RESEARCH QUESTION Investigating variation in infant mortality rates among US states

BACKGROUND The Measure of America (MOA) dataset was first compiled for the American Human Development Project, an initiative launched in 2007 by the Social Science Research Council (SSRC).[1] The project's objective was to introduce to America the use of the human development index (HDI), a statistic composite index of wellbeing and opportunity. Drawn from observational data collected or produced by governmental agencies, the dataset includes a large selection of quantitative variables for all 50 states, ranging from life expectancy to unemployment, from the number of teen pregnancies to energy consumption.

Of all the variables, the most troubling are perhaps the numbers for infant mortality. In the United States, infant mortality is defined as the death of an infant before his or her first birthday.[i] For a family, the death of an infant is a tragedy; for society, it is a mark of failure. Infant mortality can be taken as an indicator of overall health in a community, it is "an important summary reflecting social, political, health care delivery and medical outcomes in a geographic area."[ii]

If infant mortality rates (IMR) serves as a summary for overall health, the United States underperforms relative to national wealth, usually a strong determinant for health. Reviewing infant mortality rates from The World Bank[iii] for the OECD club of 32 mostly rich countries shows that the United States (IMR 6.25) falls between Slovakia (IMR 5.77) and Latvia (IMR 6.52), and has a rate that is more than double that of countries such as Japan (IMR 2.36) and France (IMR 3.11) (Fig 1).

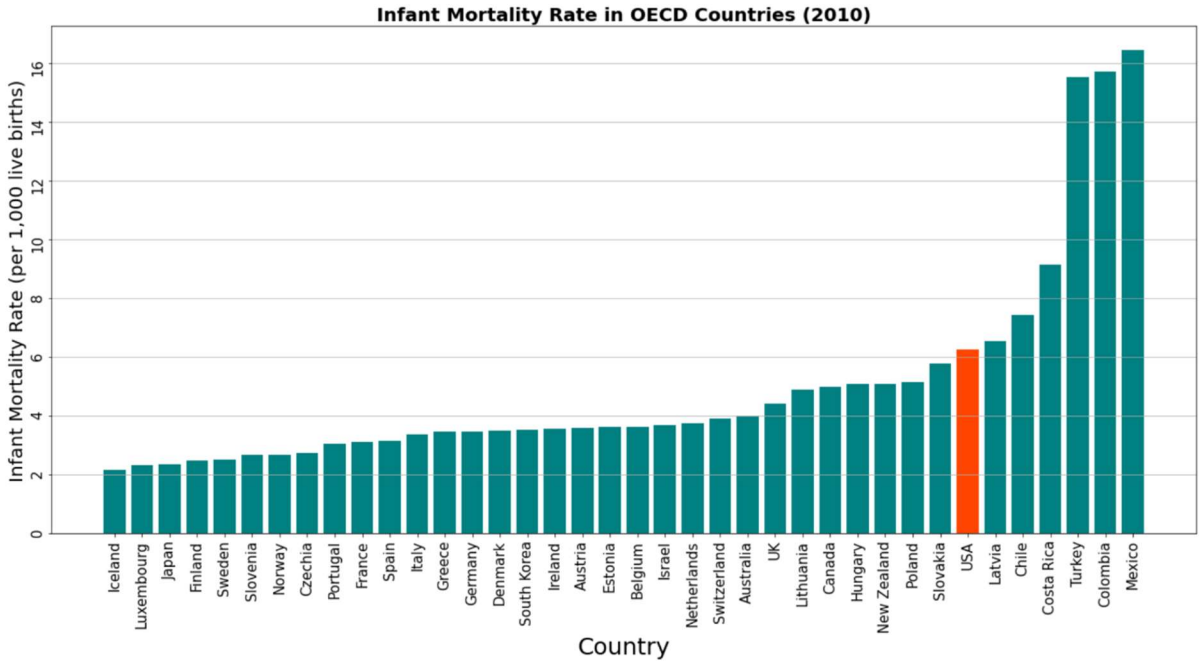

Figure 1. Infant Mortality Rate (IMR) in OECD countries (2010)

[1] The Social Science Research Council (SSRC), a US-based nonprofit organization with the mission of advancing research in the social sciences and related disciplines.

The United States' underperformance is hardly surprising considering its paltry health insurance coverage. In 2010, 16% of Americans were uninsured (Fig 2).[iv] And even among American women with employer-provided health insurance, more than half (62%) had no maternity coverage (Fig 3).[v]

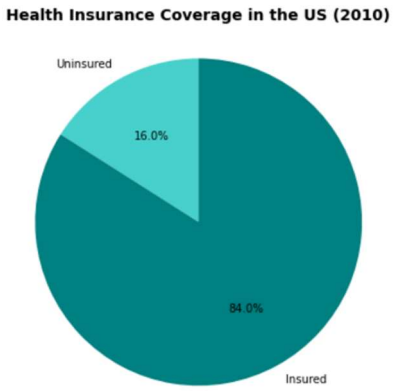

Figure 2. Health insurance coverage in the US (2010)

