

Supplementary Material

Figure S1: Flowchart detailing sample exclusions

Figure S2: US County Map showing geographic distribution of sample counties

Figure S3: County-level factors associated with higher or lower than predicted excess mortality using a Negative Binomial model

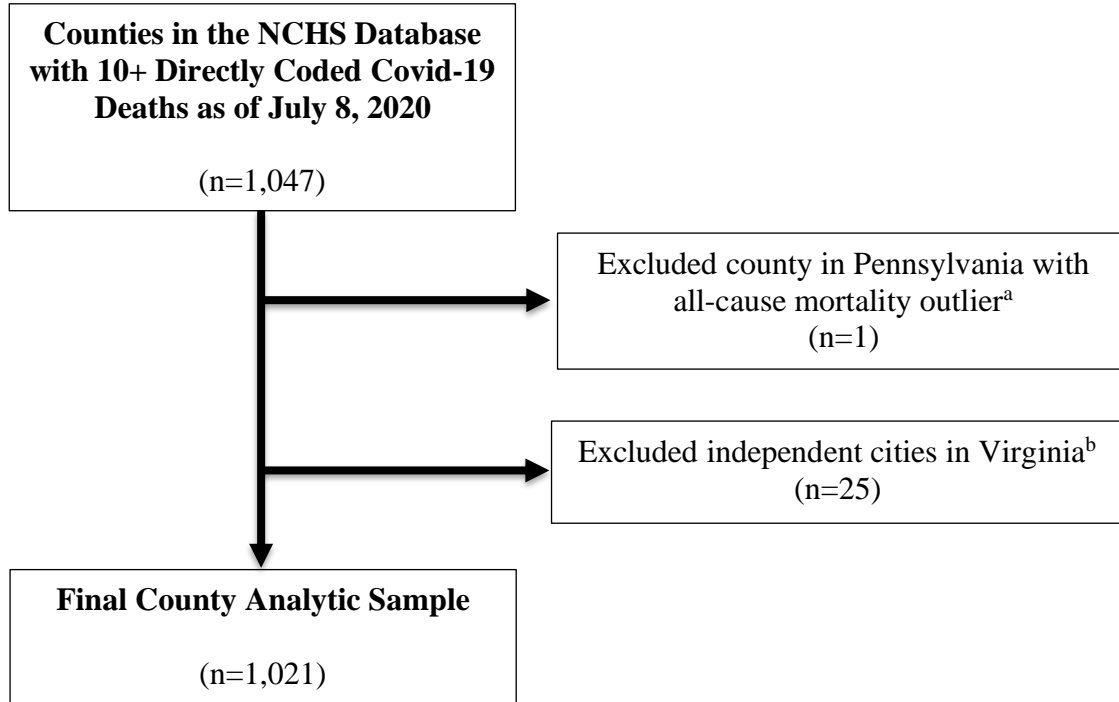
Figure S4: Indirectly age-standardized excess death rates in the lowest and highest quintiles of demographic and structural factors

Table S1: Data sources and years for county-level factors

Table S2: Characteristics of counties included in the analysis

Table S3: Comparison of OLS and Negative Binomial models with indirect age standardization

