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 Data and Methods

To explore if income level affects vaccination rates,we compare the GDP to total doses of vaccines administered. GDP demonstrates the economy and development level of an area, and thus we think it is a creditable proxy for the individual income level in a state. Data is drawn from two separate sources. One data set contains the Current-Dollar Gross Domestic Product (GDP) by state from 2020Q1 to 2021Q3; while the other contains total vaccination rates by state. Also, in order to get a better understanding of how changes of a state's GDP affects vaccination rates, we also draw each state's total population as a baseline comparison. In general, there are 51 observations in both datasets, including 50 states and 1 district. We chose to use a simple linear regression to test the relationship between GDP and COVID-19 vaccination rates at a state level within the United States. In this analysis, the explanatory variable was GDP, and the outcome variable was COVID-19 vaccination rates. Note that in order to get a better sense of this relation, we divide both each observation value in both datasets by the total population of each state. The name of two datasets after we divide those values by population are "income ratio" and "vaccination ratio." We created a scatter plot to model the relationship as displayed below.

#Finding From the scatter plot, we can see that there is a positive correlation between a state's GDP and vaccination rates. We run the simple linear regression analysis to get a better understanding of how the two variables affect each other. In our analysis, we set the null hypothesis as that a state's GDP has no influence over that state's vaccination rates.

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

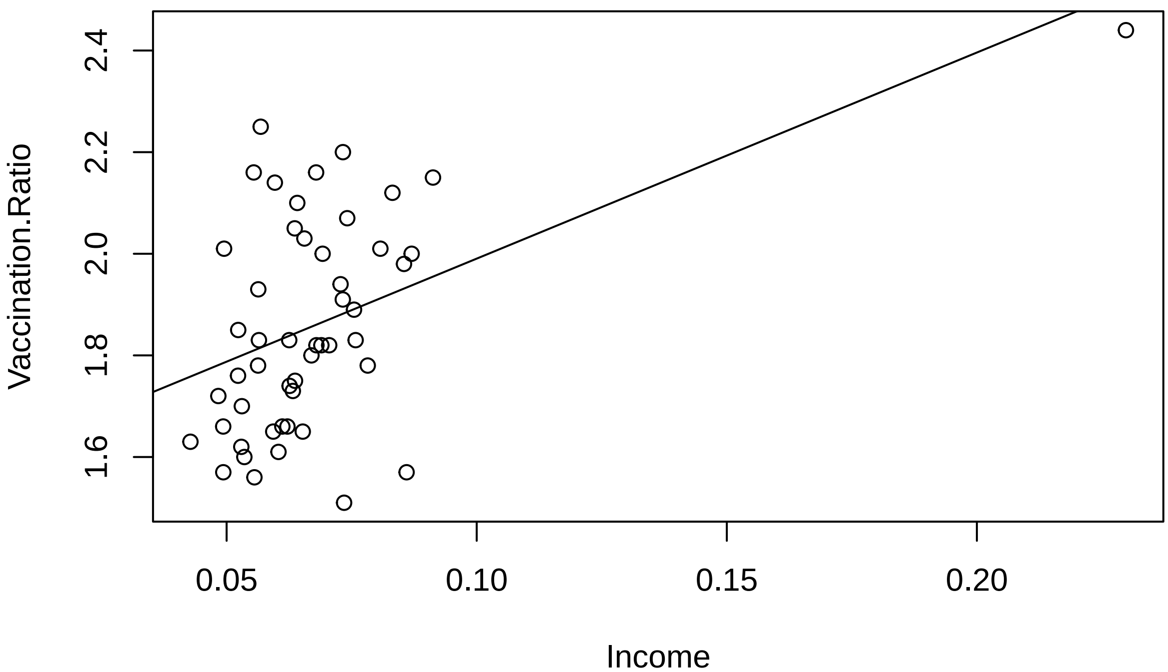
```
##   State.Name     Income Vaccination.Ratio
## 1    Alabama 0.04931529              1.66
## 2     Alaska 0.07579234              1.83
## 3    Arizona 0.05628906              1.78
```

*American University
†American University
‡American University

```
## 4   Arkansas 0.04834675              1.72
## 5 California 0.08546529              1.98
## 6   Colorado 0.07322607              1.91
```