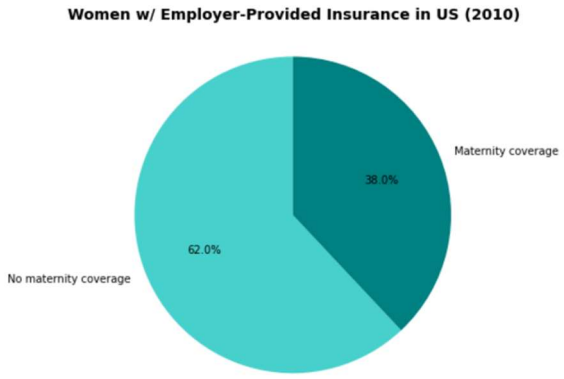

Figure 3. Maternity coverage for women w/ employer-provided insurance in the US

ANALYSIS The United States, of course, is not uniform given its vast size. From the MOA data, infant mortality across all 50 states ranges from a low of 3.75 for the state of Alaska to a high of 9.67 for the state of Mississippi. The bar chart below shows each states' IMR and the 50-state median (6.32) indicated by the dashed orange line (Fig 4).

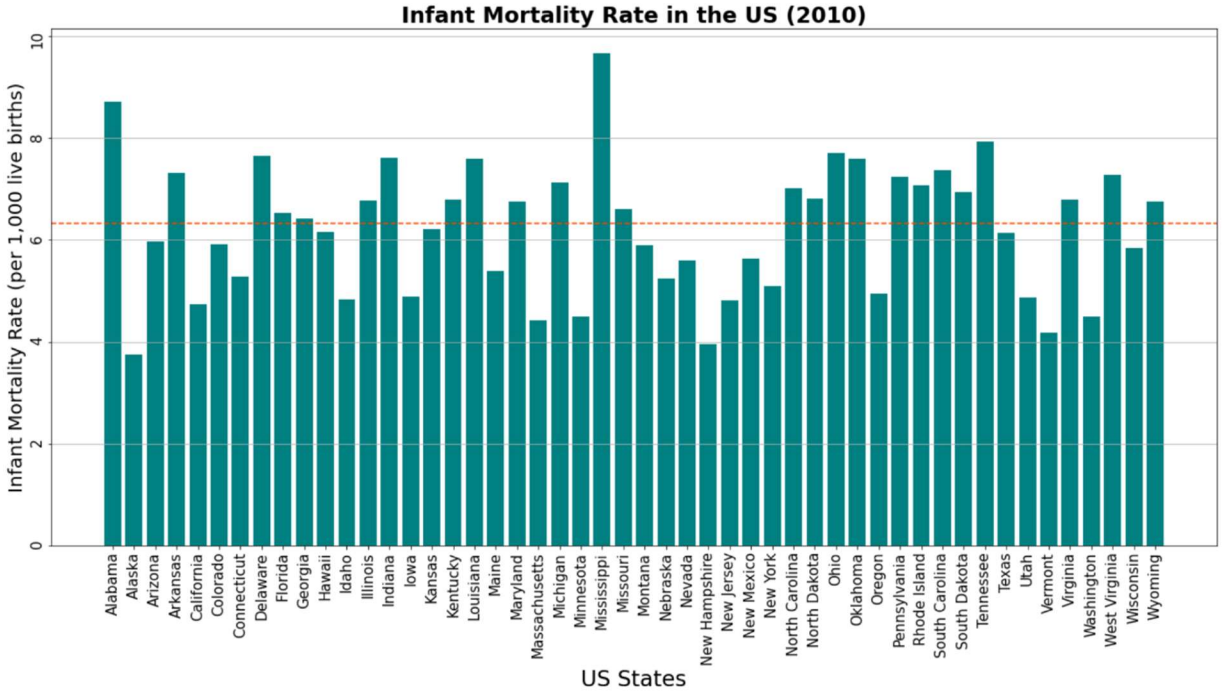

Figure 4. Infant Mortality Rate (IMR) in the US by state.

Plotting those same data in a histogram (Fig 5) and a boxplot (Fig 6) shows the distribution across all 50 states. Both the histogram and the boxplot show that the distribution of IMR is skewed.

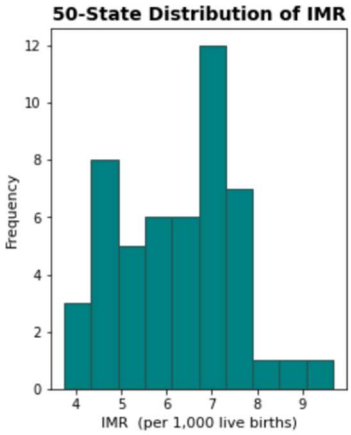

Figure 5.

