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 [Context]

(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 51 observations, representing each of the 50 united states, and 1 district.

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.

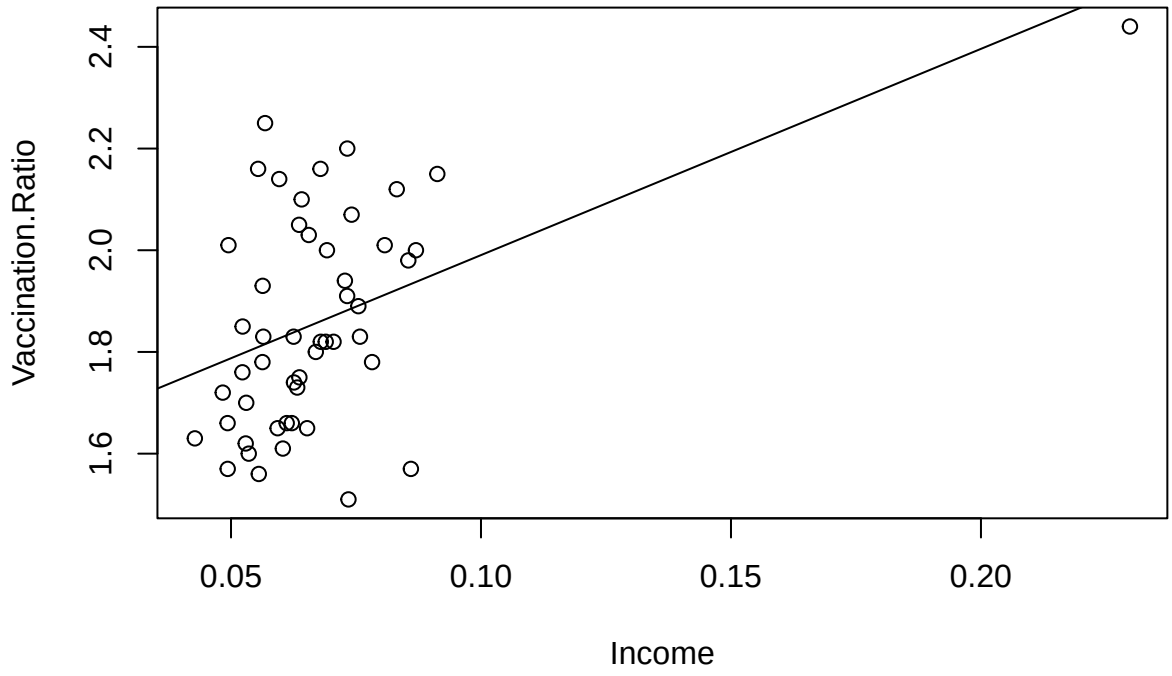
```
dat <- merge(income, vaccination)
head(dat)
```

```
##   State.Name     Income Vaccination.Ratio
## 1    Alabama 0.04931529              1.66
```

*American University
†American University
‡American University

```
## 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
```

```
plot(Vaccination.Ratio~Income, dat)
abline(lm_out)
```



% latex table generated in R 4.1.2 by xtable 1.8-4 package % Sun Jan 9 12:58:18 2022

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	1.58	0.08	20.99	0.00
Income	4.06	1.04	3.91	0.00

Table 1: Regression result

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

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

```
##   State.Name     Income Vaccination.Ratio
## 1    Alabama 0.04931529              1.66
## 2     Alaska 0.07579234              1.83
```

