**HHS 2500**

**Data Science I**

**Fall Semester 2021**

**Final Project**

**Title: An Examination of the Relationships between Vaccination, COVID-19 Case Rates and Excess Mortality**

**Report Author: Hoang Nhat Duy Le**

**Language: SAS**

**Date due \_\_12/14/2021\_\_\_\_\_\_**

**Date submitted \_\_12/12/2021\_\_\_\_\_\_**

**Grade \_\_\_/10**

Table of Contents

1. Introduction………………………………… 3
2. Procedure…………………………………... 3
3. Control Variables…………………………... 3
4. Regression….………………………………. 4
5. Appendix…………………………………… 6
6. **Introduction**

A COVID‑19 vaccine is a vaccine intended to provide acquired immunity against severe acute respiratory syndrome coronavirus 2 (SARS‑CoV‑2), the virus that causes coronavirus disease 2019 (COVID‑19). Prior to COVID‑19, a vaccine for an infectious disease had never been produced in less than several years – and no vaccine existed for preventing a coronavirus infection in humans. However, vaccines have been produced against several animal diseases caused by coronaviruses and in this paper, we will research about the effect of the vaccine to determine whether it helps to reduce the virus.

1. **Procedure**

* Research question: Are COVID-19 vaccination rates associated with fewer cases and deaths (excess mortality)?
* Using data that were obtained from Our World in Data and the Worldbank International Database, the purpose of this study was to determine whether vaccination rates significant reduced new COVID-19 cases or excess mortality in approximately 165 nations during fall, 2021. Two linear regressions (Tables 2 and 3 below) were conducted while including controls for national wealth (gross domestic product), medical expenditures, etc. Etc

1. **Control Variables**

Table 1. Descriptive Statistics.

Variable Mean Standard Deviation N

New Case Rate (per million) 175.5 234.2 173

Fully Vaccinated Rate (per hundred) 32.65 25.13 173

Gross Domestic Product 14021 19641 173

Excess Mortality Rate 75.68 149.5 36

Include all other relevant variables

The descriptive statistics in Table 1 indicate that the average case rate across all nations during September and October of 2021 was 175.5. The average full vaccination rate was 32.65, indicating that most adults among the 173 nations were fully vaccinated by the end of October 2021. The average gross domestic product was 14021USD and excess mortality rate was 75.68. These are the variables that we have from the original Covid data (in fact, we need to merge the excess mortality from the GitHub page of Covid Vaccine Data).

New control variables added when merging all the supportive datasets:

* Current Health Expenditure
* Total Population
* Population Over 64
* Unemployment Rate

1. **Regression on New Case Rate and Excess Mortality**
   * + 1. New Case Per Million

Table 2. Regression: New Case Rate as Predicted by Control Variables During Fall, 2021.

Variable Coef. S.E. Sig.

Fully vaccinated rate (per hundred) 2.19 0.98 .027

Current Health Expenditure -12.89 9.08 .1583

Unemployment Rate 7.00 2.83 .015

Include all other relevant variables 11.11 1.11 .999

N=154 R2 = 0.1623 Adj R2 = 0.1221

Contrary to expectations, the regression coefficients from Table 2 indicate that that countries with higher vaccination rates experienced higher overall new cases of COVID-19. It can explain that in countries with high rate of new cases, people are more encouraged to take the vaccine which helps to boost the rate of people fully vaccinated. Additionally, nations with greater Unemployment Rate had higher new case rates. Even though there is not an obvious correlation, but we can interpret that countries with high unemployment rate which results in lower GDP which results in lower vaccine availability to prevent virus from spreading. Finally, nations that invested more in medical care also experienced lower new COVID case rates. This is understandable since these countries are willing to buy vaccines for their residents which may help to reduce the number of new cases. Nevertheless, these cross-sectional regressions cannot scientifically demonstrate cause-effect relationships. Consequently, the results in Table 2 could also be interpreted as indicating nations with high COVID-19 exposure were more likely to encourage vaccination.

* + - 1. Excess Mortality:

Table 3. Regression: Excess Mortality Rate as Predicted by Control Variables During Fall, 2021.

Variable Coef. S.E. Sig.

Fully vaccinated rate (per hundred) -1.61 1.71 .35

Domestic Health Expenditure -26.39 17.61 .14

Total Population - .0334 1.10 .02

Include all other relevant variables 11.11 1.11 .999

N=36 R2 = 0.41 Adj R2 = 0.27

As we expected, the regression coefficients from Table 3 indicate that that countries with higher vaccination rates experienced lower excess mortality from COVID-19. Additionally, nations with greater Domestic Health Expenditure had lower excess mortality as well. Finally, something unexpected occurs when nations with higher population had lower excess mortality rate. This can explain that countries with high population are usually developed countries and they have high expenditure on Health and Education. Therefore, their residents may have better protection from Covid 19. Negative coefficients lead to inverse correlation.

1. **Conclusion**
   * + 1. Did higher vaccination rates reduce COVID-19 cases? (yes, no, unclear).

Even though our data show that higher vaccine rate leading to more new cases, but it is only temporary, as proved by the health expenditure of each country. Therefore, while our research is not time-variant, it somehow shows the benefits of vaccines to fight against COVID-19.

* + - 1. Did higher vaccination rates reduce excess mortality? (yes, no, unclear).

Yes, it was showed in Table 3 and it also shows that more health expenditure helps to reduce number of excess mortality.