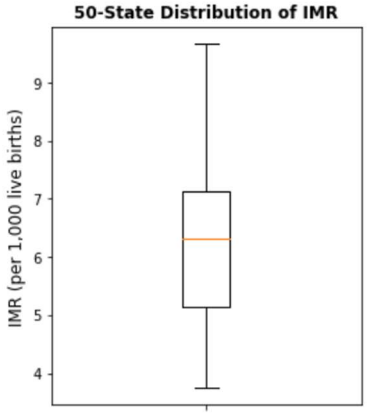

Figure 6.

The histogram of 50-state IMR is positively skewed, with a small number of states with high IMR. Distribution appears to be bi-modal, with one cluster of states measuring around 5 infant deaths per 1,000 live births and a second cluster measuring around 7 infant deaths.

The boxplot of 50-state IMR has a much longer "whisker" above Q3 (IMR 7.25) than the whisker below Q1 (IMR 5.09), showing that there are more states with higher-than-median IMR than there are states with lower-than-median IMR. Indeed, the underperformance of some states has pulled the median line (indicated by the yellow line) up by 0.1032 from a mean of 6.2168 and to the median of 6.32. Which states account for the lower-than-median whisker and which for the higher-than-median whisker? Alaska, New Hampshire, Vermont, Massachusetts and Minnesota are the five states with the lowest IMR, with a group mean of 4.16. Delaware, Ohio, Tennessee, Alabama, Mississippi are the five states with the highest IMR, with a group mean of 8.336, more than double that of the lowest-IMR group.

Why such variation? The MOA data can be used to investigate variables responsible for such variation in IMR. Of course, given that the MOA data are drawn from observational studies and confounding variables are not averaged out, any explanatory variables thus identified would, at best, be potential explanatory variables. There would still be value in this, as IMR could never ethically be part of a randomized experiment.

Among the variables, a good place to start the investigation is wealth.[vi] As economists L. Pritchett and L. Summers wrote in an influential paper on public health, "wealthier is healthier." The correlation between the two is found across countries, and within countries, across individuals.[vii] The MOA data include the variables of per capita GDP (USD) and median earnings (USD), which are compared with IMR in scatter plots (Fig 7 and Fig 8, respectively).

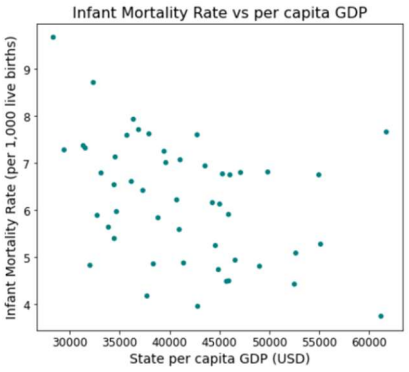

Figure 7.

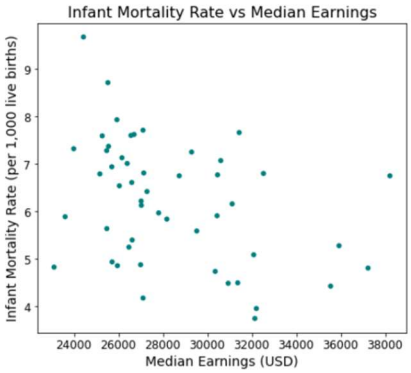

Figure 8.

The scatter plots of IMR vs per capita GDP and of IMR vs median earnings both show negative correlations, meaning that higher GDP and higher median earnings are both correlated to lower IMR. The Pearson correlation coefficients are calculated to be −0.3989 and −0.3784, respectively, both considered weak.

If "wealthier is healthier" is generally true, it would be logical that the reverse should also be true. The MOA data also include the variables of poverty rate (% below federal poverty threshold) and child poverty rate (% living in families below the poverty line), which are compared with IMR in scatter plots (Fig 9 and Fig 10, respectively).

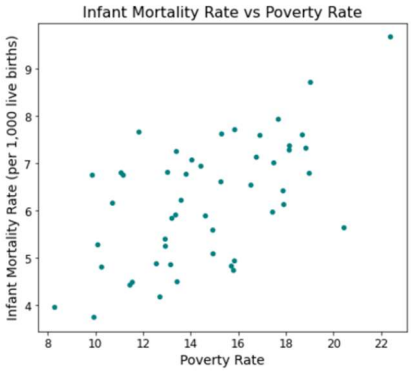

Figure 9.

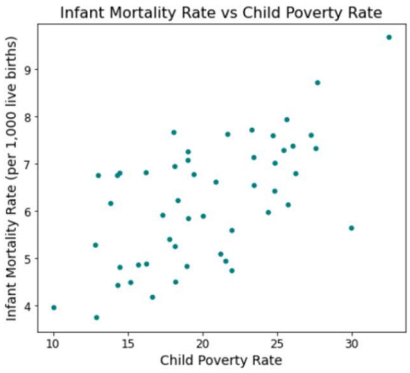

Figure 10.

