

# Report: Comparative Analysis of Wages and Employment-to-Population Ratio by Education in the US

Nhat Khang Vo - 22663032 - of27eruw

## 1 Introduction

This report presents a comparative analysis of wages and the employment-to-population ratio in the United States from 1979 to 2022. The primary objective is to address the Question: Is there a direct correlation between wages and employment rates across different education levels in the US (1979–2022)? Specifically, the analysis examines the relationship and potential dependency between wages and employment rates within distinct educational groups, providing insights into the labor market dynamics and the role of education in shaping economic outcomes.

## 2 Data Sources

For this report two data sources were chosen which represent each type: 1. [Wages by Education](#), detailing wages categorized by education levels. 2. [Employment-to-Population Ratio](#), providing employment rates across educational groups. These datasets were selected to analyze potential correlations between wages and employment levels. Both datasets are provided by the same author, [Asaniczka](#), who has ensured they are well-documented, contain no missing data, and are sourced from reliable references.

### 2.1 Data Structure

The datasets are extracted using an ETL pipeline implemented via Kaggle and Bash scripts. Since both datasets were created by the same author, the necessary categories for this analysis share consistent column names, simplifying the integration and subsequent analysis. For this project, the selected categories include Advanced Degree, Bachelor's Degree, Some College, High School, and Less than High School. The "Employment-to-Population Ratio" dataset represents its data as percentages, while the "Wages by Education" dataset adjusts its figures using the [CPI-U-RS \(Consumer Price Index Research Series\)](#). The CPI-U-RS accounts for inflation over the years by standardizing the value of wages in terms of constant dollars. This adjustment ensures that the analysis accurately reflects real purchasing power and economic trends, making it more suitable for comparing wages over time. By eliminat-

ing the distortion caused by inflation, the CPI-U-RS provides a clearer picture of wage changes and their potential relationship with employment trends across different educational attainment levels. The datasets used in this analysis are complete, containing no missing data, and are sourced from reputable and reliable providers. For the purpose of this report, the analysis is limited to the timeframe between 1979 and 2022. This period provides a robust span for examining long-term trends and relationships in wages and employment, offering valuable insights into labor market dynamics over multiple economic cycles.

### 2.2 License

For both datasets, the author has applied the [CC0: Public Domain](#), which grants unrestricted rights for use, modification, and distribution. While crediting is not required under this license, it will be provided as a courtesy for this project.

## 3 Analysis

In this section, we analyze both wages and employment rates across various educational categories. The first part of the analysis compares the changes in wages and employment rates, highlighting trends and differences among the different educational levels. The second part employs regression analysis to identify potential patterns and relationships between wages and employment rates, offering a quantitative assessment of how these two variables interact within each educational category.

### 3.1 Wage and Employment Change

The following graphs illustrate year-over-year percentage changes, showing how the respective values increased or decreased compared to the previous year. This section focuses on identifying any discernible patterns or trends in employment and wage changes over time.

In the first figure (1), which depicts the changes in wage and employment rates for individuals with an "Advanced Degree," there appear to be no noticeable similarities between the two line graphs. The fluctuations in both graphs are substantial and do not exhibit any consistent or discernible trends over time.

