Decomposing Black-White Wage Gaps Across Distributions: Young U.S. Men and Women in 1990 vs. 2011

Richey, Jeremiah and Tromp, Nikolas

Office of the Comptroller of the Currency, Kyungpook National University

1 September 2016

Decomposing Black-White Wage Gaps Across Distributions: Young U.S. Men and Women in 1990 vs. 2011

Jeremiah Richey[1]
Credit Risk Analysis Division
Office of the Comptroller of the Currency
Jeremiah.richey@occ.treas.gov

Nikolas Tromp
Kyungpook National University
nikolas.tromp@gmail.com

Abstract

We investigate changes in black-white wage gaps across wage distributions for young men and women in the U.S. between 1990 and 2011. Gaps are decomposed into composition and structural effects using a semi-parametric framework. Further, we investigate the roles of occupational choice and self-selection. We find a fall in the composition effect shrinks the wage gap at the lower end of the distribution for men and women. Conversely, an increase in the composition effect for men, and an increase in the structural effect for women, drives a widening of the wage gap at the upper end of the wage distribution.
(C14, J31, J71)

Key words: Black-white wage gaps, Discrimination, Decompositions

[1]The views expressed in this paper are those of the authors alone and do not necessarily reflect those of the Office of the Comptroller of the Currency or the US Department of the Treasury. The authors take responsibility for any errors.

Decomposing Black-White Wage Gaps Across Distributions: Young U.S. Men and Women in 1990 vs. 2011

Abstract

We investigate changes in black-white wage gaps across wage distributions for young men and women in the U.S. between 1990 and 2011. Gaps are decomposed into composition and structural effects using a semi-parametric framework. Further, we investigate the roles of occupational choice and self-selection. We find a fall in the composition effect shrinks the wage gap at the lower end of the distribution for men and women. Conversely, an increase in the composition effect for men, and an increase in the structural effect for women, drives a widening of the wage gap at the upper end of the wage distribution.
(C14, J31, J71)

Key words: Black-white wage gaps, Discrimination, Decompositions

1. Introduction

More than half a century has passed since the Civil Rights Act of 1964 prohibited racial discrimination in the U.S. labor market. During this period, initial reductions in the black-white wage gap were followed by stagnation, meaning average wages of blacks continue to lag behind those of their white counterparts (Smith and Welch, 1989; Blau and Beller, 1992; Card and Lemieux, 1994; Albrecht, van Vuuren, and Vroman, 2015). Research suggests the declining wage gap until the 1980's was due to relative improvements in the quantity and quality of education obtained by blacks, occupational upgrading of blacks, and anti-discriminatory legislation (Freeman, 1973, Smith and Welch, 1989; Card and Krueger, 1992). Since the 1980's, however, an increase in the returns to education (benefitting relatively higher-educated whites more) and occupational differentiation, coupled with a retreat of anti-discrimination initiatives and increased discrimination appear to have driven the stagnation (Anderson and Shapiro, 1996; Cancio, Evans, and Maume, 1996).

Previous research provides valuable insights into what explains the black-white gap and how the determinants of the gap may have changed overtime. However, a shortcoming of this research is it only examines the wage gap at the mean or median and thus overlooks potential differences across the wage distribution. This paper supplements the existing literature by analyzing the black-white wage gap across the entire wage distribution and at two points in time. More precisely, we explore what drives the black-white wage gap for young men and women in the U.S. across the entire wage distribution and how this has changed between 1990 and 2011. This allows us to explore whether the explanations for the stagnation in the declining wage gap found previously are heterogenous across the wage distribution. We adopt the approach developed by DiNardo, Fortin, and Lemieux (1996) in order to separate wage gaps into parts ascribed to differences in productivity characteristics (composition effects) and parts ascribed to differences in returns to productivity characteristics (structural effects) at various quantiles across the wage distribution. Furthermore, we investigate the role of occupational choice and control for self-selection into employment by adopting a novel method to impute wages.

Our results reveal that between 1990 and 2011 the black-white wage gap for both men and women narrowed at the lower end and widened at the upper end of the wage distribution. This highlights the importance of looking beyond simple summary measures such as the mean or median when examining wage gaps. Decompositions show that a fall in the composition effect drives the narrowing of the wage gap at the lower end of the wage distribution for both men and women. Alternatively, a rise in the composition effect for men, and a rise in the structural effect for women, drive the widening of the wage gap at the upper end of the wage distribution. For men we suggest these results match a scenario of a rise in the skill premium which may have amplified the effect of differences in productivity characteristics at the upper end of the distribution and compressed them at the lower end. We also conjecture that with women's advancement in the labor force to higher paying professions (Blau and Kahn, 2016), there may be greater room for racial discrimination at the upper end of the distribution. The inclusion of occupational choice and controlling for self-selection bias appear to have heterogenous effects both across the wage distribution and between 1990 and 2011. However, these adjustments do not alter the qualitative nature of our main findings.

2. Methods and Data

2.1 Decomposition Method

Originally introduced by Oaxaca (1973) and Blinder (1973), decomposition techniques are often used to explain differences in wages between groups. In this paper, we use a generalization of the standard Oaxaca-Blinder (OB) method developed by DiNardo, Fortin, and Lemieux (1996) (henceforth DFL). This technique estimates counterfactual wage densities by using reweighting functions. These counterfactual wage densities are those that black workers would exhibit (holding productivity characteristics constant) if they were paid like white workers. Thus, the difference between the counterfactual and actual black wage density is the structural effect - i.e., the component of the wage gap due to differences in returns to productivity characteristics. The remaining part of the wage gap is ascribed to the composition effect - i.e., the component of the wage gap due to differences in productivity characteristics.[2]