The scatter plots of IMR vs poverty rate and of IMR vs child poverty rate show positive correlations, meaning that higher poverty rate and higher child poverty rate are both correlated to higher IMR. The Pearson correlation coefficients are calculated to be 0.5780 and 0. 6152, respectively, both considered moderate. From this analysis, poverty—whether among the population overall or among children—is more strongly correlated with IMR than measures of wealth.

To examine further the relationship between IMR and child poverty, states were then categorized by whether they had above- or below-mean IMR, and by whether they had above- or below-mean child poverty rate (20.30). Two-way frequency tables were generated for the actual number of states and the proportions of all 50 states (Tables 1 and 2).

ChildPov_Comparata	Above	Below
IMRComparata		
Above	16	10
Below	7	17

Table 1.

ChildPov_Comparata	Above	Below
IMRComparata		
Above	0.695652	0.37037
Below	0.304348	0.62963

Table 2.

The proportion of states with above-mean IMR having above-mean child poverty rate was 0.6957 (95% CI = [0.51, 0.88]) and the proportion of states with above-mean IMR having below-mean child poverty rate was 0.3703 (95% CI = [0.19, 0.55]) (Fig 11). A chi-squared test was conducted to compare the proportions of above- and below-mean child poverty rates with above-mean IMR. There was not a significant difference in the proportions of the states with above- and below-mean poverty rates (p=0.0444 and chi^2 = 4.0424). The average difference in proportions was −0.3253 (95% CI = [−0.59, −0.06]). Despite the moderate correlation strength between states with above-mean IMR and above-mean child poverty rate, the data does not provide support that that states with above-mean IMR have a greater chance of having an above-mean child poverty rate.

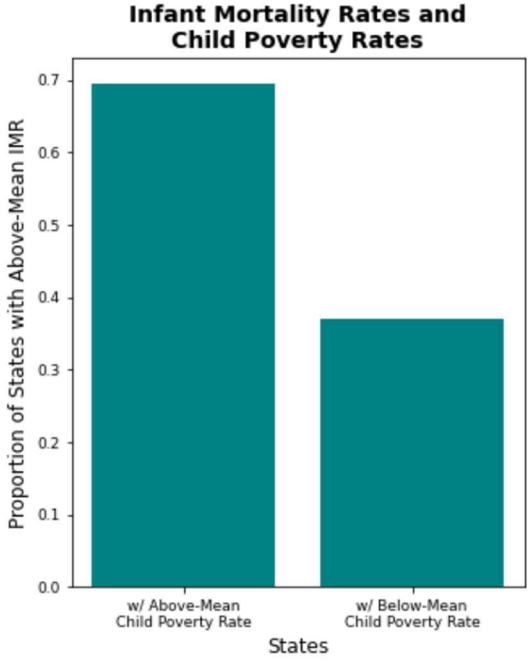

Figure 11.

In reality, poverty is not simply the reverse of wealth. Poverty is complex, caused by a number of factors that may or may not be closely intertwined. The same can be said of infant mortality. While research has identified socioeconomic status as having a significant association with birth outcomes, it is not the only factor. One factor was identified by researchers P. Blumenshine, S. Egerter, C. Barclay, C. Cubbin, and P. Braveman[viii] in 2010 through a review of literature. They concluded that there was substantial variation in IMR among racial or ethnic subgroup, with non-Hispanic Black mothers experiencing the highest IMR. Other researchers have arrived at the same conclusion. The Peterson-KFF Health System Tracker points out that "researchers focusing on the substantial Black-White infant mortality gap have generally found that controlling for maternal background factors also does not fully explain the disparity." That report offers a possible explanation for the disparity: "Notably, recent research exploring U.S. maternal and infant health disparities discusses structural racism as a primary risk factor for African-American mothers and their infants, largely due to the complex stress it places on mothers throughout their life." [ix]

Unfortunately, structural racism is not one of the variables included in the MOA data. So how can the way in which structural racism influences IMR be explored? It could be argued that structural racism in America is the most enduring legacy of slavery. Therefore, the legality of slavery in 1860 on the eve of the Civil War was added as a variable for each state in the MOA dataset (Fig 12).