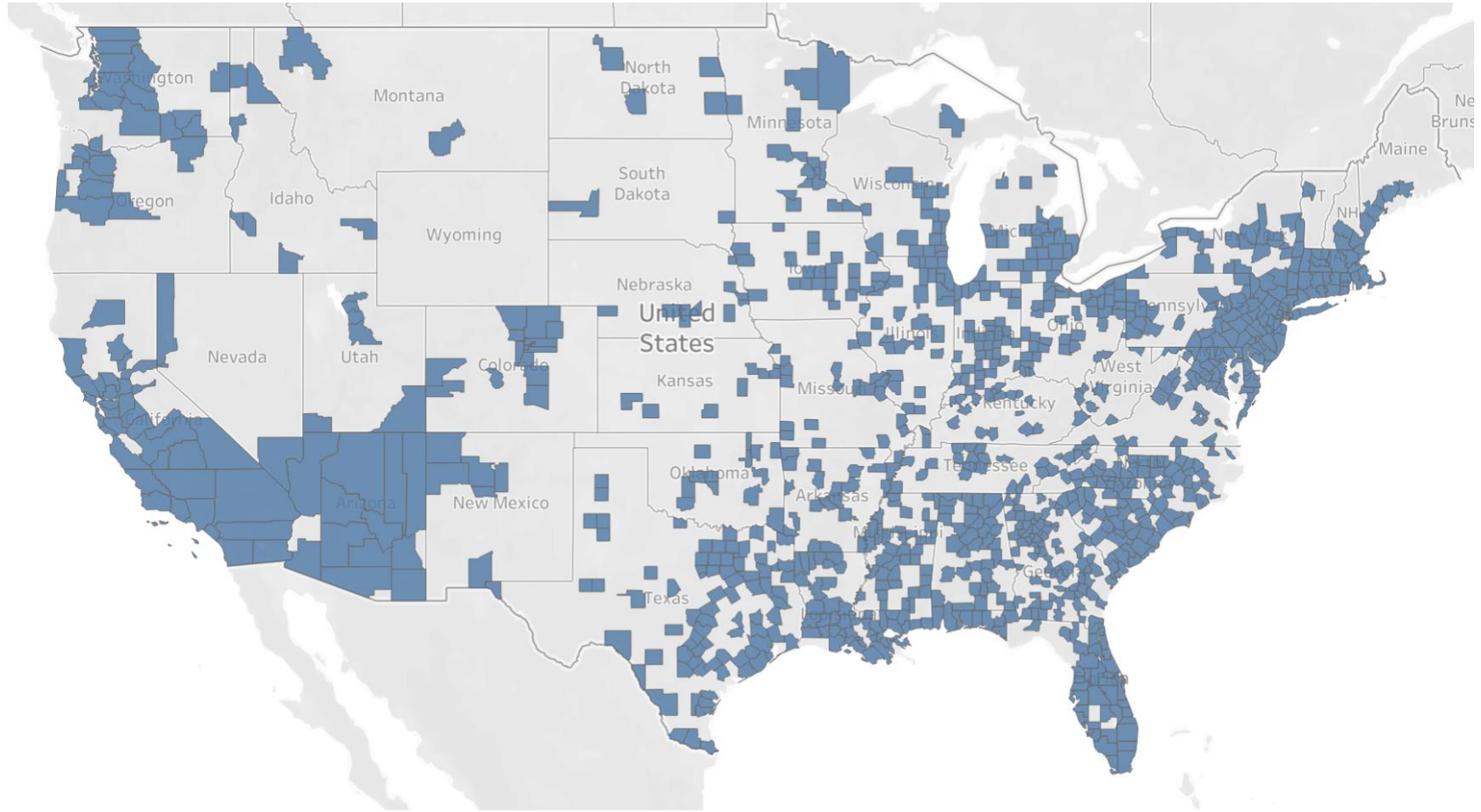
Figure S1: Flowchart detailing sample exclusions



a. All-cause mortality in this county was more than 200% greater than the county with the next highest all-cause mortality.