DFL start by noting a wage density, $f_{s|c}(w)$, is defined as the integral of a density of wages conditional on productivity characteristics, $f_s(w|x)$, over a distribution of productivity characteristics, $f_c(x)$. The subscripts s and c define the wage structure and distribution of covariates (i.e., $s, c = w$ or b for white or black) and thus the wage density is defined as:

$$f_{s|c}(w) = \int f_s(w|x) f_c(x) dx. \tag{1}$$

The actual wage densities for whites and blacks are $f_{w|w}$ and $f_{b|b}$. Therefore, the black-white wage gap is written as $f_{w|w} - f_{b|b}$. The black-white wage gap is decomposed by subtracting and adding the counterfactual wage density $f_{w|b}$ as follows:

$$f_{w|w} - f_{b|b} = \underbrace{(f_{w|w} - f_{w|b})}_{\text{Composition Effect}} + \underbrace{(f_{w|b} - f_{b|b})}_{\text{Structural Effect}}. \tag{2}$$

DFL note that the counterfactual wage density can be written as:

$$\begin{aligned} f_{w|b}(w) &= \int f_w(w|x) f_b(x) dx && (3) \\ &= \int f_w(w|x) f_w(x) \frac{f_b(x)}{f_w(x)} dx && (4) \\ &\equiv \int f_w(w|x) \psi(x) f_w(x) dx, && (5) \end{aligned}$$

where the "reweighting" function $\psi(x)$ is

$$\psi(x) = \frac{f_b(x)}{f_w(x)} \equiv \frac{f(x|b)}{f(x|w)}.^{3} \tag{6}$$

Using Bayes' rule, this ratio can be rewritten as:

$$\psi(x) = \frac{Pr(b|x)}{Pr(w|x)} \cdot \frac{Pr(w)}{Pr(b)}. \tag{7}$$

[2]Decompositions can be written by switching the reference group for the structural and composition effects. In other words, the counterfactual wage density could also be that which white workers would exhibit (holding productivity characteristics constant) if they were paid like black workers. The choice ultimately comes down to the assumption regarding whether the white or black wage would exist in the absence of discrimination. We assume the white wage structure would prevail as is common in this line of research.

[3]We rewrite $f_b(x)$ as $f(x|b)$ to make the conditioning explicit.

$Pr(w)$ and $Pr(b)$ can be estimated with sample proportions and $Pr(b|x)$ and $Pr(w|x)$ can be easily estimated parametrically with a probit model. Equation (5) shows that the counterfactual wage density is identical to the white wage density "reweighted" by the function $\psi(x)$. Thus, after estimating the weights, $\psi(x)$, we can obtain the counterfactual wage density and point estimates using simple weighted quantile functions.

For the decompositions to be valid, in the sense that they identify a meaningful structural and composition effect, the ignorability assumption must hold.

Assumption - Ignorability: *Let D_g denote race for $g = w$ or b and ϵ denote unobservables that affect wages conditional on X. Also let (D_g, X, ϵ) have a joint distribution. For all x in X: ϵ is independent of D_g given $X = x$ (Fortin, Lemieux, and Firpo 2010).*

The ignorability assumption states that unobservables, ϵ, conditional on x are independent of race; $\epsilon \perp D_g | x$. As Fortin et al. (2010) explain, the correlation between unobserved ability and measured education likely produces inaccurate estimates for the return to education and thus invalidates *detailed* decompositions. This clearly applies to our analysis as education is one of our major covariates. However, as long as the distribution of unobserved ability conditional on observed characteristics is the same across races, *aggregate* decompositions are valid. We therefore restrict our analysis and interpretations to aggregate decompositions focusing on the total structural and composition effects rather than those for each individual covariate.[4]

2.2 Wage Imputations to Correct for Self-Selection Bias

Sample selection issues are potentially important components of measured wage gaps. These issues may be especially pertinent for female labor supply decisions (Chandra, 2003; Neal, 2004; Olivetti and Petrongolo, 2008; Albrecht et al., 2015). If the decision to work is correlated with the error term in the wage equation, estimates of the wage gap, the wage structure, and thus the decomposition will be biased. Some authors have attempted to correct for these issues while conducting decompositions across the distribution by using a method, suggested by Buchinsky (1998), which resembles the Heckman Two-Step approach (Albrecht et al., 2009). In this paper, rather than relying on restrictions necessary for such approaches, we propose a method of wage imputations based on the same reweighting scheme as in our decompositions. In particular, we assume that if the base wage structure is supplemented with key variables the selection issue can be controlled

[4]See Fortin et al. (2010) for a discussion of other technical assumptions for decompositions to be valid.

for, i.e., the decision to work is independent of the error term in the supplemented wage structure.[5] This supplemented model is then used to impute wages for non-workers via the same reweighting method used for our decompositions. This yields the 'full' distribution of wages (in the sense that it incorporates non-workers' hypothetical market wages). Once this is obtained we can proceed as before with our decompositions. Thus, this imputation-decomposition method is essentially a double-weighting scheme.

In particular, we begin with a supplemented wage structure, $f(w|z)$, where z is the set of productivity characteristics plus our additional covariates controlling for selection. These include parental characteristics (education levels, income levels, and age of mother at birth) and additional demographic characteristics (marital status, number of children, and a binary indicator if a child under 5 is present in the household). Given this, we wish to know the distribution of wages for non-workers if they were working. Thus we utilize the same counterfactual estimation as the decompositions where the subscripts w and nw now denote worker and non-worker. Thus, the counterfactual wage density for non-workers if they were working is

$$\begin{aligned} f_{w|nw}(w) &= \int f_w(w|z) f_{nw}(z) dz && (8) \\ &= \int f_w(w|z) f_w(z) \frac{f_{nw}(z)}{f_w(z)} dz && (9) \\ &\equiv \int f_w(w|z) \psi(z) f_w(z) dz. && (10) \end{aligned}$$

If our assumption of selection on observables holds once we supplement our wage model, $f(w|z)$ is a valid model for both workers' and non-workers' wages and the distribution of non-workers' wages are simply a reweighted distribution of workers' wages. To get a 'full' distribution of wages we create weights defined as: $W_i = Pr(working) + [1 - Pr(working)]\psi_i$ which are applied to workers' wages. Then we proceed as before to decompose the wage gaps but using our now preweighted data.

