

Data Analysis of Race-based and Sex-based Employment Discrimination

Junchi Tian, M.S.

Nima Zahadat, Ph.D.

The George Washington University

ABSTRACT

Although under the Civil Rights Act of 1964, employment discrimination is not overt, race-based and sex-based discrimination privately exists in the U.S. workplace, and it has been pervasive for a long time. The minorities and women receive unfair evaluation and disparate treatment so that they face higher unemployment rates, lower earnings with the same performance and fewer opportunities to be promoted to managerial or corporate positions. To quantify and present employment discrimination would be key for public policy makers to understand how the allocation mechanism operates. Through the raw data, this thesis aims to use Python to analyze the data of employment status and visualize the differences between race, sex and age. The raw data collected by the Census Bureau includes the following demographics: population over 16, labor force, not in labor force, employed and unemployed persons, employment and unemployment rate etc. and is divided by race, gender and age from 1948 to 2017.

Zenodo Link: <https://zenodo.org/record/2187328#.XA80JPZFyIU>

Keywords: employment status, employment discrimination, race, sex, data mining

1. INTRODUCTION

Since the 1960s, progressivism has led to a huge impact on various aspects of American society including feminism, civil rights movement and LGBT movement. Especially the Civil Rights Act of 1964, which is a milestone of the social movement, any discrimination is prohibited on the basis of race, skin color, religion, sex in the workplace. However, in the real world, race and sex still be significant factors when the employee makes decisions of job offer and promotion. A well-known study organized by the University of Michigan Law school showed that regarding earning, after graduation the gap between men and women were very small. But it widened in 15 years and made the point that women's salary is only 60 percent of that of men. Fix, Galster and Struck (1993) have controlled other variables including same age group and same gender and tested with Hispanic, white and African American men. They were required to send the same fabricated resumes but they were easy to recognized their racial identities. The result indicated that the black men and Hispanic men are three times more likely to be refused compared to white man. It is obvious that employment discrimination is pervasive in the U.S. workplace.

The mechanism of discrimination is also involved in other intermediate variables. Roehling and Pichler (2007) mentioned the weight discrimination is also significantly related to race and sex. Roberto and Isabel (2006) noticed that minorities also lack enough network to get good job opportunities. Many studies built significant statistics models to quantify employment discrimination. Emilio (2008) developed a performance rating model to show the status of weak: minorities and women. Lisa (2003) used a race-gender composition to describe the details of change on pay deterioration in occupations from 1971-1981 and 1982-1992.

The employment discrimination brings adverse effect on society. It violates people's pursuit to a fair society, produces conflicts and violence, shaped communities without trust and interferes with democratic institutions and economic development. Also, studies need to be updated to face with new statistical caliber. More kinds of data are available including sex-race composition and specific Hispanic-Latino ethnicity for further data analysis.

Thus, it is necessary to quantitative employment discrimination based on race and sex, and better understands the current situations for the policymakers. Through the data from Census Bureau and data mining programmed by python, it is possible to get an objective view of how race and sex influence the employment status.

2. LITERATURE REVIEW

2.1 Model building

There are a lot of research papers that authors developed a statistical model or used a quantitative method to analyze the impact of employment based on race, gender or the composition. Emilio (2008) developed and tested a model based on performance rating and demographics including gender, age, education attainment and salary and finally found that with same performance rating score, women and minorities receive less compensation than the white man. Thus, there is a race-based bias in the workplace. Lisa (2003) used a race-gender composition to analyze change on pay deterioration in occupations from 1971–1981 and 1982–1992. The result clearly showed that the composition influences the wage. Wage erosion took place obviously with a large percentage of white women and black men from early 1970 to early 1990 and then

the high percent of black women also involved in the next decade.

Budig (2002) investigated gender-based differences both in wages and in wage growth by using data from the NLSY and detailed measure of gender composition on jobs. The result clearly show that men are advantaged, net of controls, in both pay levels and wage growth in all jobs, regardless of gender composition. Roth, Philip, Huffcutt, Allen Bobko (2003) conducted a new meta-analysis of ethnic group differences in job performance. More homogeneous categories of job performance are analyzed on Black-White difference. Analysis of Hispanic-White standardized differences describe that they are generally lower than Black-White differences in various categories.