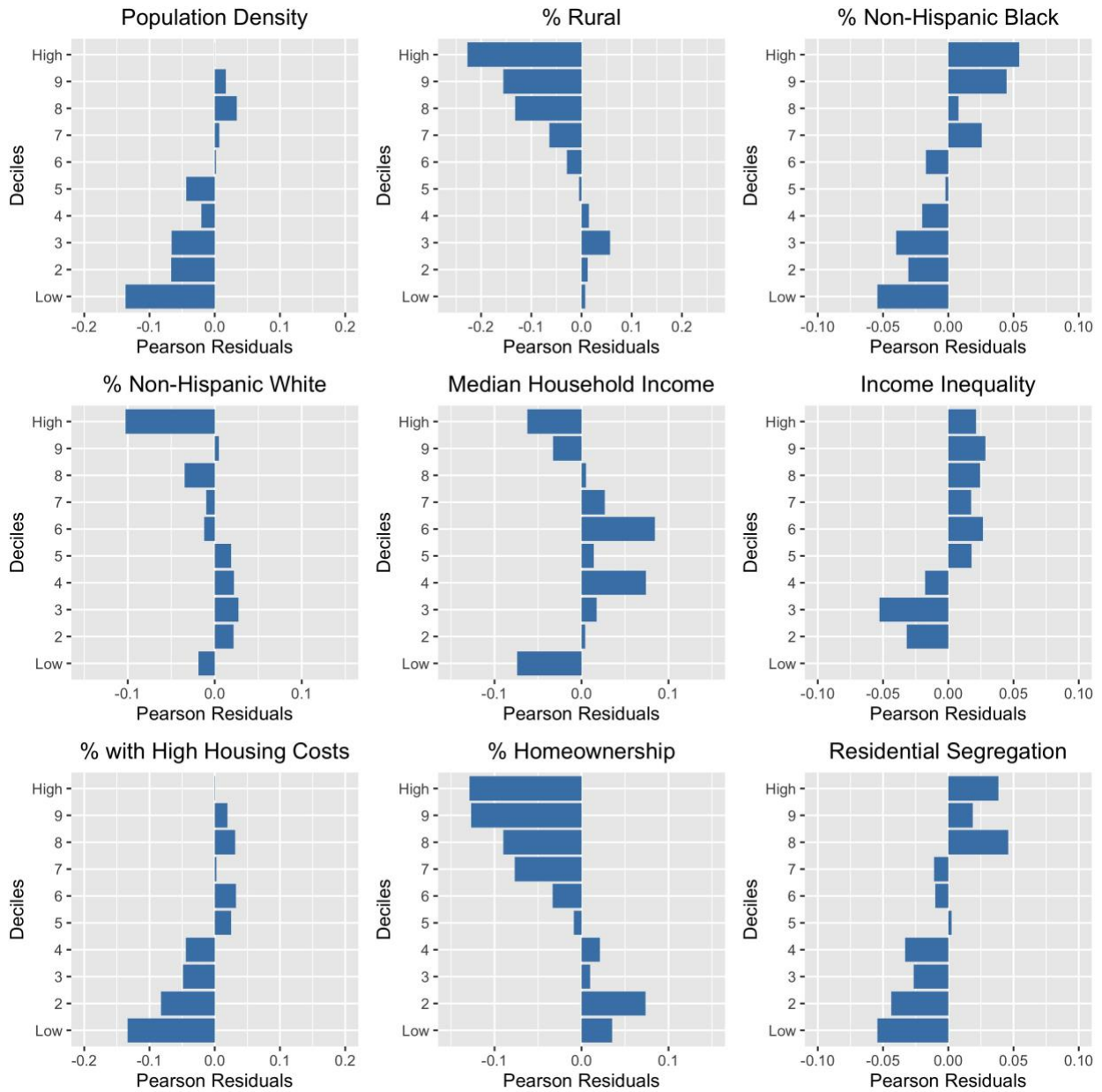
b. We chose not to incorporate independent cities into neighboring counties because several of the neighboring counties were not included in the 1,021 counties represented in the dataset.

Figure S2: US County Map showing geographic distribution of sample counties (n=1,021)



a. Counties in Hawaii and Alaska not pictured.

Figure S3. County-level factors associated with higher or lower than predicted excess mortality using a Negative Binomial model^{a,b,c,d}



