IZA DP No. 8005

Selection and the Measured Black-White Wage Gap Among Young Women Revisited

James Albrecht
Aico van Vuuren
Susan Vroman

February 2014

Selection and the Measured Black-White Wage Gap Among Young Women Revisited

James Albrecht
*Georgetown University
and IZA*

Aico van Vuuren
Free University of Amsterdam (VU)

Susan Vroman
*Georgetown University
and IZA*

ABSTRACT

Selection and the Measured Black-White Wage Gap Among Young Women Revisited

Derek Neal (JPE 2004) used the NLSY79 to show that the observed median log wage gap between young white and young black women in 1990 underestimated the true, selection-corrected gap, i.e., the gap we would have expected to see had all of these women been employed in 1990. In this paper, we use the NLSY97 to update his analysis. The observed median log wage gap increased substantially between 1990 and 2011, as did the selection-corrected gap. These increases are explained to a considerable extent by changes in the distribution of educational attainment across young white and black women.

Corresponding author:

Susan Vroman
Department of Economics
Georgetown University
Washington, DC 20057
USA
E-mail: vromans@georgetown.edu

1 Introduction

In an influential paper, Derek Neal (2004) used data from the National Longitudinal Survey of Youth, 1979 (NLSY79) to show that in 1990, even though the employment rates of young (aged 25-33) white and black women were quite similar, the difference between their median log wages was substantially less than it would have been had all white and black women in that age group been employed. The pattern of selection into employment across the distribution accounted for the fact that the observed median log wage gap between the two groups of women substantially underestimated the true, selection-corrected gap. In particular, black women with relatively little education, typically single mothers receiving government aid, were underrepresented among the employed while at the same time white college graduates were less likely to work than were black college graduates because the white college graduates were more likely to stay at home to raise children with support from high-income husbands.

The pattern of selection into employment in 1990 was driven to a considerable extent by government aid programs, and there have been substantial changes in those programs since that time. As a result of welfare reform, more transfers now go to workers and fewer to nonworkers, and more transfers go to married couples and fewer to single mothers (Moffitt and Scholz 2010). Given these changes, and other developments in the labor market since 1990, it is natural to ask how the observed median log wage gap and the selection-corrected median log wage gap among young women have changed since Neal's (2004) analysis.

In this paper, we use data from the 2011 wave of the National Longitudinal Survey of Youth, 1997 (NLSY97) to address these questions. The individuals in the NLSY97 sample were aged 12 - 17 at the time of their first interview in 1997; thus, the women interviewed in 2011 were aged 26 - 31 at that time. To make our 1990 sample comparable to our 2011 sample, we redo Neal's (2004) analysis looking only at the 26 - 31 year olds in the NLSY79. Our results for 1990 using women in the restricted age range line up quite well with Neal's (2004) results.

We then reestimate the median regressions presented in Neal (2004) using the data from the NLSY97. We find that the raw (with no controls other than race) median log wage gap between young white and young black women in 2011 is 0.25 log wage points, which is considerably larger than the raw gap in 1990 of 0.17 log wage points. When we add in marginally attached women,

1

i.e., those who didn't report a wage in the survey year, but did report a wage in one of the previous four years, the estimated gap at the median increases in both years, but is still higher in 2011. (It goes from 0.21 log wage points in 1990 to 0.28 log wage points in 2011.) When we adjust for selection using the same approach as Neal (2004), the gap rises further in both years, although it increases by more in 1990. This adjustment primarily entails imputing wages for low-educated women who had government aid in the survey year and the four previous years and had no market work or spousal support in those years. Adding imputed wages for these women has a similar effect in 1990 and 2011; it raises the median log wage gap. That is, the differential pattern of selection into employment for young black versus young white women is essentially the same in 2011 as it was in 1990. Despite the change in the welfare programs, these imputations are predominantly adding black women who would have earned low wages in both years.

When we add controls for education and potential experience, a different picture emerges. As Neal (2004) found, controlling for potential experience and education did not have much effect on the median log wage gap in 1990. We find that it stays roughly the same at 0.17 log wage points. When we adjust for selection using wage imputations like those in Neal (2004) and we control for the human capital variables, the estimated median log wage gap in 1990 rises only slightly to 0.19, much below the 0.25 that we found adjusting for selection without controlling for these variables. This is consistent with Neal (2004) and implies that a large part of the selection is accounted for by the observables, i.e., potential experience and education. The results are not the same for 2011. When we control for the human capital variables, the median log wage gap among those who reported a wage in 2011 falls from 0.25 to 0.14. This is due to the considerable increase in college graduates among the young white women as well as to the fact that the employment rate increased for these women and fell for young black college graduates. This may be the result of a differential effect of the recession in 2011. When we adjust for selection as well as controlling for human capital variables, the median log wage gap remains essentially the same rather than rising as it did for the 1990 data. In sum, once we control for the most basic human capital variables, the estimated median log wage gap among women is somewhat lower in 2011 than it was in 1990. In addition, there is no evidence of selection in the 2011 data, other than what is explained by education and potential experience.

In the next section, we discuss the data. In Section 3, we then redo the

analysis of the median log wage gaps in Neal (2004), restricting attention to women aged 26 - 31. We compare these results with the corresponding estimates for 2011. In Section 4, we use Neal's imputation rules and estimate the median regressions for the restricted 1990 sample and the 2011 sample. In the final section, we conclude.

