Replication code for Cronin and Evans, “Excess Mortality from COVID and non-COVID Causes ion Minority Populations’

The files are in three subfolders. Each is described in the text below

\Main\_results -- this generates the results for Tables 1 and 2 and Figure 1

lyl\_by\_age\_race\_sex.csv – CSV file that has expected remaining life years left by age, sex, and race/ethnicity

AH\_Excess\_Deaths\_by\_Sex\_\_Age\_\_and\_Race.csv – data from the CDC COVID web page that has COVID and non-COVID deaths from 2015.

pop\_20xx.csv – CSV data for US population, January of 20xx where xx=2015-2020. The data is by age, sex, race/ethnicity.

read\_pop\_data\_for\_groups\_1.do.txt – stata do file that reads the pop\_20xx data and produces stat data pop1520\_yy.dta -- that have population by single ages for 2015-2020 for group yy. You should run this first.

hf – Hispanic female

hm – Hispanic male

nhbf – non-Hispanic black female

nhbm – non-Hispanic black male

nhwf -- non-Hispanic white female

nhwm -- non-Hispanic white male

read\_pop\_data\_for\_nation\_1.do.txt – reads pop\_2020.csv and produces a stata data set, reshaped, that has single age population for the country.

read\_mort\_data\_nation\_4.do.txt – stata do file produces results for Table 1.

read\_mort\_data\_fig\_1\_standard\_1.do.txt – stata data file that produces the results for Figure 1

read\_mort\_data\_by\_group\_2.do.txt – stata data file that produces the results for Table 2

\RDD -- program to generates Figure 2 in the paper

deaths\_by\_single\_age\_2\_22\_2021.csv – CSV data that has deaths by single year ages for males and females – these are summed together to get total deaths by age in the RDD program

pop\_2020.csv – csv data that has population by single year of age for January 2020. We use the 1st column, total population

rdd\_at\_65.do.txt – stata do file that produces the RDD estimates.

\Appendix\_results – subfolders that that produces the results in the appendix

\Figure S1 – compare 2019 w COVID age distribution

Deaths\_by\_single\_age\_2019.xlsx – excel file w/ data from CDC wonder multiple cause of death data that has deaths by single age in 2019 and COVID deaths by single age.

\Figure S2 – unemployment rates. This subfolder uses data from regular monthly CPS from IPUMS.org

cps\_00104.dat – deliminted data from IPUMS that has individual level data

cps\_00104.do – stata do file that reads in data from IPUMS and saves it as the stata file cps\_basic.dta

unemployment\_rates\_by\_month\_sex\_race\_1.do – stata file that reads cps\_basic.dta and generates monthly unemployment rates for the six groups in our analysis by month – outputs them for use in excel.

\Place-of\_death

Deaths\_by\_pace\_65+\_2019.xlsx -- excel file that has place of death for people aged 65 and older in 2019. We obtained this from CDC Wonder multiple cause of death data.

\Table S1 – this tables examines the sensitivity of the results in Table 1 in the paper to sing different sets of years. You will need the basic data from \main\_results to run these programs

There are five programs: read\_mort\_data\_nation\_zz\_19.do.txt where zz=15 through 19 and represents using years 2015 through 2019, 2016-2019, etc., to calculate excess deaths.

\Table S2 – this table examines the sensitivity of the estimates in Table 2 to how we calculate life years lost for a particular 5-year age band. You will need the basic data from \main\_results to run these programs

read\_more\_data\_by\_group\_median\_age\_1.do.txt – stata do file. Here, we use the life years left for the median age in an age band. As an example, for ages 20-24, we would use age 22.

read\_more\_data\_by\_group\_age\_average\_1.do.txt – stata do file. Here, we use the simple average life years lost for all ages in the age band.