Research Article

Thomas Patten Krumel Jr.*, Corey Goodrich, Rui Sun and Nathan Fiala

Linking Employment and Death: Measuring the Structural Disparity in COVID-19 Deaths for Non-telework Essential Workers

https://doi.org/10.1515/bejeap-2021-0332
Received September 8, 2021; accepted June 24, 2022

Abstract: The intensity of the early months of the COVID-19 pandemic were a surprise to many people and functioned as an unexpected disturbance where individuals and businesses were slow to adapt their behavior. This event allows us to explore pre-pandemic structural differences in employment and estimate the public health impacts of these first few months of the pandemic. Novel datasets provided by the Connecticut Department of Health and the Massachusetts Department of Health enable us to link deaths to industry and occupation directly at the individual level. A significant number of working-aged people died from COVID-19, with black and Hispanic populations dying at much higher rates. Linking individual deaths with employment, we find that nearly half of these deaths come from people working non-telework essential jobs. Black and Hispanic non-telework essential workers died at a rate 14-percentage points higher than white individuals employed in these same occupations.

Keywords: COVID-19, health disparities, essential workers, ethnic disparities, pandemics, coronaviruses

JEL Classification: I14, J15, J81

The views expressed herein do not necessarily represent the positions or policies of the Wisconsin Department of Workforce Development.

***Corresponding author: Thomas Patten Krumel Jr.**, North Dakota State University, Fargo, USA, E-mail: thomas.krumel@ndsu.edu
Corey Goodrich, Wisconsin Department of Workforce Development, Madison, USA, E-mail: corey.goodrich@dwd.wisconsin.gov
Rui Sun, Mercer University, Macon, USA, E-mail: sun_r@mercer.edu
Nathan Fiala, University of Connecticut, Storrs, USA; Makerere University, Kampala, Uganda; and RWI – Leibniz Institute for Economic Research, Essen, Germany, E-mail: nathan.fiala@uconn.edu

1 Introduction

The abrupt outbreak of COVID-19 and subsequent lockdowns in the US created significant health and economic issues across the country. The early outbreak was especially impactful, and unexpected,[1] for the New England region. At the end of 2020, Connecticut and Massachusetts both still ranked in the top-5 states for COVID-19 deaths per capita, though most of their respective cumulative death totals came from the first few months of the pandemic. The first death from COVID-19 in Connecticut was registered on March 17, 2020, with Massachusetts following the next day. Over the next two and a half months, the two states' combined death toll would increase to just over 10,000. On March 23 and 24, 2020, Connecticut and Massachusetts imposed strict lockdowns in response to rising cases. All businesses and activities not deemed to be essential were closed.

However, the employment impact of the pandemic and lockdowns was not as high as it could have been, given that many people were able to work from home or continue their in-person work on the frontline. These frontline workers included health and emergency personnel, as well as those working in other critical industries. Grocers, construction workers, those in transportation and warehouse industries, as well as many others continued working. Across the two states, 4.5 million people were still employed in the month of May 2020, down from 5.3 million in early March 2020.

Novel datasets provided by the Connecticut Department of Health (CTDoH) and Massachusetts Department of Health (MADoH) allow us to better understand the public health and economic impacts of COVID-19 by directly linking individual deaths to a person's occupation. Since both states analyzed were hit hard, and for the most part unexpectedly so, during the early months of the pandemic, our research exploits the initial surprise disturbance to improve our understanding of structural differences in employment concentration before individuals and businesses were able to adapt their behavior.

We focus our analysis on Connecticut and Massachusetts, since these states and populations were not fully aware of the dangers of COVID-19 or the impact it could have on their economies and so represent an ideal case study for understanding how preexisting structural differences in non-telework essential (NTE) workers may have resulted in a disproportionate impact.[2] Additionally, since

1 The unexpectedness of the outbreak allows us to avoid concerns about differences in ideological behaviors in response to the outbreak.

2 While it is likely that some of our conclusions will generalize to states that had very different pandemic experiences, if a state experienced their peak later in the pandemic after the risks were

we are analyzing two neighbor states that had similar experiences, we do not have concerns about pooling our sample to increase our power. There exist small idiosyncratic differences when examining the individual states, but our main results are consistent across both.

Our main analysis performed over deaths from March 2020 through May 2020 finds that a significant number of working-aged people died from COVID-19 as the primary cause of death, and these numbers are especially pronounced for black and Hispanic individuals. Using these two datasets, we can create a database of nearly all individual deaths in each state that link directly to the job of the deceased, as well as providing detailed demographic information on each individual. We find that just over 55 percent of working aged black and Hispanic deaths come from people employed in NTE jobs,[3] a point estimate that is 38 percent higher than the relative population share of individuals estimated to be working these jobs. This is compared to nearly 37 percent of working age whites who were employed in non-telework essential jobs, or only 24 percent higher than the relative population share of individuals working these jobs. We also find evidence that 83 percent of the evaluated deaths experienced by black and Hispanic workers were from those in NTE occupations. The timing of these impacts is most pronounced in April and May 2020 and taper off in the subsequent months until October 2020. It is important to note that since more aggregated statistics do not allow for this type of linking, we believe that we are among the first to analyze such individual-level data directly relating COVID-19 health outcomes and employment types and we are the first to perform such analysis over an approximate for in person work employment. We conclude our analysis by estimating the combined dollar value for NTE workers under 60 years of age that died from COVID-19 in both states during the first two and half months of the pandemic to be just over $2.3 billion, though with large differences across racial and ethnic groups.

This paper proceeds as follows. In Section 2 we highlight the related literature on the topic. Section 3 discusses the data that we use in this analysis. Section 4 presents the results of the analysis. Section 5 we assess the relative mortality cost using a standard economic calculation. Section 6 discusses some additional mechanisms that could lead to our results. We conclude with a discussion of the implications of our analysis in Section 7.

better understood, then disentangling the impact from non-telework essential employment and behavioral adjustments becomes less straightforward.

3 It should be noted that we are using the definition of non-telework essential jobs as an approximation for frontline employment.

2 Literature Review

Couch, Fairlie and Xu (2020) have documented that the employment impacts were in fact smaller than might have been expected for blacks, compared to previous economic recessions, at least initially. They attribute this to a "favorable industry distribution [that] partly protected black employment during the early stages of the pandemic". Their analysis does not incorporate the disproportionate health impacts blacks experienced during the early stages of the pandemic, especially at younger ages. 53 percent of the initial COVID-19 deaths in Connecticut and Massachusetts from working aged people (we define as below 60 years old)[4] were from blacks and Hispanics, while representing a combined total of just 22 percent of the population in these two states.