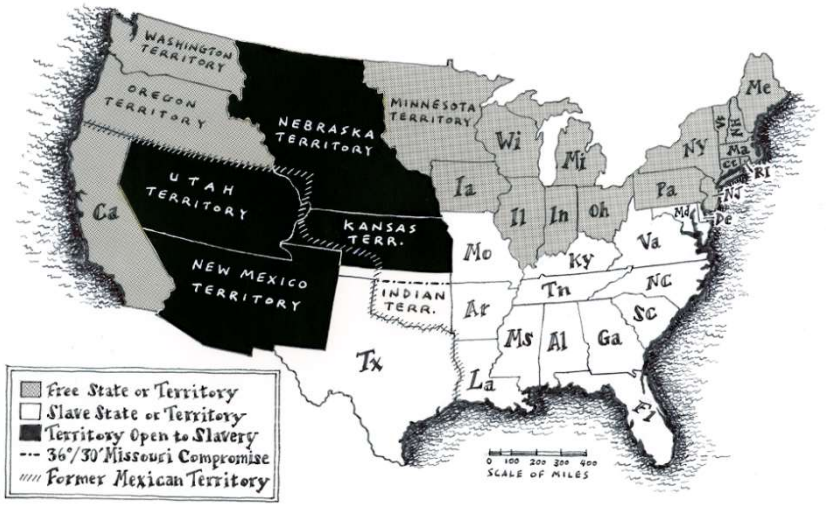

Figure 12

States were divided into two[2] groups: those where slavery was legal (26 states) and those where slavery was illegal (24 states).[x] The average IMR for slave states was calculated to be 6.78 (95% CI = [6.38, 7.18]), and the average IMR for free states was 5.61 (95% CI = [5.12, 6.10]). A two sample t-test was conducted to compare the average IMR for slave states and free states. There was a significant difference in IMR of these two sets of states (p=0.0009). The average difference in IMR was 1.17 (95% CI = [0.54, 1.81]). This analysis demonstrates that average IMR for slave states is higher than the average IMR for free states. This same conclusion can be drawn from plotting the data in a boxplot (Fig 13).

Slavery in the United States was not uniform. The large, single-crop agricultural operations of plantations required a system of slavery that transformed those societies. "The plantation system evolved into the prevailing institution that penetrated all aspects of social, cultural, economic and political life […] Since this form of slavery was the dominant social institution of the region, the Antebellum South also evolved into a peculiar society, characterized by a rigid system of racial stratification in the distribution of wealth, power, status and privilege."[xi]

Given this, states were divided anew, this time into three groups: those where slavery was legal and plantations were predominant[xii] (12 states), those where slavery was legal and plantations were not predominant (14 states), and those where slavery was illegal (24 states). The average IMR for plantation slave states was calculated to be 7.36 (95% CI = [6.82, 7.91]); average IMR for slave states was 6.28 (95% CI = [5.85, 6.71]); and the average IMR for free states was 5.61 (95% CI = [5.12, 6.10]). A two sample t-test was conducted to compare the average IMR for plantation slave states and free states. There was a significant difference in IMR of this pair of states (p= 0.0001). The average difference was 1.7558 (95% CI = [1.02, 2.49]). This analysis demonstrates that average IMR for plantation slave states is higher than the average IMR for free states. This same conclusion can be drawn from plotting the data in a boxplot (Fig 14).

[2] For the purposes of this analysis, states were considered slave states if slavery was legal in more than half of its 1860 territory or if it was open to slavery. Alaska and Hawaii, which were not yet part of the United States, were considered free.

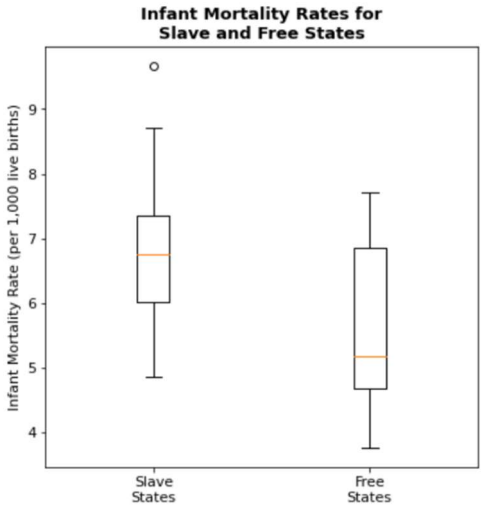

Figure 13.

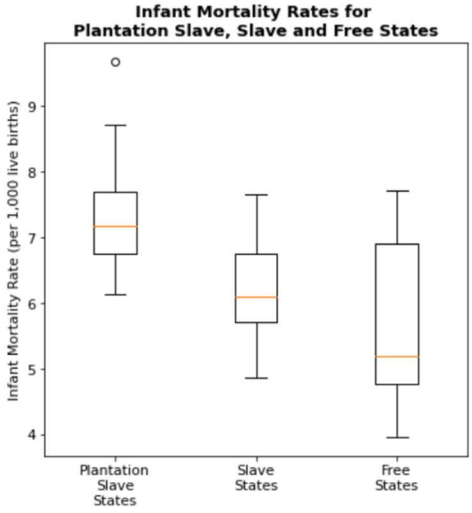

Figure 14.

The ways in which structural racism may be contributing to IMR can be further explored by returning to the MOA data and considering the variable of states' rate of incarceration.[3] As was the case of IMR, the United States compares poorly with fellow OECD countries, with an incarceration rate that is six times higher than the OECD average (Fig 15).[xiii] As was the case of IMR, there is substantial variation in incarceration rates among racial or ethnic subgroups, with non-Hispanic Black men incarcerated nearly seven times more than white non-Hispanic men.[xiv] Therefore, as a thought exercise, the relationship between IMR and the incarceration rate is explored.