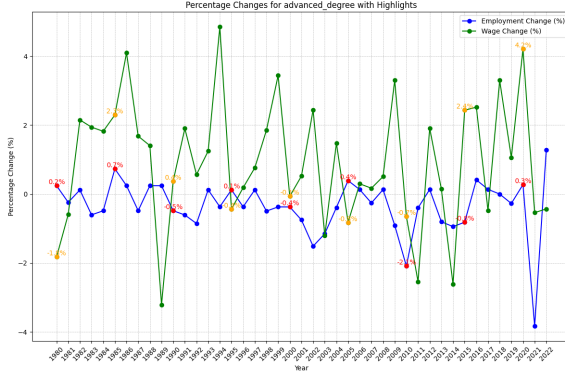


Figure 1: Wage and Employment Change (%) of advanced degree

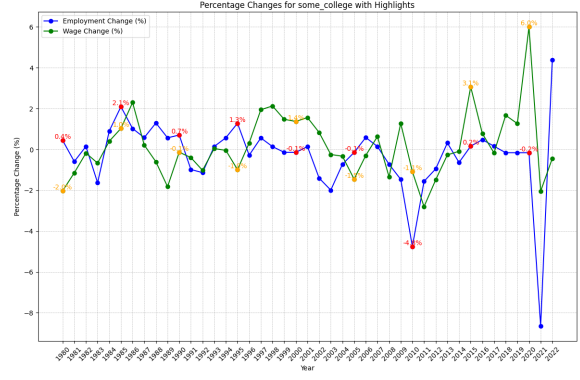


Figure 3: Wage and Employment Change (%) of some college

While figure (2) for "Bachelors Degree" the line-graphs start to align more to each other (from here on out, look at the timeframe 1990 - 2005), there are still high variance between the two graphs.

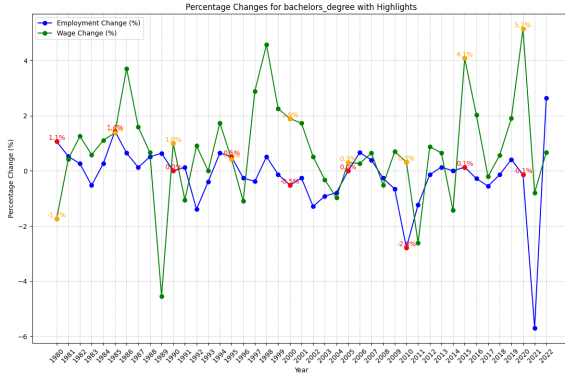


Figure 2: Wage and Employment Change (%) of bachelors degree

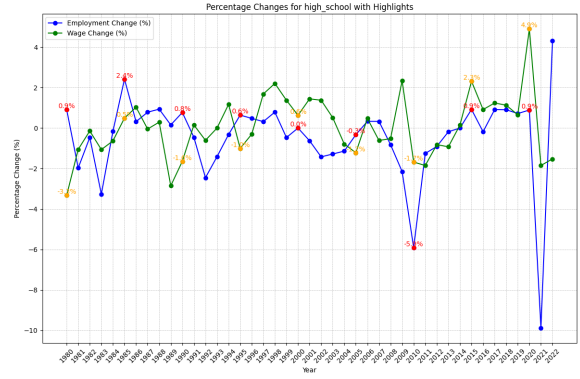


Figure 4: Wage and Employment Change (%) of high school

With lower educational level ("Some College" (3), "High School" (4) and "Less than High School" (5)) the graphs exhibit a more consistent trend, with the Employment and Wage Change lines showing less variability compared to those with higher educational levels. Specifically, during the 1990-2005 period, the graphs follow a similar trajectory, indicating a more uniform relationship between wage and employment changes for their respective group.

This section suggests that, at lower levels of education, there is a stronger correlation between wage and employment changes. The following subsection delves deeper into this relationship by applying a linear regression model for a more detailed analysis.

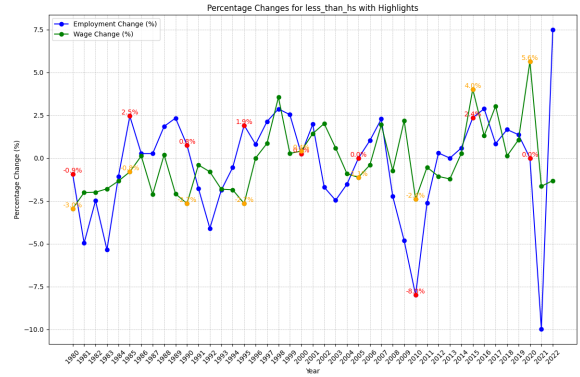


Figure 5: Wage and Employment Change (%) of less than HS

### 3.2 Regression Model Analysis

To explore the relationship between wage and employment rate, we will use linear regression and a scatterplot. The **R-squared** value indicates how well the model explains the data. It ranges from 0

to 1, with values closer to 1 suggesting a strong relationship and values closer to 0 indicating a weak one. The **P-value** helps assess whether the relationship is statistically significant. A P-value below 0.05 generally means the relationship is meaningful and not due to random chance, while a P-value above 0.05 suggests that the observed relationship might be due to randomness.