2.2 Sample Selection

Stephane and Debra (1997) used a nationally representative sample of U.S. workers and their organizations. Based on the concept of ‘homosocial reproduction’, authors designed a race-specific model to find whether the gap in Black-White attainment exists. Finally, they found that the determinants of promotion systematically differ for Blacks and Whites. Using Swedish register data from 2001 to 2003, Magnusson (2013) described the detailed shape of the relationship between wages and occupational sex composition and conducted a non-linear relationship that highest wages were earned by both of men and women in sex-integrated occupations, and both genders gain by transferring to relatively sex-integrated occupations. From the research that female-dominated occupations widely generate lower wages than male-dominated occupations from U.S., Ruijter, Doorne-Huiskes (2002) etc. focused on the gender wage-gap in the Dutch labor market and finally found that both men and women earn lower wages if they are employed in female-dominated occupations.

2.3 Social Impact

The social impacts also be a key area for employment discrimination. Loafman and Little (2014) conducted that although under Title VII of the Civil Rights Act of 1964, the EEOC and government have done an influential job to weaken the discrimination based on criminal record and race, this issue the employment of ex - convicts still created too many tensions and problems in the business operation and legal justification. John and Peter (1991) focused on nature and trend in the future of employment discrimination litigation. Authors also specifically analyze cases from various perspectives. Donald, Melvin and Kecia (2005) applied a model on human capital theory to racial earnings inequality and suggested black and

Hispanic men's trajectories were more limited than white men and well-educated career inequalities between each race were more significant.

2.4 Other Factors

Some authors introduced more factors to research how the mechanism of discrimination form. Roehling and Pichler (2007) noticed the impact of weight discrimination related to sex and race in the workplace. Their results showed that women take 16 times weight discrimination more than men and there is huge weight-related employment discrimination between overweight and normal weight respondents. Another important reason for race-based differences in the workplace is networks mentioned by Roberto and Isabel (2006). Without the network, African Americans and Hispanics were isolated from good job opportunities. Authors delineated the different mechanisms and how these mechanisms operate at several stages of the recruitment process. From personal data of U.S. service organization during the period 1978-86 with managerial, administrative, and professional employees, Trond and Ishak (2004) defined three type of discrimination: allocative discrimination, within-job wage discrimination and valuateive discrimination, and conducted that valuateive discrimination is a major reason of wage differences compared to with-in job discrimination.

3. RESEARCH METHDOLOGY

Based on the literature review, both of employment discrimination model and its social impact have been detailed discussed. Many exploratory studies existed including more intermediate variables and more accurate and complicated models. However, from 2000 to now, data is collected more particularly and specifically. For example, there is no data about Asian before 2000. Also, the specific Hispanic and Latino ethnicities, including Mexican, Cuban, Central & South American are collected after 2000. The sex-race composition also is witnessed new change in the new century. Thus, this study is valuable to show new trend on employment discrimination. The result will be summarized by and measured by two explanatory factors: employment-population ratio and unemployment ratio and be recognized by race and sex.

4. DATA

The raw data from official website of Census Bureau is collected by Census Bureau and are excel files, which consist of five sheets. The first sheet is the number of employed person and labor force and their ratio of

population by men and women and by White, Black or African American, Asian, Native American, Native Hawaiian and other Pacific Islander and detailed Hispanic or Latino ethnicity 2000-2016. The second sheet is the overall population over 16, labor force, employed person, unemployed person from 1948 to 2017. The third and fourth separately focused on male and female divide by age group from 16-24 years to 65 years and older. The fifth sheet show the comparison of employment status of native born and foreign, citizen and non-citizen.

5. DATA ANALYSIS

In order to better analyze data, data cleaning is an important step. There is missing value like the employment status of Asian before 2000 because of the change of statistical standards. There is no need to deal with outliers because the data is continuous and pretty flat. For different data sheets, different periods will be selected.

