Disproportionate Decline In American Life Expectancy in 2020*

Russell Luchin, Michael Shmelev

February 20, 2024

Introduction

Data

Methodology

Data was acquired from Nuzzo and Ledesma, 2023 (CITE). Raw data was imported from Arias et al, 2022 (CITE) and The World Bank, 2022. This data was acquired from a replication package for Nuzzo and Ledesma, 2023 (CITE). The data was cleaned and replicated as per the instructions of Nuzzo and Ledesma, 2023 (CITE) and the associated replication package.

Data used to create Figures X and X were further cleaned and aggregated to isolate American life expectancy by ethnic group, separated between two groups "Non-Hispanic White" and "Non-White". As the raw data separates ethnic groups per the US Census (Hispanic, Non-Hispanic AIAN, Non-Hispanic Black, Non-Hispanic Asian, Non-Hispanic White) all groups that were outside the "Non-Hispanic White" category were averaged by summarizing their life expectancy at birth into a single variable. "Non-Hispanic White" was untouched beyond the initial preparation from Nuzzo and Ledesma, 2023 (CITE).

Figure X was isolated from Figure 4(a) from Nuzzo and Ledesma, 2023 (CITE) to highlight expected American mortality rate against observed American mortality rate (all causes). This data comes from The World Bank, 2022 (CITE) and is cleaned according to the replication

---

*Code, data, and .QMD available at: LINK. All code and analysis done with R programming language (R Core Team, 2023). Credit to Jennifer B. Nuzzo and Jorge R. Ledesma for the replication package that was used as a basis for this project.

1

package from Nuzzo and Ledesma, 2023 (CITE). Other than isolating the American mortality rate from a larger figure, the data is left untouched and is accurate to the cleaned data retrieved from Nuzzo and Ledesma, 2023 (CITE).

Raw Data

Table 1: Life Expectancy by Racial Group

year_id	All	Hispanic	Non-Hispanic AIAN	Non-Hispanic Asian	Non-Hispanic Black	Non-Hispanic White
2006	77.8	80.3	NA	NA	73.1	78.2
2007	77.9	80.9	NA	NA	73.2	78.2
2008	78.1	81.0	NA	NA	73.7	78.4
2009	78.5	81.2	NA	NA	74.2	78.7
2010	78.7	81.4	NA	NA	74.7	78.8
2011	78.7	81.6	NA	NA	74.9	78.8
2012	78.8	81.9	NA	NA	75.1	78.9
2013	78.8	81.9	NA	NA	75.1	78.8
2014	78.9	82.1	NA	NA	75.3	78.8
2015	78.7	81.9	NA	NA	75.1	78.7
2016	78.7	81.8	NA	NA	74.9	78.6
2017	78.6	81.8	NA	NA	74.9	78.5
2018	78.7	81.8	NA	NA	74.7	78.6
2019	78.8	81.9	71.8	85.6	74.8	78.8
2020	77.0	77.9	67.1	83.6	71.5	77.4
2021	76.1	77.7	65.2	83.5	70.8	76.4

Table 1 prints the raw data from the data-frame "life", created from the raw data in "life_table.csv" acquired from Arias et al, 2022. This data shows average life expectancy, per year from 2006 to 2021, separated by racial/ethnic group.

Table 2: All Cause Death Rates

Country Name	2020	2021	2022
Aruba	9.690000	11.235000	NA
Africa Eastern and Southern	7.620548	8.036899	NA
Afghanistan	7.113000	7.344000	NA
Africa Western and Central	10.601297	10.678952	NA
Angola	7.821000	8.009000	NA
Albania	10.785000	11.325000	NA

Table 2 shows raw data from the data-table "dths", created from raw data in "all_cause_dth_rates.csv" acquired from The World Bank, 2022. This table shows the first column of the table (country name) alongside the last 3 columns of the total table (all cause death rate per 1000 people, separated by year) for the first 6 rows of the total data-table.

