

Racial Wage Gap: Analysis of difference in wages between black and white Americans

Tamires Amorim

Abstract

This paper aims to analyze the racial wage gap between black and white American workers. The difference in earnings was evaluated by using data from 2014 to 2019 acquired from the Current Population Survey (CPS). It was performed linear regression models and the Blind-Oaxaca Decomposition in R. The models estimated the earning gap for the male population of workers according their age, education attainments and marital status. From the models, it was possible to observe differences in earnings for white and black workers, although further investigation is necessary to obtain a complete understanding if differences in wages can be explained by discrimination in the labor force.

1. Introduction

The racial wage gap is a frequent topic of research in the United States, black workers have been treated differently in the labor market for a long time due to the country history of discrimination and economic inequalities in between white and African Americans. Furthermore, the present paper wants to evaluate the recent trends in the differences in earnings between white and black workers, by using the wage decomposition method and evaluating if those differences are driven by discrimination.

According to the Federal Reserve Bank of San Francisco (2017), it is challenging to fully interpret the racial wage inequality by looking at age, education, job type or location. Although, the goal of this paper is to observe how those variables and interactions might explain the differences in earnings between the workers.

The study starts by evaluating papers on the subject to draw conclusions on what aspects of their research might be useful for the current study, then goes to model simple regressions to establish comparisons across different races and educational attainments. After, it puts together the individual regressions into the decomposition method and finally the interpretation of the results obtained.

The main methodology for evaluating discrimination is the Blind-Oaxaca Decomposition, which is broadly used in the academia to measure the difference in wages earned in groups distinguished by race or gender, where the regressions wants to determine how much of the earning gap is due to discriminations and how much is due to distinct characteristics.

Acquired from the Current Population Survey (CPS), the sample accounts from 2014 to 2019. The main variables used were related to labor force, education, age, location and marital status. The first four models were designed to help understand the differences results of the linear regression for both races, the first two took age, race and paid by the hour variables in the subset, and the other models took age, race and male into the subset.

For the two Oaxaca models the subset followed the same design of the last two models described above and differ from each other by the variables used in the Oaxaca formula, where the second evaluation took into consideration the marital status as independent variable. By following those steps, the goal is to perform the examination of the wage gap between black and white workers in the United States.

2. Literature Review

There are many studies about wage gap in between groups, most of them are based on race or gender. The papers on racial wage gap presented different views and methods on how to best evaluate this issue. As pointed by Wilson and Rodgers (2016) “there is no single African American labor market narrative (...) the African American experience is not monolithic...”, it represented one of the biggest challenges when evaluating the models. Although, it is important to stress the similar results when observing the differences that educational attainments have on workers of different races. By reading the following papers, the present analysis aimed to capture the common variables and methods used to draw similar conclusions in those characteristics that seems to repeat.

The paper “black-white wage gaps expand with rising wage inequality” by Valerie Wilson and William M. Rodgers III (2016), brings an interesting evaluation of the wage differences across time, the study was able to draw a timeline of pay disparities from 1979 to 2015, although it was not able to capture social differences that has great effects in the labor market composition, for instance the mass incarcerations that occurred in between that period.

In “The Structure of disadvantage: Individual and Occupational Determinants of the Black-White wage gap” by Eric Grodsky and Devah Pager (1990), by using regression models they were able to show that income inequality between races in same and different occupations are not due to skill demand for the job, neither individual characteristics, status or demographic composition. It also compared private and public sectors, where the first have restrictions in the movement of black men in higher positions in a company, limiting their earnings, and the second more “related to human capital and occupational placement” (Grodsky, Pager 1990). Their main conclusion was that in the private sector the wage inequality was greater for those in higher-earning occupations, although it could not conclusively explain what they believe to be a “social segregation of black and white professionals”.

