Race, Gender and the Future of Work*

Pegah Moradi

May 19, 2017

Abstract

Leading up to and following the 2016 American presidential election, "white working class" unemployment has become particularly politically salient. A simultaneous discussion on the role of automation in unemployment complicates the political narrative; by one estimate, 47% of American jobs are at risk of computerization (Frey and Osborne, 2013). This study analyzes how computerization corresponds with racial, ethnic, and gender demographics within occupational groups. Using an effect factor based on occupational demographic and automatability data, we found that White workers are more heavily affected by automation than other racial groups. Conversely, however, we found that the proportion of white workers in an occupation is negatively correlated with an occupation's automatability, suggesting that White workers may simply be over-represented in small, automation-resistant occupations.

*All reproducible code and data for this study can be found at `https://github.com/pegahmoradi/pegahmoradi.github.io/tree/master/automation_race`

Introduction

Following the proliferation of Donald Trump's anti-immigration and protectionist rhetoric in the 2016 American presidential election, the state of the "white working class" in America became extraordinarily politically salient. Amidst the myriad retrospective analyses of the election, two primary narratives came forward with regard to the white working class: Firstly, the white working class is experiencing major job loss due to immigration and free trade, explaining the demographic's shift towards Trump's rhetoric. In opposition, the plight of the white working class is a myth; white workers are experiencing the same economic difficulties as non-white workers of similar socioeconomic standing.

Literature on the topic is abundant. A *New York Times* article connected Donald Trump's victories to areas experiencing major job loss (Porter, 2016). The Wonkblog of *The Washington Post* cited job loss and lowered wages for the "revenge" of the white working class, notably mentioning, "[the] workers increasingly came to see trade deals as the culprit" (Tankersley, 2016).

A concurrent, yet seemingly isolated discussion has been buzzing around the American technology sector on the role of automation in American society and how machine learning and artificial intelligence will inevitably cause mass job loss sooner rather than later. The most recognized of the literature on automation and employment comes from Carl Frey and Michael Osborne's authoritative 2013 study titled, "The Future of Employment," which calculated automation probabilities for American occupations and determined that approximately 47% of jobs are at risk of automation. In their notable 2014 text, *The Second Machine Age: Work, Progress, and Prosperity in a Time of Brilliant Technologies*, Andrew McAfee and Erik Brynjolfsson take automation as unavoidable and instead focus on a path forward. *Los Angeles Times* columnist David Horsey began the connection of these two disjoint conversations on employment, arguing that "Robots, not immigrants, are taking American jobs" (2017). Both *The Second Machine Age* and Horsey's article make similar assertions for the future: Governmental policies must engage with the reality of automation to protect workers.

The main purpose of this study is to intersect these two conversations that are currently occurring in parallel. That is, this paper seeks to quantitatively analyze the relationship between race and gender (the "white working-class man"), and the probability of computerization. This paper has two primary goals: The first is to determine the relationship between racial demographics in each individual occupation and that occupation's probability of automation. The second major goal is to consider the effect of automation on each racial, ethnic, and gender groups in the aggregate. In other words, which racial groups will be most affected by automation? We expect that white workers and male workers will be the racial and gender groups respectively that will be affected the highest by computerization.

Methodology and Findings

The occupational demographic data for this analysis was taken from the Bureau of Labor Statistics (BLS) Labor Force Statistics from the 2016 Current Population Survey, a monthly survey of employment patterns often referred to as the Household Survey. The particular dataset used in this paper was the annual averages of employed persons by detailed occupation, sex, race, and Hispanic of Latino ethnicity. BLS reports the number individuals in an occupation, and the percentage that are Asian, Black, Hispanic/Latino, and women. The percentage of White workers is determined from subtracting the percentages of Asian and Black workers from 100. (Hispanic/Latino workers can be of any race and are therefore not included in the calculation.) The proportion of male workers is similarly determined from the proportion of female workers.

This data was compared to the computerisation probabilities determined by Frey and Osborne (2013).[1] Frey and Osborne considered "Perception and Manipulation," "Creative Intelligence," and "Social Intelligence" as the three primary categories of skills necessary to offset computerization. The degree to which an occupation required skills in these three

[1]Thank you to Professor Michael Osborne of Oxford University for providing access to the datasets from his 2013 study with Professor Carl Frey.

categories corresponded to the probability that the occupation would be computerized.

Because the occupational data from Frey and Osborne was based on the O*NET categorizations and the data from BLS was classified according to the Standard Occupational Classification (SOC), we used the SOC codes provided in the Frey and Osborne data to match the data sets. Notably, because only occupations from Frey and Osborne that matched with SOC data were considered in this study, the occupations analyzed were fewer in number than the BLS provides demographic data for. Furthermore, BLS does not report demographic data foe occupations with fewer than 50,000 employees.

