AGender Wage Gap: A Panel Data Analysis of OECD Countries

NAME

DATE

**Abstract**

There are several negative effects of the gender wage difference, including lower living standards, less savings, and less money for retirement for women. Additionally, it reduces women's motivation to participate in the job market and discourages them from investing in education and training, which hinders economic growth and productivity. In order to understand the gender wage gap and guide policy responses, this study will use panel data analysis to identify significant factors.

The fixed effects model and the random effects model are the two econometric models used in the study. To account for stable, unobserved traits that could affect the dependent variable, the fixed effects model contains individual-specific dummy variables. With the help of this model, it is possible to examine how independent variables have an impact on the pay gap while taking individual differences into account.

In contrast, the random effects model treats effects that are specific to a particular person as random variables linked to the independent components. It accounts for both individual variations and components with time-varying characteristics while capturing the average link between the independent and dependent variables.

The fixed effects model should be used in the analysis of the wage\_gap dataset, according to the Hausman test results. This method successfully tackles the effects that are unique to each individual, offering insightful information about the causes of the gender wage gap and providing direction for policy measures meant to close it.

**Introduction**

Women being paid less than males for doing the same work, which is a serious issue that has negative effects like lower living standards, less savings, and lower retirement income. This inequality discourages women from investing in education and training and lessens their incentives to participate in the labor market, which results in a cycle of poverty for women and their children. It also has a detrimental effect on economic growth and productivity. The ideals of equality and human rights are also violated by gender-based pay discrimination. In order to solve this problem, it is crucial to examine the gender wage gap using panel data analysis in order to pinpoint the underlying causes and guide the development of policy initiatives aimed at lowering it. Low p-values for the SUBJECT and TIME variables support the main hypothesis addressed in this study using the fixed effects model, which implies a substantial association between individual-specific qualities and the dependent variable. The wage gap is mostly explained by individual variations, according to the random effects model, which also emphasizes the major role of individual-specific effects. Both the estimated variance and the proportion of the total variance are both significant. The random effects model's statistically significant coefficients for the SUBJECT and TIME variables highlight the importance of both personality traits and temporal circumstances in determining the observed salary disparity. In summary, both models emphasize the importance of individual-specific effects and offer useful insights into the variables causing the pay gap, directing appropriate solutions to the problem.

**Literature review**

The research examines the gender wage disparity in OECD nations from 1980 to 2010. The authors estimate the gender wage gap in each nation and pinpoint the causes of the disparity using panel data from the OECD's Labor Force Statistics.

According to the authors, there is a sizable gender wage disparity in every OECD nation. Women made 82% of what men made in 2010, while the gender wage gap was at a median of 18%. The differences were greater in some nations than others, with Turkey (43%) and Mexico (38%), respectively, having the greatest discrepancies.

The authors point out a number of causes for the gender pay gap, including occupational segregation, discrepancies in training and experience, discrimination, and family obligations. Occupational segregation draws attention to the propensity of women to concentrate in lower-paying professions. In addition, compared to men, women are less likely to hold a college degree or have a wealth of professional experience. Another important aspect is discrimination, which manifests itself in uneven remuneration for equal labor. In addition, women frequently experience professional interruptions and lower incomes due to their increased involvement in family duties.

The gender wage gap has been decreasing over time in most OECD nations, according to the research, but it still remains sizable, underscoring the need for more initiatives to address this issue fully. In the report's concluding section, numerous policy solutions are discussed for closing the gender wage gap. These possibilities include funding early childhood education and care, establishing paid parental leave policies, passing anti-discrimination legislation, and advancing gender equality at work.

This research review mentions the impact of gender stereotypes, gender bias, and unconscious bias on the gender wage gap in addition to the factors mentioned by the authors. Lower expectations and hence lower compensation can result from stereotypes that portray women as less competent. The manifestation of gender bias may take the shape of unfair job promotions or task assignments. Individuals' unconscious prejudices can also play a role in decision-making procedures that disfavor women.

In conclusion, this thorough examination of the gender wage disparity in OECD nations offers insightful information for decision-makers working to address the problem. In order to promote gender equality in labor markets and work toward eliminating the gender wage gap, policymakers can create effective policies and interventions by understanding the factors that contribute to the disparity.