Prepared Data

Table 3: Life Expectancy by Racial Group (Cleaned)

year_id	race	life_exp	Racial-ethnic group
2006	All races and origins	77.8	Non-White
2007	All races and origins	77.9	Non-White
2008	All races and origins	78.1	Non-White
2019	Non-Hispanic White	78.8	Non-Hispanic White
2020	Non-Hispanic White	77.4	Non-Hispanic White
2021	Non-Hispanic White	76.4	Non-Hispanic White

Table 3 shows the first and last 3 rows of the prepared life expectancy data-frame. All racial groups outside of "Non-Hispanic White" have been aggregated into a single variable per year from 2006, while "Non-Hispanic White" remains untouched from the raw data.

Table 4: All Cause Death Rate (Prepared)

location_name	iso3	year	dth_rate	year_id	t	preds
United States	USA	year_2013	8.215	2013	1	8.214250
United States	USA	year_2014	8.237	2014	2	8.304786
United States	USA	year_2015	8.440	2015	3	8.395321
United States	USA	year_2016	8.493	2016	4	8.485857
United States	USA	year_2017	8.638	2017	5	8.576393
United States	USA	year_2018	8.678	2018	6	8.666929
United States	USA	year_2019	8.700	2019	7	8.757464
United States	USA	year_2020	10.300	2020	8	8.848000
United States	USA	year_2021	10.400	2021	9	8.938536

Table 4 shows the prepared data from "dths", from "all_cause_death_rates.csv". In the preparation of this data, the same regression used to predict all cause death rates from Nuzzo and Ledesma, 2023 (CITE) was used to create the "preds" column. This regression used to compute expected death rates comes from running country specific linear regressions where observed death points from 2013 to 2019 were inputted, as described from Nuzzo and Ledesma, 2023 (CITE).

Measurement

Data was measured through 4 independent variables. These variables are expected death rate, measured death rate, life expectancy, and total change in life expectancy. The latter 2 variables were separated by race, between "Non-White" and "Non-Hispanic White" in our prepared data. Measured data was nominal, collected and categorized by racial/ethnic group, year, and # of deaths. Linear regression was used to calculate expected death rate, using the same linear regression replicated from Nuzzo and Ledesma, 2023 (CITE).

Quantitative data was collected to create these variables, with Nuzzo and Ledesma, 2023 (CITE) being the foundational source for the measured variables. All 4 variables were predominantly measured in years, with the notable exception being where observed/expected death rate was also quantified with a ratio per 1000 (deaths per 1000 population). Measured data was prepared as per the replication package from Nuzzo and Ledesma, 2023 (CITE), with minor variations to the "life" and "dths" data-frames.

Results

Life Expectancy
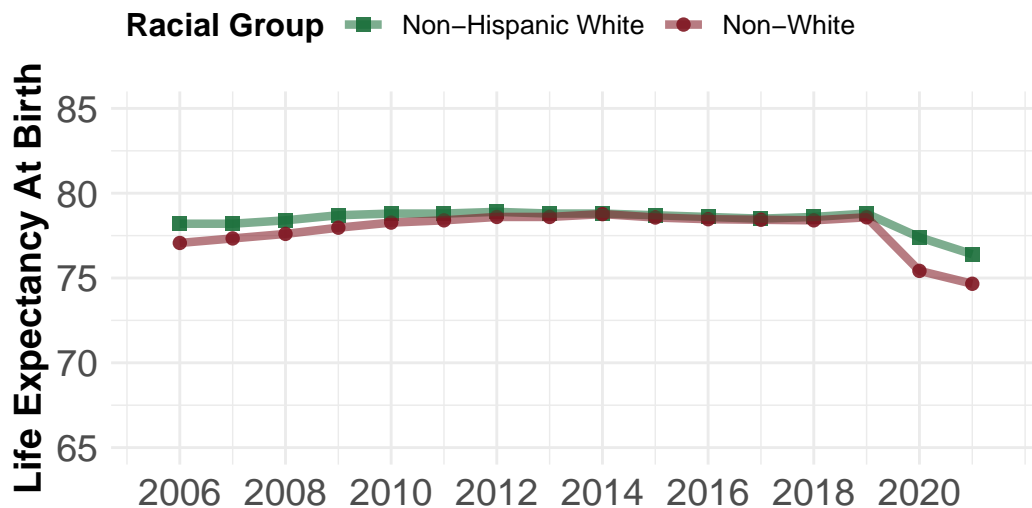
Plotting our prepared life-expectancy data-frame we are able to identify significant trends in American life expectancy regarding racial disparity. Between 2006 and 2012 we observe a