The disproportionate health impacts of COVID-19 by race and ethnicity has been well documented in public health research. However, less is known about how this spread happened. In Massachusetts specifically, Figueroa et al. (2020) analyze total COVID-19 cases by town, merging demographic variables from Census data to help explain COVID-19 spread. Because the authors were unable to directly observe who got COVID-19 in those towns, the need for statistical inference was required. They find that, based on the town-level characteristics, "a 10 percentage point increase in the Black population was associated with a 312.3 increase in COVID-19 cases per 100,000, while a 10 percentage point increase in the Latino population was associated with an increase of 258.2 cases per 100,000". This study helps to contextualize our own as they also analyze one of the two states examined in this paper; however, the authors did not identify a mechanism for the higher spread in these communities, in addition to not directly observing their outcome variable. We believe we extend this research significantly.

To our knowledge, no one has been able to directly link the racial differences in health impacts of the pandemic directly to NTE employment at a disaggregated level. Aleligne, Duke, and Ebong (2021) suggests that the workplace was a likely contributing factor for explaining the racial disparities in COVID-19 health outcomes, citing lack of ability to social distance, a higher proportion of individuals living paycheck to paycheck, a relative lack of sick leave, and higher share of essential employment as primary reasons for this difference (though it

4 We use age 60 as our cutoff for our working-age population because of the difficulty in defining individuals who are retired in our sample, discussed further in our data description section. Additionally, by examining individuals over the age of 60, we increase our total deaths significantly, and it becomes even more challenging to disentangle deaths arising from frontline employment from old age. We believe this is a conservative choice, as this decision's impact is a dampening of our major conclusions, thus functioning as an effective lower bound.

is important to mention that the authors do not provide empirical analysis supporting these statements). Papageorge et al. (2021) examine behavioral predictors during the pandemic and find that low-income, less work flexible and lack of outside space are negatively correlated with behaviors that limit the spread of disease and argue their results are consistent with previously found relationships between health and socio-economic status. Using county-level Census data, McLaren (2021) demonstrates that the sharp disparity in COVID-19 mortality peaked in the spring and summer of 2020, dissipating by the fall. McLaren also importantly suggests that the concentration of workers in essential services might help to explain this disparity, though given the use of aggregate data, this is could not be more formally tested in his analysis. Again, using Census data (but this time employing the state as the unit of analysis), Rogers et al. (2020) find a correlation between existing structural differences in essential employment concentration and the racial and ethnic death disparity from COVID-19. Krumel and Goodrich (2021) imply the same at the county by examining the meatpacking industry, with outbreaks driven by the extreme physical proximity of workers in the industry, an industry that employs a disproportionate number of Hispanic immigrants. Hawkins (2020) uses the Current Population Survey (CPS) and concludes that "[p]eople of color were more likely to be employed in essential industries and in occupations with more exposure to infections and close proximity to others. Black workers in particular faced an elevated risk for all of these factors". Combining the CPS and O*NET, Asfaw (2021) finds evidence that black individuals are overrepresented in occupations with a "high potential risk of exposure to disease and infection at work and inability to maintain physical distancing at work"; whereas, Hispanic workers were overrepresented in occupations "where potential risk of inability to work from home was the highest". Dobis, Krumel, and Sanders (2022) argue that in specific cases, "the mix of jobs across industries played a direct role" in COVID-19 spread within persistently poor rural counties, geographic locations that "coincide with high-minority counties in most regions of the country" (Dobis et al. 2021). Finally, Hawkins, Davis, and Kriebel (2021) conduct a similar approach to our analysis in just Massachusetts. They are unable to group deaths by NTE as we do in this manuscript and instead categorize a set of broad occupation categories where essential employment and in-person status can only be inferred. We are thus able to expand on this important work by directly linking NTE employment to deaths outcomes.[5]

5 Additionally, it should be noted that the authors omitted important confounders such as comorbidities, education, and whether the individual died in a hospital from their analysis, as well as grouping individuals by age categories, rather than using the level. Another obvious extension to their work is we include a second state in the analysis, which increases power, and

We expand on the above research by combining the preexisting structural employment conditions, specifically for NTE employment, and health implications of the COVID-19 pandemic, by using individual-level death data over multiple states. The economic impacts have been shown to be large for minorities, though not as large as they have historically been. We find this is likely because racial and ethnic minority groups are placed in high risk, NTE employment, probably directly on the frontline. Concerns about the long-run effects of COVID-19 on economic outcomes for minorities such as unemployment and income loss are important, but also mask that the occupations they were involved in has led to an incredible loss of life.

3 Data

To conduct our analysis, we employ novel datasets provided by the CTDoH and the MADoH. This data is collected from death certificates from both states and includes information on all individuals who died in either state, including their usual occupation.[6] Our analysis excludes non-residents since we are not able to determine if they worked in either state and the data for these individuals is significantly less populated than residents.[7,8] Both of these datasets also contain fully populated information on the characteristics (including personal demographics, such as age, gender, ethnicity, race, and town of residence) of those that died, as well as employment information, such as occupation and industry of usual employment. Related to death itself, we can observe the primary cause of death, all associated comorbidities, the condition under which the individual died, date of death, and location of death.

As mentioned above, the first death from COVID-19 in Connecticut was on March 17, 2020 with Massachusetts registering their first death a day later. By the end of May 2020, these two states had reached a total of 10,155 deaths, of which residents of Massachusetts accounted for nearly two thirds of the deaths. The four months succeeding May saw a substantial decrease in COVID-19 deaths, with 1114 deaths occurring over this time period before the start of a second wave in October

the inclusion of value of statistical life calculations, which puts a dollar estimate on the loss of life.

6 It is important to note that we only know the "usual" occupation of the individual and not necessarily the occupation at the time of death. That being said, it is extremely likely that the occupation reported was the occupation at time of death for most observations.

7 However, if an individual died in Connecticut and was a resident of Massachusetts, or vice versa, they were included in our analysis.

8 With the inclusion of non-residents in our data sample, our results prove to be robust.

(a pattern seen in both states, when evaluated separately).[9] The timing of these deaths allows us to treat the COVID-19 pandemic as largely a surprise disturbance since, for many individuals, the pandemic came about unexpectedly, especially prior to June. Similarly, many occupations and industries likely had delayed responses in implementing social distancing policies (Krumel and Goodrich 2021). This delayed response likely caused increased exposure among frontline workers during the earlier months of the pandemic.