**Data**

The Labour Force Statistics database of the OECD provides the information on the gender wage gap. The database compiles information from OECD member nations on employment, unemployment, salaries, and other labor market variables. Based on the difference between the median wages of men and women in comparison to the median earnings of males, data on the gender wage gap is collected. Both full-time employees and self-employed individuals have access to the data. There are 8 variables and 1310 observations in the data on the gender wage gap. Location, Indicator, Subject, Measure, Frequency, and Time are the variables.

**Data Transformation**

I examined the role of data transformation in analyzing gender wage discrepancies, employing methods that improve comprehension of underlying trends and patterns and, ultimately, produce more insightful conclusions.

**Checking Missing Values and Dropping the Flag Code Column**

The "Flag Codes" column will not be included in the data set because it is not significant to our study and does not include any values. The column contains no data and makes no contribution to the variables we are evaluating because it is empty. Because of this, deleting it streamlines the data set and concentrates our study on the important factors. The was also no missing values in the data.

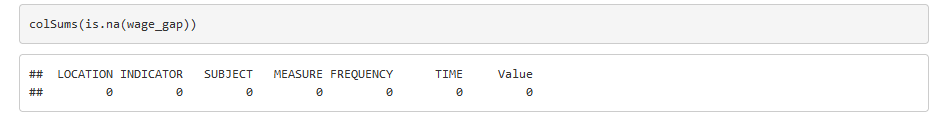


Figure 1: missing value indicator

**Data Visualization**

**1.A bar chart to compare wages across different subjects.**

This bar chart compares the wages across different subjects (employee and self-employed).

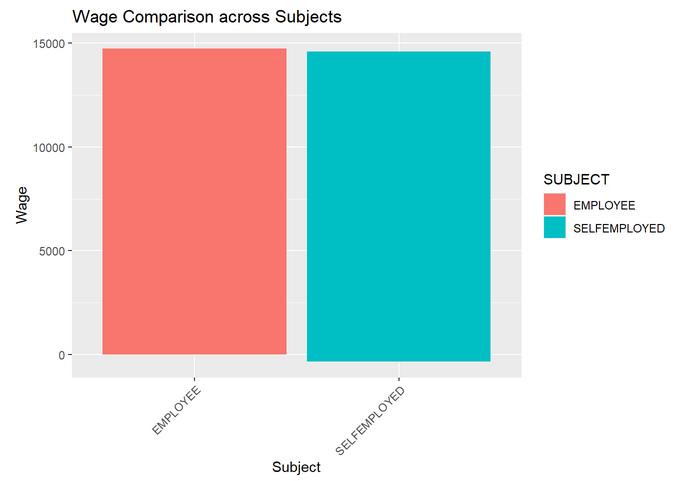


Figure 2:Bar plot of wage comparison across subject

Each The dataset includes information on wages and gender for each individual.The bar plot represents the wage values for different subjects.The x-axis of the plot represents the subjects in the dataset. The height of each bar corresponds to the wage value for that particular individual.

**2. Line plot**

A line graph to show the wage trend over time for a specific location.

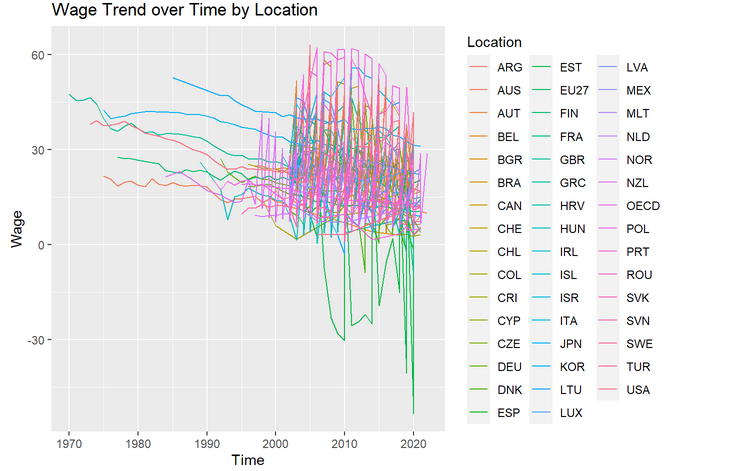


Figure 3:Line plot of wage trend over time by location

This The line plot includes multiple lines, where each line represents a specific location.The x-axis represents time, typically measured in years or other relevant time units.The y-axis represents the wage values for each location.

**3. Box plot**

A boxplot to compare the wage distributions across different indicators.