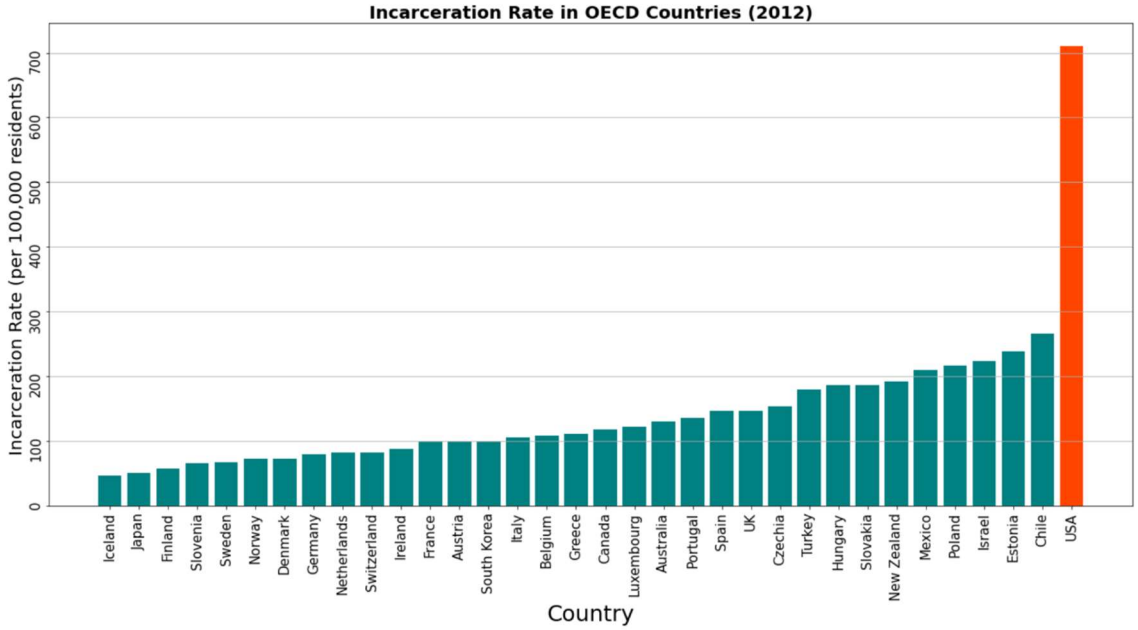

Figure 15.

[3] The rate of incarceration for Illinois was missing from the MOA dataset. A rate for Illinois for 2010 was found here: https://bjs.ojp.gov/content/pub/pdf/p10.pdf.

A simple linear regression analysis was conducted to examine the relationship between these two variables. The scatter plot (Fig 16) shows a positive correlation, where a higher incarceration rate is correlated with higher IMR. The Pearson correlation coefficients are calculated to be 0.6015 (P=3.8414 e-06), which is considered moderate. A significant equation was found IMR = Incarceration Rate * 0.0052 + 4.0643. For every additional person per 100,000 inhabitants incarcerated, an additional 0.0052 infants do not live past their first birthday, per 1,000 live births (95% CI = [0.003, 0.007]).

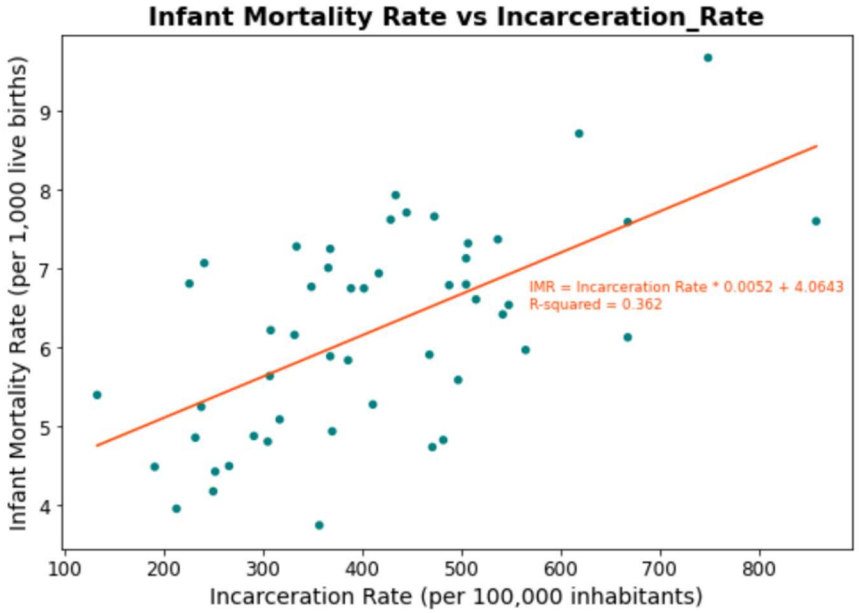

Figure 16.

The author of this paper is not suggesting that a cause and effect relationship exist between IMR and incarceration rate. Rather, the foregoing statistical analysis suggests the possibility of a lurking variable.