For the data mining, Python and Python packages including Numpy, Pandas and Matplotlib will be used. First, the basic exploratory data analysis will be completed including the overall employment situations during past 70 years and the trend of employment status for each race. In addition, it is possible to test whether there is difference among the race and the sex and how big the difference is. And then, the race-sex combination will also be related to the employment status to further refine result. Finally, the age group will be considered as a new variable.

6. KEY FINDINGS

Understanding the distribution of race population is the first step. From figure 1, white (69%) is unquestionable biggest race. African American (11%) and Hispanic or Latino (14%) follow. In contrast, Asian (5%) has smallest ratio of population. They are main races in USA. Thus, as minorities, Asian, African American and Hispanic or Latino will be compared to White.

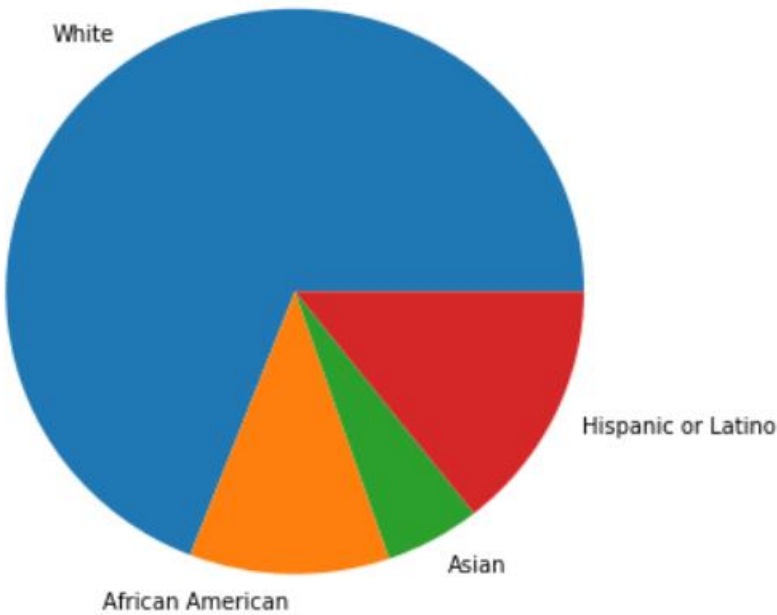


Figure 1

To further quantify employment discrimination, two important indicators will be used: employment-population ratio and unemployment ratio. For employment-population ratio (figure 2), it is significant that the difference between men and women is huge and decline very soon from 50% in 1948 to 10% in 2017. In terms of race, as we can see, the figure 3 also reflects in narrowed gap among four main races. However, although the number of Asian, White and Hispanic or Latino are pretty close, African American is still hard to find a job compare with other races.

