Title

County-level socioeconomic disparities in COVID-19 mortality in the United States

Denys Dukhovnov[a,1] and Magali Barbieri[a,b]

[a] Department of Demography, University of California, Berkeley

[b] French Institute for Demographic Studies, France

Abstract

We estimated COVID-19 mortality rates and the number of years of life lost to the virus during each of the three epidemic waves in 2020 for all US counties grouped into five socioeconomic categories, using preliminary data from the Johns Hopkins University and from the Centers for Disease Control. Our results demonstrate clear inequalities in COVID-19 mortality emerging throughout the course of the epidemic. During March-May 2020, COVID-19 mortality was highest in the quintile of counties most socioeconomically advantaged and lowest in the two most disadvantaged quintiles but the pattern reversed during June-August and widened substantially by September-December, such that COVID-19 mortality rates were three times higher in the bottom quintile of counties than in the top quintile. Similarly, an increasingly clear gradient of years of life lost (YLL) to COVID-19 by socioeconomic quintile emerged through 2020. The population in the top socioeconomic quintile would have lost 2.52 years if the mortality conditions of March through May had persisted throughout the entire year, 1.13 years under the conditions of June-August, and 1.33 years under those of September-December. The corresponding figures for the bottom (least affluent) quintile would have been 1.20, 1.64 and 2.15 years.

Introduction

Prior research has demonstrated the disproportionate mortality impact of recent natural disasters (heat waves, hurricanes, floods, etc.) on the most socioeconomically disadvantaged populations in the United States (1, 2, 3, 4). Analyses based on preliminary data indicate that this is also likely the case for COVID-19. Studies conducted on incomplete data or specific

geographic locations have showed that COVID-19 cases and deaths are concentrated in areas characterized by low median household income and high levels of income inequality among other factors (5, 6, 7, 8). More generally, a growing body of evidence has indicated a higher-than-average level of mortality from COVID-19 in some communities within the US (9, 10, 11, 12, 13, 14).

Most of the factors identified in prior studies to explain differentials in COVID-19 mortality pertain to socioeconomic characteristics. Disadvantaged populations tend to have greater exposure to the virus because they are more likely to be over-represented in essential activities and within the economic sectors least amenable to remote work (15); they rely on public transportation more often than the general population; they live in overcrowded and multi-generational households (16). Once infected with SARS-CoV-2, they are at an increased risk of hospitalization, intensive care, and death. Individuals without paid sick leave or health insurance may delay testing or seeking treatment (17). Those with lower socioeconomic status also suffer from higher rates of pre-existing conditions (e.g. obesity, diabetes, cancer, chronic obstructive pulmonary diseases or cardiovascular diseases) that are associated with increased risks of severe symptoms and deaths from COVID-19 (18, 19). However, due to limitations in the statistical information currently available, no comprehensive study on socioeconomic differentials in COVID-19 mortality has been published as yet. This paper is the first to provide a country-wide assessment of the disproportionate vulnerability of economically disadvantaged communities to COVID-19 mortality.

Discussion

We found evidence of increasing variations in COVID-19 mortality across socioeconomic groupings of counties throughout the course of the epidemic (Figure 1). At first, deaths were concentrated in a small number of counties within each socioeconomic quintile. The virus had not fully spread across the US, affecting primarily densely populated, high economic activity, and cosmopolitan areas. The burden of COVID-19 mortality was borne disproportionately by a small number of wealthy urban areas, especially those located in the Connecticut-New Jersey-New York region, and, to a smaller extent, to more disadvantaged urban areas such as those of

Chicago and Detroit (Figure 3). The 5 New York City boroughs (of which only Bronx county falls into the most deprived quintile) alone accounted for more than 22,000 COVID-19 deaths, at the time representing 21.2% of all deaths nationwide, with approximately 8,000 additional deaths elsewhere across New York state. If they had experienced the mortality conditions of the initial wave of the epidemic during the entire year 2020, Americans in the top socioeconomic quintile would have lost 2.52 years of life expectancy at birth.