Table 2: regression result

	Vaccination ratio
	Vaccination.Ratio
Income	4.06^{***}
	(1.04)
Constant	1.58^{***}
	(0.08)
Observations	50
R^2	0.24
Adjusted R^2	0.23
Residual Std. Error	0.19 (df = 48)
F Statistic	15.28^{***} (df = 1; 48)
Note:	*p<0.1; **p<0.05; ***p<0.01

```
## 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.

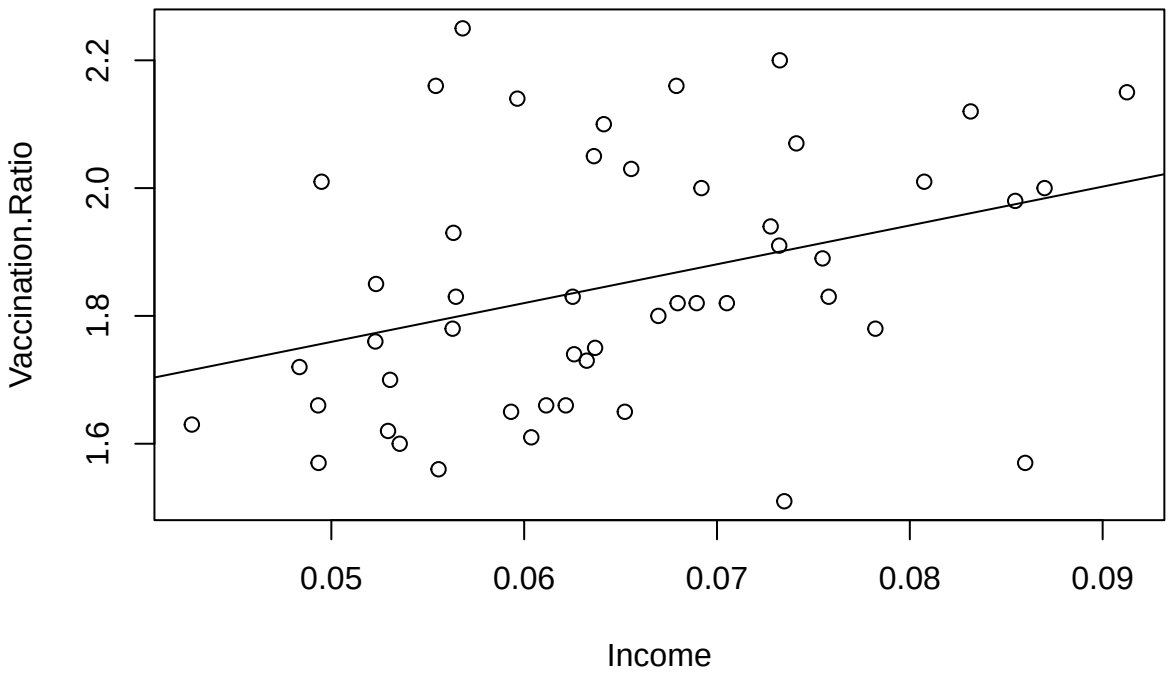
Note that this section may be integrated into Section 3, if joining the two improves the overall presentation.

Our results for the `cars` data include estimating the linear model There are 51 observations in both datasets, including 50 states and 1 district. Our null hypothesis is that changes in percentages of income do not affect vaccination rates. We can get a clear picture of how income affects vaccination rates from this regression result. As the coefficient of income equals to 4.60, it represents that there is a positive correlation between vaccination rate and the changes of income. And, as one percentage of income increases, the vaccination rate increases by 4.06.

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

5 Extract the coefficient on Income:

```
plot(Vaccination.Ratio~Income, dat)
abline(lm_out)
```



% latex table generated in R 4.1.2 by xtable 1.8-4 package % Sun Jan 9 12:58:18 2022

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	1.46	0.16	9.30	0.00
Income	6.07	2.38	2.55	0.01

Table 3: NO DC Regression result

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

5 Reference

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.

6 References

Table 4: regression result

	NO DC Vaccination ratio
	Vaccination.Ratio
Income	6.07**
	(2.38)
Constant	1.46***
	(0.16)
Observations	49
R^2	0.12
Adjusted R^2	0.10
Residual Std. Error	0.19 (df = 47)
F Statistic	6.52** (df = 1; 47)
Note:	*p<0.1; **p<0.05; ***p<0.01