We analyze the mortality rate between non-telework and all other working age individuals and evaluate racial and ethnic disparities in COVID-19 deaths. Our analysis focuses on the three largest racial and ethnic groups (white, Hispanic, and black), which constitute approximately 92% of Connecticut's and Massachusetts's population, respectively (we exclude demographic groups with less than 5% of the total population, as analysis of these small samples will be subject to noise).[10] When evaluating NTE workers, we aggregate Hispanic and black deaths into a single "minority" variable in order to avoid issues with small sample sizes.

3.1 COVID-19 Deaths

To determine the cause of death of the individual, we use the International Classification of Diseases 10th Revision (ICD-10) codes which are provided on both datasets. We then use two different codes to determine if the death was from COVID-19, U07.1 and B34.2. U07.1 is defined as COVID-19, virus identified and B34.2 is defined as coronavirus infection, unspecified site. While B34.2 could include other coronaviruses, it is very likely that the deaths we are measuring from this code are directly related to the COVID-19 pandemic. Specifically, prior to March 2020, dating back to 2015, there were a total of 12 deaths reported with this code, in months after March 2020 there were a total of 142 deaths. Our analysis of COVID-19 deaths is also restricted to deaths that had COVID-19 listed as the primary cause. During March–May 2020 there were 41 deaths in which COVID-19 was listed as a comorbidity but not primary cause of death for individuals under the age of 60. Because of the relatively small number, we exclude COVID-19 as

9 Of the 1114 deaths, 733 occurred in June, 186 in July, 99 in August, and 96 in September.

10 Of the 558 COVID-19 deaths in the months of March, April, and May 2020, 22 were Asian or Pacific Islander and 11 were Native American. Due to the small sample sizes in the available data, it makes sense to exclude both groups due to the inability to verify similar COVID-19 outcomes as blacks and Hispanic. Confirming similar outcomes is important when in our later analysis minority groups are combined.

a comorbidity and focus strictly on observations with COVID-19 as the primary cause of death.

3.2 Employment Data

Our main contribution to the literature is our ability to observe employment information, along with detailed demographic characteristics, connected to each individual death. Our primary analysis of NTE COVID-19 deaths (which we believe we are the first to analyze) is done via occupational and industry information provided by both datasets. This data is collected from an interview after the individual's death with descendants. Since this data is not entered into a standardized form (e.g. North American Industry Classification System codes), the data needed to be hand coded to determine whether the occupation was NTE essential. To determine if the occupation was NTE, we used the New York definition of essential employment, as used in Forsythe et al. (2020), with the important caveat that we considered all manufacturing workers essential in contrast to the Forsythe et al. paper. The inclusion of manufacturing industries within our definition of essential is due to the following change of phrasing within executive orders issued in New York on and after April 9th "[e]ssential manufacturing including…any parts or components necessary for essential products that are referenced within this guidance". While it is possible that some manufacturing firms would have been deemed unessential, it is infeasible to determine essential firms given data constraints. It is important to note however, the Delaware definition of essential, as used in Couch, Fairlie, and Xu. (2020) defines all manufacturing industries as essential. Additionally, we control for whether or not the individual was likely to be in person by using the definition of telework created by Dingel and Neiman (2020) to create our preferred classification of NTE. Due to the non-standardized form of which industry and occupation data was inputted into our dataset, our analysis has the limitation that there exists a small level of uncertainty for a handful of individuals' essential and telework classification. For our preferred classification, we erred on being conservative (i.e. classifying a job as remote, if it was likely moved to telework).[11]

11 We should note that over the time period of March-July 2020 there were 23 individuals in which a classification could not be determined based on the information provided. 19 of these individuals stated some variation of "unknown" for the occupation and industry and thus were dropped from the sample. The remaining 4 individuals had occupation or industry information but were unable to be classified as either essential or telework capable occupations. We categorized these individuals as NTE in order to be conservative in our estimation and to avoid upward bias in our results.

Our analysis of NTE work is restricted to individuals that are in the working age population, which we define to be under the age of 60. The primary reason for this restriction, is that the occupational data does not consistently identify whether an individual is retired in the data. Therefore, there are some concerns that individuals identified over the age of 60 will be retired, when the data reports that they are working.[12]

Since we cannot consistently verify the retirement status of older workers using our data, we make the conscious decision to drop these individuals from our sample. The decision to limit our analysis to individuals under the age of 60 dampened our aggregate results. This is possibly occurring because an individual who is over 60 and working a non-telework essential job likely has more risk factors than a comparable younger individual.

4 Results

This study is motivated by examining two simple figures demonstrating the scale of the racial disparity in COVID-19 deaths in Connecticut and Massachusetts at the beginning of the pandemic. After illustrating how wide this disparity was, especially for the working-age population, we will attempt to isolate this difference's likely mechanism.

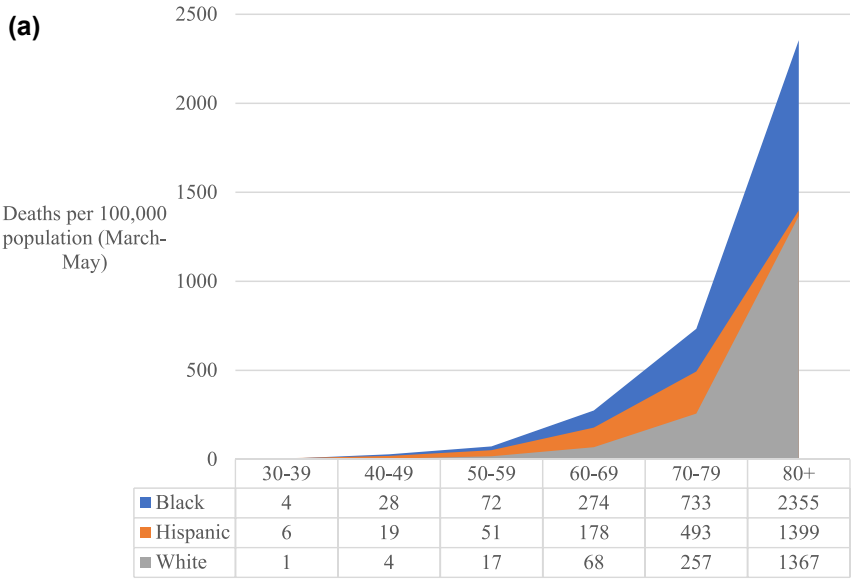
We conclude this section by attaching a monetary estimate of the NTE workers' COVID-19 deaths' total loss using a standard economic calculation.

4.1 Race and Ethnicity Gaps in COVID-19 Deaths

Our data uniquely allows us to observe individual COVID-19 deaths by race, ethnicity, and age. As a first pass at describing our data, it makes sense to aggregate deaths by age categories for the different racial and ethnic groups we are analyzing.