The paper “Performance Pay and White-Black wage gap” by John S. Heywood and Daniel Parent (2012), approached the issue by evaluating if there is a difference in payment by job performance between white and black Americans. The research strongly believed that performance pay jobs are more related to white workers than the black workers. The work used data from Current Population Survey (CPS) for the years 1976–99, and observed variation in the raw log

wage, regression in the log wage. The results were like the previous studies when comparing public and private sector since the private sector pays more for white workers in higher performance work than its black counterparts, although it was not able to reach definite conclusions on explaining this trend because this kind of occupation is evaluated in a subjective way.

3. Data

The data extracted comes from the IPUMS/CPS database, with a sample from 2014 to 2019, chosen this way to observe the most recent trend on racial wage gap. The age group was restricted to the ages of 25 to 64, according to the Organization for Economic Co-operation and Development, people in this group are in “their prime working lives” and from 55 to 64 “passing the peak of their career and approaching retirement” (OECD, 2020). The analysis is based on this group, given the bias that usually by this time in the life of a worker the wages are expected to be higher than in previous years and they will be active in the labor force.

As dependent variable the present work used the natural log of hourly wages, where it measured the percent disadvantage of black workers’ wages relative to whites, such that a smaller estimate indicates a smaller wage gap and more equity (Bound and Freeman 1992). Hourly wage was the most used methodology in papers that evaluate racial wage gap, that might differ between workers that do not earn by hour worked, it was extracted by the CPS database and the amounts were expressed as they were reported to the interviewer.

As independent variables, the set of educational attainments ranging from high school to advanced degree, age, and marital status, the models were also subset differently starting with only age, race and different salaries payment system, to models that male were specified, as well if they worked in the private sector, lived in metropolitan areas, since as indicated by Wilson and Rodgers (2016), could indicate important differences given that wage levels in the cities differ from countryside or suburbs.

The intentional restriction to the Male population, is due to the complexities to establish wage gap comparisons for women because of variables related to “family structure and corrections for racial variation in women’s labor force participation” (Grodsky and Pager, 2001).

The wage decomposition model used the package Blinder-Oaxaca Decomposition in R, the variables for model 1 and model 2 were set as follow:

$$\ln W_w - \ln W_B = (X_w \beta_w - X_B \beta_B) + (Y_w \beta_w - Y_B \beta_B) + (Z_w \beta_w - Z_B \beta_B) + \sigma_w (\theta_w - \theta_B)$$

Where we see the natural logarithm of hourly earnings for white workers minus the natural logarithm of hourly earnings for black workers equals the sum of the variables of observed attributes as educational attainments, age and marital status (X, Y and Z) for whites minus blacks and the standardized residual (meaning that it is distributed with a mean of zero and variance of one) and the residual standard deviation of wages.

4. Results

The first set of models 1,1.1,2 and 2.2 served as baseline to evaluate the differences between some interactions and draw important conclusions before starting the main decomposition model. The main differences were in the subsets and dependent variable, where the first set of models focused on earnings by the week restricted for workers that were paid by hour worked and for the labor force without considering gender.

The first model is about black workers it has 7149 observations, which is the lowest compared to the models 1.1, 2 and 2.2, the first model implies that age and lower levels of education are not as significant as higher education attainments. The possible conclusion is that higher education attainments hold a greater significance level for the change in earnings for black workers.

The second model (1.1) is about white workers, it has 40215 observations, some of the differences between the two models is that, when evaluating the relationship between earnings and education attainments for white workers the higher level of education did not hold as much significance as age and lower education attainments, which is the opposite from what was observed in the first model. Another interesting difference was the negative relation between the independent variables and dependent variables. Those observations might be due to differences in the earning system, probably white workers are not paid by the hour, or not all the white workers are driven to further their education given that age and lower levels of education already have high returns in earnings.