After wrangling the data, we conducted some exploratory data analysis on computerization and the proportions of certain demographics in a given occupation, as seen in figure 1. A basic linear curve was fit to the graph using the ggplot package in R. Because of the

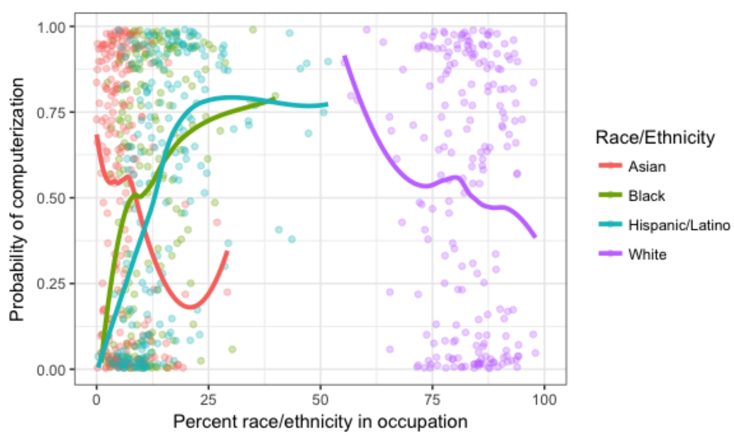
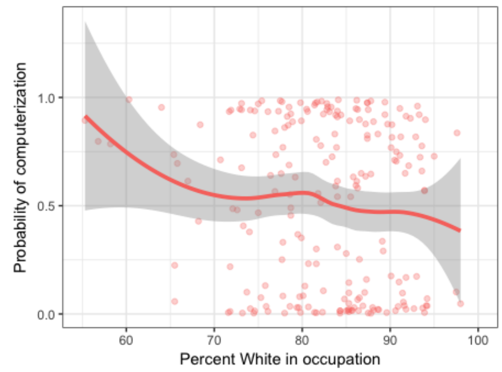


Figure 1: Occupational automatability by proportion of each racial/ethnic group.

large discrepancy between the proportion of white workers and minority workers in each occupation, minority workers were also considered independently, as seen in figures 2a and 2b. Women and men were similarly considered, as seen in figures 3a and 3b.

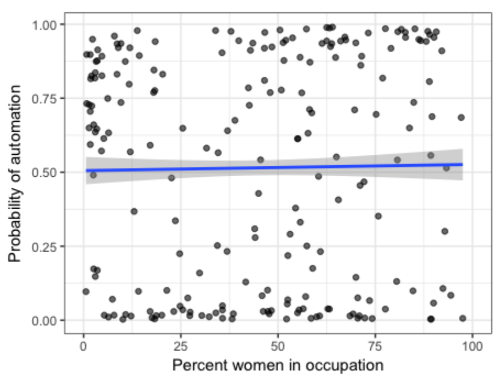
Because demographic data is limited by proportion, we also considered automatability by the number of individuals of each race. BLS does not report exact numbers by race/ethnicity, so these numbers were generated by multiplying the proportions of individuals with the total number of people employed in an occupation and rounding down to generate a natural

(a) Minority workers

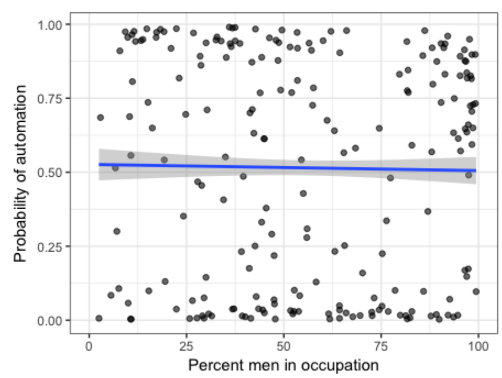


(b) White workers

Figure 2: Occupational automatability by proportion of workers in racial/ethnic group. Shading indicates standard error.



(a) Female workers



(b) Male workers

Figure 3: Occupational automatability by proportion of workers in gender. Shading indicates standard error.

number. The probabilities of computerization was transformed into a categorical variable by creating bins of size 0.1 (from 0.0 - 0.9), inclusive of the lower bound. Jitter was added to the resulting values to mitigate the effects of categorizing the probabilities and a regression model was fit to the data, as in figure 4. The same process was applied to the gender data, as in figure 5.