Establishing a baseline of comparison, Figure 1(a) presents the pooled death rates per 100,000 population for both states by race as a relative density from

12 There were over 500 deaths from COVID-19 over March-July 2020 of individuals aged 60–64, of those individuals just 7 were identified as being retired, or just over 1% of the entire sample. Of the nearly 10,500 individuals that died from the pandemic over the age of 60 during the same time period, 146 were listed as retired, well below the retirement rate of either state. Since the data appears to be drastically under counting individuals who are retired, it makes sense to limit our sample to the under 60 population since retirement will be far less prevalent. This also provides another important contrast to Hawkins, Davis, and Kriebel (2021), as they assume that all individuals under 65 were working at the time of death.



	30-39	40-49	50-59	60-69	70-79	80+
Black	4	28	72	274	733	2355
Hispanic	6	19	51	178	493	1399
White	1	4	17	68	257	1367

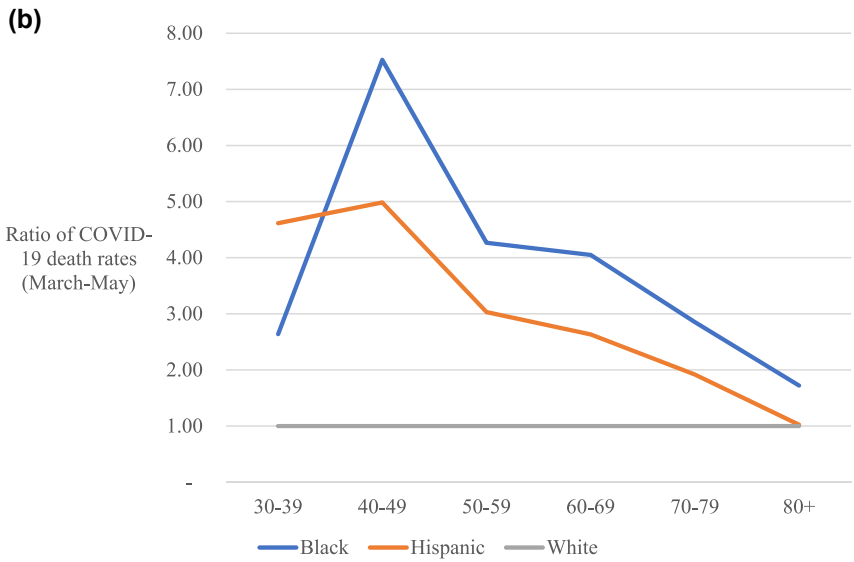


Figure 1: (a) Pooled COVID-19 deaths per 100,000 by race and age category (March–May, 2020). (b) Pooled relative COVID-19 deaths per 100,000 by race and age category (relative to whites, March–May, 2020).

March–May 2020. As Figure 1(a) illustrates, there is an exponential increase in the death rate per 100,000 people as the population ages. Working-age individuals, on the aggregate, appear to have a minimal risk of dying from COVID-19, with a significant jump occurring for individuals over 70. This result is consistent with what is known about the susceptibility to the virus for elderly populations. Given the magnitude of deaths in the oldest age bracket in Figure 1(a), it is difficult to observe the extent to which the racial disparity exists in the younger populations because of the scaling used.

Figure 1(b) presents the relative differences between black and Hispanic individuals compared to whites from March–May 2020. Under the age of 60, depending on the specific age category, black individuals are between two and a half and eight times more likely to die from COVID-19 than whites, as a share of their relative population. Hispanics are more than four and a half times more likely to die from COVID-19 between the ages of 30 and 49, leveling off to three times for the remainder of our working-aged sample. Working up the age categories, a convergence occurs. Minorities are still dying at higher rates than whites, but in the 80 and older category, the difference is less than two times higher for blacks and is negligible for Hispanics.

Combined, Figure 1(a) and (b) suggest that older people die from COVID-19, regardless of race or ethnicity, but minorities are significantly more likely to die from COVID-19 at the younger age brackets. This result should raise the obvious question as to why working-age minority individuals are dying at far higher rates in these two states at the start of the pandemic.[13]

4.2 The Relationship Between Non-telework Essential Employment and COVID-19 Deaths

Given that individuals in our sample under the age of 60 are dying from COVID-19 at rates significantly less than 1 percent per their population, we are effectively using deaths to approximate for exposure levels. If a group of individuals is exposed to COVID-19 at a higher rate, then it follows that they should also be dying at a higher rate, statistically speaking. Assuming that there are not systematic differences that make certain groups more likely to die after contracting the virus, higher exposure rate could help to explain the disparity in deaths for working-aged individuals that we are observing. Our hypothesized mechanism for increased

13 Almost identical results are produced using a national dataset in a Brookings analysis titled "Race gaps in COVID-19 deaths are even bigger than they appear", suggesting our additional analysis might not be unique to the two states we are examining and could be generalizable. Though, additional research should be conducted to confirming this.

exposure to COVID-19 is through employment, specifically in person essential jobs.

As described above in the data section, the individual death record files provided by the CTDoH and the MADoH also include the employment of the deceased. This feature allows us to examine whether there is a correlation between non-telework employment and death from COVID-19 at the individual-level. This is an important extension on the previous literature. When McLaren describes the decision to use county-level data, he states, "[t]he ideal would be individual data but that is not publicly available, so the present goal is to work with the finest possible geographic units". We start our analysis by examining the percent deaths from COVID-19 that we classify as working an NTE job.

As is illustrated in Table 1, there were just over 500 total deaths under the age of 60 from COVID-19 in Connecticut and Massachusetts from March 17 through the end of May 2020. This figure represents around 5 percent of total COVID-19 deaths during this timeframe. According to the 2019 American Community Survey (ACS) Census estimates, there are in excess of four times the number of working-aged whites in these two states than there are blacks and Hispanics; however, blacks and Hispanics have a nearly identical number of deaths from the virus for individuals under 60 (blacks and Hispanics have 11 percent higher deaths under 60, overall) compared to whites, and under the age of 50, they actually have 222 percent more total deaths.[14] There exists a disparity when examining just the levels of death, without factoring in relative population size, which provides context as to why there is such a large disparity when population adjustments are made in the previous figures.

Of the working-aged individuals who died from the virus, we find that just over half of blacks and Hispanics were employed in an NTE occupation. Our results are robust across age categories, being bounded between 52 and 59 percent. For working-aged whites, we see a smaller share of deaths being frontline workers; however, 37 percent of these deaths were also attributed to an individual who was working in an NTE position, being bounded between 31 and 38 percent across age categories.[15] These results are suggestive that NTE

14 We aggregate black and Hispanic working age individuals into a single category for the remainder of the analysis to increase our sample size.