2.3 Data

The data used in our analysis come from the 1979 and 1997 cohorts of the National Longitudinal Survey of Youth (NLSY). The NLSY79 panel survey follows US youths aged 14-22 starting in 1979 and the NLSY97 panel survey follows U.S. youths aged 12-16 starting in 1997. While the 1979 cohort oversamples poor whites and military, both the 1979 and 1997 cohorts oversample minorities. We include the oversample of minorities for both cohorts but exclude the oversampled military and poor whites from the 1979 cohort which were

[5] Neal (2004), Olivetti and Petrongolo (2008), and Albrecht et al. (2015) similarly rely on an imputation method. However, their focus on the median allows for less restrictive assumptions on the imputations.

discontinued in 1984 and 1990.

We use similar procedures as Neal (2004) and Albrecht et at. (2015) in constructing our dataset so as to align our results with previous research. One issue is that the age distribution is 25-33 years old for the NLSY79 cohort in 1990 and 26-31 years old for the NLSY97 cohort in 2011.[6] Therefore to compare workers of the same age range for 1990 and 2011 we reconstruct the NLSY79 sample to parallel the NLSY97 sample. This means matching workers aged 25, 32, and 33 in 1990 with their data from when they were within in the 26-31 age range. We use log of hourly wages in 1990 dollars for both cohorts, limiting our sample to those who report an hourly wage of greater than 1 and less than 100. Our independent variables for workers are similar to those used by Anderson and Shapiro (1996) and much of the literature. These include years of education, experience, experience squared, cognitive ability, and dummy variables for residence in urban areas and the south. In addition we fully interact all these variables in our estimation. Table 1 provides summary statistics for these variables. Our sample sizes for working white and black men in 1990 are 1863 and 1031 and in 2011 are 1350 and 535. These increase by 75, 76, 149, and 100 respectively with imputations to include non-workers. For women, our sample sizes for working whites and blacks in 1990 are 1738 and 1008 and in 2011 are 1210 and 611. These increase by 159, 103, 200, and 123 respectively with imputations to include non-workers.

Education is measured as years of schooling. Experience is measured as full-time equivalent (FTE) experience rather than potential experience. One benefit of using the NLSY is that it provides detailed data on weekly hours worked from which we can construct our FTE experience variable. This is important as potential experience can overestimate the true level of experience (Antecol and Bedard, 2004). Cognitive ability is measured using Armed Forces Qualification Tests (AFQT) scores constructed from the Armed Services Vocational Aptitude Battery (ASVAB). Comparability of original test results for the NLSY79 and NLSY97 cohorts is difficult as the ages which the two cohorts took the ASVAB and the format of the tests both differ. Thus, we use AFQT scores reconstructed by Altonji, Bharadwaj, and Lange (2012). This was achieved by mappings across ages and formats (see Segall, 1997; Altonji, Bharadwaj, and Lange, 2012 for details).[7]

In wage decomposition studies, serious consideration must go into which variables are included. In reality, discrimination may occur in both market and pre-market settings, and by including a variable in the

[6] At the time of writing the paper 2011 was the latest wave available.

[7] We are grateful to Altonji, Bharadwaj, and Lange for making the constructed scores publicly available on Fabian Lange's website http://www.econ.yale.edu/ fl88/.

analysis we may incorrectly attribute 'pre-market' discrimination related to that variable to the composition effect. For example, if we include occupation dummy variables, then discrimination is measured as different returns to the same occupation, but discrimination in entering an occupation (or the expectation of future discrimination in certain occupations) is omitted from the structural effect and subsumed in the composition effect. Though our base specification omits occupation dummy variables, we include them in an auxiliary analysis to investigate the sensitivity of our results to occupational choice. The occupation dummy variables include manager, professional, sales, service, farming, craft, operatives, with clerical as the reference group. These are constructed from 3-digit occupation codes in the NLSY79 cohort and 4-digit occupation codes in the NLSY97 cohorts.[8]

3. RESULTS

3.1 Decomposition Results for Men

The baseline wage gaps and decomposition results for men are presented in Table 2. Row 1 shows the log wage gap in 1990 at the mean and the 10^{th}, 25^{th}, 50^{th}, 75^{th}, and 90^{th} quantiles. The wage gap is 0.18 log wage points at the 10^{th} quantile, rises to approximately 0.31 at the 25^{th}, 50^{th}, and 75^{th} quantiles, and then falls to 0.24 at the 90^{th} quantile. This inverted U-shape can be seen in Figure 1(a). We use the DFL method to decompose the wage gaps into structural and composition effects, as presented in rows 2 and 3. At the 10^{th} and 25^{th} quantiles the composition effect explains the entire wage gap and more, which is offset by a negative structural effect. Moving up the wage distribution the structural effect increases and composition effect decreases, each explaining roughly 50% of the wage gap at the 90^{th} quantile. Figure 1(a) provides a visual representation of these patterns across the distribution. Overall, the results show that differences in productivity characteristics explain the entire wage gap at the lower end of the distribution, while differences in returns to productivity characteristics increase in importance at higher quantiles.[9]

Row 4 of Table 2 presents the log wage gap for men in 2011. The log wage gap widens as we move up the wage distribution, from 0.17 at the 10^{th} quantile to 0.34 at 75^{th} quantile, and then falls slightly to 0.31 at 90^{th} quantile. As can be seen by comparing figures 1(a) and 2(a), there is a steeper rise in the wage gap

[8] Despite the codes differing slightly between the 79 and 97 cohorts, the fact we are analyzing the aggregate impact of the inclusion of occupation choices makes this largely irrelevant.