By analyzing these two statistics, we will determine if a significant trend exists between wage and employment rate.

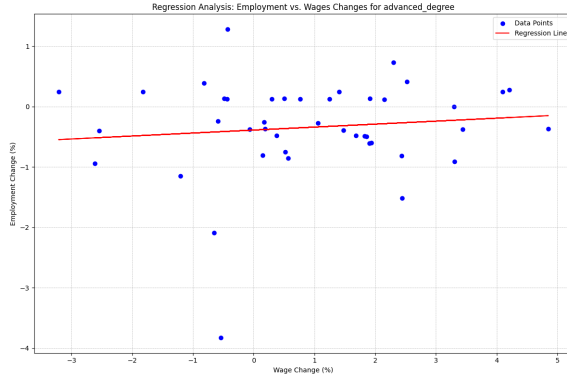


Figure 6: Regression plot category "Advanced Degree" — R-squared: 0.0126 — P-value: 0.4730

From the previous section, we observed that for the "Advanced Degree" category, there is no meaningful relationship between wage and employment rate. This is supported by a very low R-squared value of 0.0126 (1.26%), indicating almost no correlation, and a high P-value of 0.4730, showing that the relationship is not statistically significant (6).

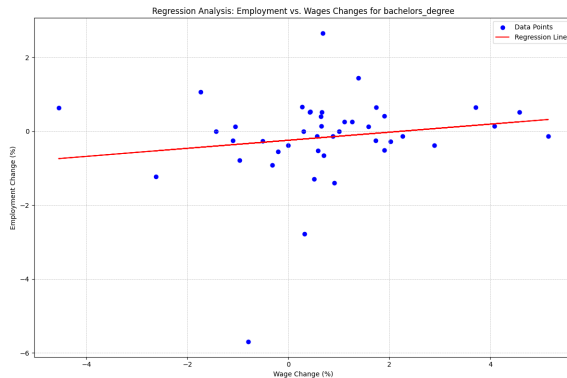


Figure 7: Regression plot category "Bachelors Degree" — R-squared: 0.0266 — P-value: 0.2963

While the R-squared and P-value improve as educational levels decrease (7)(8)(9)(10), no category demonstrates a strong or meaningful relation-

ship. All categories exhibit low R-squared values, indicating weak explanatory power, and P-values that are significantly above the standard threshold of 0.05 in most cases, further confirming the lack of statistical significance in these relationships.

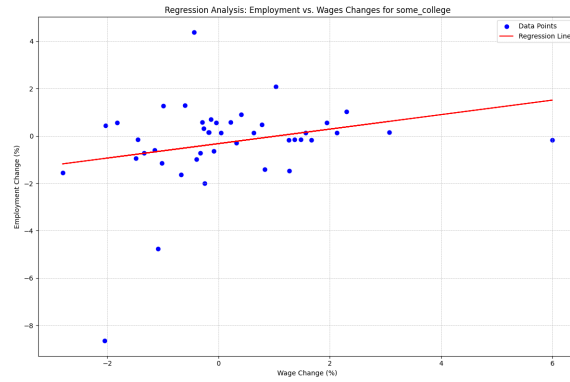


Figure 8: Regression plot category "Some College" — R-squared: 0.0683 — P-value: 0.0906