However, during June-August 2020, the epidemic contracted in the areas that had initially been most affected and it started spreading through bursts and spurts through the rest of the country. The infection reached further and further into the interior of the country and, especially, towards the poorer South and Southwest. Nearly half of all deaths from this period occurred in Arizona, the Carolinas, Florida, Georgia, Nevada, New Mexico and Texas. The summer also marked a turning point for the West Coast. While a few scattered areas in Northern California and Washington State had made the headlines with initial epidemic clusters during the early months of 2020, case counts began to increase during the summer and particularly after October throughout California, especially in and around the San Francisco and Los Angeles metropolitan areas.

The Northeastern coastal regions effectively curbed the rampant spread of the epidemic following the devastation of the spring, possibly thanks to the selective shutdowns and social distancing policies implemented before and during the summer. By contrast, the dynamic in the Southern states and in the West endured and further increased in magnitude during the last three months of 2020. It was accompanied by a dramatic increase in the number of deaths recorded, particularly in the most socioeconomically disadvantaged rural areas (Figure 3). While the mortality rate from COVID-19 more than halved during the second wave compared to the first wave in the most prosperous areas of the country, rates increased in the two least affluent quintiles. They deteriorated for all quintiles between the second and third wave but much more so in the least affluent quintiles. The changing geography of COVID-19 in September-December 2020 thus amplified the emerging and distinctive pattern of the summer, with an association between COVID-19 mortality and area-level socioeconomic deprivation similar to the pattern observed for mortality from all causes during prior years. The situation became such that

Americans in the bottom quintile of the socioeconomic distribution would have lost 2.15 years of life under the mortality conditions of the period from September to December, compared with 1.33 years for those in the top quintile. This last figure is very close to the number of years of life expectancy at birth lost to COVID-19 in the overall US population for 2020 as a whole. Our estimate is somewhat higher than that from other studies, but most of these studies relied on preliminary data which did not include the increased death counts from the third wave of the epidemic (9, 20, 21). This is for instance the case with the CDC estimate of a 1.0 year loss of life expectancy for the national population, which was based on death records from the first half of 2020 (22). At this time, different sources indicate disparate number of COVID-19 deaths. The John Hopkins dataset we used in this paper was no exception, as it reported close to 25,000 fewer COVID-19 deaths in 2020 than the CDC preliminary data for the corresponding geographic areas. In addition, the homogeneity of misclassification and underreporting in data is questionable. Preliminary research suggests that 30 to 40 percent of the excess deaths might have been correctly (e.g. cancer deaths resulting from delayed care) or incorrectly (e.g. misclassified deaths) attributed to causes other than COVID-19 (23, 24). However, preliminary research also suggests that such effects were disproportionately borne by the most vulnerable segments of the population (25, 26). We thus believe that taking into account the overall mortality excess of 2020 (directly and indirectly attributable to the epidemic) could have only magnified our results.

A limiting factor specific to our analysis was the absence of detailed mortality data at the county-level. To address this concern with the minimum compromise we had to assume a uniform age schedule of COVID-19 mortality (i.e., an age pattern similar across all counties though varying from one week to the next). There are some indications of a greater susceptibility of under-privileged populations (ethnic minorities for instance) to COVID-19 at younger ages (9). Since a death at a young versus old age results in more years of life lost, then this would, again, only further strengthen our findings on the differential impact of the epidemic on the expectation of life at birth by socioeconomic category.

Another important point to consider is that people who died of COVID-19 would have had higher mortality risks than the general population if they had survived, given the greater

lethality of the infection for those individuals with common underlying health problems which also increase the odd of dying from other causes (obesity, hypertension, Type 2 diabetes, cancer and a range of cardiovascular diseases). It is highly likely that such individuals were at increased risks of death than healthier individuals even before they became infected (27). Would we have been able to take these increased risks into account and the fact that people with comorbidities are disproportionately concentrated in the most socioeconomically disadvantaged communities, differences in the number of years of life lost across socioeconomic quintiles would have been lower than we found.