In order to determine the total effect of computerization on each racial group, an Effect Factor (ϵ) was created using the following formula:

$$\epsilon_{occupation} = Pr(Computerization) \times n_{group}$$

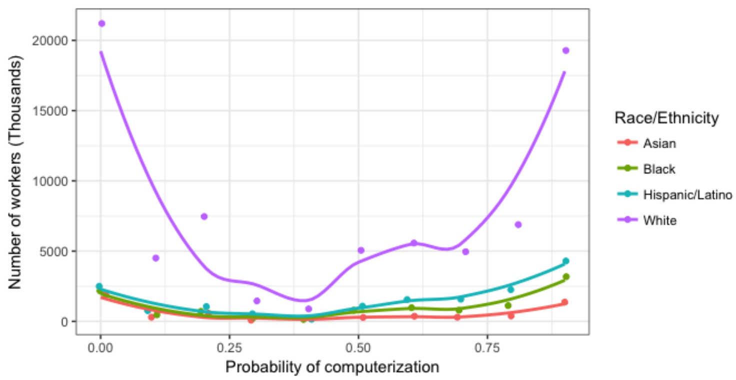


Figure 4: Number of workers at risk of computerization by race and ethnicity.

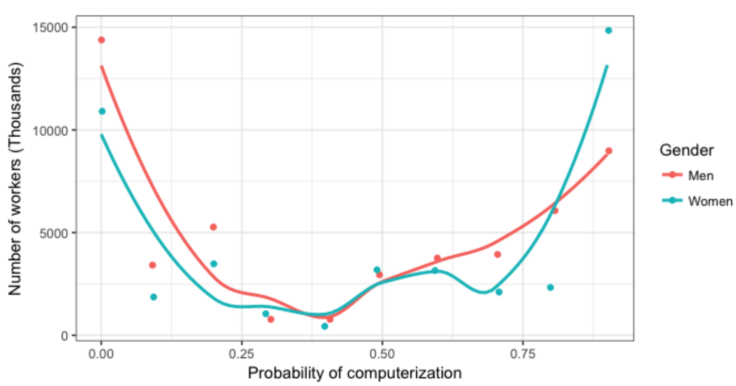


Figure 5: Number of workers at risk of computerization by gender.

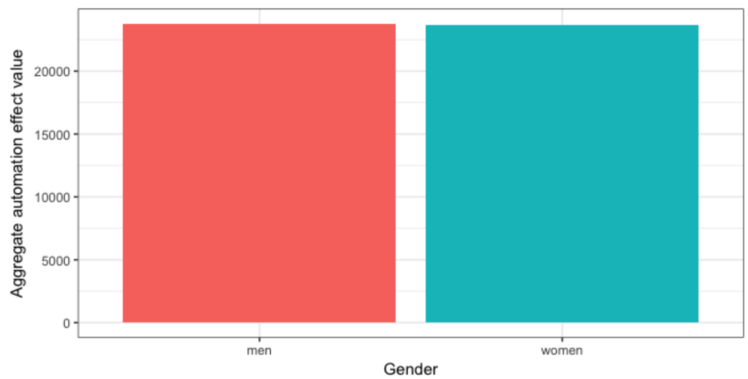


Figure 6: δ by gender.

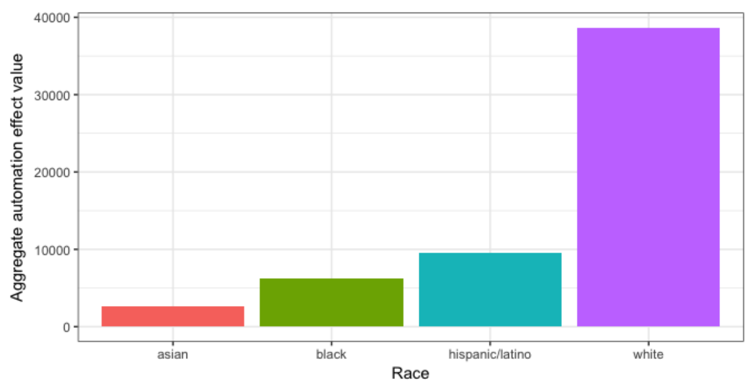


Figure 7: δ by race/ethnicity.

where n_{group} is the number of individuals in a given demographic group. The aggregate effect factor for each race (δ) was then:

$$\delta_{group} = \sum_{occupations} \epsilon$$

The resulting δ values were graphed for each demographic group, as in figure 6 and 7. In order to further consider the effects of automation on each racial group individually, the δ values were scaled by the proportion of the total employed population each demographic group held, as in figure 8 and 9.

Discussion

The major finding of this exploratory study was that automatability of occupation affects White workers to a much larger degree than non-White workers (figure 8), yet as the propor-

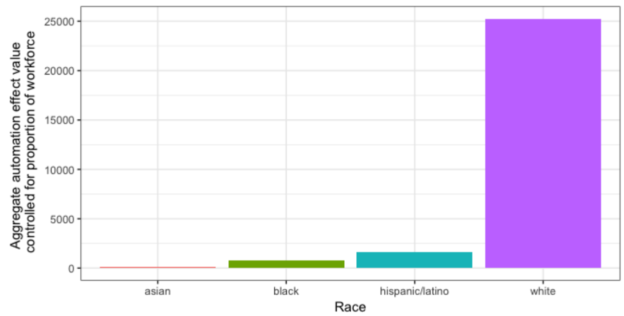


Figure 8: δ by race/ethnicity, controlled for different proportions of the total workforce held by each group.

