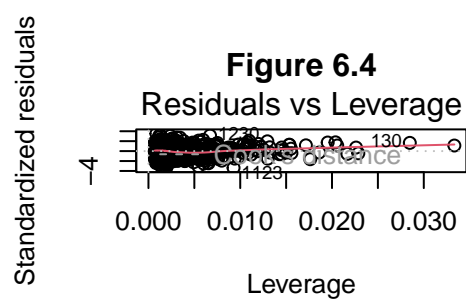
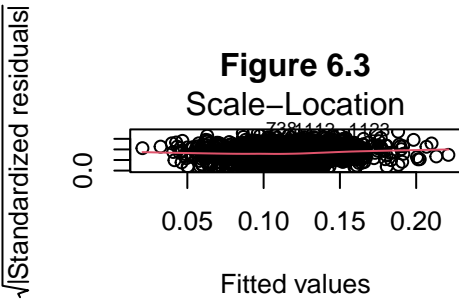
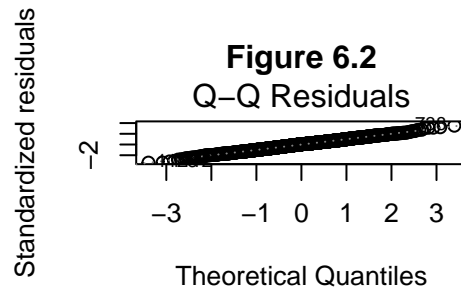
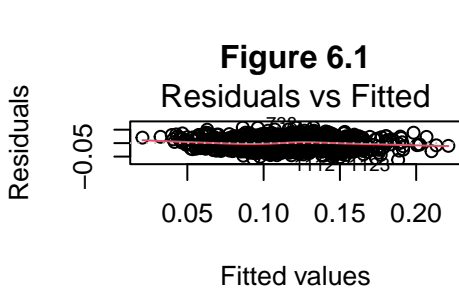
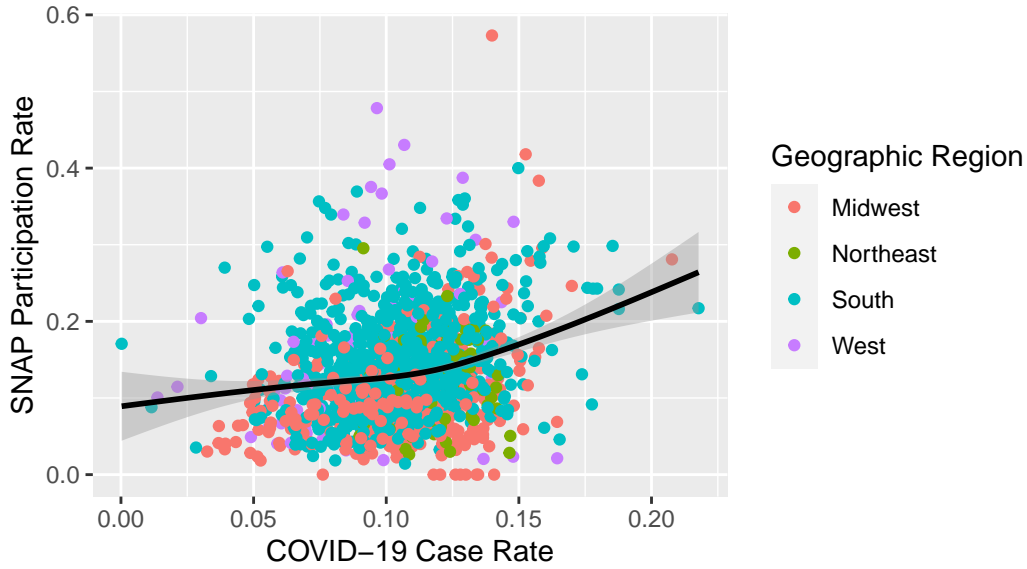


Table 1: Linear Model Summary

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	0.0925192	0.0057294	16.148174	0.0000000
case_rate	0.0374319	0.0201183	1.860588	0.0630029
poverty_rate	0.3876474	0.0159183	24.352339	0.0000000

	Estimate	Std. Error	t value	Pr(> t)
median_income	-0.0000006	0.0000000	-11.589599	0.0000000

Table 2: Negative Binomial Model Summary

	Estimate	Std. Error	z value	Pr(> z)
(Intercept)	3.0013275	0.3091609	9.7079778	0.0000000
case_rate	-0.3104221	1.0966181	-0.2830722	0.7771215
snap_rate	3.6205128	0.6154573	5.8826388	0.0000000
poverty_rate	13.2238182	1.0671998	12.3911361	0.0000000
median_income	0.0000699	0.0000027	26.1441385	0.0000000