15 A second data limitation we faced is that retirement is not well reported in our data. Employment in our dataset is identified based on reporting by the descendants at the time of death. It is entirely possible that rather than informing that the deceased was retired, they communicate the former career. A restriction we implemented to provide additional confidence in the reported results is we categorized individuals who died in a nursing care facility as not in the workforce (93 percent of our dropped observations are from the 50–59 age category, which is unsurprising). Our results change very marginally for black and Hispanic from the full sample

Table 1: Deaths by age, race and ethnicity, and frontline status (March to May, 2020).

Ages	Total covid deaths		Non-telework essential covid deaths		Percent non-telework essential		Employed non-telework essential	
	Black and Hispanic	White	Black and Hispanic	White	Black and Hispanic	White	Black and Hispanic	White
18–39	31	14	16	5	51.61%	35.71%	38.74%	30.44%
40–49	71	32	42	10	59.15%	31.25%	43.03%	29.64%
50–59	163	193	88	73	53.99%	37.82%	39.53%	28.14%
18–59	265	239	146	88	55.09%	36.82%	39.83%	29.64%

workers, in general, were more susceptible to dying from COVID-19 in the younger age categories, regardless of race or ethnicity during the first months of the pandemic.

We can further confirm the conjecture that NTE workers were more susceptible to die from COVID-19 by using the relative population, with the ACS's 2019 industry and occupation data. Applying the same essential and telework definition as used in our deaths data, we can provide employment estimates. Estimates for the number of black and Hispanic individuals employed in NTE positions in Connecticut and Massachusetts during the pandemic is 40 percent of their total population, or 330,949 total individuals. This is compared to 30 percent of whites who were estimated to be employed in non-telework essential jobs in these two states, or 1,368,413 total individuals. Against the estimated share of the relative population employed in NTE jobs, working-age blacks and Hispanics died at a 38 percent higher rate than would be predicted. In contrast, white NTE workers died at a 24 percent higher rate than would be predicted. While both are in excess of their relative population, the difference being more pronounced for black and Hispanic NTE workers further implies these individuals might have been disproportionately on the frontline, though we are unable to confirm this is the case internally using our data. However, a New York City Comptroller Report suggests this conjecture has merit (Stringer 2020).

We view the estimated large gap between black and Hispanic NTE workers and whites as our primary result. By taking the difference in the share of COVID-19 working-aged deaths attributable to non-telework employment, our estimates suggest that blacks and Hispanics are 14-percentage points more likely to die from COVID-19 in the working-age category employed also employed in NTE occupations. We believe our analysis helps to contextualize the large observed disparities in deaths in the younger age categories which decrease when working up the age categories. Older individuals were dying at high rates, regardless of race or ethnicity. Younger individuals were dying at much lower rates, meaning those who did die likely had a higher likelihood of exposure to the virus. The occupational distribution that insulated certain individuals from losing employment at the

---

and more significantly for whites (a total of 42 deaths are from nursing homes, with 66.67% of these deaths being white individuals). It is important to note that our subsequent results are not impacted by the restriction and are robust when examining the full sample. A second sensitivity analysis we perform is to drop everyone over 54 from our analysis because of the possibility that we cannot accurately identify who is retired, and our results become more pronounced, since we observe more significant disparities in the younger population. Since we observe greater disparities existing within younger population, including retired individuals in our sample, if anything, seems to bias our results toward zero.

pandemic height also likely drove the higher death rates experienced by minority groups.[16]

Extending this analysis, we can further infer to what degree the elevated death rates for blacks and Hispanics are explained by occupations. From our Census data, we know how many NTE workers there are in the population and can therefore measure the overall death rate among all NTE workers and all non-NTE workers in the data. We also know the fraction of blacks and Hispanics workers and white workers in NTE occupations. These values allow us to explain the portion of the excess mortality of blacks and Hispanics by occupation alone. Performing this calculation, we estimate that deaths are elevated for NTE blacks and Hispanics individuals by 12 percent, which is line with our other previously presented calculations.

A different way to approach this same question is to compare the death rates between blacks and Hispanics and white NTE workers. An estimated 0.04 percent of NTE blacks and Hispanics workers died from COVID-19 during these initial months, compared to 0.006 percent of NTE whites. NTE black and Hispanic workers were 6 times more likely to die from COVID-19 compared to white NTE workers than their population share would suggest. For the entire population of black and Hispanic workers, 0.03 percent died from COVID-19 compared to 0.004 percent of all white workers. This value estimates that black and Hispanic workers were 7 times more likely to die from COVID-19 than white workers. Combining these two calculations suggests that 83 percent of the evaluated deaths experienced by black and Hispanic workers was from those in NTE occupations.

5 Value of Statistical Life Calculations

Value of statistical life calculations (VSL) are typically used in cost-benefit analyses to value mortality risks. This calculation allows government agencies to evaluate proposed government regulations (or deregulations) using an agreed-upon economic parameter. In this manuscript, we are not attempting to estimate any benefits, but we do believe it is crucial to estimate the disproportionate burden placed on individuals who continued to work on the frontline at the pandemic's height. While stating that more than 500 working-age lives were lost to COVID-19 in just two small states working NTE positions over a two-and-a-half-month span at the pandemic's initial apex is striking, attaching a monetary value to this loss makes this more impactful.

16 Appendix A examines some robustness checks to enhance confidence in the correlation explored in this manuscript.

In this section, we will calculate the VSL to assess the relative mortality cost. The value of VSL is reported in dollars per statistical life, which can be estimated using the wage premiums that workers received for reducing occupational mortality risks. Due to the individual level heterogeneity, VSL, by its nature, is not constant over time and across individuals. To address this issue and get better estimates of it for different age groups in our sample, we adopt a widely used approach that uses the implied relationship between the value of a statistical life year (VSLY) and VSL to estimate the age group-specific VSL values. The implied VSL can be calculated based on the following relationship:

$$\text{VSLY} = \frac{r\text{VSL}}{1-(1+r)^{-L}}$$

where r is the discount rate that people use to link VSL and the trajectory of VSLY and L is the remaining life expectancy. In our calculations, we use 3%, i.e. the real interest rate, for the discount rate and the remaining life expectancy data from Arias, Xu, and Kochanek (2019). For the baseline value of VSLY, we use the estimate reported in Viscusi and Hersch (2008), where they estimate an average VSLY of $496,000 in 2019 dollars. This value is comparable to the values used by the U.S. Department of Health and Human Services. Table 2 summarizes the age-group specific VSL values.