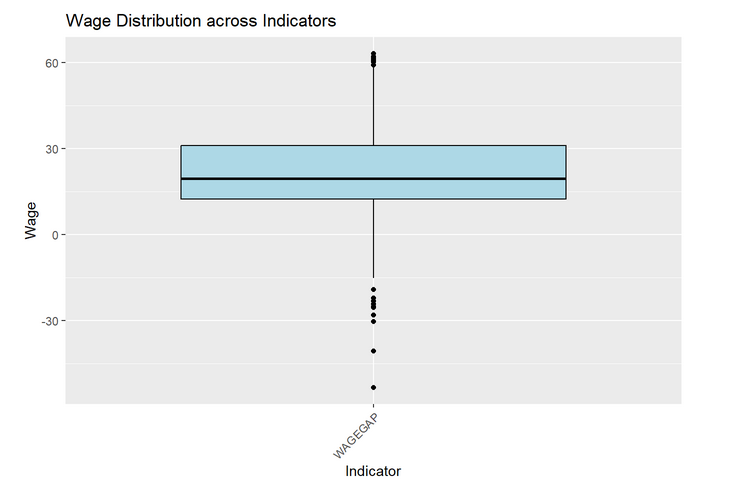


Figure 4:Boxplot of Wage distribution across indicators

This The boxplot includes multiple boxes, where each box represents a specific category. The y-axis represents wage values. By comparing the boxplots, we can identify any variations in wage distributions across different indicators.

**4. Histogram**

This histogram plot illustrates the distribution of wages in your dataset. The x-axis represents the wage values, divided into several bins.



Figure 5:Histogram of wage distribution

By examining the histogram, we can gain insights into the wage levels and the gender wage gap in the data. The visualization allows you to observe the distribution of wages for the subjects identify any differences or similarities in wage levels, and assess the central tendency of each group. This information can provide valuable insights into the gender wage gap and the distribution of wages among different genders in your dataset.

**Method**

The following were the models I used

1. **Fixed Effects Model:**

In a fixed effects model, the equation can be represented as follows:

Value = SUBJECT + TIME + LOCATION + ε\_i

where:

-**Value** represents the dependent variable of the outcome been predicted.

**Predictors**

-**SUBJECT** is a categorical variable representing different subjects, the employed and self-employed category.

-**TIME** represents the temporal aspect of your data. It could be measured in units such as years.

-**LOCATION** denotes a categorical variable that represents different categories of the location.

- ε\_i is the error term or residual.

In panel data, the individual-specific features that are expected to be consistent across time or among entities are captured by the fixed effects (\_i). Unobserved heterogeneity among the entities or people in the data is accounted for by these effects. Different methods, such as within-group transformation or adding dummy variables for each item in the regression, can be used to estimate the fixed effects.

The fixed effects model is used to take into account impacts that are unique to a certain person and are stable throughout time. We can account for unobserved traits that might have an impact on the dependent variable by incorporating individual-specific dummy variables in the model. The fixed effects approach enables us to separate the within-individual variations in the independent variables and analyze their impact on the dependent variable while adjusting for individual-specific characteristics in our analysis of the wage\_gap dataset.

1. **Random Effects Model:**

In a random effects model, the equation can be represented as follows:

Value ~ SUBJECT + TIME + ε\_i

where:

-**Value** represents the dependent variable, which in this case is the wage or earnings of individuals.

**-SUBJECT** is a categorical variable representing different subjects, the employed and self-employed category.

**-TIME** is a continuous variable that represents the temporal aspect of the data. It indicates the year when the wage data was collected.

- ε\_i is the error term or residual.

Individual-specific effects are viewed by the random effects model as random variables that could be associated with the independent factors. The average relationship between the independent and dependent variables is captured while still allowing for individual variances. The random effects model, used in our analysis of the wage\_gap dataset, takes into account the heterogeneity across people and sheds light on how the independent variables affect the dependent variable while taking into account both individual-specific effects and time-varying components.

We may acquire a thorough knowledge of the causes impacting the wage gap by employing both the fixed effects and random effects models in our study, taking into account both individual-specific features and time-varying factors.

**Performing the Hypothesis**

In the wage\_gap data set, there is a substantial correlation between the time impact (TIME) and the individual-specific effect (SUBJECT) on the dependent variable (VALUE).

Therefore, the results indicate that the value of the relevant variable (VALUE) in the wage\_gap data set is statistically significantly influenced by both being a self-employed individual (SELFEMPLOYED) and the passage of time (TIME).