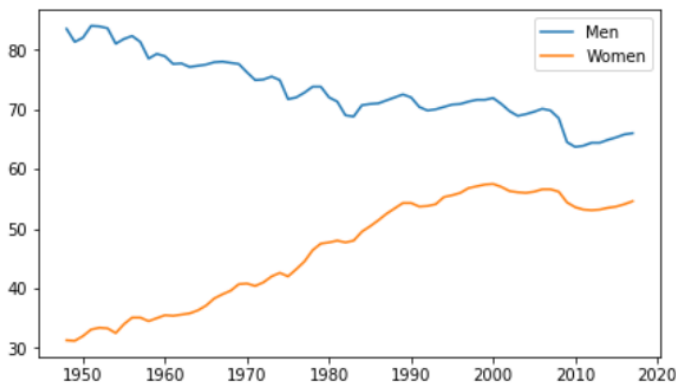


Figure 2

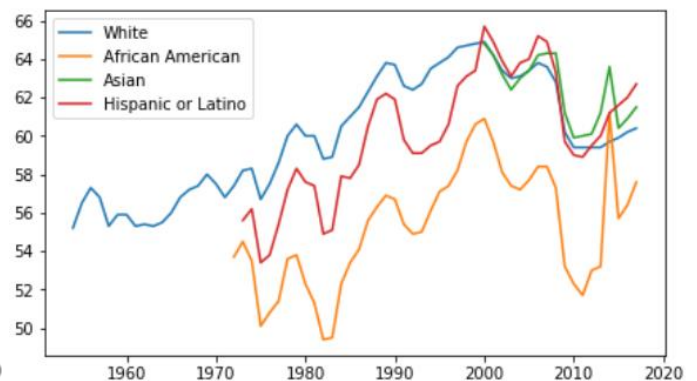


Figure 3

For specific Hispanic or Latino ethnicities, figure 4 indicates that Central & South American keeps the best performance. Puerto Rican ranks the last. Mexican, Cuban and other Hispanic or Latino are in the medium level. The differences among them don't noticeable change for 16 years.

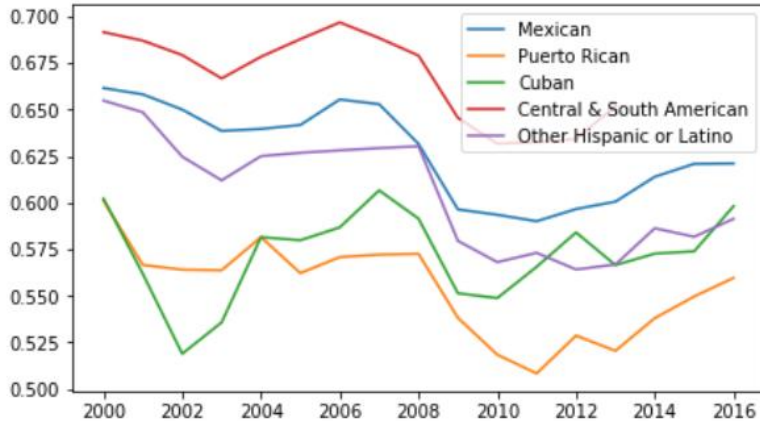


Figure 4

Figure 5 and figure 6 describe unemployment ratio. The situation of race is similar to employment-population ratio. African American is likely difficult to find a job. Asian and White are easier and also don't tend to lose a job. The ratio between men and women is pretty similar. Probably in the 60s and 70s, men are easier to get job opportunities, but now they are almost equal.

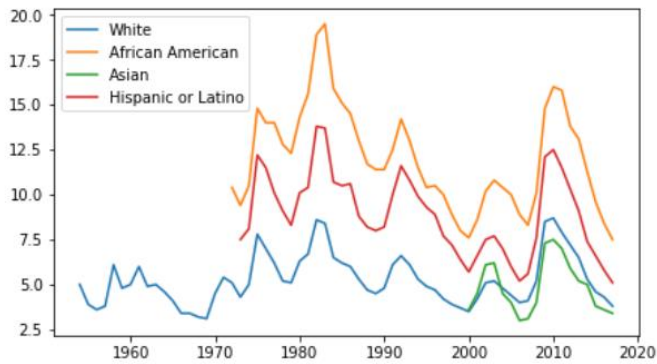


Figure 5

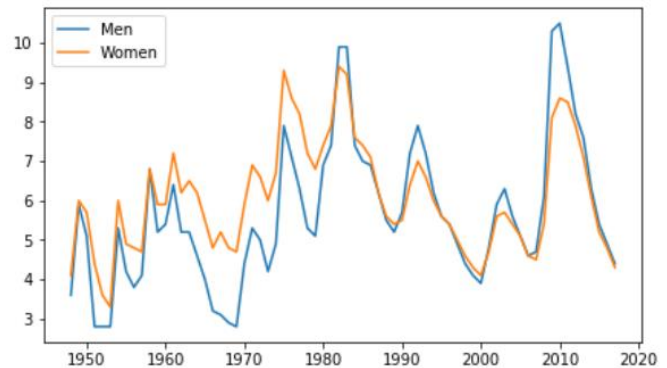


Figure 6

Two indicators based on race show similar information. However, the employment-population ratio is different from unemployment ratio based on sex. But this will not influence the result of sexual employment discrimination, because these women who do not work and also are not unemployed

population are unwilling to work and prefer to stay at home. The stereotype that women can only be housewife is also a kind of employment discrimination. Therefore, it will not influence our conclusions.