Table 2 estimates that the cost of COVID-19 in terms of NTE workers lost in Connecticut and Massachusetts from March–May 2020 is just over 2.3 billion dollars. For the level amount, nearly 1.48 billion of the total VSL lost is contributed from blacks and Hispanics and just under 0.85 billion comes from whites. Before adjusting for population, the relative cost of mortality for our black and Hispanic category is 74 percent higher than whites. This number jumps to 5 times higher when controlling for population sizes.

Our VSL estimates mimic the descriptive analysis of overall deaths per 100,000 population presented in Figure 1(a) and (b). The differences are most pronounced in the younger age categories, with significant convergence

Table 2: Value of statistical life calculations for frontline workers who died from COVID-19 (March to May, 2020).

	Blacks and Hispanics	Whites	Total
18–39	201,346,595.00	62,343,538.00	263,690,133.00
40–49	459,021,234.00	109,290,770.00	568,312,004.00
50–59	815,245,288.00	676,283,023.00	1,491,528,311.00
18–59	1,475,613,117.00	847,917,331.00	2,323,530,448.00

occurring as we work up the age distribution. Nearly 90 percent of the total contribution to our VSL calculation is coming from individuals over the age of 40, but in this age category, blacks and Hispanics are estimated to have lost over 3 times as much as whites. This number increases to more than 4 times more in the 40–49 age category. In our oldest age category, there is a much less stark difference between our black and Hispanic category and whites, with blacks and Hispanics estimated to have lost 21 percent more.

6 Additional Mechanisms

Hawkins, Davis, and Kriebel (2021) listed several limitations to their study. These were primarily related to their empirical analysis and they made no mention of other mechanisms that could not be ruled out other than their occupational categories. Their listed limitations included an inability to observe comorbidities, concerns over the accuracy of the employment data in the death certificates data, and sample size. In our appendix, we have reduced concerns on all of these by including analysis over comorbidities, dropping likely retired individuals from our sample and by adding a second state to increase our sample size.

While we have improved confidence over the empirical analysis from the previous literature, we are missing an important counterfactual: those that got COVID-19 but did not die. This means we are limited in understanding all of the potential mechanisms for our findings. In addition to the hazards of being an NTE worker during a pandemic, we present here four additional potential mechanisms for the deaths we observe that we cannot fully refute.

First, blacks and Hispanics could be less healthy than whites in general. We present evidence in our appendix that suggests this is not the case, but we cannot fully rule out the possibility that COVID-19 disproportionally killed the unhealthy NTE workers or that the unhealthy became NTE workers.

Second, minorities could have less access to adequate healthcare. Most blacks and Hispanics dying in these states lived in cities, which have better equipped hospitals (Orgera et al. 2020).[17] However, we don't know the amount of time a person waited for treatment. As discussed in the appendix, there was no specific COVID-19 treatment early in the pandemic, meaning access to a hospital with ventilators was vital for severe cases of COVID-19 (NIH, 2020).

Third, we don't know the immigration status of people in our data. If people were worried about their immigration status, they may delay treatment. Again,

17 A study of health infrastructure in rural and metro locations found that while there was similar hospital beds per capita, metro areas had slightly over 1 more ICU bed per 10,000 individuals.

treatment early in the pandemic was largely dependent upon ventilator access which would have likely been a late-stage option for treatment, and the population of undocumented residents in Connecticut and Massachusetts is very small at about 3% in each state.

Finally, as we mention above, we are attempting to estimate the direct effect of COVID-19 and NTE work. There is also an indirect effect we are not able to pick up, which is if they died from COVID-19 as results of contracting the virus from someone else in the household or they interact with who works in an NTE job. Since we are excluding everyone over the age of 60 from our primary calculations, we are hypothetically estimating a lower bound of the overall effect.

7 Discussion

We believe that our study is the first to directly disentangle the relationship between NTE employment and working-age deaths from COVID-19. Previous research has explored the economic impacts of the pandemic on different ethnicities and races, suggesting "a favorable industry distribution partly protected black employment during the early stages of the pandemic" (Fairlie, Couch and Xu 2020). We believe that our results help to contextualize why there was such a large health disparity from COVID-19, especially at the beginning of the pandemic, a result that could only be inferred and suggested at using aggregate data previously. A slightly favorable industry distribution from an economic standpoint plausibly resulted in higher death rates from increased exposure. Once behavioral adjustments were made, our results also suggest that NTE workers were exposed to the virus less frequently than they were initially, at least over the remainder of our study period, and curbed the number of working-age individuals who died from the virus significantly. However, even where measurable improvements have been made, NTE workers still appear to be at higher risks than the general population.

Appendix A: Robustness Checks

In the main manuscript, we have demonstrated that there exists a large death disparity in COVID-19 between blacks and Hispanics and whites from Marchi – May 2020, which is especially pronounced in the working-age population. We have shown that there appears to be a strong correlation to working non-telework essential jobs in the COVID-19 deaths for this age bracket, which is a possible mechanism for this disparity. In this appendix, we will examine a set of alternative hypotheses that could plausibly be driving our results.

Table A1 provides our summary statistics by non-telework essential status and ethnic and racial category. The table is broken into three panels, demographic, comorbidities, and education indicators.[18] It also tests whether the estimated means for blacks and Hispanics employed in non-telework essential occupations are statistically different from two sets of comparison groups. The first comparison group is whites who were also working in non-telework essential occupations at the time of death. The second is blacks and Hispanics under the age of 60 who either worked telework or non-essential occupations at the time of death. In conducting this analysis, we want to show the areas of imbalance between black and Hispanic non-telework essential workers from similarly situated white workers, and against the same demographic background, but removing non-telework essential status. Combined, we believe that these two comparison groups generate an imperfect counterfactual as we can hold constant both the ethnic and racial background one and non-telework essential status in the other.[19]

Starting with the demographic panel, we find evidence that black and Hispanic non-telework essential workers are younger than white workers by an average of three years, and this difference is statistically meaningful. Given what is known about the virus, the fact that we observe black and Hispanic workers being younger, on average, further supports the conjecture that these individuals are being exposed to COVID-19 at a higher rate.[20] It could also be that non-telework essential workers are less healthy on average; however, we do work to rule this out in Panel b. To preview these results, we find that there exist no statistically significant evidence that black and Hispanic individuals who work in non-telework essential occupations are less healthy than those who can either telework or are employed in a non-essential industry. Furthermore, we find no evidence that black and Hispanic non-telework essential workers are unhealthier than non-telework essential whites.