gradual closing of life expectancy disparities between white and non-white Americans. From 2012 to 2019 we observe life expectancy to be not significantly different between white and non-white Americans. This trend could be due to certain low-income healthcare programs being implemented in 2010, such as Obamacare (CITE). However, following the COVID-19 pandemic in 2020, we observe a dramatic decrease in life expectancy across both our measured categories. White Americans observe about half the decrease from non-white Americans, who see a far more dramatic decline in average life expectancy than white Americans.

It should also be noted that both racial groups see a reversal in trends from 2006 to 2019 following the 2020 pandemic. From 2020 onwards, we observe a steady decline in average life expectancy when there seemed to be a steady increase up until 2019. This could point to a greater failure of the American healthcare system, although data for this is inconclusive as there is very little to extrapolate from 2020 onwards in this dataset.

Change in Life Expectancy (2019-2020 and 2020-2021)
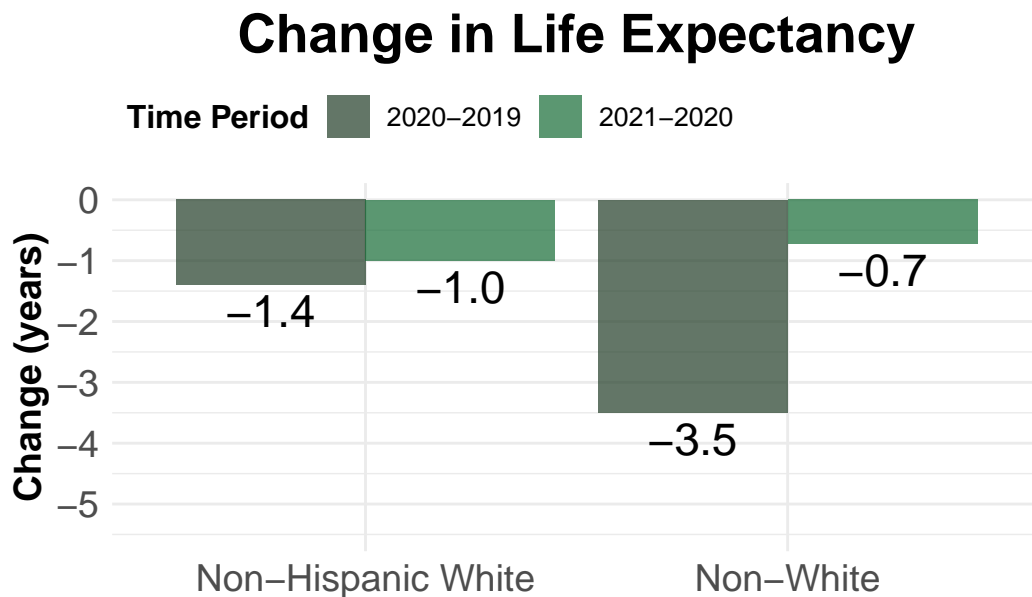
Graphing change in life expectancy we observe a very concerning disparity between white and non-white Americans. While average life expectancy only dropped by about a year for white Americans, non-white Americans observed nearly a 4 year cut (averaged) to their life expectancy. This is a disparity of over 200% when, prior 2020, observed life expectancy was minimally different between white and non-white Americans.