**Results**

**Hypotheses 1**

According to the fixed effects model's findings, the p-values for the SUBJECT and TIME variables are incredibly low, providing strong evidence that the coefficients for these variables are not zero.

In particular, the p-value for SUBJECT and the p-value for TIME are both less than 2.2e-16. These extremely low p-values imply that the dependent variable, VALUE, is statistically significantly affected by both SUBJECT and TIME.

As a result, we may say that in the fixed effects model, there is a substantial link between SUBJECT, TIME, and VALUE.

.

**Hypotheses 2**

The random effects model's capture of the individual-specific effects makes a major contribution to understanding the pay difference in the dataset.

The "individual" component of the random effects model's estimate of the individual-specific factors' influence on the pay gap is that it has a significant bearing. The estimated variance of the individual-specific effects is 1838.505, which highlights the significant individual heterogeneity. The estimated proportion of individual-specific effects in the total variance, which is 0.969, indicates that individual differences may account for a sizable percentage of the pay gap's fluctuation.

Additionally, the SUBJECT and TIME independent variable coefficients are statistically significant with very low p-values. This suggests that the wage disparity is significantly influenced by both the individual-specific traits, such as topic, and the time-varying elements represented by TIME. We can therefore draw the conclusion that individual-specific effects, together with other factors represented by the independent variables, play a substantial role in explaining the pay difference seen in the dataset based on the findings of the random effects model.

Therefore, based on the results of the random effects model, we can conclude that individual-specific effects, along with other factors represented by the independent variables, play a significant role in explaining the wage gap observed in the dataset.

**Hausman Test**

In this study, we use both the fixed effects model and the random effects model to assess the pay disparity in a panel dataset. While the random effects model implies that these effects are random variables correlated with the independent variables, the fixed effects model adjusts for individual-specific effects. The Hausman test, which compares the two models and evaluates the correlation between the individual-specific effects and the independent variables, is used to decide which model is the most appropriate. We can choose the model that produces more accurate and reliable estimates by using the Hausman test. This analysis strengthens the validity and reliability of our data and enables us to reach solid conclusions regarding the factors affecting the wage disparity.

**Findings**

A statistically significant test statistic of 19.86 with 2 degrees of freedom and a p-value of 4.869e-05 is obtained using the Hausman test. This provides substantial evidence that the random effects model's consistency and effectiveness are not true.

As a result, we can draw the conclusion that in this instance, the fixed effects model is chosen over the random effects model. It is discovered that the independent variables and the individual-specific effects in the random effects model are associated, indicating that the fixed effects model is preferable for the study of the wage\_gap dataset.

It is crucial to keep in mind that employing the fixed effects model enables the control of person-specific effects that persist over time, resulting in more precise and reliable estimation of the correlations between the independent variables and the dependent variable.

In conclusion, based on the findings of the Hausman test, we advise utilizing the fixed effects model to take individual-specific effects into account when analyzing the wage\_gap dataset.