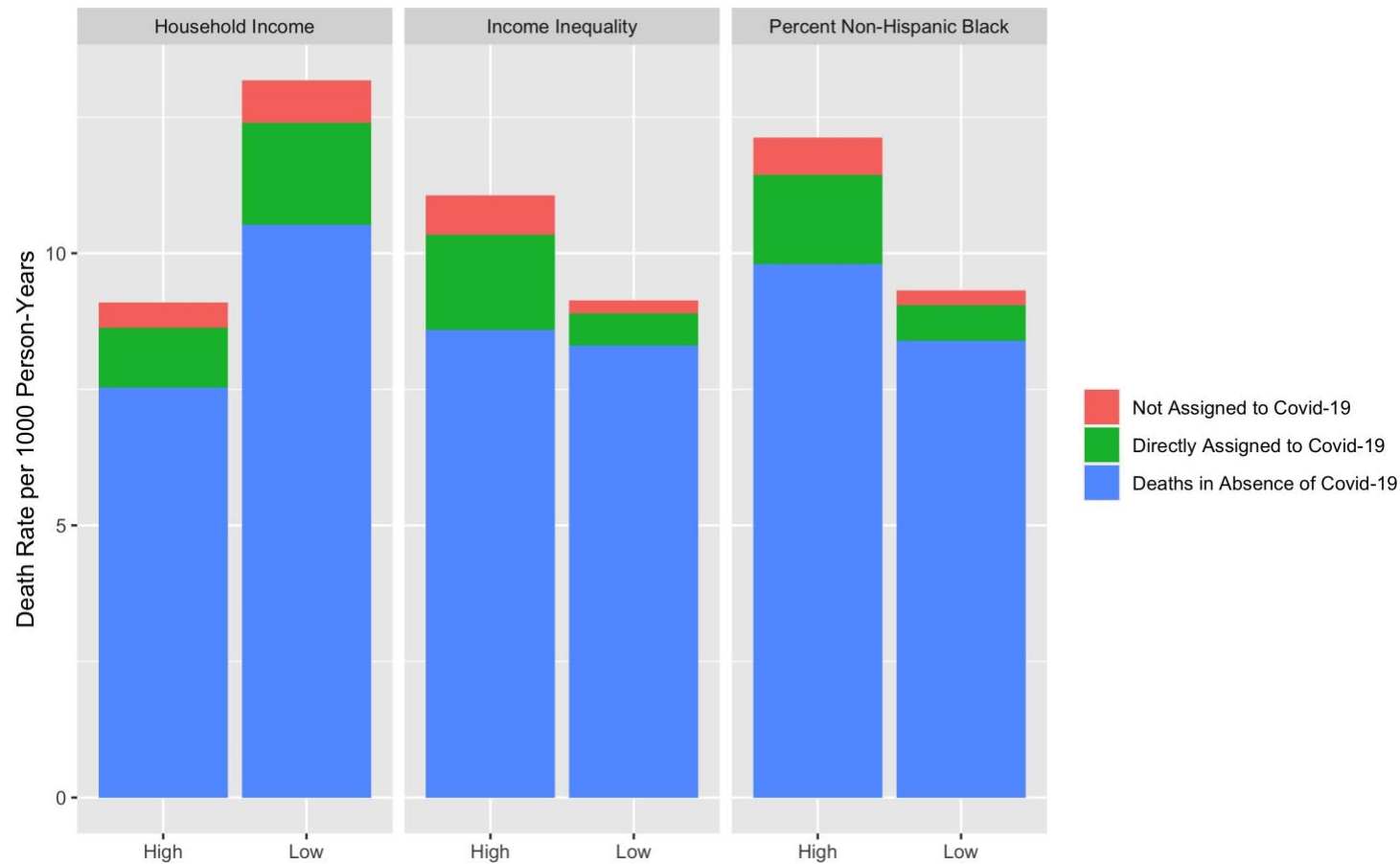
a. $n = 1,021$ counties

b. Pearson residuals calculated from the model: $M(i) = \alpha + \beta_1 M^*(i) + \beta_2 C(i)$, where $M(i)$ = deaths from all-causes in county i in 2020 with the 2020 population used as an offset, $M^*(i)$ = Death rate from all causes, county i in 2013-2018, and $C(i)$ = Covid-19 death rate in county i in 2020. Model weighted by the 2020 population.

c. Counties stratified into deciles by each demographic or structural factor. Weighted means were calculated for the residuals in each decile.

d. A Poisson model was tested prior to the Negative Binomial model but was rejected due to poor goodness of fit.

Figure S4. Indirectly age-standardized excess death rates in the lowest and highest quintiles of demographic and structural factors^{a,b,c}



a. $n = 1,021$ counties

b. Counties in the “high” category represent counties in the upper 20% of values for each demographic and structural factor. Counties in the “low” category represent counties in the lower 20% of values.

c. Crude death rates calculated from primary model: $M(i) = \alpha + \beta_1 M^*(i) + \beta_2 C(i)$, where $M(i)$ = Death rate from all causes in county i in 2020, $M^*(i)$ = Death rate from all causes, county i in 2013-2018, and $C(i)$ = Covid-19 death rate in county i in 2020. Model weighted by the 2020 population. Death rates were indirectly age-standardized.