[9] We find that a potentially large bias can occur in OB decompositions as a result of the linear assumption. Although we do not report OB decomposition results in the tables, we note that the composition effect and structural effect are 0.273 and -0.003 respectively. These results differ considerably to the semi-parametric DFL decomposition at the mean, where the composition effect and structural effect are 0.216 and 0.054.

in 2011. The higher wage gap at higher quantiles are commonly referred to as a "glass ceiling" in this line of research (Albrecht et al., 2009). Between 1990 and 2011, the wage gap fell at the lower end and median and rose at the upper end of the wage distribution. Rows 5 and 6 present the structural and composition effects for 2011. Between 1990 and 2011, a fall in the composition effect outweighs a rise in the structural effect, causing the smaller wage gap at the 10^{th} and 25^{th} quantiles. At the 50^{th} quantile, both the composition and structural effects fell. At the 75^{th} and 90^{th} quantiles, on the other hand, a rise in the composition effect outweighs a fall in the structural effect causing the larger wage gap. The changes in the wage gap across the distribution between 1990 and 2011 match a scenario of a rising skill premium over the last few decades. In other words, increasing returns for high-skilled workers may have amplified the effect of existing differences in black and white productivity characteristics causing an increase in the composition effect at higher quantiles. In contrast, lower returns for lower-skilled workers may have compressed the effect of existing differences in black and white productivity characteristics causing a decrease in the composition effect at lower quantiles. Unfortunately many of the changes in wage gaps, structural effects, and composition effects between 1990 and 2011 are not statistically different from zero (except for the rise in the structural effect at the 10^{th} quantile, fall in the composition effect at the 10^{th} and 25^{th} quantiles, and rise in the composition effect at 90^{th} quantile).

In Table 3 we present the wage gap and decomposition results with the inclusion of occupation dummy variables. In 1990, adding these variables leads to a fall in the composition effect (rise in the structural effect) of 0.05 at the 10^{th} quantile and a rise in the composition effect (fall in the structural effect) of 0.11 at the 75^{th} quantile. The changes at the remaining quantiles are extremely small. In 2011, adding occupation dummy variables increases the composition effect (decreases the structural effect) across the entire wage distribution, most notably by 0.11 at the 25^{th} quantile. As a result, the composition effect explains almost the entire wage gap at the lower end of the wage distribution, as shown in Figure 2(b). This suggests differences in occupations between blacks and whites may play an important role in wage differentials across the two groups in 2011, particularly at the lower end of the distribution. Overall, however, the inclusion of occupation dummy variables does not alter the nature of our main findings for men.

Lastly, we present the wage gap and decomposition results corrected for self-selection bias using our proposed imputation method in Table 4. In 1990, adding imputed wages has a minor impact on the wage gap, slightly raising it by 0.02, 0.03, and 0.02 log wage points at the 10^{th}, 50^{th}, and 90^{th} quantiles. The structural effect increases slightly at every quantile except the 75^{th}, while the composition effect falls at the 10^{th} and 25^{th} quantiles. In 2011, adding imputed wages increases the wage gap by between 0.01 to 0.03

at every quantile. The structural effect increases at the lower end of the distribution and decreases at the upper end. The composition effect, on the other hand, falls at the 25^{th} and rises at the 50^{th}, 75^{th} and 90^{th} quantiles. The above results suggest that the black-white wage gap for men is likely to be somewhat underestimated when self-selection bias is ignored. However, when comparing Figures 1(a) and 2(a) with 1(c) and 2(c) we see that the magnitude of the underestimation may be minimal, leaving our main findings unchanged.

3.2 Decomposition Results for Women

The baseline wage gap and decomposition results for women are presented in Table 5. The wage gap at the 10^{th} quantile is 0.09 log wage points, increases to 0.16, 0.19, and 0.21 at the 25^{th}, 50^{th}, and 75^{th} quantiles, and then falls to 0.17 at the 90^{th} quantile. Despite showing a similar trend to men, as can be seen in Figure 3(a), the wage gap for women is considerably smaller. In fact, at the 10^{th} and 25^{th} quantiles the black-white wage gap for women is approximately half of that for men in 1990. Using the DFL method to decompose the wage gap shows that the composition effect explains the whole wage gap and more across the entire distribution. This is offset by the negative structural effect across the distribution in 1990 which implies black women may be receiving a wage premium for their productivity characteristics.

In 2011, the black-white wage gap for women also shows a "glass ceiling effect" increasing from 0.02 at the 10^{th} quantile to 0.27 at the 90^{th} quantile. This is illustrated in Figure 4(a). Between 1990 and 2011, the wage gap fell at the lower end of the wage distribution and rose at the upper end, as it did for men. A fall in the composition effect between 1990 and 2011 drives the reduction in the wage gap at the 10^{th} quantile and a fall in both structural and composition effects causes the reduction in the wage gap at the 25^{th} quantile. In contrast to men, the increase in the wage gap at the 50^{th}, 75^{th}, and 90^{th} quantiles is a result of a rise in the structural effect outweighing a fall in the composition effect. These changes can be seen in Figure 4(a). We conjecture that as women have entered more managerial and professional jobs (Blau and Kahn, 2016), they may find themselves in relatively higher paying jobs where there is greater room for discrimination, causing the rise in the structural effect at the upper end of the distribution. Unfortunately many of the changes in wage gaps, structural effects, and composition effects between 1990 and 2011 are not statistically different from zero (except for the fall in the wage gap at the 10^{th} quantile, rise at the 90^{th} quantile, and the rise in the structural effect at the 90^{th} quantile).