More precise estimates will be possible when complete and detailed data become available but we believe our current findings of the emergence of a strong mortality gradient by socioeconomic category of counties will hold. They suggest that to better control the epidemic and mitigate its impact on the general US population, current policies to strengthen testing, increase vaccination efforts, or invest in public health infrastructure must proactively target the most socioeconomically disadvantaged areas that are bearing the heaviest mortality burden from the epidemic.

References

1. Fothergill A, Peek LA. Poverty and disasters in the United States: A review of recent sociological findings. *Natural hazards*. 2004 May;32(1):89-110.
2. Bolin B, Kurtz LC. Race, class, ethnicity, and disaster vulnerability. *Handbook of disaster research*. 2018:181-203.
3. Dash N. Race and Ethnicity. In: Thomas DSK, Phillips BD, Lovekamp WE, Fothergill A. ed. *Social Vulnerability to Disasters*. 2nd edition. Boca Raton (FL): CRC Press, Taylor & Francis Group. 2013. P. 113-128.
4. Laska S, Morrow BH. Social vulnerabilities and Hurricane Katrina: an unnatural disaster in New Orleans. *Marine technology society journal*. 2006 Dec 1;40(4):16-26.
5. Chen JT, Waterman PD, Krieger N. COVID-19 and the unequal surge in mortality rates in Massachusetts, by city/town and ZIP Code measures of poverty, household crowding,

race/ethnicity, and racialized economic segregation. In *Harvard Center for Population and Development Studies* 2020 May 9. Harvard Center for Population and Development Studies.

6. Chen JT, Krieger N. Revealing the unequal burden of COVID-19 by income, race/ethnicity, and household crowding: US county versus zip code analyses. *Journal of Public Health Management and Practice*. 2021 Jan 1;27(1):S43-56.
7. Mollalo A, Vahedi B, Rivera KM. GIS-based spatial modeling of COVID-19 incidence rate in the continental United States. *Science of the total environment*. 2020 Aug 1;728:138884.
8. Seligman B, Ferranna M, Bloom DE. Social determinants of mortality from COVID-19: A simulation study using NHANES. *PLoS medicine*. 2021 Jan 11;18(1):e1003490.
9. Andrasfay T, Goldman N. Reductions in 2020 US life expectancy due to COVID-19 and the disproportionate impact on the Black and Latino populations. Proceedings of the National Academy of Sciences. 2021 Feb 2;118(5).
10. Cheng KJ, Sun Y, Monnat SM. COVID-19 Death Rates Are Higher in Rural Counties With Larger Shares of Blacks and Hispanics. The Journal of Rural Health. 2020 Sep;36(4):602-8.
11. Cordes J, Castro MC. Spatial analysis of COVID-19 clusters and contextual factors in New York City. Spatial and Spatio-temporal Epidemiology. 2020 Aug 1;34:100355.
12. Fielding-Miller RK, Sundaram ME, Brouwer K. Social determinants of COVID-19 mortality at the county level. PloS one. 2020 Oct 14;15(10):e0240151.
13. Kim SJ, Bostwick W. Covid19 Social Vulnerability and Racial Inequality in COVID-19 Deaths in Chicago. Health education & behavior. 2020 Aug;47(4):509-13.
14. Pro G, Hubach R, Wheeler D, Camplain R, Haberstroh S, Giano Z, Camplain C, Baldwin JA. Differences in US COVID-19 case rates and case fatality rates across the urban-rural continuum. Rural and remote health. 2020 Aug;20(3):6074.
15. Rogers TN, Rogers CR, VanSant-Webb E, Gu LY, Yan B, Qeadan F. Racial Disparities in COVID-19 Mortality Among Essential Workers in the United States. World medical & health policy. 2020 Sep;12(3):311-27.

16. Rollston R, Galea S. COVID-19 and the social determinants of health.

17. Khalatbari-Soltani S, Cumming RC, Delpierre C, Kelly-Irving M. Importance of collecting data on socioeconomic determinants from the early stage of the COVID-19 outbreak onwards. *J Epidemiol Community Health*. 2020 Aug 1;74(8):620-3.