**APPENDIX**

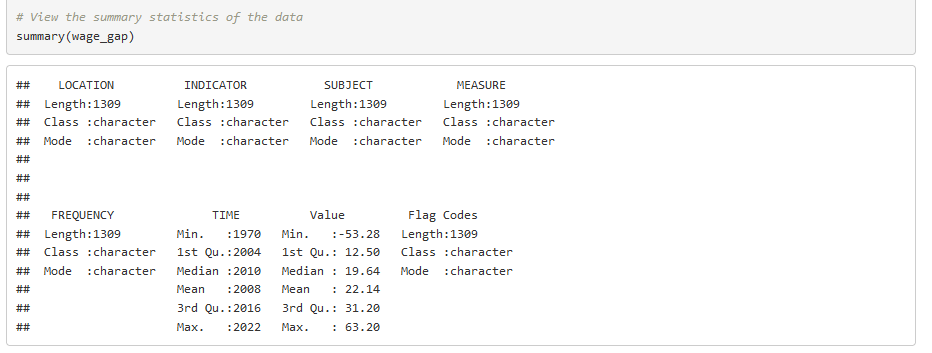


Figure 6: Data Summary

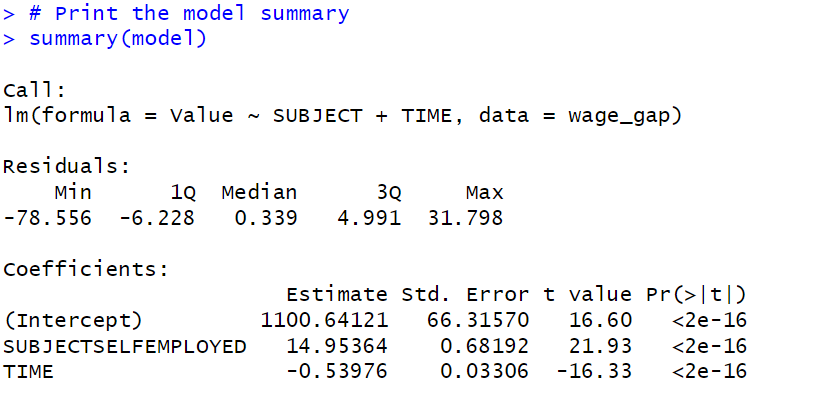


Figure 7: Regression Model

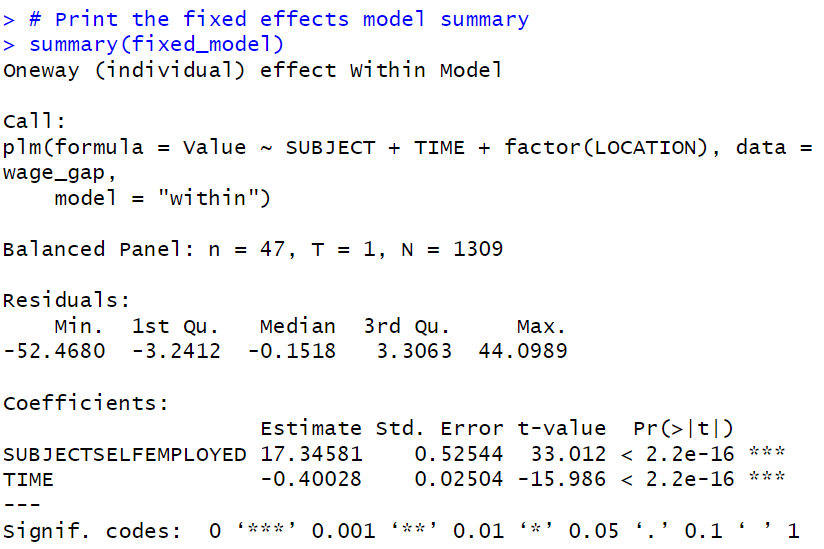


Figure 8: Fixed effect

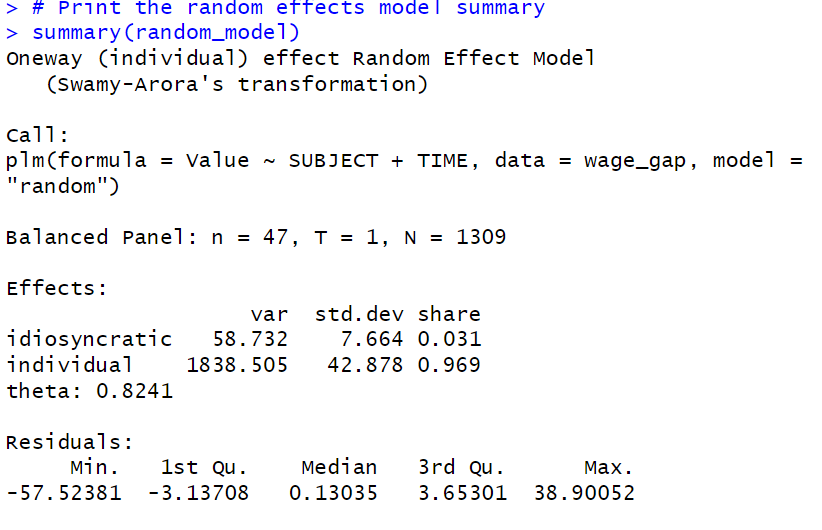


Figure 9:Random Effect

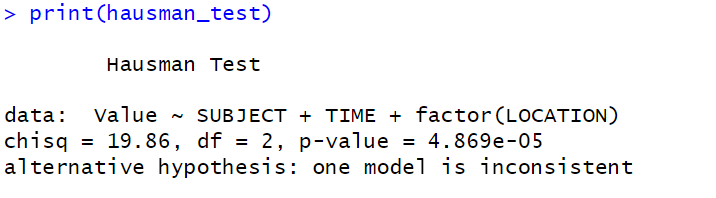


Figure 10: Hausman Test