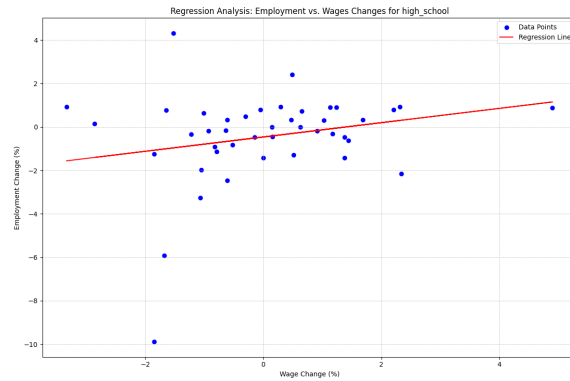


Figure 9: Regression plot category "High School" — R-squared: 0.0552 — P-value: 0.1292

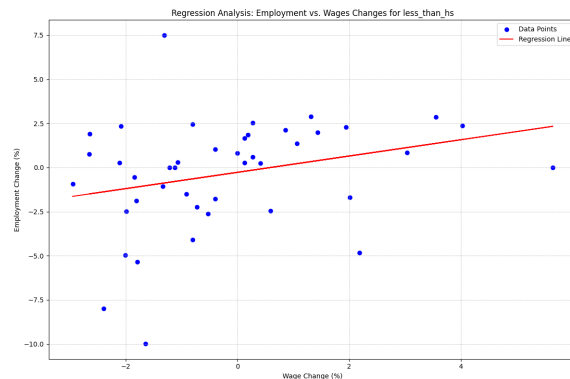


Figure 10: Regression plot category "Less than High School" — R-squared: 0.0802 — P-value: 0.0657

The category closest to showing a relationship is "Less than High School" (10), which has a low R-squared value of 0.0802, indicating that only 8.02% of the variation in the relationship between wage and employment rate is explained. Additionally, the P-value of 0.0657 is slightly above the conventional threshold of 0.05, meaning the relationship is not statistically significant. However, since the P-value is close to the threshold, we can cautiously suggest the possibility of a weak significance for the relationship.

We can also visually confirm from the scatterplots that there may be no clear relationship, as the data points form widely spread clusters. This dispersion indicates a weak correlation between the variables.

other influential factors. Variables such as time lags in wage adjustments, economic shocks, policy changes, and external influences could independently affect wages and employment rates. These factors, along with the potential for delayed effects on the relationship between wages and employment rates, warrant further research. A more comprehensive study incorporating such variables and more sophisticated modeling techniques is necessary to better understand the intricate dynamics of the labor market and the role of education within it.

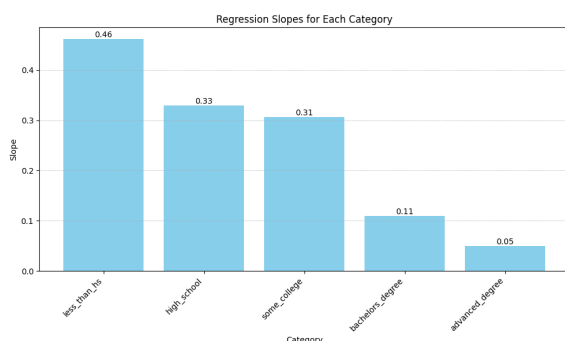


Figure 11: Regression slope barchart

Furthermore, in many cases, the slope (11) of the regression line is very small, further suggesting that any observed relationship between wage and employment rate is minimal and unlikely to be practically meaningful.

## 4 Conclusion & Discussion

The analysis reveals that there is no strong or significant relationship between wages and employment rates across different levels of educational attainment. While the correlation can be clearly dismissed for higher educational levels, a slight improvement is observed at lower educational levels. However, regression analysis indicates only a weak significance for the relationship at the "Less than High School" level, which also exhibits minimal variation.

To directly address the question posed in this report: No, there is no strong or consistent correlation between wages and employment rates across different education levels identified in this analysis.

It is crucial to note, however, that this analysis is an initial exploration and does not account for