18. Alcendor DJ. Racial disparities-associated COVID-19 mortality among minority populations in the US. Journal of clinical medicine. 2020 Aug;9(8):2442.

19. Ravi K. Ethnic disparities in COVID-19 mortality: are comorbidities to blame?. The Lancet. 2020 Jul 4;396(10243):22.

20. Goldstein JR, Lee RD. Demographic perspectives on the mortality of COVID-19 and other epidemics. Proceedings of the National Academy of Sciences. 2020 Sep 8;117(36):22035-41.

21. Heuveline P, Tzen M. Beyond deaths per capita: comparative COVID-19 mortality indicators. *BMJ open*. 2021 Mar 1;11(3):e042934.

22. Arias E, Tejada-Vera B, Ahmad F. Provisional life expectancy estimates for January through June, National Vital Statistics System Vital Statistics Rapid Release. Report 010. 2021 Feb.

23. Faust JS, Krumholz HM, Du C, Mayes KD, Lin Z, Gilman C, Walensky RP. All-cause excess mortality and COVID-19–related mortality among US adults aged 25-44 years, March-July 2020. *JAMA*. 2021 Feb 23;325(8):785-7.

24. Kiang MV, Irizarry RA, Buckee CO, Balsari S. Every body counts: measuring mortality from the COVID-19 pandemic. *Annals of Internal Medicine*. 2020 Dec 15;173(12):1004-7.

25. Chen YH, Glymour MM, Catalano R, Fernandez A, Nguyen T, Kushel M, Bibbins-Domingo K. Excess Mortality in California During the Coronavirus Disease 2019 Pandemic, March to August 2020. *JAMA Internal Medicine*. 2020 Dec 21.

26. Stokes A, Lundberg DJ, Hempstead K, Elo IT, Preston SH. Assessing the Impact of the Covid-19 Pandemic on US Mortality: A County-Level Analysis. *medRxiv*. 2020 Sep 2.

27. Cairns AJ, Blake DP, Kessler A, Kessler M. The Impact of Covid-19 on Future Higher-Age Mortality. Available at SSRN 3606988. 2020 May 19.

28. U.S. Census Bureau. (2019). *2014-2018 American Community Survey 5-year Public Use Microdata Samples*. Retrieved from https://www.census.gov/programs-surveys/acs/data.html, data downloaded May 13,2020.

29. Singh GK, Siahpush M. Widening socioeconomic inequalities in US life expectancy, 1980–2000. International journal of epidemiology. 2006 Aug 1;35(4):969-79.

30. Singh GK, Miller BA, Hankey BF, Feuer EJ, Pickle LW. Changing area socioeconomic patterns in US cancer mortality, 1950–1998: part I—all cancers among men. Journal of the National Cancer Institute. 2002 Jun 19;94(12):904-15.

31. Singh GK, Miller BA, Hankey BF. Changing area socioeconomic patterns in US cancer mortality, 1950–1998: part II—lung and colorectal cancers. Journal of the National Cancer Institute. 2002 Jun 19;94(12):916-25.

32. National Center for Health Statistics. Restricted-use Detailed Mortality File 2019, as compiled from data provided by the 57 vital statistics jurisdictions through the Vital Statistics Cooperative Program.

33. US Census Bureau, Population Estimates. Annual County Resident Population Estimates by Age, Sex, Race, and Hispanic Origin: April 1, 2010 to July 1, 2019 (CC-EST2019-ALLDATA). https://www.census.gov/data/tables/time-series/demo/popest/2010s-counties-detail.html, data downloaded on February 22, 2021.

34. Johns Hopkins Center for Systems Science and Engineering. Johns Hopkins University 2020. Covid-19 Data Repository. https://github.com/CSSEGISandData/COVID-19. Accessed April 17, 2021

35. National Center for Health Statistics. Center for Disease Control. Provisional Death Counts for Coronavirus Disease 2019 (COVID-19). Weekly Updates by Select Demographic and Geographic Characteristics. https://data.cdc.gov/NCHS/Provisional-COVID-19-Death-Counts-by-Sex-Age-and-W/vsak-wrfu. Accessed April 17, 2021.