A disparity this large is abnormal, considering the gradual trend of life expectancy becoming equal between white and non-white Americans up to 2019 (Figure 1). Figure 2 supplements the dramatic difference in life expectancy dropping between white and non-white Americans by giving a nominal value (difference in life expectancy) to these groups. A difference of around 200% implies great stratification between white and non-white Americans, in spite of regular efforts to normalize healthcare access (CITE).
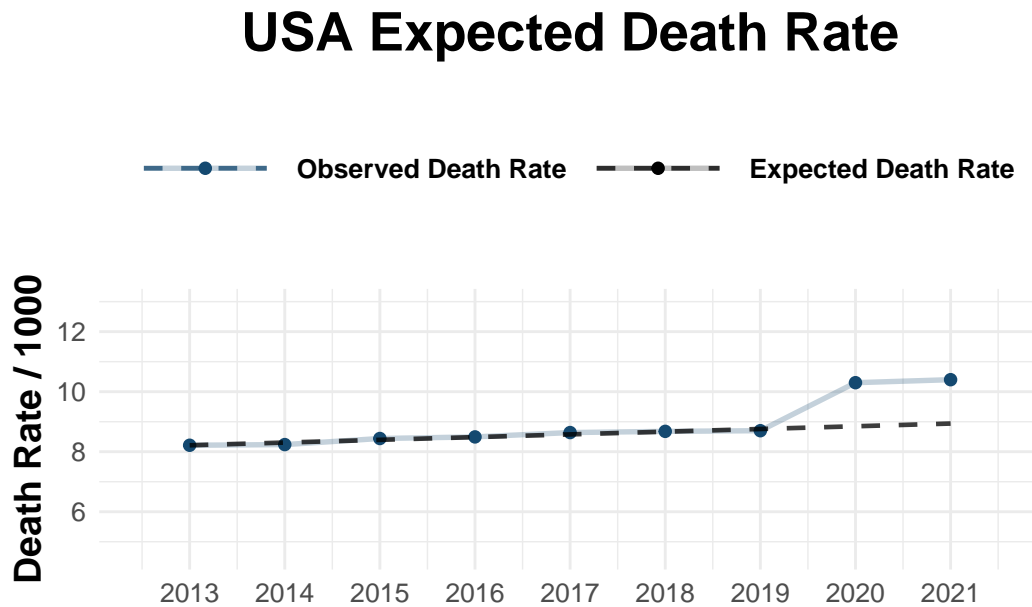
Expected vs Observed Death Rate

Figure 3 graphs our expected death rate alongside our observed death rate. As seen in Figure 3, there is a substantial difference between our predicted death rate (calculated using linear regression with death rates from 2013-2019) and our observed death rate; particularly in 2020.

What should be noted is that, in 2020, the COVID-19 pandemic struck the globe as nations struggled to mitigate the damage the infection caused. The linear regression from Nuzzo and Ledesma, 2023 (CITE) has no way to account for this, and therefore the expected death rate continues the trends observed between 2013-2019.

Cross referencing this data with Figure 1, we note some consistent data points. When comparing average life expectancy from 2013 to 2019, we note a slight decline in the expected lifespan across both white and non-white Americans. Figure 3 visualizes a slight increase in death rate, which is consistent with our findings in Figure 1.

Cross Referencing the Life Expectancy Change and Death Rate

Comparing Figures 2 and 3, we observe a significant data point to be measured. Figure 3 observes a significant difference between expected death rate and observed death rate (per 1000), the difference being 2~ deaths per 1000 people. By averaging the change in life expectancy from figure 2, we can reach a similar value. We average the change in life expectancy with this equation: **A = (x+y)/2**, where A is the average of X and Y.

Focusing on the change between 2019 and 2020, we observe and average life expectancy drop of 2.5. Cross-referencing this average with the difference between expected and observed death rate from Figure 3, we note a similar increase of about 2.2~ more deaths than expected. We see a similar pattern when observing the change between 2020 and 2021, where the average decrease in life expectancy is .85 years and the increase in death rate is about .5 years.

As these values are not identical, it is a possibility that this is a coincidence. However, it can be further hypothesized that, if there was a lower disparity between the drop in average life expectancy between white and non-white Americans, then the observed increases in death rate per 1000 would be lower. This is because the drop observed in figure 2, and the climb observed in figure 3, have very similar nominal values assigned to them (differences of .2-.3 when an average is calculate across groups in figure 2 and compared to figure 3).