Table S1: Data sources and years for county-level factors

Variable	Data Source:
% 65 Years and Older	Census Population Estimates, 2018
Population Density	Census Population Estimates, 2020 and Palewire Github (https://gist.github.com/palewire/5cf017f21730ebd8303fb51e0cc7a2cd)
% Non-Hispanic Black	Census Population Estimates, 2018
% Non-Hispanic White	Census Population Estimates, 2018
Median Household Income	Small Area Income and Poverty Estimates, 2018
Income Inequality	American Community Survey, 5-year estimates, 2014-2018
% with High Housing Costs	Comprehensive Housing Affordability Strategy (CHAS) data, 2012-2016
% Homeownership	American Community Survey, 5-year estimates, 2014-2018
Residential Segregation	American Community Survey, 5-year estimates, 2014-2018

Table S2. Characteristics of counties included in the analysis (n=1,021)^a

Characteristic	Mean	Standard Deviation	Minimum	Maximum
<u>Demographic Factors:</u>				
Population Density ^b	2,500	7,123	1.8	71,211
% Rural	12.0%	16.1%	0.0%	100.0%
<u>Structural Factors:</u>				
% Non-Hispanic Black	13.4%	12.3%	0.1%	82.0%
% Non-Hispanic White	56.9%	20.9%	2.7%	97.1%
Median Household Income	\$67,142	\$17,951	\$26,449	\$140,382
Income Inequality ^c	4.8	0.8	3.2	9.2
% with High Housing Costs ^d	15.6%	4.4%	4.8%	30.9%
% Homeownership	62.4%	10.6%	19.6%	89.8%
Residential Segregation ^e	38.2	10.7	2.7	90.4

a. Means and standard deviations are weighted with the estimated county population in 2020.

b. Population density in units of people per square mile.

c. Ratio of household income at the 80th percentile to income at the 20th percentile

d. % with high housing costs – monthly housing costs (including utilities) exceed 50% of monthly income

e. Residential segregation – non-white/white index of dissimilarity where higher values indicate greater residential segregation between non-white and white county residents.

Table S3. Comparison of OLS and Negative Binomial models with indirect age standardization

Model	Number of Excess Deaths per 1 Directly Coded Covid-19 Death	% Excess Deaths Not Attributed to Covid-19
OLS Model ^{a,b}	1.356 [95% CI (1.242, 1.471)]	26.3% [95% CI (20.1%, 32.5%)]
OLS Model, Age-Standardized ^{a,b,c}	1.415 [95% CI (1.298, 1.532)]	29.3% [95% CI (23.5%, 35.2%)]
Negative Binomial Model ^{d,e}	1.406	28.9%
Negative Binomial Model, Age-Standardized ^{c,d,e}	1.371	27.0%

a. The OLS models were specified as $M(i) = \alpha + \beta_1 M^*(i) + \beta_2 C(i)$, where $M(i)$ = Death rate from all causes in county i in 2020, $M^*(i)$ = Death rate from all causes, county i in 2013-2018, and $C(i)$ = Covid-19 death rate in county i in 2020. Model weighted by the 2020 population. For the Negative Binomial model, $M(i)$ = deaths from all-causes in county i in 2020 rather than the death rate, with the 2020 population used as an offset.

b. Number of excess deaths per 1 directly coded Covid-19 death is equivalent to the regression coefficient for directly coded Covid-19 deaths.

c. Death rates were indirectly age-standardized.

d. To calculate the number of excess deaths per 1 directly coded Covid-19 deaths, we used marginal prediction to calculate the all-cause death rate in 2020 at values of directly coded Covid-19 mortality that were +/- 0.1 deaths per 1000 people from the weighted mean of directly coded Covid-19 mortality. The change in all-cause mortality between these values was divided by 0.2 deaths per 1000 people to yield the number of excess deaths per 1 directly coded Covid-19 death.

e. A Poisson model was tested prior to the Negative Binomial model but was rejected due to poor goodness of fit.