Moving next to gender, we find no difference between black and Hispanic non-telework essential workers and white non-telework essential workers; however, we do find evidence that employed individuals who died from COVID-19 were overwhelmingly male, regardless of ethnic or racial background. This result is also consistent with what is known about the pandemic. Woman were more

18 The education variable was only available in the Massachusetts dataset, so panel c is a subset of our sample.

19 We believe this is another improvement over Hawkins, Davis, and Kriebel (2021) since the authors made no attempt to develop a viable comparison group in their analysis.

20 It could be the case that black and Hispanic workers have a higher death rate from COVID-19, however, literature such as Mackey et al. (2021) finds no evidence of a significant difference in death rates between racial and ethnic groups who have contracted COVID-19.

Table A1: Comparison of means between frontline workers by ethnic and racial groups (March to May, 2020).

Covariates	Non-telework essential		Not non-telework essential		Differences	
	Black and Hispanic (1)	White (2)	Black and Hispanic (3)	White (4)	(1)–(2)	(1)–(3)
Panel a: Demographic variables						
Age	50.05	53.28	51.54	52.48	−3.23***	−1.49
Male	73.29%	68.18%	67.80%	62.12%	5.11%	5.49%
Hospital	86.99%	96.59%	91.53%	87.88%	−9.60%**	−4.54%
Panel b: Comorbidities						
Average number of comorbidities	3.76	3.92	3.81	3.79	−0.16	−0.05
Elixhauser v1	1.53	1.51	1.29	1.27	0.02	0.25
Elixhauser v2	1.75	2.03	1.32	1.73	−0.29	0.42
Charlson	0.79	0.84	0.71	0.59	−0.05	0.08
Panel c: Education variables						
Less than high school	17.72%	1.82%	15.63%	4.65%	15.90%***	2.10%
High school diploma	54.43%	74.55%	43.75%	34.88%	−20.12%**	10.68%
Some college or two-year degree	20.25%	9.09%	12.50%	27.91%	11.16%*	7.75%
Bachelor's degree of higher	7.59%	14.55%	28.13%	32.56%	−6.95%	−20.53%***
Observations (*Panel a and b*)	146	88	59	66		
Observations (*Panel c*)	79	55	32	43		

disaffected economically from the pandemic, both in terms of lost employment impacts and lack of available of childcare and were, as a result, insulated from COVID-19 exposure (Collins et al. 2021).

We use whether an individual died at a hospital as an approximation for healthcare access. Early in the pandemic, there wasn't any specific treatment options for COVID-19 and as a result treatment was less effective than it is currently, so access might be a lower order determinant of whether an individual survived (NIH, 2020). We find evidence that black and Hispanic non-telework essential workers are statistically different from white non-telework essential workers, suggesting black and Hispanic individuals might have had less access to healthcare than similarly employed whites. Though, we find no statistical significance on hospital deaths between non-telework essential minorities and telework and non-essential black and Hispanic individuals.

As alluded to above, one possible explanation for the prevalence of non-telework essential COVID-19 deaths is that non-telework essential workers systematically have worse health conditions than individuals who are either able to telework or are non-essential workers, not that they had a higher exposure to the virus. To fully rule this alternative hypothesis out, we would need data on individuals who did not die from COVID-19 to run a regression where our left-hand side variable is whether the individual died or not.[21] Unfortunately, we do not have information on the population of individuals who did not die in our data, so this analysis is untenable. To attempt to make a valid comparison that helps to refute this alternative, we evaluate several possible hypothesis tests given our data using four different measures to proxy for the general health of the individual. The four measures are as follows: average number of comorbidities, Elixhauser index using AHRQ weighting, Elixhauser index using VW weighting, and the Charlson index. For the later three of these methods, we use the comorbidities R package to estimate our results. The AHRQ method uses methodology developed by Moore et al. 2017, the VW method was developed using definitions from van Walraven et al. 2009, and the Charlson index is derived from Charlson et al. 1987. These three indices allow us to not only measure the number of comorbidities present at death, but the severity of the comorbidities. For specifically the AHRQ method, the weighting of the comorbidities ranges from a 14 for metastatic cancer to a −7 for drug abuse. The negative score is reflective of "a protective relationship with in-hospital mortality" (Moore et al. 2017).

21 This sentiment is expressed in McLaren (2021) by stating "[i]deally, one would have a large sample of COVID-19 patients and non-COVID-19 patients with full information about employment, education, occupation, income, and so on for each individual but that is not available".

Table A1 *Panel b* presents the results from the 4 different methods we employed among our black and Hispanic and white sample, which is further divided into individuals that are employed in non-telework essential employment and those who are employed in telework or non-essential occupations. Our analysis is of every COVID-19 death under the age of 60 for the months of March to May. We find that in our total sample black and Hispanic non-telework essential workers have on average 3.76 total comorbidities, with telework or non-essential workers being slightly higher at 3.81. When we evaluate the difference between different racial and ethnic groups, we find that non-telework essential workers for the black and Hispanic population have on average less comorbidities than white individuals, however, this difference is not statistically significant.

The second, third, and fourth lines of Table A1 *Panel b* present two different methods of weighting for the Elixhauser index, AHRQ and VW, and the Charlson index. For all three of these methods, we find no statistically significant evidence that black and Hispanic non-telework essential individuals had worse health conditions than similarly employed whites. We also find no significant evidence that there are health differences between non-telework essential blacks and Hispanics and telework and non-essential blacks and Hispanics.

The results from all four of these methods suggest that black and Hispanic non-telework essential workers do not have systematically different health conditions from telework and non-essential individuals at the time of death from COVID-19 or from white individuals employed in similar occupations. We are however restricted to death data, and therefore cannot make any inference on the general health of non-telework essential workers as juxtaposed to individuals who are employed in telework or non-essential occupations.

Another alternative hypothesis is that rather than estimating the effect of non-telework essential employment we are indirectly picking up socioeconomic status. While we do not have a perfect solution to rule out this alternative, our data allows us to provide suggestive evidence to the contrary. In addition to employment, the individual deaths from Massachusetts also contains the education attainment of the deceased. Unfortunately, this variable is not available in the Connecticut data, so the analysis over this variable is from a subsample of our population. We break this variable into four indicators based on relative sample size and comparability. We combine less than high school into a single category, some college and either an Associates or certificate into another category, and bachelors and graduate education into the third. High school diploma or GED is the only category that we leave uncombined.