The most important and interesting result is from figure 7 and figure 8 when we use sex-race composition instead of single variable to measure employment-population ratio. It is obvious that sex is more influential factor than race because the different line can be concluded as difference of sex not race. But African American men is exception. They are remarkable lower than other race male and even white women in some years.

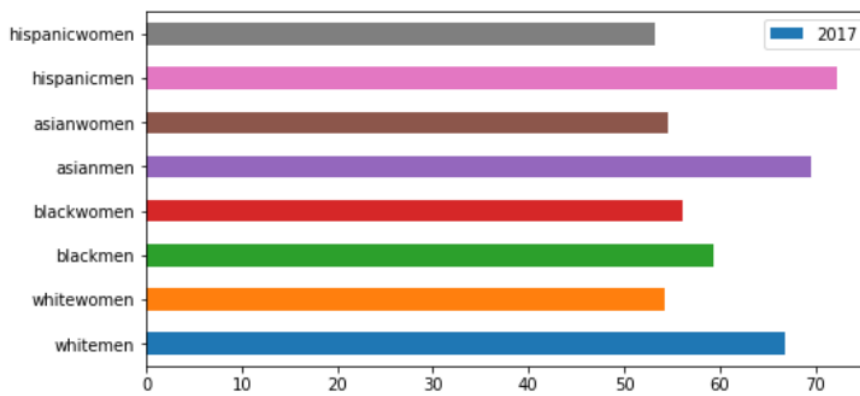


Figure 7

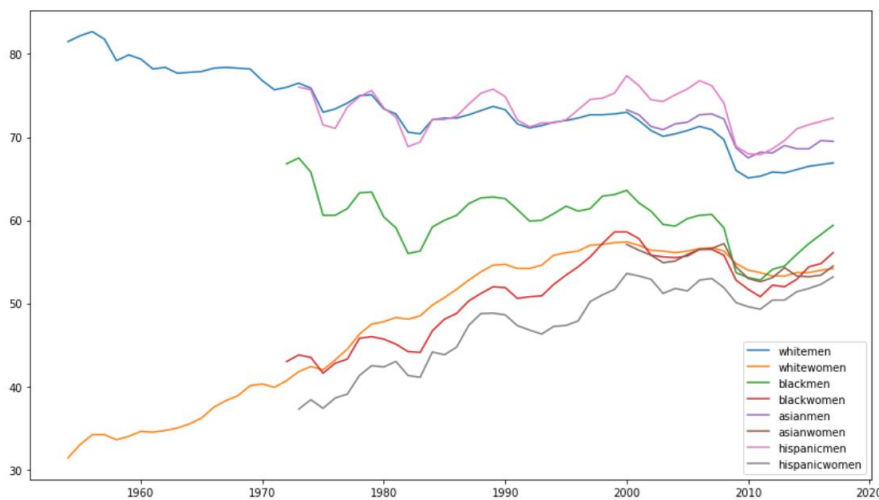


Figure 8

7. RECOMMENDATIONS

There are some points that the study can be improved. Only two indicators are not enough for a big picture of employment discrimination. More indicators, such as labor force, daily salary and education expenditures should be considered for further analysis. Age can be another important factor to measure employment status by establishing more complex compositions like sex-age and race-age. For different specific period, like 70s or 80s, analysis can be combined with the social environment at that time. Also social impact can be explicitly researched. There are some missing data: Asian before 2000, Hispanic or Latino before 1974, African American before 1973 etc. But they are also meaningful and potential fields for further employment discrimination research.

8. CONCLUSION

The study revealed that employment discrimination undoubtedly exists for a long time. In terms of race, Asian is “Model Minority” and has similar even better employment status than White. African American is under most significant discrimination regardless of male or female. Hispanic or Latino is slightly worse than White but show better trend. There isn’t huge change of employment discrimination on Hispanic or Latino ethnicity within 16 years. In terms of sex, male had huge advantage on employment in 50s but fell very soon.