Model 2 is about male black workers paid by the hour, it has 31172 observations, more than previous models, the only significant variables are age and some college education, age is significant and holds the concept indicated by the OCDE about age and labor force relation. Comparing to model 2.2 about white workers paid by the hour, there are 217818 observations, the significance level is high for all the independent variables, and as the model 2 it also holds a negative relation for lower levels of education and positive for higher levels, another interesting observation is that we expect a 19% increase in earnings for those with advanced degree compared with 7.8% for black workers with the same level of education.

From the previous model's evaluation, it was possible to draw valuable conclusions on what variables and interactions might be interesting in adding on the Oaxaca model. Models 2 and 2.2 were the ones chosen due to its higher number of observations and differences in the significance levels of the age and educational attainments variables, were chosen to further the investigation.

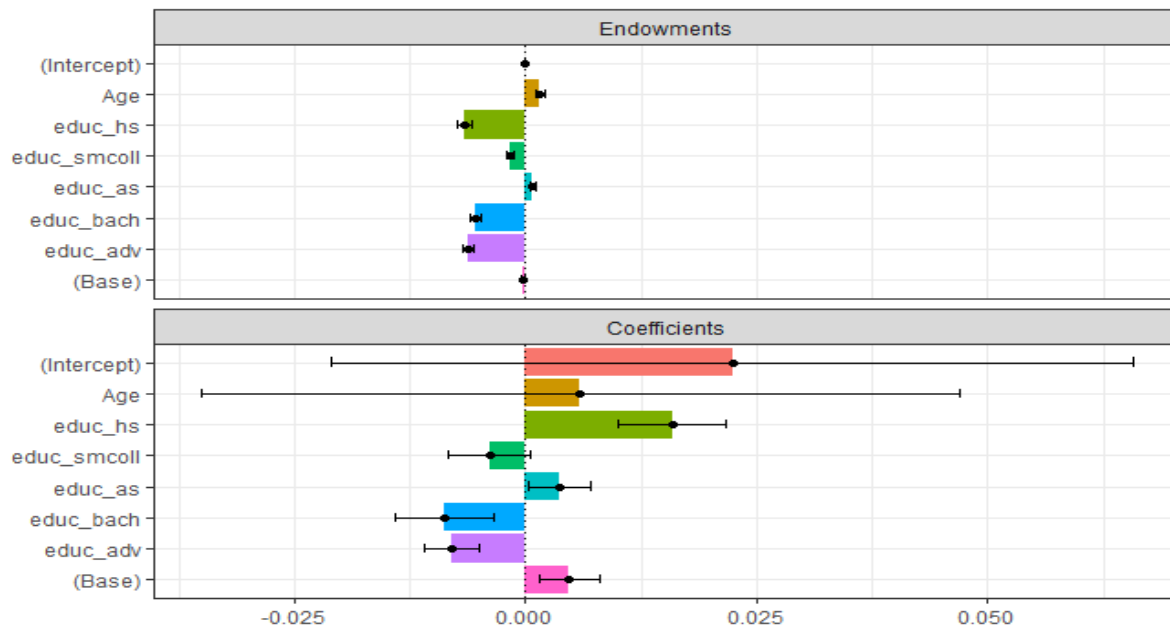
Start by decomposing the wage gap between black and white workers. The wage gap could be due to age or educational attainments. The main formula takes into consideration log hourly wages as the dependent variable and age and educational attainments as dependent variables. The bootstrapped standard errors are calculated based on 100 replicates. The races were divided in groups A and B, where group A corresponds to black workers and group B to white workers, numbers observed in model 2 and 2.2 respectively. The n pooled represents the total of observations.

```
## $n.A
## [1] 31172
##
## $n.B
## [1] 217818
##
## $n.pooled
## [1] 248990
```