We find evidence of significant imbalance occurring across our set of education indicators. Succinctly, blacks and Hispanic non-telework essential workers appear to be less well educated when compared to similarly employed whites,

having a higher share of individuals with less than a high school diploma or GED and a lower percentage of high school graduates. Of the individuals who have received a bachelor's degree or more, there is no statistical difference between the two groups. Additionally, black and Hispanic individuals have a higher and statistically significant rate of some college education compared to whites. Black and Hispanic telework and non-essential workers appear to be better educated when compared to the individuals in non-telework essential positions with the same ethnic and racial background. This can be attributed to individuals who are able to telework being in high-education occupations (Dingel and Neiman 2020). We observe a similar pattern when performing this comparison for white individuals. Taken together, these results do suggest that socioeconomic status might be different between our comparison groups, and as such, it makes sense to control for non-telework essential status conditional on education attainment.

While this analysis does not fully rule out these mechanisms driving our results, it does push back against these alternatives.[22] Even with the inability to assign causality in this research, our descriptive findings are an important contribution to the literature.

References

Aleligne, Y. K., A. Duke, and I. A. Ebong. 2021. "Racial Disparities in Coronavirus Disease 2019 (COVID-19) Outcomes." *Current Opinion in Cardiology* 36 (3): 360–6.

Arias, E., J. Xu, and K. D. Kochanek. 2019. "United States Life Tables, 2016." In *National Vital Statistics Reports*, 68. Hyattsville, MD: National Center for Health Statistics.

Asfaw, A. 2021. "Racial Disparity in Potential Occupational Exposure to COVID-19." *Journal of Racial and Ethnic Health Disparities* 1–14. https://doi.org/10.1007/s40615-021-01110-8.

Charlson, M. E., P. Pompei, K. L. Ales, and C. R. MacKenzie. 1987. "A New Method of Classifying Prognostic Comorbidity in Longitudinal Studies: Development and Validation." *Journal of Chronic Diseases* 40 (5): 373–83.

Collins, C., L. C. Landivar, L. Ruppanner, and W. J. Scarborough. 2021. "COVID-19 and the Gender Gap in Work Hours." *Gender, Work and Organization* 28: 101–12.

Couch, K. A., R. W. Fairlie, and H. Xu. 2020. "Early Evidence of the Impacts of COVID-19 on Minority Unemployment." *Journal of Public Economics* 192: 104287.

Dingel, J. I., and B. Neiman. 2020. "How Many Jobs Can Be Done at Home?" *Journal of Public Economics* 189: 104235.

Dobis, E. A., T. Krumel, and A. Sanders. 2022. "Persistently Poor Rural Counties Experienced Lower Employment Impacts from the Coronavirus (COVID-19) Pandemic but Have Higher Cumulative COVID-19 Case Rates." In *Amber Waves: The Economics of Food, Farming,*

22 Our results are also robust to a conditional correlation that was removed for space considerations.

Natural Resources, and Rural America. Washington D.C.: Economic Research Service, USDA.

Dobis, E. A., T. P. Krumel, J. Cromartie, K. L. Conley, A. Sanders, and R. Ortiz. 2021. *Rural America at a Glance: 2021 Edition (No. 1962-2021-3215)*. Washington D.C.: Economic Research Service, USDA.

Figueroa, J. F., R. K. Wadhera, D. Lee, R. W. Yeh, and B. D. Sommers. 2020. "Community-Level Factors Associated with Racial and Ethnic Disparities in COVID-19 Rates in Massachusetts: Study Examines Community-Level Factors Associated with Racial and Ethnic Disparities in COVID-19 Rates in Massachusetts." *Health Affairs* 39 (11): 1984–92.

Forsythe, E., L. B. Kahn, F. Lange, and D. Wiczer. 2020. "Labor Demand in the Time of COVID-19: Evidence from Vacancy Postings and UI Claims." *Journal of Public Economics* 189: 104238.

Hawkins, D., L. Davis, and D. Kriebel. 2021. "COVID-19 Deaths by Occupation, Massachusetts, March 1–July 31, 2020." *American Journal of Industrial Medicine* 64 (4): 238–44.

Hawkins, D. 2020. "Differential Occupational Risk for COVID-19 and Other Infection Exposure According to Race and Ethnicity." *American Journal of Industrial Medicine* 63 (9): 817–20.

Krumel, T., and C. Goodrich. 2021. *COVID-19 Working Paper: Meatpacking Working Conditions and the Spread of COVID-19. No. 1962-2021-3009*. Washington D.C.: Economic Research Service, USDA.

Mackey, K., C. K. Ayers, K. K. Kondo, S. Saha, S. M. Advani, S. Young, H. Spencer, M. Rusek, J. Anderson, S. Veazie, M. Smith, and D. Kansagara. 2021. "Racial and Ethnic Disparities in COVID-19–Related Infections, Hospitalizations, and Deaths: A Systematic Review." *Annals of Internal Medicine* 174 (3): 362–73,

McLaren, J. 2021. "Racial Disparity in COVID-19 Deaths: Seeking Economic Roots with Census Data." *The B.E. Journal of Economic Analysis & Policy* 21 (3): 897–919.

Moore, B. J., S. White, R. Washington, N. Coenen, and A. Elixhauser. 2017. "Identifying Increased Risk of Readmission and In-Hospital Mortality Using Hospital Administrative Data." *Medical Care* 55 (7): 698–705.

Orgera, K., D. McDermott, M. Rae, G. Claxton, W. Koma and C. Cox. 2020. "Urban and Rural Differences in Coronavirus Pandemic Preparedness." *Health System Tracker* (cited March 22, 2021). https://www.healthsystemtracker.org/brief/urban-and-rural-differences-in-coronavirus-pandemic-preparedness/.

Papageorge, N. W., M. V. Zahn, M. Belot, E. Van den Broek-Altenburg, S. Choi, J. C. Jamison, and E. Tripodi. 2021. "Socio-demographic Factors Associated with Self-Protecting Behavior during the Covid-19 Pandemic." *Journal of Population Economics* 34 (2): 691–738.

Rogers, T. N., C. R. Rogers, E. VanSant-Webb, L. Y. Gu, B. Yan, and F. Qeadan. 2020. "Racial Disparities in COVID-19 Mortality Among Essential Workers in the United States." *World Medical & Health Policy* 12 (3): 311–27.

Stringer, S. M. 2020. "New York City's Frontline Workers." In *Report*. New York: Bureau of Policy and Research.

van Walraven, C., P. C. Austin, A. Jennings, H. Quan, and A. J. Forster. 2009. "A Modification of the Elixhauser Comorbidity Measures into a Point System for Hospital Death Using Administrative Data." *Medical Care* 626–33. https://doi.org/10.1097/mlr.0b013e31819432e5.

Viscusi, W. K., and J. Hersch. 2008. "The Mortality Cost to Smokers." *Journal of Health Economics* 27 (4): 943–58.