As a comparison, the same analysis was conducted to examine the relationship between IMR and births to teen girls. Many studies have shown that infants born to teens are more likely to die in their first year of life compared to women aged 20 and over. Moreover, IMR is highest for infants of non-Hispanic black teenage girls compared with infants of non-Hispanic white and Hispanic teenage girls. [xv]

A simple linear regression analysis was conducted to examine the relationship between these two variables. The scatter plot (Fig 17) shows a positive correlation, where a higher number of births to teen girls is correlated with higher IMR. The Pearson correlation coefficients are calculated to be 0.5934 (P=5.5866 e-06), which is considered moderate. A significant equation was found IMR = Birth to Teen Girls * 0.0758 + 3.6379. For every additional birth to a teen girl per 100,000 inhabitants, an additional 0.0758 infants do not live past their first birthday, per 1,000 live births (95% CI = [0.046, 0.106]).

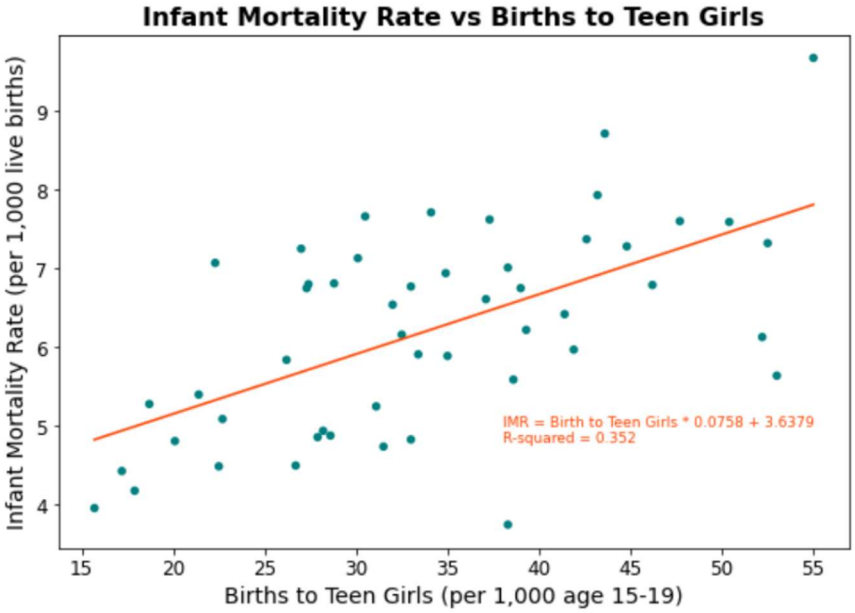

Figure 17.

Statistical analyses for IMR vs incarceration rate and for IMR vs births to teen girls (Tables 3 & 4) produced only small coefficients of determination (0.362 and 0.352, respectively), but both had significant linear regressions. What is interesting is the similarities between the two sets of analysis: one set for two variables that, superficially, should have nothing in common; the second set for two variables that have been linked as having a causal relationship in many studies.

Dep. Variable:	Infant_Mortality_Rate	**R-squared:**	0.362
Model:	OLS	**Adj. R-squared:**	0.349
Method:	Least Squares	**F-statistic:**	27.21
Date:	Sat, 30 Apr 2022	**Prob (F-statistic):**	3.84e-06
Time:	15:47:56	**Log-Likelihood:**	-71.931
No. Observations:	50	**AIC:**	147.9
Df Residuals:	48	**BIC:**	151.7
Df Model:	1		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
const	4.0643	0.438	9.277	0.000	3.183	4.945
x1	0.0052	0.001	5.217	0.000	0.003	0.007

Omnibus:	4.502	**Durbin-Watson:**	2.031
Prob(Omnibus):	0.105	**Jarque-Bera (JB):**	1.924
Skew:	-0.021	**Prob(JB):**	0.382
Kurtosis:	2.040	**Cond. No.**	1.30e+03

Table 3.

Dep. Variable:	Infant_Mortality_Rate	**R-squared:**	0.352
Model:	OLS	**Adj. R-squared:**	0.339
Method:	Least Squares	**F-statistic:**	26.09
Date:	Sat, 30 Apr 2022	**Prob (F-statistic):**	5.59e-06
Time:	16:32:40	**Log-Likelihood:**	-72.308
No. Observations:	50	**AIC:**	148.6
Df Residuals:	48	**BIC:**	152.4
Df Model:	1		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
const	3.6379	0.526	6.913	0.000	2.580	4.696
x1	0.0758	0.015	5.107	0.000	0.046	0.106

Omnibus:	0.195	**Durbin-Watson:**	1.866
Prob(Omnibus):	0.907	**Jarque-Bera (JB):**	0.272
Skew:	-0.137	**Prob(JB):**	0.873
Kurtosis:	2.765	**Cond. No.**	126.

Table 4.

