Does Covid-19 have the same impact on the employments in the developed and developing countries?

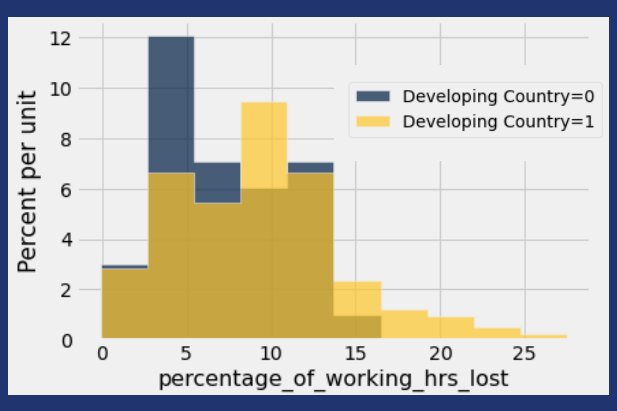
Name: Jiayi Zhou (Joy) | Course Title: DIGHUM 100 | Instructor: Adam Anderson | Date: 2021/6/27

Introduction

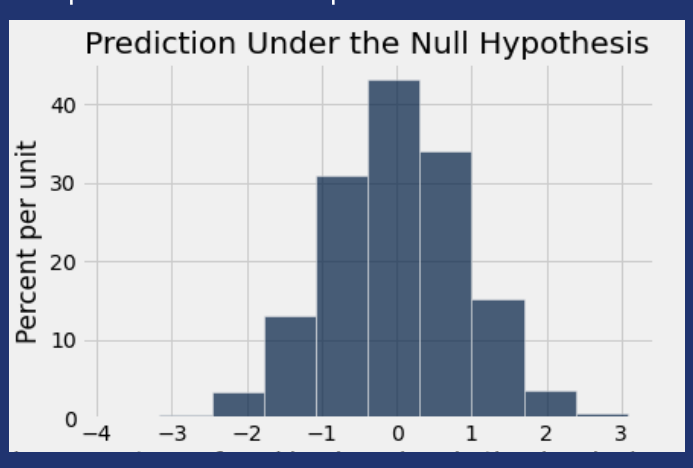
Starting from January 2020, the world has experienced the one of the greatest pandemic in the human history. The covid-19 virus not only jeopardized the global health conditions, challenged the medical infrasturctures, the subsequent quarantine also disrupted the employment conditions in countries around the globe. According to report titled "COVID-19: Tackling the Jobs Crisis in the Least Developed Countries," produced by the International Labour Organizations, the contagion was less severe in the least developed countries because of the early loackdowns and "limited connectivities." It was the employment and job markets that absorbed most of the shocks in these countries.

The research question in this project would be "Does Covid-19 have the same impact on the emplyment of developed and developing countries?" Our initial hypothesis is that the developing countries would be impacted more negatively. The data analysis we run in this project intends to provide numerical evidences to support, or perhaps against, our initial prediction. In addition, we would consider the distinct social and economic characteristics of the developed and developing countries, producing a potential explanation to explain what we find.

Question No.1: Are the percentages of working hour lost caused by Covid-19 differ between the Developing and the Developed Countries?



Basic visualization: The color-coded histogram above takes 189 countries into account. The x-axis represent the percentage of working hours lost. The bar area present the percentage of these 189 countries which has lost x percentage of working hours.The tail of the yellow bars, which represent the developing countries, have further spread to the right. The tail of the blue bars, which represent the developed countries, concentrate more on the left. Graphically speaking, the developing countries seem to lose larger percentage of working hours compare with the developed countries.



Dataset and country categorizations

The dataset that is used in this project is titled as "Impact of Covid-19 on employment - ILOSTAT," which is published on Kaggle.

Added dummy variable "Developing Country": The original dataset doesn't contain the information on whether a country is developed or developing. This project has edited the dataset by adding a dummy variable "Developing Country." If a country belongs to developed country, this indicator equals to 0; otherwise it equals to 1. This country categorization is based on the United Nation's report titled "World Economic Situation Prospects." The UN has provided a list of transitioning countries that is neither developing or developed. In this project, all of the transitioning countries are considered as developing countries.