In Table 6 we present the wage gap and decomposition results for women with the inclusion of occupation dummy variables. In 1990, adding these variables results in minimal changes as is made clear by comparing Figures 3(a) and 3(b). The composition effect falls (structural effect rises) by 0.03 at the 10^{th} quantile and the structural effect falls (composition effect rises) by 0.01 at the 25^{th} and 90^{th} quantiles. In 2011, adding occupation dummy variables results in a decline in the composition effect (rise in the structural effect) of 0.03 and 0.07 at the 25^{th} and 50^{th} quantiles. Overall, the effect of including occupational choice is minimal and differs both across the wage distribution and between 1990 and 2011.

Lastly, Table 7 presents the results corrected for self-selection bias using our proposed imputation method. In contrast to men, the inclusion of imputed wages for women in 1990 leads to a minor reduction in the black-white wage gap by 0.04, 0.02, and 0.01 at 10^{th}, 25^{th}, and 90^{th} quantiles. Adding those with imputed wages tends to increase the structural effect and decrease the composition effect across the distribution. We conjecture the relatively higher imputed wages for non-working black women may be the result of the wage premium black women received in 1990. In 2011, adding those with imputed wages results in small reductions in the wage gap at the 10^{th} and 25^{th} quantiles and small increases at the 75^{th} and 90^{th} quantiles. The composition effect falls at the lower end of the distribution and increases at the upper end. The structural effect rises at the 25^{th} and falls at the 50^{th} quantile. The higher imputed wages for white women at the upper end of the distribution may be a result of white women with high productivity characteristics being married to high-earning white men and thus opting out of employment (Neal, 2004).

4. CONCLUSION

In this paper, we explore changes in the black-white wage gap for young men and women in the U.S. across the entire wage distribution between 1990 and 2011. Using a semi-parametric reweighting method we decompose wage gaps into the effect of differences in productivity characteristics and differences in returns to productivity characteristics. In addition, we investigate the importance of the inclusion of occupational choice and correct for self-selection into employment utilizing a novel imputation method based on the same reweighting procedure used in our decompositions.

Our results show a “glass ceiling” effect where the black-white wage gaps for both men and women widen as we move up the wage distribution. Between 1990 and 2011, the wage gap fell at the lower end and rose at the upper end of the distribution for both men and women. Decompositions shows that the narrowing of the wage gap at the lower end of the distribution is generally driven by a reduction in the composition effect for both men and women. This may be the result of a decline in wages for low skilled workers compressing the effect of differences in productivity characteristics. A rise in the composition effect at the upper end of the wage distribution for men drives the widening of the wage gap. This matches a scenario where a rise in the skill premium amplifies differences in productivity characteristics between black and white men. For women, on the other hand, an increase in the structural effect drives the widening of the wage gap at the upper end of the wage distribution. We suggest that as women enter more managerial and professional professions there may be more room for discrimination.

The inclusion of occupational choice has varying impacts both across the wage distribution and between 1990 and 2011. Overall, the differences in occupational composition between blacks and whites seem to be more important for men than for women. Controlling for self-selection bias by imputing wages leads to small increases in the wage gap for men in both 1990 and 2011. For women, the wage gap fell both in 1990 and at the lower end of the distribution in 2011 but increased at the upper end in 2011.

References

Albrecht, James, Aico van Vuuren, and Susan Vroman. 2009. Counterfactual distributions with sample selection adjustments: Econometric theory and an application to the Netherlands. *Labour Economics*, 16(4), pp. 383-396.

——. 2015. The Black-White Wage Gap Among Young Women in 1990 vs. 2011: The Role of Selection and Educational Attainment. *Labour Economics*, 33, pp. 66-71.

Altonji, Joseph G., Prashant Bharadwaj, and Fabian Lange. 2012. Changes in the Characteristics of American Youth: Implications for Adult Outcomes. *Journal of Labor Economics*, 30(4), pp. 783-828.

Anderson, Deborah and David Shapiro. 1996. Racial Difference in Access to High-Paying Jobs and the Wage Gap between Black and White Women. *Industrial and Labor Relations Review*, 49(2), pp. 273-286.

Antecol, Heather and Kelly Bedard. 2004. The Racial Wage Gap: The Importance of Labor Force Attachment Difference across Black, Mexican, and White Men. *The Journal of Human Resources*, 39(2), pp. 564-583.

Blau, Francine D. and Andrea H. Beller. 1992. Black-White Earnings Over the 1970s and 1980s: Gender Differences in Trends. *The Review of Economics and Statistics*, 74(2), pp. 276-286.

Blau, Francine and Lawrence Kahn. 2016. The Gender Wage Gap: Extent, Trends, and Explanations. NBER Working Paper No. 21913

Blinder, Alan S. 1973. Wage Discrimination: Reduced Form and Structural Variables. *Journal of Human Resources*, 8(4), pp. 436-455.

Buchinsky, Moshe. 1998. The Dynamics of Changes in the Female Wage Distribution in the USA: a Quantile Regression Approach. *Journal of Applied Econometrics*, 13, pp. 1-30.

Cancio, A. Silvia, T. David Evans, and David J. Maume, Jr. 1996. Reconsidering the Declining Significance of Race: Racial Differences in Early Career Wages. *American Sociological Review*, 61,(4), pp. 541-556.

Card, David and Alan B. Krueger. 1992. School Quality and Black-White Relative Earnings: A Direct Assessment. *The Quarterly Journal of Economics*, 107(1), pp. 151-200.

Card, David and Thomas Lemieux. 1994. Changing Wage Structure and Black-White Wage Differentials. *The American Economic Review*, 84(2), pp. 29-33.

Chandra, Amitabh. 2003. Is the convergence in the racial wage gap illusory? NBER working paper 9476.

DiNardo, John, Nicole M. Fortin, and Thomas Lemieux. 1996. Labor Market Institutions and the Distribution of Wages, 1973-1992: A Semiparametric Approach. *Econometricia*, 64(5), pp. 1001-1044.

Fortin, Nicole, Thomas Lemieux and Sergio Firpo. 2010. Decomposition Techniques in Economics. *NBER Working Paper Series*, No. 16045.