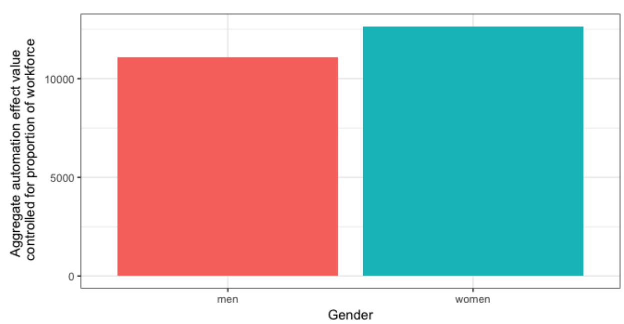


Figure 9: δ by gender, controlled for different proportions of the total workforce held by each group.

tion of white individuals in an occupation increases, the automatability of that occupation tends to decrease. Similarly gendered differences in automation effect may exist, although further analysis is necessary to determine significance.

One possible explanation for the discrepancy between δ for white workers and the trends in computerization by racial demographic proportions is that fewer workers overall are in occupations with a low probability of computerization compared to occupations with a relatively high probability of computerization. That is, white workers are overrepresented in smaller, less automatable occupational groups.

An indicator of this phenomenon would be higher percentages of White people with advanced degrees, which tend to correlate negatively with automatability. According to the U.S. Census, however, this is not necessarily the case; Asian individuals obtained advanced degrees in much higher proportions than White individuals, although a larger proportion of White individuals obtain advanced degrees than Black individuals (Ryan and Bauman, 2015). Still, the U-shape curve in figure 4 is demonstrative of the suggested effect; occupations with very high and very low automatability tend to have the largest number of white workers in them. Furthermore, the visualization of δ is inherently relative and the analysis of trends with regard to the proportions in each occupation is inherently absolute; the effect on white workers is large in comparison to that of other races, while the trends in automatability with regard to proportion of white workers stands on its own.

Ultimately, the narrative of the "white working class" experiencing economic anxiety at higher levels than other groups may be more convincing when viewed in the lens of automation and job loss. We suggest for further literature on the economic plight of the "white working class" to heavily consider automation as a factor in addition to the current paradigmatic factors, such as international trade and immigration.

References

Brynjolfsson, E., & McAfee, A. (2016). *The Second Machine Age: Work, Progress, and Prosperity in a Time of Brilliant Technologies.* New York London: W.W. Norton & Company. (OCLC: 937374818)

Frey, C., & Osborne, M. (2013, September). The Future of Employment: How suscenptible are jobs to computerisation? *Oxford Martin School*.

Horsey, D. (2017, March). Robots, not immigrants, are taking American jobs. *Los Angeles Times*. Retrieved 2017-05-19, from `http://www.latimes.com/opinion/topoftheticket/la-na-tt-robots-jobs-20170330-story.html`

Porter, E. (2016, December). Where Were Trumps Votes? Where the Jobs Werent. *The New York Times*. Retrieved 2017-05-19, from `https://www.nytimes.com/2016/12/13/business/economy/jobs-economy-voters.html`

Ryan, C., & Bauman, K. (2016, March). *Educational Attainment in the United States: 2015* (Tech. Rep.). U.S. Census Bureau. Retrieved from `https://www.census.gov/content/dam/Census/library/publications/2016/demo/p20-578.pdf`

Tankersley, J. (2016, November). *How Trump won: The revenge of working-class whites.* Retrieved 2017-05-19, from `https://www.washingtonpost.com/news/wonk/wp/2016/11/09/how-trump-won-the-revenge-of-working-class-whites/`

U.S. Department of Labor, B. o. L. S. (2016). *Employed persons by detailed occupation, sex, race, and Hispanic or Latino ethnicity* (Tech. Rep.). Bureau of Labor Statistics. Retrieved from `https://www.bls.gov/cps/cpsaat11.htm`

Table 1: Automation by Race

	Dependent variable:
	Probability of Automation
Pct Ocuppation	-0.020^{***}
	(0.005)
Black	-0.321^{***}
	(0.064)
Latino	-0.393^{***}
	(0.063)
White	0.362
	(0.263)
Pct Occ:Black	0.038^{***}
	(0.007)
Pct Occ:Latino	0.038^{***}
	(0.006)
Pct Occ:White	0.015^{**}
	(0.006)
Constant	0.634^{***}
	(0.040)
Observations	848
R^2	0.091
Adjusted R^2	0.083
Residual Std. Error	0.361 (df = 840)
F Statistic	12.015^{***} (df = 7; 840)
Note:	$^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$