In conclusion, Sexism is more significant than racism in the workplace except for African American, which means women is discriminated more than race minority. But employment discrimination is the process of getting weaker and weaker from 50s, particularly on the sex.

The results provide us a dropping tendency of employment discrimination, which means the government has taken effective strategies and actions on this problem. However, government and society still should pay more attentions to potential sexism and unequal employment status of African American on the workplace.

BIOGRAPHY

Junchi Tian studied Business Administration in his undergraduate degree in Dongbei University of Finance and Economics. He has joined Data Science Program at The George Washington University since August 2018.

Dr. Nima Zahadat is a professor of data science, information systems security, and digital forensics. His research focus is on studying the Internet of Things, data mining, information visualization, mobile security, security policy management, and memory forensics. He has been teaching since 2001 and has developed and taught over 100 topics. Dr. Zahadat has also been consultant with the federal government agencies, the US Air Force, Navy, Marines, and the Coast Guard. He enjoys teaching, biking, reading, and writing.

REFERENCES

- Baldi, Ste'phane, and Debra B. McBrier. (1997). "Do the Determinants of Promotion Differ for Blacks and Whites?" *Work and Occupations* 24 (4): 478–97.
- Budig, Michelle J.(2002), Male Advantage and the Gender Composition of Jobs: Who Rides the Glass Escalator? , *Social Problems*, Vol.49(2), pp.258-277
- Bureau of Labor Statistics, Labor Force Statistics from the Current Population Survey
Website: <https://www.bls.gov/cps/lfcharacteristics.htm>
- Catanzarite, Lisa, (2003), Race-Gender Composition and Occupational Pay Degradation, *Social Problems*, Vol.50(1), pp.14-37
- Castilla, Emilio j., (2008), Gender, Race, and Meritocracy in Organizational Careers, *American Journal of Sociology*, Vol.113(6), pp.1479-1526
- Donohue, John J. ; Siegelman, Peter, (1991), The Changing Nature of Employment Discrimination Litigation, *Stanford Law Review*, 1 May 1991, Vol.43(5), pp.983-1033
- Fernandez, Roberto M., and Isabel Fernandez-Mateo. (2006). "Networks, Race and Hiring." *American Sociological Review* 71 (1): 42–71.
- Ford, J. K; Kraiger, K; Schechtman, S L.; House, Betty J. (1986). Study of Race Effects in Objective Indices and Subjective Evaluations of Performance: A Meta-Analysis of Performance Criteria. *Psychological Bulletin*, Vol.99(3), pp.330-337

- Magnusson, Charlotta, (2013), More women, lower pay? Occupational sex composition, wages and wage growth, *Acta Sociologica*, Vol.56(3), pp.227-245
- Petersen, Trond, and Ishak Saporta. (2004). "The Opportunity Structure for Discrimination." *American Journal of Sociology* 109 (4): 852–902.
- Reskin, Barbara,(2004), The proximate causes of employment discrimination, *Contemporary Sociology*, Vol.29(2), pp.319-328
- Roehling, Mark V. ; Roehling, Patricia V. ; Pichler, Shaun,(2007), The relationship between body weight and perceived weight-related employment discrimination: The role of sex and race, *Journal of Vocational Behavior*, Vol.71(2), p.300-318
- Roth, Philip L. ; Huffcutt, Allen I. ; Bobko, Philip, (2003), Ethnic Group Differences in Measures of Job Performance: A New Meta-Analysis, *Journal of Applied Psychology*, Vol.88(4), pp.694-706
- Ruijter, Judith M. P. de etc. (2013), Size and Causes of the Occupational Gender Wage-gap in the Netherlands, *European Sociological Review*, Vol.19(4)
- Schmitt, Neal, and Martha Lippin. (1980). "Race and Sex as Determinants of the Mean and Variance of Performance Ratings." *Journal of Applied Psychology* 65(4):428–35.