Freeman, Richard B. 1973. Changes in the Labor Market for Black Americans. *Brookings Papers on Economic Activity*, 4(1), pp. 67-132.

Neal, Derek. 2004. The Measured Black-White Wage Gap among Women Is Too Small. *Journal of Political Economy*, 112(S1), pp. S1-S28.

Oaxaca, Ronald L. 1973. Male-Female Wage Differentials in Urban Labor Markets. *International Economic Review*, 14(3), pp. 693-709.

Olivetti, Claudia and Barbara Petrongolo. 2008. Unequal Pay or Unequal Employment? A cross country analysis of the gender gap. *Journal of Labor Economics*, 26(4), pp. 621-654.

Segall, Daniel O. 1997. Equating the CAT-ASVAB. In *Computerized Adaptive Testing: From Inquiry to Operation*, ed. William A. Sands, Brian K. Waters, and James R. McBride, Washington, DC: American Psychological Association, pp. 181-198.

Smith, James P and Finis R. Welch. 1989. Black Economic Progress After Myrdal. *Journal of Economic Literature*, 27(2), pp. 519-564.

Table 1: Descriptive Statistics

	Men - 1990		Men - 2011		Women - 1990		Women - 2011	
	White	Black	White	Black	White	Black	White	Black
Median Log Wage	2.35	2.04	2.26	1.98	2.12	1.92	2.15	1.90
Mean Log Wage	2.33	2.06	2.28	2.02	2.09	1.92	2.14	1.97
	(0.49)	(0.47)	(0.54)	(0.52)	(0.54)	(0.49)	(0.60)	(0.51)
Years of Education	13.33	12.49	14.26	13.04	13.50	13.12	15.09	14.20
	(2.48)	(2.01)	(2.65)	(2.34)	(2.28)	(1.96)	(2.60)	(2.58)
FTE Experience	8.15	7.01	7.64	6.61	7.10	5.82	7.21	6.59
	(2.72)	(2.84)	(2.69)	(2.70)	(2.70)	(2.95)	(2.44)	(2.57)
AFQT	0.45	-0.70	0.36	-0.56	0.52	-0.52	0.45	-0.35
	(0.93)	(0.87)	(0.92)	(0.97)	(0.80)	(0.76)	(0.77)	(0.93)
Urban	0.73	0.82	0.71	0.80	0.74	0.83	0.70	0.82
South	0.30	0.57	0.30	0.61	0.32	0.63	0.31	0.66
Sample Size	1863	1031	1350	535	1738	1008	1210	611

Note: Standard deviations are in parenthesis. Wages are in 1990 dollars. AFQT scores are standardized.

Table 2: Decomposition Results for Men

		Mean	10^{th}	25^{th}	50^{th}	75^{th}	90^{th}
	Wage Gap	0.27***	0.18***	0.31***	0.31***	0.30***	0.24***
		(0.02)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
NLSY79	Structural	0.05	-0.12	-0.03	0.07	0.13	0.12*
- 1990	Effect	(0.07)	(0.09)	(0.09)	(0.07)	(0.15)	(0.06)
	Composition	0.22***	0.30***	0.34***	0.24***	0.17	0.12**
	Effect	(0.07)	(0.09)	(0.09)	(0.07)	(0.15)	(0.06)
	Wage Gap	0.26***	0.17***	0.26***	0.28***	0.34***	0.31***
		(0.03)	(0.03)	(0.03)	(0.04)	(0.04)	(0.07)
NLSY97	Structural	0.08**	0.05	0.11**	0.06	0.11***	0.06
- 2011	Effect	(0.04)	(0.04)	(0.05)	(0.06)	(0.04)	(0.08)
	Composition	0.18***	0.12***	0.15***	0.22***	0.23***	0.26***
	Effect	(0.03)	(0.03)	(0.05)	(0.05)	(0.04)	(0.05)

Note: Bootstrapped standard errors (based on 1,000 replications) are reported in parentheses. Statistical significance is denoted by *** for the 1% level, ** for the 5% level, and * for the 10% level.

Table 3: Decomposition Results for Men - Occupation Dummy Variables

		Mean	10^{th}	25^{th}	50^{th}	75^{th}	90^{th}
	Wage Gap	0.27***	0.18***	0.31***	0.31***	0.30***	0.24***
		(0.02)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
NLSY79	Structural	0.03	-0.07	-0.02	0.06	0.02	0.12*
- 1990	Effect	(0.04)	(0.06)	(0.06)	(0.05)	(0.09)	(0.07)
	Composition	0.24***	0.25***	0.33***	0.25***	0.28***	0.13**
	Effect	(0.04)	(0.06)	(0.06)	(0.04)	(0.09)	(0.06)
	Wage Gap	0.26***	0.17***	0.26***	0.28***	0.34***	0.31***
		(0.03)	(0.03)	(0.03)	(0.03)	(0.04)	(0.07)
NLSY97	Structural	0.04	0.01	0.00	0.03	0.08	0.03
- 2011	Effect	(0.04)	(0.04)	(0.07)	(0.07)	(0.05)	(0.08)
	Composition	0.22***	0.16***	0.26***	0.25***	0.25***	0.29***
	Effect	(0.04)	(0.04)	(0.07)	(0.07)	(0.05)	(0.06)

Note: Bootstrapped standard errors (based on 1,000 replications) are reported in parentheses. Statistical significance is denoted by *** for the 1% level, ** for the 5% level, and * for the 10% level.

