Personal Income and COVID-19 Vaccination Rates

-Na Li*-Reed Weiler†-Yun-Jen Tsai‡

2022-01-09

1 Introduction

There is a growing media consensus that income might have something to do with vaccination rates, so we wanted to investigate whether there is any merit to these claims. Research on this topic could potentially aid in understanding how economic inequality poses a barrier to effective strategies for combating the pandemic, and inform future pandemic response strategies.

2 [Our Substance and Context Section Title Here]

(Jenell) Talk about why we choose the two data sets and what methods that we plan to use to see the relations that we are trying to describe.

(Emphasize personal health benefits to boost COVID-19 vaccination rates: https://www.pnas.org/content/pnas/118/32/e2108225118.full.pdf)

(Correlation Between Health and Wealth: https://militaryfamilieslearningnetwork.org/2019/08/08/correlations-between-health-and-wealth/)

3 Data and Methods

Data is drawn from two separate sources. One data set contains personal income level (GDP) by state, the other contains total vaccination rates by state. Our combined dataset has 50 observations, representing each of the 50 united states.

We chose to use a simple linear regression to test the relationship between income per capita and COVID-19 vaccination rates at a state level within the United States. In this analysis, the explanatory variable was income level, and the outcome variable was COVID-19 vaccination rates. We created a scatter plot to model the relationship between income and vaccination rate by state, displayed below.

```
library(readxl)
income <- read_excel("income Ratio.xlsx")
colnames(income) <- c("State.Name", "Income")
head(income)
```

*American University
†American University
‡American University

```
## # A tibble: 6 x 2
##   State.Name Income
##   <chr>       <dbl>
## 1 Alabama    0.0493
## 2 Alaska     0.0758
## 3 Arizona    0.0563
## 4 Arkansas   0.0483
## 5 California 0.0855
## 6 Colorado   0.0732
```

```r
vaccination <- read.csv("vaccination ratio.csv", fileEncoding = "UTF-8-BOM")
head(vaccination)
```

```
##   State.Name Vaccination.Ratio
## 1    Alabama              1.66
## 2     Alaska              1.83
## 3    Arizona              1.78
## 4   Arkansas              1.72
## 5 California              1.98
## 6   Colorado              1.91
```

```r
dat <- merge(income, vaccination)
head(dat)
```

```
##   State.Name     Income Vaccination.Ratio
## 1    Alabama 0.04931529              1.66
## 2     Alaska 0.07579234              1.83
## 3    Arizona 0.05628906              1.78
## 4   Arkansas 0.04834675              1.72
## 5 California 0.08546529              1.98
## 6   Colorado 0.07322607              1.91
```

4 [Our Results Section Title Here]

Here, we explain and interpret our results. We try to learn as much as we can about our question as possible, given the data and analysis. We present our results clearly. We interpret them for the reader with precision and circumspection. We avoid making claims that are not substantiated by our data.

Note that this section may be integrated into Section 3, if joining the two improves the overall presentation.

Our results for the `cars` data include estimating the linear model

$$\text{Vaccination.Ratio}_i = \beta_0 + \beta_1(\text{Speed}_i) + \epsilon_i.$$

5 Extract the coefficient on speed:

```r
income_coef <- coef(lm_out)["Income"]
```

Below we show the model estimates. The first table uses `xtable()`, the second uses `stargazer()` (Hlavac 2018).

Using the `cars` data, we find that each unit of speed is associated with 4.1 more units of distance.

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	1.59	0.07	21.22	0.00
Income	4.06	1.03	3.95	0.00

Table 1: regression result

Table 2: regression result

	Vaccination ratio
	Vaccination.Ratio
Income	4.06^{***}
	(1.03)
Constant	1.59^{***}
	(0.07)
Observations	51
R^2	0.24
Adjusted R^2	0.23
Residual Std. Error	0.19 (df = 49)
F Statistic	15.60^{***} (df = 1; 49)
Note:	*p<0.1; **p<0.05; ***p<0.01

6 Discussion

One limitation of our data is that the CDC doesn't specify whether the "total vaccination" data counts each dose of the vaccine as a separate instance of "vaccination," nor does it specify how it would account for booster shots, etc. So, there is some ambiguity as to how to interpret the results of our analysis, since the outcome variable could potentially be measuring a variety of different scenarios. Future research should strive to distinguish between instances of single-dose vaccination, double-dose vaccination, inclusion of a booster shot, and/or non-vaccination status as separate categories so as to better isolate the statistical impact of income level on each distinct outcome.

References

Hlavac, Marek. 2018. *Stargazer: Well-Formatted Regression and Summary Statistics Tables.* Bratislava, Slovakia: Central European Labour Studies Institute (CELSI). https://CRAN.R-project.org/package=stargazer.