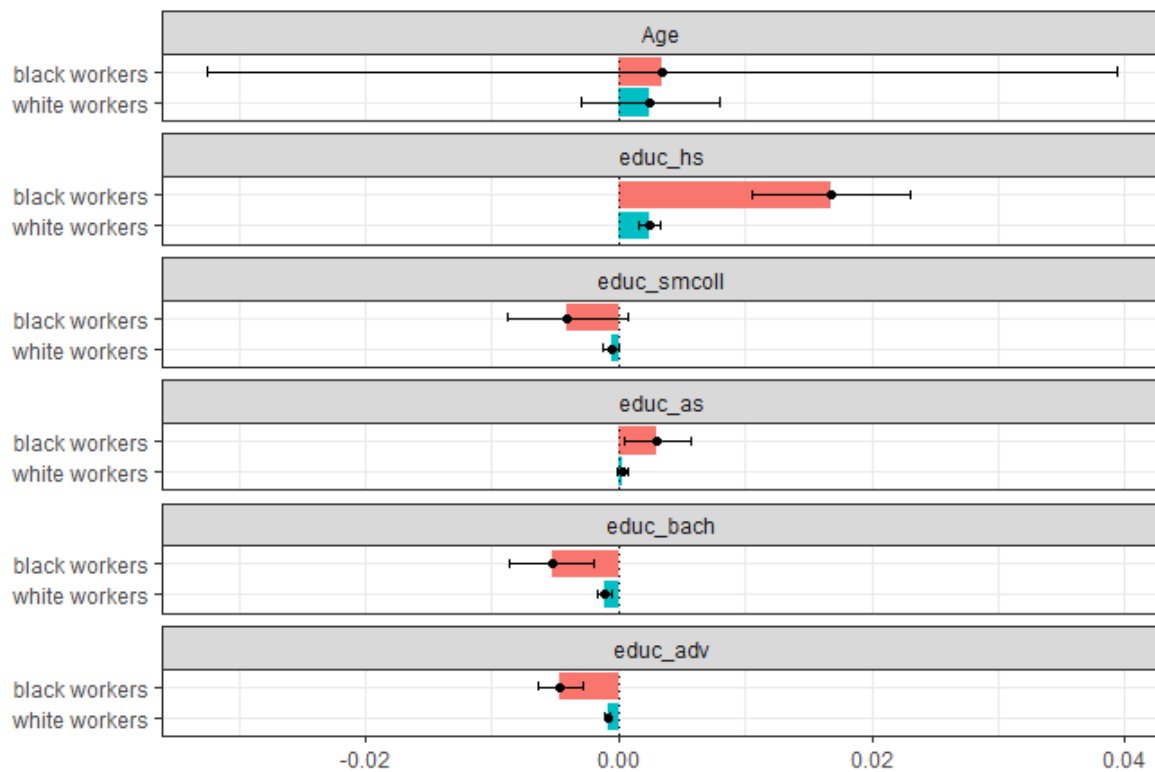
Following, the mean hourly wage for each group, where black workers earns 0.022 cents more than white workers, although as observed before the number of observations pooled for black workers were less than white workers, which might have brought the unexpected return.

The threefold Blinder-Oaxaca decomposition suggests that approximately 0.017 cents can be attributed to group differences in age or education (endowments), 0.03 due to differences in the

coefficients and 0.008 cents can be accounted by the interaction of the two. In the plot it is observed the estimation results for age and education attainments with the error bars that indicate 95% confidence intervals.



