The Gender Unemployment Gap

Yetong Xu

Executive Summary

In recent years, following the COVID-19 pandemic, many countries have experienced periods of high unemployment rates. This has made me curious: could there be gender differences in unemployment rates due to various factors? If you are someone who has experienced unemployment, are interested in this topic, a student looking to conduct related research or exploration, or a government policy-maker, then congratulations, you are the intended audience for this project! Below is the link to this project: https://xuyetongseattle1900.shinyapps.io/Unemployment_Gender_Gap/

Background and Context

UNCTAD analysis(Zarrilli, 2021) shows that the COVID-19 pandemic has had a negative impact on both women's and men's employment – but at different stages of the crisis due to the gender segregation of economic activities in many countries. Early measures to curb the spread of the virus first hit jobs held predominantly by women, such as personal services. At the outset of the pandemic, a higher prevalence of the virus correlated with a higher rate of female unemployment. But as the crisis worsened and disrupted cross-border value chains, the impact on men’s employment increased because they tend to work in sectors and jobs that are more dependent on international trade. Even more worrisome, though, than how the pandemic has affected unemployment rates is its impact on women's participation in the labour market. The available data reveals that even in countries where men’s unemployment rate outpaced that of women, more women left the labour market entirely in 2020. The significant drop in the number of women actively searching for a job threatens to reverse decades of progress on women’s empowerment.

However, the issue of the gender gap in unemployment rates is not just a contemporary concern. Even before the pandemic, various factors contributed to this disparity. As mentioned in the study of Stefania(2013), the unemployment gender gap, defined as the difference between female and male unemployment rates, was positive until 1980. This gap virtually disappeared after 1980--except during recessions, when men’s unemployment rates always exceed women’s. At the cyclical frequency, they find that gender differences in industry composition are important in recessions.

Data

This project incorporates a variety of data sources to ensure a comprehensive examination of unemployment gender gap in the United States. I mainly use four data sets: **1.childcare_costs.csv, 2.counties.csv, 3.State Shapefiles, 4.County Shapefiles.**

The first data set (https://raw.githubusercontent.com/rfordatascience/tidytuesday/master/data/2023/2023-05-09/childcare_costs.csv) is sourced from the National Database of Childcare prices, the one I use in this project is the condensed/separated files via Tidy Tuesday project (https://github.com/rfordatascience/tidytuesday/tree/master), a weekly social data project organized by the R4DS community.

The second data set (https://raw.githubusercontent.com/rfordatascience/tidytuesday/master/data/2023/2023-05-09/counties.csv) is a file that lists the county FIPS code, county name, state, and state abbreviation for all counties in the U.S. If you're interested in exploring the documentation for the Tidy Tuesday dataset, including variable types and definitions, please find it here: https://github.com/rfordatascience/tidytuesday/blob/master/data/2023/2023-05-09/readme.md

The third data sets are geographic shapefiles for U.S. states. You can read in the data by using the `tigris::states()` function from the `tigris` package in R. This function accesses the United States Census Bureau's TIGER/Line Shapefiles, providing detailed and up-to-date geographic information on state boundaries.

Similarly, the fourth data sets are geographic shapefiles for U.S. counties. You can obtain the data by using the `tigris::counties()` function from the `tigris` package. These shapefiles offer precise boundary outlines and additional geographic details for each county in the United States.

Fortunately, all the above four data sets don't have licenses or permission limitations, and all of them are collected by the original sources. And all the data used in this project don't include any personal or privacy information. However, even if the dataset itself does not contain personal information, it is still necessary for us to consider whether the processing and dissemination of the data could inadvertently expose or infer individual identities. For example, even if the data is anonymized, individuals who was under unemployment might still be identifiable in cases of small sample sizes or specific regions.

At the same time, it is important to note that while the childcare_costs.csv file contains 34,567 observations and 61 variables, and the counties.csv file has 3,144 observations and 4 variables, this project only includes data from the years 2008 to 2018. This means that we can only explore the changes in unemployment rates and gender disparities in unemployment rates during this ten-year period. I think this is one of the biggest limitations of the data. It means that the unemployment outcomes we observe are influenced by the global financial crisis of 2008. We cannot consider the changes in unemployment rates and gender disparities over a longer time frame. This will impact the analysis of this project. Also, the most important variables in this project are gender and age. After cleaning the data, males and females each account for about 50% of the data, and the age groups 16 and over and 24 to 64 each account for about 50%.

For data cleaning, I selected all the variables related to unemployment rate and year from childcare_costs.csv firstly. Then I used `pivot_longer` function to transform it into Long Format, and re-coded the variables which represent gender and age. After that, I used `inner_join` function to merge childcare_costs.csv and counties.csv. At last, while constructing the Shiny App in another file, I've imported the state and county shapefiles and incorporated visualization charts. However, when merging the county shapefiles with the previously cleaned unemployment data, this project lost data for several areas (California, Alaska, Arizona). This issue will be addressed in future research.