What factors lead to the variation in IMR among US states? What might result in a significant difference in means for IMR between states where slavery was once legal and states where slavery was always illegal? What lurking variable might cause IMR and incarceration rates to demonstrate moderate correlation? The foregoing analyses are fairly rudimentary and they are used to explore what are highly complex relationships. It would be a stretch to claim that structural racism is a partial answer to any of the previous three questions. Yet consider that only two average American lifetimes have passed since the end of slavery. Would it be surprising that we can still feel the looming presence of this institution?[xvi] At the very least, it is possible to conclude that the differences in IMR between states are most accurately described not as variation, but disparity.

WORKS CITED

[i] Centers for Disease Control and Prevention. "Infant Mortality." *Reproductive Health*. 8 Sep 2021, https://www.cdc.gov/reproductivehealth/maternalinfanthealth/infantmortality.htm. Accessed 30 Apr 2022.

[ii] March of Dimes. "Mortality and Morbidity." *Peristats*. https://www.marchofdimes.org/peristats/data?top=6. Accessed 30 Apr 2022.

[iii] The World Bank. "Mortality rate, infant (per 1,000 live births)." https://data.worldbank.org/indicator/SP.DYN.IMRT.IN. Accessed 30 Apr 2022.

[iv] Cohen, Robin A., Ward B.W., and Schiller J.S. "Health Insurance Coverage: Early Release of Estimates from the National Health Interview Survey, 2010." National Center for Health Statistics. Centers for Disease Control and Prevention. June 2011, https://www.cdc.gov/nchs/data/nhis/earlyrelease/insur201106.pdf. Accessed 30 Apr 2022.

[v] Rosenthal, Elisabeth. "American Way of Birth, Costliest in the World." *The New York Times*. 30 Jun 2013, https://www.nytimes.com/2013/07/01/health/american-way-of-birth-costliest-in-the-world.html. Accessed 30 Apr 2022.

[vi] Prittchett, Lant, L. H. Summers. "Wealthier is Healthier." *World Development Report Policy Research Working Papers*. 1993. https://documents1.worldbank.org/curated/en/684651468741004317/pdf/multi0page.pdf. Accessed 30 Apr 2022.

[vii] Baird, Sarah, J. Friedman, and N. Schady. "Aggregate Income Shocks and Infant Mortality in the Developing World." *The Review of Economics and Statistics*, 93(3): 847–856. Aug 2011, https://documents1.worldbank.org/curated/en/830521468129272116/pdf/946670JRN0Box30LIC00130rest0a000084.pdf. Accessed 30 Apr 2022.

[viii] Blumenshine, Philip, S.Egerter, C.J. Barclay, C. Cubbin, and P.A. Braveman. "Socioeconomic disparities in adverse birth outcomes: a systematic review." *American Journal of Preventive Medicine*, 39(3):263-72. Sep 2010, https://pubmed.ncbi.nlm.nih.gov/20709259/. Accessed 30 Apr 2022.

[ix] Kamal, Rabah, J. Hudman, and D. McDermott. "What do we know about infant mortality in the U.S. and comparable countries?" *Health & Wellbeing*. Peterson-KFF Health System Tracker. 18 Oct 2019, https://www.healthsystemtracker.org/chart-collection/infant-mortality-u-s-compare-countries/. Accessed 30 Apr 2022.

[x] Brown, Joshua. "Map of Free and Slave States in 1860" *SHEC: Resources for Teachers.* American Social History Project/Center for Media and Learning. https://shec.ashp.cuny.edu/items/show/2008.

[xi] https://www.nationalgeographic.org/article/plantation-system/. Accessed 30 Apr 2022.

[xii] Wandrei, Kevin. “What Were the Biggest Slave States?” *Classroom*. https://classroom.synonym.com/what-were-the-biggest-slave-states-12083927.html. Accessed 30 Apr 2022.

[xiii] Kearney, Melissa S., B. H. Harris, E. Jácome, and L. Parker. “Ten Economic Facts about Crime and Incarceration in the United States.” *The Hamilton Project*. May 2014, https://www.brookings.edu/wp-content/uploaxds/2016/06/v8_thp_10crimefacts.pdf. Accessed 30 Apr 2022.

[xiv] Guerino, Paul. Harrison, Paige M., Sabol, William J. “Prisoners in 2010.” *Bureau of Justice Statistics Bulletin*. December 2011, https://bjs.ojp.gov/content/pub/pdf/p10.pdf. Accessed 30 Apr 2022.

[xv] Woodall, A.M., Driscoll A.K. “Racial and ethnic differences in mortality rate of infants born to teen mothers: United States, 2017–2018.” *NCHS Data Brief*, no 371. National Center for Health Statistics. 2020. Accessed 30 Apr 2022.

[xvi] Desmond, Matthew. “In order to understand the brutality of American capitalism, you have to start on the plantation.” *The 1619 Project, Slavery and Capitalism*. The New York Times Magazine. Aug 14, 2019, <https://www.nytimes.com/interactive/2019/08/14/magazine/slavery-capitalism.html> Accessed 30 Apr 2022.