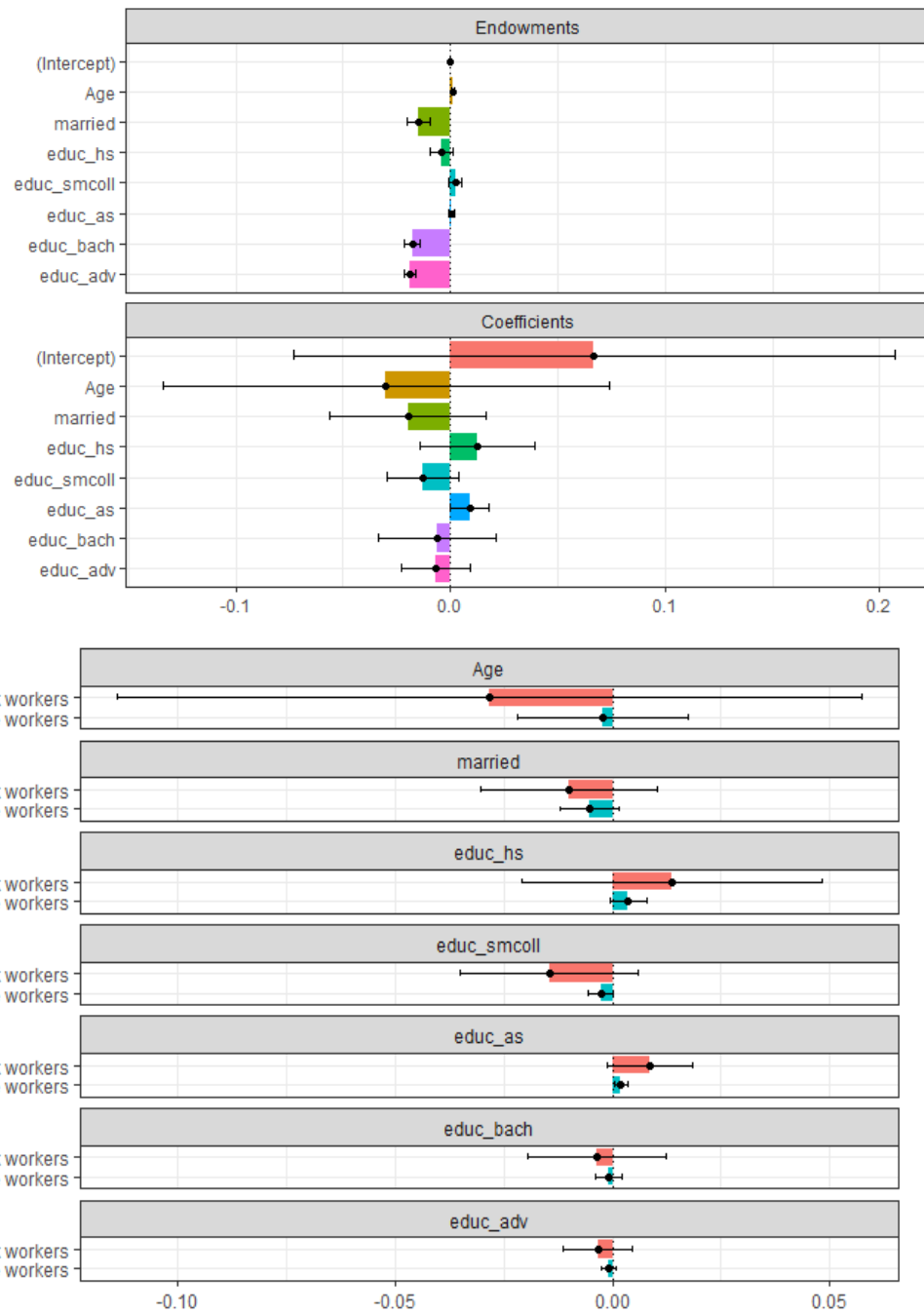
From the plot endowments only education in the bachelor, advanced degree and high school seems statistically significant for the conclusion in the wage gap differences, which is understandable given that higher education attainments imply higher earnings across different groups. When analyzing differences in ages between white and black workers the difference in earnings, it is not possible to find conclusive information given that it is extremely low portion in the explanation of the gap. To confirm the differences in wages explained by the educational attainments. The twofold decomposition confirmed what was observed in the plot above, where high school education seems to be the main driver in the difference in wages between black and white workers.



The second Oaxaca decomposition model takes into consideration only workers in the private sector that lives in the city, in this subset there is a lower number of observations. In the same as the first Oaxaca model, it had a lower number of black workers compared to white workers, although the difference in earnings is higher and the white workers earn more than the black employees.