Table 4: Decomposition Results for Men - Imputed Wages

		Mean	10^{th}	25^{th}	50^{th}	75^{th}	90^{th}
	Wage Gap	0.28***	0.20***	0.31***	0.34***	0.30***	0.26***
		(0.02)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
NLSY79	Structural	0.06	-0.08	-0.02	0.09	0.13	0.14**
- 1990	Effect	(0.07)	(0.09)	(0.10)	(0.07)	(0.15)	(0.07)
	Composition	0.22***	0.28***	0.33***	0.24***	0.17	0.12**
	Effect	(0.07)	(0.09)	(0.09)	(0.07)	(0.15)	(0.06)
	Wage Gap	0.28***	0.18***	0.28***	0.30***	0.35***	0.34***
		(0.03)	(0.03)	(0.04)	(0.04)	(0.04)	(0.08)
NLSY97	Structural	0.09**	0.06	0.16***	0.05	0.10**	0.04
- 2011	Effect	(0.04)	(0.04)	(0.05)	(0.06)	(0.05)	(0.09)
	Composition	0.19***	0.12***	0.12**	0.25***	0.25***	0.31***
	Effect	(0.03)	(0.04)	(0.05)	(0.06)	(0.04)	(0.05)

Note: Bootstrapped standard errors (based on 1,000 replications) are reported in parentheses. Statistical significance is denoted by *** for the 1% level, ** for the 5% level, and * for the 10% level.

Table 5: Decomposition Results for Women

		Mean	10^{th}	25^{th}	50^{th}	75^{th}	90^{th}
	Wage Gap	0.17***	0.09***	0.16***	0.19***	0.21***	0.17***
		(0.02)	(0.03)	(0.03)	(0.02)	(0.03)	(0.03)
NLSY79	Structural	-0.11**	-0.05	-0.12***	-0.12*	-0.07*	-0.10***
- 1990	Effect	(0.04)	(0.16)	(0.04)	(0.07)	(0.04)	(0.04)
	Composition	0.28***	0.14	0.27***	0.31***	0.28***	0.27***
	Effect	(0.04)	(0.16)	(0.04)	(0.07)	(0.03)	(0.03)
	Wage Gap	0.17***	0.02	0.11***	0.25***	0.25***	0.27***
		(0.03)	(0.03)	(0.03)	(0.03)	(0.05)	(0.04)
NLSY97	Structural	-0.04	-0.03	-0.15***	-0.04	-0.01	0.07
- 2011	Effect	(0.04)	(0.02)	(0.05)	(0.09)	(0.05)	(0.05)
	Composition	0.21***	0.06**	0.25***	0.30***	0.26***	0.20***
	Effect	(0.04)	(0.03)	(0.05)	(0.09)	(0.05)	(0.04)

Note: Bootstrapped standard errors (based on 1,000 replications) are reported in parentheses. Statistical significance is denoted by *** for the 1% level, ** for the 5% level, and * for the 10% level.

Table 6: Decomposition Results for Women - Occupational Dummy Variables

		Mean	10^{th}	25^{th}	50^{th}	75^{th}	90^{th}
	Wage Gap	0.17***	0.09***	0.16***	0.19***	0.21***	0.17***
		(0.02)	(0.03)	(0.03)	(0.02)	(0.03)	(0.03)
NLSY79	Structural	-0.11**	-0.02	-0.13***	-0.12	-0.07	-0.11***
- 1990	Effect	(0.05)	(0.11)	(0.04)	(0.09)	(0.05)	(0.04)
	Composition	0.28***	0.11	0.29***	0.31***	0.28***	0.28***
	Effect	(0.04)	(0.11)	(0.04)	(0.09)	(0.04)	(0.04)
	Wage Gap	0.17***	0.02	0.11***	0.25***	0.25***	0.27***
		(0.03)	(0.03)	(0.02)	(0.03)	(0.04)	(0.04)
NLSY97	Structural	-0.02	-0.03	-0.12**	0.02	-0.01	0.06
- 2011	Effect	(0.04)	(0.04)	(0.05)	(0.08)	(0.06)	(0.05)
	Composition	0.19***	0.06	0.22***	0.23***	0.26***	0.20***
	Effect	(0.04)	(0.05)	(0.05)	(0.08)	(0.05)	(0.05)

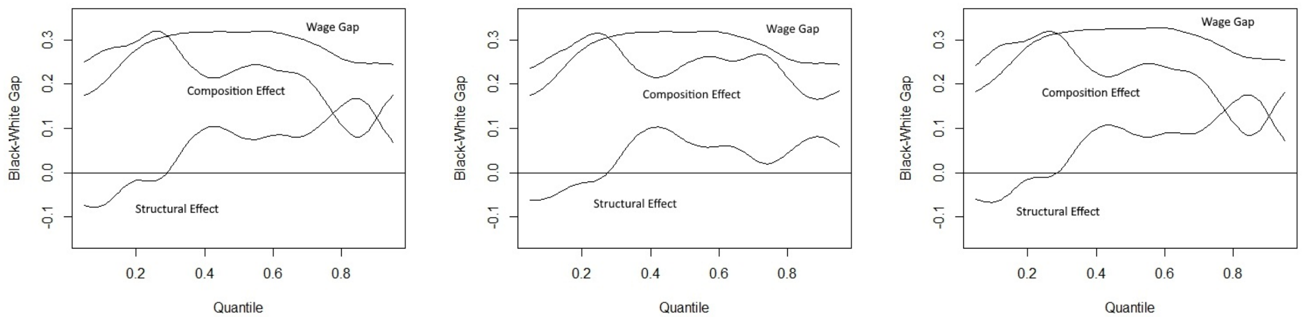
Note: Bootstrapped standard errors (based on 1,000 replications) are reported in parentheses. Statistical significance is denoted by *** for the 1% level, ** for the 5% level, and * for the 10% level.

