ON THE ASSOCIATION BETWEEN ESTIMATED COVID-19 CASE FATALITY RATES AND THE SOCIOECONOMIC AND HEALTH FACTORS IN THE US STATES OF FLORIDA AND NEW YORK

By Mansi Joshi‡ and Yanming Di‡ and Sharmodeep Bhattacharyya‡ and Shirshendu Chatterjee†,*

City University of New York †
Oregon State University, Corvallis ‡

Background. Being the leading country worldwide in both infections and deaths in the COVID-19 pandemic, the United States has struggled exceptionally last year. It is of great interest to understand the association between various socioeconomic and health related factors to the severity of COVID cases and the related fatality. In this work, we examine how the association patterns changed over the time in two states that were hit hard by the pandemic: New York and Florida. Amid the COVID-19 pandemic, based on CDC data as of Feb 10th, 2021, 81.2% (359,956 deaths) deaths were reported in older population of ages 65 and up in the United States [2]. It is natural to think that when looking at data aggregated at the county level, the case fatality rates should be associated with the variables reflecting the age structures of the counties. Indeed, we observed that the percentage of population ages 65 and older was highly associated with estimated case fatality rate in the beginning of the pandemic. As the pandemic progresses, however, especially in the later stages, the association between the percentage of population ages 65 and older and the estimated case fatality rate turned insignificant in NY state counties. Instead, factors such as median household income became significantly associated with case fatality rate at the county level. In this paper, we summarize changes in the association patterns and explore possible explanations in the states of Florida and New York.

Methods. We examined county level reported counts of COVID-19 cases and deaths in the states of New York and Florida, two of the states that were hit hard by the pandemic. For each state, we divided the data into three seasonal phases based on observed waves of COVID-19 outbreak, and computed case fatality rate for each phase using the raw count of reported deaths divided by the count of reported cases at the county level. For each phase, we used univariate regressions analysis to explore marginal association between each potential covariate and the case fatality rate and used graphical models to further clarify direct or indirect associations in a multivariate setting. Due to the observational nature of publicly available data sets, it is not possible to draw definitive causal conclusions. The aim here is to show how the association changed as the pandemic progressed and present a few possible explanations.

Results. Our association study highlights that the percentage of population ages 65 and older was the strongest associated factor with case fatality rate in the first and second phases in Florida and New York but became insignificant in the third phase. As pandemic progresses especially in the later stage of fall and winter waves, median household income became a significantly associated factor. One explanation of the change in the association pattern is that early on during the pandemic, when counties were unprepared, age structure

*The research of S. C. is supported in part by the NSF grant DMS 1812148.
played the most significant role in determining the case fatality rate, as the pandemic progress, as measures were implemented to protect the elderly population, other socioeconomics factors became more significantly associated with the case fatality rate.