Differently from the first model the threefold decomposition suggests that approximately 0.05 cents can be attributed to group differences in age, education, or marital status (endowments), 0.012 due to differences in the coefficients and 0.014 cents can be accounted by the interaction of the two. In the plot it is observed the estimation results for age and education attainments with the error bars that indicate 95% confidence intervals.

Now the in the endowments plot marital status and advanced educational attainments have a higher level of significance than the first model, probably since the private sector is seem as more discriminatory in the higher education level. For the two folded decomposition it returns different results from the threefold because Age is more significant than the other variables



5. Conclusion

In Economics theory the marginal productivity of labor is one of the main determinants of wage, although empirically that assumption might not hold. There are many variables that comes into account when evaluating wages, for instance the racial wage gap, is not due to differences in productivity of the labor, but in differences that most of the time the theory and numbers might not be able to fully understand because it holds the human factor, the subjectivity of decisions in how to the employers reward their employees.

From this research the results the variables with higher returns in the explanation of wage differences, were education attainments and age. When observing the interactions with black workers in the age group from 25 to 64 were expected to have higher earnings, although it returned a bigger wage gap when compared with white workers in the private sector. In the case of education, workers in the high school level as observed in the age variable, can also explain the difference in earnings.

Choosing the right variables, cleaning the data, performing, and correctly understanding the Blind-Oaxaca decomposition composed one of the main challenges of the analysis, because it was hard to avoid the bias obtained from the readings in the different papers on the subject. Another challenge that is important to stress was the extremely small differences in earnings observed in the first model, which were not only small, but also in favor of black workers, which was not expected, despite the contradictory result, the choice to keep it in the final submission comes from the believe that not all analysis need to have the right answers for every question.

Although, the current research was not able to fully explain the difference in wages between black and white American workers, because many were the outliers, for instance, the lower representativity in the number of black workers in the sample, and other restrictions that unable the data to fully represent what was expected as most of the coefficients were not as significant, and the correlation not as strong, which was surprising but at the same time reassuring that this subject needs more investigation due to its intrinsic complexities.

References

The data extracted comes from the IPUMS/CPS database. Courtesy of Sarah Flood, Miriam King, Renae Rodgers, Steven Ruggles and J. Robert Warren. Integrated Public Use Microdata Series, Current Population Survey: Version 8.0 [dataset]. Minneapolis, MN: IPUMS, 2020. <https://doi.org/10.18128/D030.V8.0>

OECD (2020), "Employment rate by age group" (indicator), <https://doi.org/10.1787/084f32c7-en> (accessed on 12 December 2020).

Grodsky, Eric, and Devah Pager. "The Structure of Disadvantage: Individual and Occupational Determinants of the Black-White Wage Gap." *American Sociological Review*, vol. 66, no. 4, 2001, pp. 542–567. JSTOR, www.jstor.org/stable/3088922. Accessed 10 Nov. 2020.

Daly, Mary, Hobijn, Bart and Pedtke, Joseph H., (2017), Disappointing Facts about the Black-White Wage Gap, FRBSF Economic Letter, issue , number 26, <https://EconPapers.repec.org/RePEc:fip:fedfel:00141>.

Suen, Wing. "Decomposing Wage Residuals: Unmeasured Skill or Statistical Artifact?" *Journal of Labor Economics*, vol. 15, no. 3, 1997, pp. 555–566. JSTOR, www.jstor.org/stable/10.1086/209872. Accessed 17 Nov. 2020.

Heywood, John S., and Daniel Parent. "Performance Pay and the White-Black Wage Gap." *Journal of Labor Economics*, vol. 30, no. 2, 2012, pp. 249–290. JSTOR, www.jstor.org/stable/10.1086/663355. Accessed 17 Nov. 2020.

Hlavac, Marek (2018). *oaxaca: Blinder-Oaxaca Decomposition in R*. R package version 0.1.4. <https://CRAN.R-project.org/package=oaxaca>