Developed Countries: Canada, United States, Australia, Japan, New Zealand, Austria, Belgium, Denmark, Finland, France, Germany, Greece, Ireland, Italy, Luxembourg, Netherlands, Portugal, Spain, Sweden, United Kingdom, Bulgaria, Croatia, Cyprus, Czechia, Estonia, Hungary, Latvia, Lithuania, Malta, Poland, Romania, Slovakia, Slovenia, Iceland, Norway, Switzerland

Variables included in the dataset: There are 10 total variables included in the dataset, but we wound't need all of it in this project. Variables that are USED are explained below (quoted directly from Kaggle):

- # total_weekly_hours_worked(in_thousands): Total weekly hours worked of employed persons in 2020
- # percentage_of_working_hrs_lost(%) : percentage of hours lost compared to the baseline (4th quarter of 2019)
- # labour_dependency_ratio : Ratio of dependants (persons aged 0 to 14 + persons aged 15 and above that are either outside the labour force or unemployed) to total employment.
- # employed_female_25+_2019(estimates in thousands) - Employed female in 2019 who, during a specified brief period, were in one of the following categories: a) paid employment (whether at work or with a job but not at work); or b) self-employment (whether at work or with an enterprise but not at work).
- # employed_male_25+_2019(estimates in thousands) - Employed male in 2019 who, during a specified brief period, were in one of the following categories: a) paid employment (whether at work or with a job but not at work); or b) self-employment (whether at work or with an enterprise but not at work).

AB Testing:

Unlike the basic visualization, the primary goal of this test is to investigate whether the percentage of working hours lost in the developed countries come from the same underlying distribution with the developing countries. Our test-statistic would be the mean of percentage working hour loss in the developing countries minus the mean of percentage working hour loss in the developed countries. **The null hypothesis would be these averages have no difference** -- zero test statistic favors the null hypothesis. **The alternative hypothesis would be the percentage working hour loss is higher in the developing countries** -- positive test statistic favors the alternative hypothesis.

This hypothesis test has shuffled the dummy variable for country categorization, and repeated 5000 trials. The blue histogram on the bottom left displays the visualization when we assume the null hypothesis is true. That is, the difference between the percentage of working hour loss in the developed and developing country would center at zero. The observed statistics is positive and it is around 1.85. Although the shuffling process has introduced randomness, the p-value was around 1.7% and was never greater than 2%. If 5% is the significance level, this hypothesis testing would reject the null hypothesis. **It concludes a statistically significant difference between the percentage of working hour loss in the developing and the developed countries.**

Question No.2: Are the workers in the developing countries already working longer hours before the pandemic?

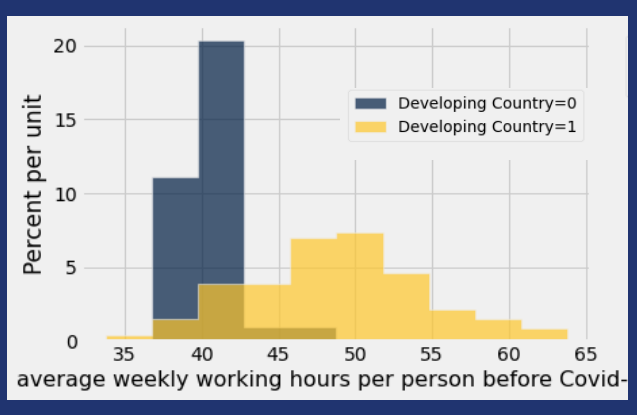
Knowing the employed population (females + males) in each country, the total working hours after the pandemic, and the percentage of working hours lost, this project has calculated the average weekly working hour per person before the pandemic. This variable is added as an extra column in the dataset.