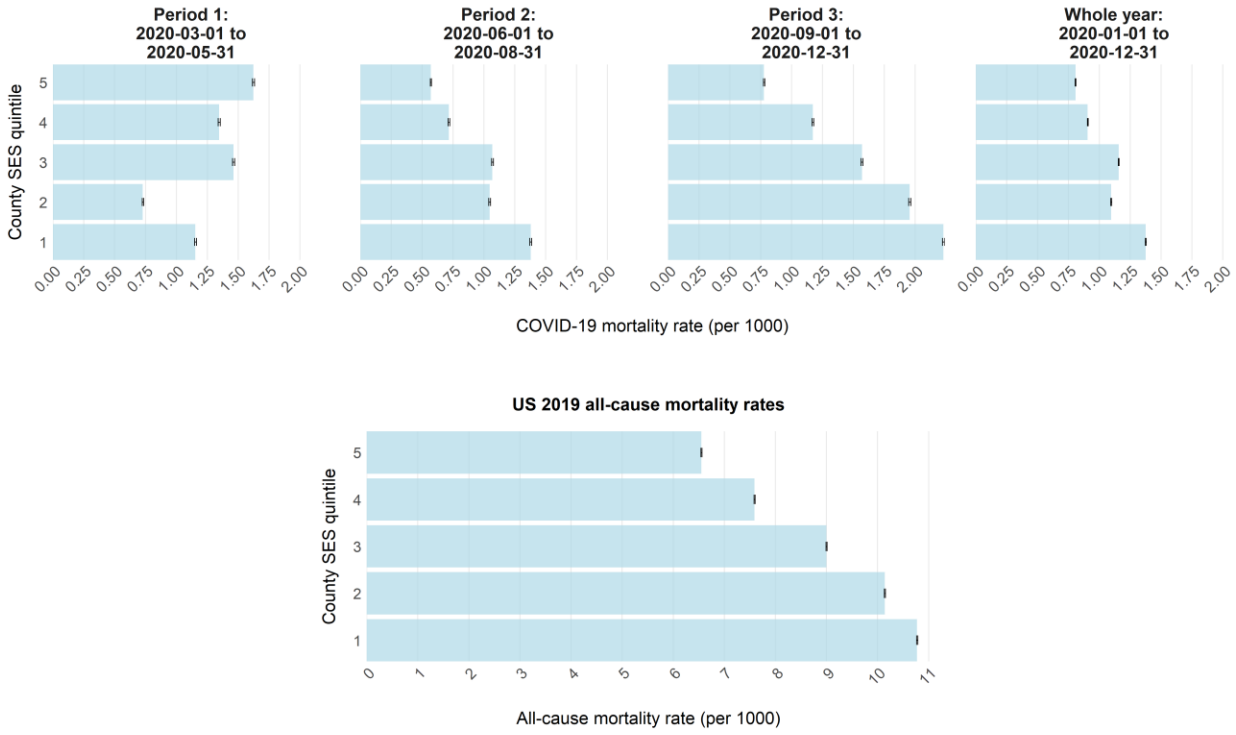
36. Keyfitz N, Frauenthal J. An improved life table method. Biometrics. 1975 Dec 1:889-99.

37. Chiang CL. *The life table and its applications*. Malabar, FL: Krieger; 1984 Dec.

38. Lo E, Vatnik D, Benedetti A, Bourbeau R. Variance models of the last age interval and their impact on life expectancy at subnational scales. Demographic Research. 2016 Jul 1;35:399-454.