Technology&Platform

In this project, I used ggplot2 and tmap to do the data visualization. ggplot2 is a widely used data visualization package in R. We can use it to create complex and aesthetically pleasing graphics. In this project, ggplot2 has been utilized to create the line chart and the bar chart which help in understanding the trends and patterns of the unemployment rates' gender gap. Another R package, tmap, is specifically tailored for thematic maps. It's an essential tool for geospatial data visualization used in this project. With tmap, I was able to effectively map unemployment data, providing a geographical perspective to the gender gap analysis.

Also, the platform for this project is a Shiny web application, which is an R package that makes it easy to build interactive web apps straight from R. This allows for a dynamic and user-interactive way to present data and findings, making the information more accessible and engaging. The final product of this project is hosted on shinyapps.io, a platform for deploying Shiny applications on the web. This hosting service is specifically designed for R-based applications.

Analysis

In this project, I first used line charts to display the trends of average unemployment rates over time for different genders and age groups. I attempted to initially observe whether there were different overall trends in average unemployment rates for different genders across various age groups.

Subsequently, I calculated the unemployment gender gap for each age group in different years, by subtracting the female unemployment rate from the male unemployment rate for the same year and age group. Then, I used bar charts to present these results. At this point, I've observed the unemployment gender gap over time. But I believe that, given we only have data for ten years, merely observing the changes in unemployment rates over time does not yield many strong conclusions.

Therefore, I then used a static map to show the unemployment gender gap across different states in the United States. We can observe that, over these ten years, the unemployment gender gap vary considerably between different states. Thus, finally, in this project, I provided an interactive map. If you are interested in the unemployment gender gap of any state, you can select the state you are interested in for a more detailed observation at the county level. For this interactive map, I've also used k-means for clustering.

Findings and Discussion

The goal of this study is to observe whether there are significant differences in the unemployment gender gap in the United States between the years 2008 and 2018, across different years or states.

Initially, from the line charts, we can see that, for both men and women, the trend in the average unemployment rates is roughly the same. From 2008 to 2013, the average unemployment rate continuously rose, peaking in 2013, and then showed a downward trend. Also, throughout these ten years, the average unemployment rate for men was consistently higher than that for women. This period coincides with the global financial crisis and subsequent recovery

According to the bar charts, the age group 16 and over has a higher unemployment gender gap compared to the 20 to 64 age group. Additionally, in each age group, the peak in the unemployment gender gap was reached in 2013.

Subsequently, based on the results shown in the static map, we can observe that the unemployment gender gap varies significantly between different states, with Michigan and West Virginia, among others, having a higher unemployment gender gap during the period from 2008 to 2018.

Finally, following the results displayed on the previous static map, we can opt to conduct a more detailed examination of the interactive maps for states like Michigan and West Virginia. We find that there are differences between different counties too. Some counties have a noticeably higher male unemployment rate compared to the female rate, while in other counties, the female unemployment rate exceeds that of males.

What's more, I think the higher unemployment gender gap in state such as Michigan and West Virginia can be attributed to their economic structures and the impact of the 2008 financial crisis. Michigan's economy, heavily reliant on the automotive industry, and West Virginia's focus on coal mining. Both of these two industries are predominantly male-dominated sectors, and experienced significant job losses during the crisis. This disproportionately affected male employment, widening the unemployment gender gap. Additionally, these states' recovery patterns after the financial crisis, influenced by industry-specific challenges and workforce participation trends, further contributed to these disparities. Occupational segregation, which means men and women are concentrated in different job types, coupled with state-specific economic and policy factors, played an important role in shaping the unique unemployment landscapes in these states.

From the findings above, we can ask policy-makers to develop policies which aim to encourage gender diversity in traditionally male or female-dominated sectors, and implement educational initiatives that encourage young people, regardless of gender, to pursue a variety of career paths, including those in STEM fields. Meanwhile, policy-makers should recognize that the economic structure of each state is unique, recovery strategies during economic downturns should be state-specific. For example, they should focus on supporting industries which are most affected in each state, giving economic incentives for companies that actively reduce the gender gap in employment, and providing better access to childcare and other support services, enabling more individuals, especially women, to participate in the workforce.

References

Zarrilli, S., & Luomaranta, H. (2021, April 8). Gender and unemployment: Lessons from the COVID-19 pandemic. UNCTAD. Retrieved from https://unctad.org/news/gender-and-unemployment-lessons-covid-19-pandemic

Albanesi, S., & Şahin, A. (2013). The Gender Unemployment Gap (Staff Report No. 613). Federal Reserve Bank of New York. Retrieved from https://www.newyorkfed.org/medialibrary/media/research/staff_reports/sr613.pdf