Narrowing down the concept of employment and research question

Since employment is such a broad concept, this project narrowed down this term and focused on

percentage loss of working hours

average working hours per person before Covid-19



The two-color histogram on the top displays different distributions of average weekly working hours per person in the developed and developing country. It is notable that the distribution of the developed country, coded as blue, concentrated significantly on the left. The workers in these developed countries tend to work 40 hours per week. Compare with the developed countries, the workers in the developing countries (coded as yellow) tend to work up to 65 hours per week before the pandemic.

	coef	std err	t	P>|t|
const	40.4414	0.886	45.663	0.000
x1	8.2324	0.984	8.363	0.000

The above image presents a part of the result from Regression Analysis 1 - the indepedent variable is the dummy variable "Developing Country", and the dependent variable is average weeking working hours per person before Covid-19. The employees in the developing countries work 8.2324 hours more than the employees in the developed countries on average. Since the p-value for the coefficient x1 is 0, it indicates that the coefficient x1 is statistically significant. In other words, the difference between the average working hours in the developing and developed countries is statistically significant.

Links to the project: Github , Slides, and the dataset

Question 3: Did the developing countries lose a larger percentage of weekly working hours because they were already working longer hours before the pandemic?

	coef	std err	t	P>|t|	[0.025	0.975]
const	9.8729	2.703	3.653	0.000	4.541	15.205
x1	-0.0224	0.057	-0.394	0.694	-0.135	0.090

Omnibus:	23.998	Durbin-Watson:	1.764
Prob(Omnibus):	0.000	Jarque-Bera (JB):	30.228
Skew:	0.828	Prob(JB):	2.73e-07
Kurtosis:	4.048	Cond. No.	364.

It may be possible that there is a universal pattern that the more hours people work before the pandemic, the larger percentage of working hours they could lose after the pandemic, regardless of whether they are in developing or developed countries. The above regression test is to see whether the percentage of working hour lost and the pre-pandemic average working hour is correlated. **It turns out that this is not the case.** According to the p-value associated with the x1 coefficient above, the x1 coefficient is statistically insignificant. In other words, the regression analysis above does not provide evidences that the more people work before the pandemic, the larger percentage of working hours they would lose.

Our findings and the Story

Based on the analysis presented, this project suggests that:

1. The developing countries tend to lose larger percentage of working hours due to covid-19. This effect is further amplified by the fact that the workers in the developing countries already have higher average of weekly working hours before the pandemic.
2. It may be possible that there is a max threshold that people could work per week after the pandemic occur, causing the percentage loss of working hours to be larger in countries with higher average of pre-pandemic weekly working hours. However, the second regression analysis in this project has proven that this is not the case. There were no statistically significant correlation between pre-pandemic working hours and the percentage of working hour loss.
3. The developing countries suffered from larger percentage of working hour loss because: a) their fragile economies focus on manufacturing, construction, services, and minning. These jobs are easily replaceable. During the pandemic period, the unemployment rates associated with these job could be high, bringing up a large percentage of working hour loss countrywise; b) The economies of the developing countries reply heavily on the global economy, and Covid-19 pandemic is a global event that would affect us all. The developing countries don't have sufficient abiities to buffer the pandemic shock within their country.

Anknowledgement and works cited

This project has referenced the employment studies in the report titled “Covid-19: Tracking the Jobs Crisis in the Least Developed Countries” produced by International Labour Organization.

It also referenced codes and statistical knowledge from UC Berkeley Data 8 course demo, taught by Professor Ramesh Sridharan and Swupnil Sahai in Spring 2020. The academic and internet sources of the codes in this project are precisely cited in the Jupyter notebook.

Images included in this poster:

- https://keydifferences.com/difference-between-developed-countries-and-developing-countries.html
- https://www.financialexpress.com/economy/job-crisis-unemployment-rate-at-a-45-year-high-of-6-1-per-cent/1594824/
- https://www.businesstoday.in/latest/world/story/covid-19-drives-italy-to-recession-economy-nosedives-128-in-q2-271885-2020-09-01