Table 7: Decomposition Results for Women - Imputed Wages

		Mean	10^{th}	25^{th}	50^{th}	75^{th}	90^{th}
	Wage Gap	0.16***	0.05	0.14***	0.19***	0.21***	0.16***
		(0.02)	(0.03)	(0.03)	(0.02)	(0.03)	(0.03)
NLSY79	Structural	-0.09*	-0.04	-0.06	-0.10	-0.04	-0.11**
- 1990	Effect	(0.05)	(0.15)	(0.05)	(0.07)	(0.04)	(0.04)
	Composition	0.25***	0.09	0.20***	0.29***	0.25***	0.26***
	Effect	(0.05)	(0.15)	(0.05)	(0.07)	(0.04)	(0.04)
	Wage Gap	0.19***	0.00	0.10***	0.25***	0.29***	0.32***
		(0.04)	(0.04)	(0.03)	(0.04)	(0.06)	(0.06)
NLSY97	Structural	-0.02	-0.03	-0.09	-0.09	0.00	0.08
- 2011	Effect	(0.05)	(0.03)	(0.06)	(0.10)	(0.06)	(0.06)
	Composition	0.20***	0.03	0.19***	0.34***	0.29***	0.24***
	Effect	(0.05)	(0.04)	(0.06)	(0.10)	(0.06)	(0.06)

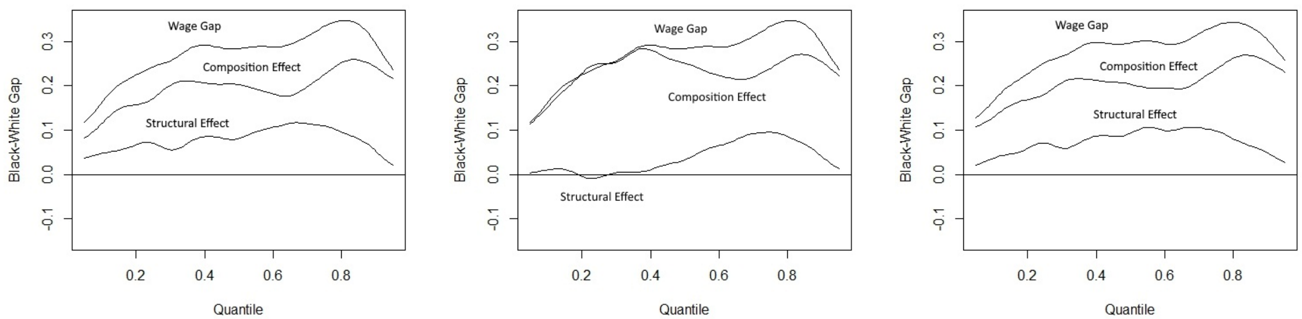
Note: Bootstrapped standard errors (based on 1,000 replications) are reported in parentheses. Statistical significance is denoted by *** for the 1% level, ** for the 5% level, and * for the 10% level.

Figure 1: Black-White Wage Gaps for Men in 1990



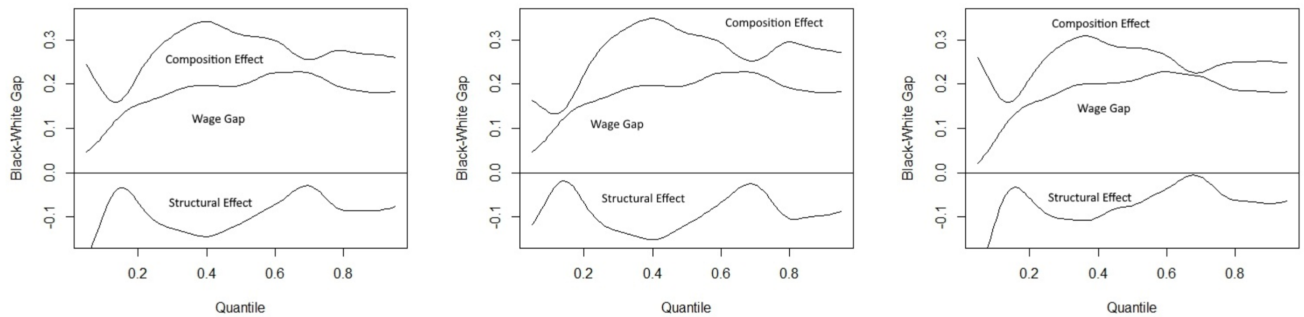
(a) Base Specifications (b) With Occupations (c) With Imputations

Figure 2: Black-White Wage Gaps for Men in 2011



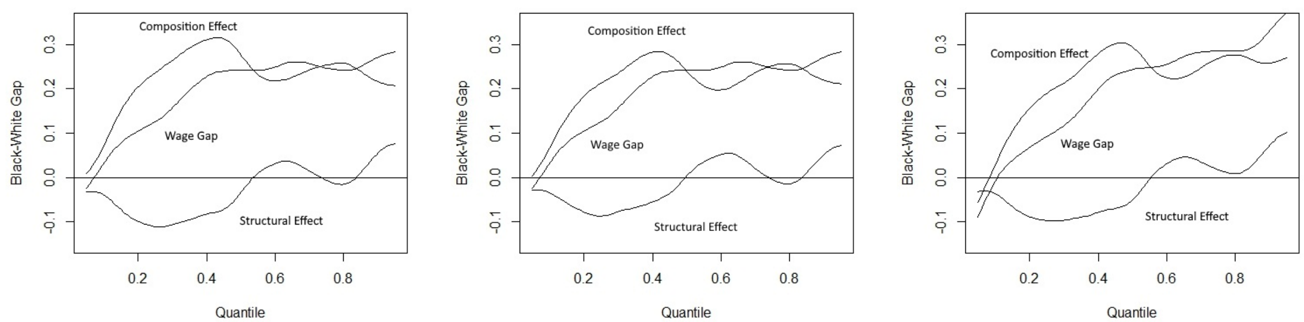
(a) Base Specifications (b) With Occupations (c) With Imputations

Figure 3: Black-White Wage Gaps for Women in 1990



(a) Base Specifications (b) With Occupations (c) With Imputations

Figure 4: Black-White Wage Gaps for Women in 2011



(a) Base Specifications (b) With Occupations (c) With Imputations