As Table 1 demonstrates, we get a statistically significant prediction of states' GDP and vaccination rates. The fitted regression model was: vaccination ratio = 4.006*(income ratio) + 1.59. The overall regression was statistically significant since (R^2 = 0.24, F(1,49) = 15.6. The income ratio is less than 0.05, so there is significant evidence to suggest that the income ratio is not 0 ($p < 0.01$). The most important information from Table 2 is the R^2 value of 0.24, which shows that 24% changes in vaccination rates can be explained by a state's GDP level.

```
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 14:39:10 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

Additionally as we see in the scatter plot 1, there is an outlier. District of Columbia is the outlier since it has the general income ratio of approximately 0.23, and 2.44 vaccination ratio. As a result, we conduct another simple linear regression analysis to see if DC has a significant impact on the regression model.

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

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

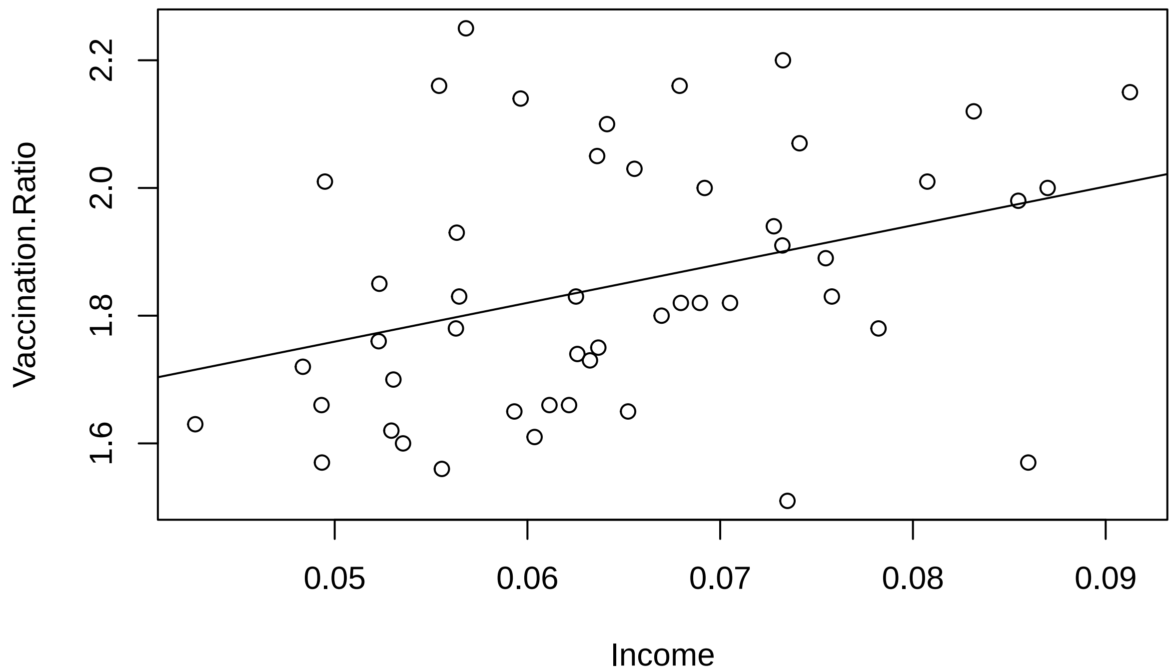
```r
dat <- merge(income, vaccination)
head(dat)
```

```
##   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
```

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.

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

```r
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 14:39:10 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

Here, compared to the regression result in Table 1, the coefficient jumps from 4.06 to 6.07 as the intercept decreases from 1.59 to 1.46. Removing DC demonstrates that income has a larger influence on the vaccination rate to a state.

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

As we see from the second scatter plot, as the outlier is removed, the regression line is tilted flatter, and the observations are evenly distributed. Hence, we can say that DC does impact the overall regression analysis. We further run another simple linear regression analysis to see how much DC impacts the overall result.

3 Conclusion

Our study attempts to explore the relationship between income and vaccination rates. From our analysis, we conclude that a state with higher GDP tends to have a higher vaccination rate. We think this may be that states with higher GDP have more resources to administer more vaccines to people in those states. Nonetheless, we do encounter some challenges. 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,"

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

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.

4 References

Cohen, Elizabeth. “Parents’ income might predict covid-19 vaccination choices for their children.” CNN Health. November 11th, 2021. https://www.cnn.com/2021/11/10/health/parents-covid-19-vaccine-choices-income/index.html

“Gross Domestic Product by State, 3rd Quarter 2021.” Bureau of Economic Analysis. https://www.bea.gov/data/gdp/gdp-state

Data Table for COVID-19 Vaccinations in the United States https://covid.cdc.gov/covid-data-tracker/#vaccinations_vacc-total-admin-count-total

Data for total population (by state) https://www.census.gov/data/tables/time-series/demo/popest/2020s-national-total.html#par_textimage

(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/)