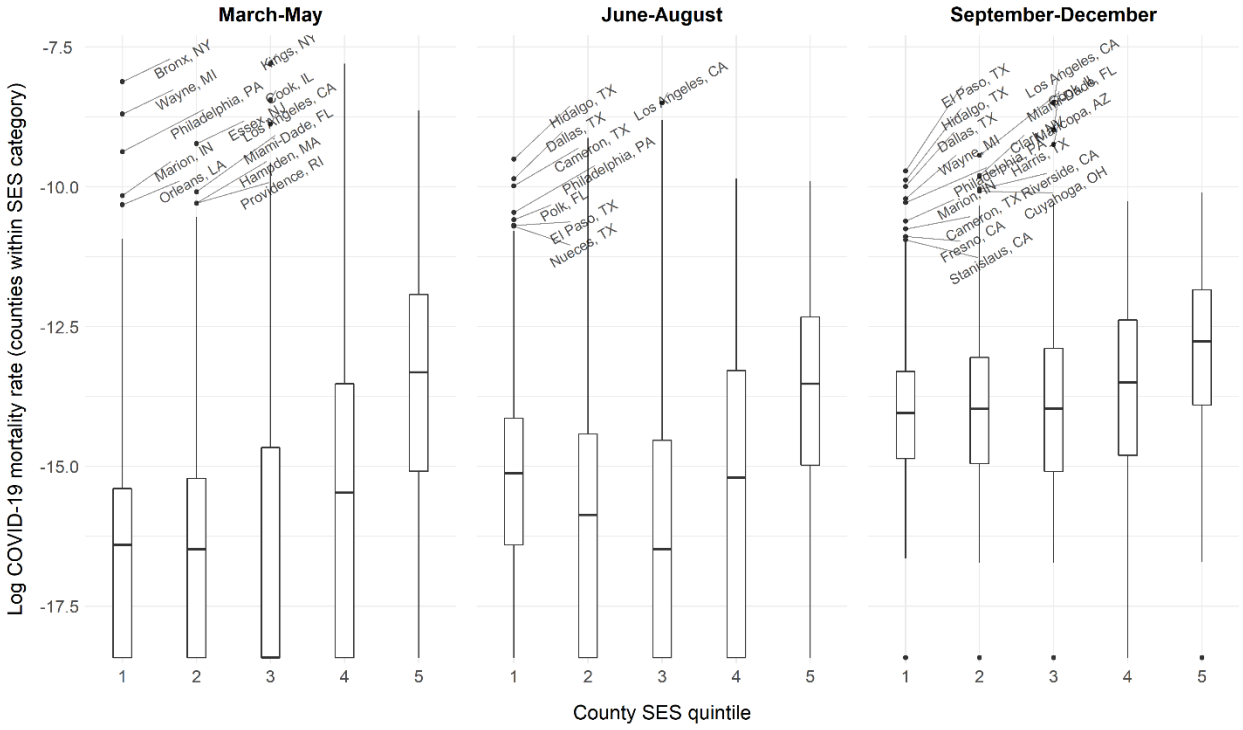
39. Silcocks PB, Jenner DA, Reza R. Life expectancy as a summary of mortality in a population: statistical considerations and suitability for use by health authorities. Journal of Epidemiology & Community Health. 2001 Jan 1;55(1):38-43.

Figure 1: Age-standardized COVID-19 mortality rates (per 1,000) by socioeconomic quintile (1 = lowest SES, 5 = highest SES) of counties and time period in 2020 and in comparison to pre-epidemic 2019 all-cause mortality rates (per 1,000), with 95% confidence intervals.



Note: 95% confidence intervals are narrow due to the large number of deaths in each SES category. All inter-quintile differences are statistically significant.

Figure 2: Natural logarithm of intra-quintile COVID-19 mortality rate by county-level SES quintile (1 = lowest SES , 5 = highest SES) and time period in 2020.



Note: 1) The box includes half of the observations, i.e. those falling between the first and third quartiles of the distribution within each socioeconomic quintile and time period. The thick black line within each box represents the median. The log mortality rate for the third quintile in the first time period overlaps with the bottom of the box. The whiskers above and below the box correspond to the lower and upper limits, called "adjacent values". The adjacent values are calculated from the interquartile range. High and low values not included in this range, i.e. the outliers, are represented by dots positioned below or above the whiskers.

2) Counties were compared against one another within each SES category and time period. Rates shown were therefore based on the aggregate population within each SES category.

Figure 3: County map of COVID-19 deaths by time period in 2020, all counties, low SES counties (quintiles 1-2), and high SES counties (quintiles 3-5).