2   Data

The NLSY79 and NLSY97 are panels designed to track the transition of young people from school to work. The NLSY79 was designed to be representative of the US population aged 14 - 21 as of the start of 1979 while the NLSY97 was designed to be representative of the US population aged 12 - 16 at the end of 1996. There has been considerable discussion about the representativeness of the NLSY97 and about the comparability of the two panels. There are three major issues. First, questions have been raised about the response rate to the NLSY97 – fewer families reported children in the 12 - 17 age range than expected. Second, the attrition rate was substantially higher in the NLSY97 than in the NLSY79. Finally, there are questions about the comparability of the aptitude tests administered in the two surveys. The respondents in the NLSY97 took these tests at a much younger age than did their NLSY79 counterparts, and the test format changed from pencil and paper in the NLSY79 to computerized in the NLSY97. See Altonji et al. (2012) for a clear and detailed discussion of these and other issues surrounding the two panels. Although the representativeness of the NLSY97 and the comparability of the NLSY79 and NLSY97 are important issues, what matters for our analysis is whether these potential problems affect young black and white women differently, and there is no strong evidence that this is the case. The fact that the labor market was particularly weak in 2011 is also an issue, but again what matters for our analysis is whether the soft labor market affected the two groups of women differently. Unemployment rates rose for both young black and young white women, but rose proportionately more for the young white women.[1]

We follow the procedures described in the Appendix to Neal (2004) to construct the datasets used in our analysis. Since the women in the sample

---

[1]The unemployment rate for black women aged 25 - 34 was 11.75 in 1990 and 16.53 in 2011. The corresponding unemployment rates for white women were 4.50 in 1990 and 7.65 in 2011.

that we use from the NLSY97 are aged 26 - 31 in 2011, we have reconstructed the sample for 1990 including only this age group.

Table 1 presents some basic descriptive statistics for the 1990 and 2011 samples. The only striking change between 1990 and 2011 is the increase in educational attainment, which in turn explains the lower level of experience among women in 2011. The percentage of young black and young white women with a college degree or more essentially doubled between 1990 and 2011, although the change for young white women started from a much higher base. The fraction of married women barely changed over this time period, and this result holds even if we look at the black and white women separately. Similarly, there is little change in the fraction of women whose youngest child is between ages 5 - 11 and between ages 11-18. There is a drop in the fraction of both black and white women with a youngest child of age 5 or less.

Table 1: Descriptive Statistics						
	1990			2011		
	All	Black	White	All	Black	White
Total number of observations	2890	1042	1848	2172	916	1256
Percent		0.38	0.62		0.41	0.59
Age at date of interview						
26 years old	0.17	0.17	0.17	0.04	0.03	0.04
27 years old	0.17	0.17	0.17	0.20	0.20	0.21
28 years old	0.17	0.19	0.16	0.20	0.21	0.19
29 years old	0.18	0.17	0.19	0.20	0.19	0.21
30 years old	0.17	0.16	0.18	0.22	0.20	0.23
31 years old	0.14	0.15	0.13	0.14	0.17	0.13
Married	0.53	0.33	0.64	0.51	0.31	0.65
Highest Degree Obtained						
High school or less	0.57	0.60	0.54	0.35	0.43	0.30
Some College	0.23	0.26	0.21	0.26	0.32	0.21
College +	0.21	0.13	0.25	0.39	0.25	0.49
Age of youngest child						
Less than or equal to 5 years old	0.45	0.47	0.44	0.38	0.39	0.36
Between 5 and 11 years old	0.18	0.24	0.14	0.20	0.28	0.15
Between 11 and 18 years old	0.02	0.03	0.01	0.03	0.05	0.01
Experience (in 1000 hours)	13.3	11.0	14.5	11.7	10.8	12.4

Table 2 shows the fraction of women aged 26 to 31 who worked in the year previous to the survey by race and educational attainment. The total number of women in the category is given in the parentheses. There is a reduction in the employment rates among young black women across all levels of education and among young white women except for those with a college degree or more. This general decline in employment rates reflects the weak state of the labor market in 2011. Note, however, that there was a slight increase in the overall employment rate among young white women between 1990 and 2011, which reflects the higher employment rate for the

highly educated young white women. Thus, in 2011, there was a substantial gap in employment rates between young black and white women; in 1990, this gap was much smaller.

Table 2: Employment Rates by Race and Education				
	High School or less	Some College	College+	All levels
1990				
Black	0.70	0.87	0.94	0.77
	(630)	(274)	(138)	(1042)
White	0.76	0.86	0.89	0.81
	(1005)	(381)	(462)	(1848)
2011				
Black	0.59	0.80	0.85	0.72
	(394)	(291)	(231)	(916)
White	0.64	0.82	0.92	0.82
	(371)	(267)	(626)	(1264)

Table 3 gives log wages by race and educational attainment for the young women in the two years. Looking at the last column, the median log wage gap increased from 17 to 25 percent between 1990 and 2011, while the overall mean log wage gap increased somewhat less rising from 17 to 20 percent over the 21 year period. Looking at education levels separately, the mean log wage gap decreased substantially for those with the lowest education, was unchanged for the intermediate group, and fell slightly for the highest education group. The increase in the overall log wage gap reflects a change in the educational composition of the sample.

	HS or less	Some College	College+	All levels	All levels
	Mean	Mean	Mean	Mean	Median
1990					
Black	1.74	1.99	2.27	1.90	1.91
	(0.44)	(0.44)	(0.38)	(0.47)	(0.69)
White	1.89	2.10	2.40	2.08	2.08
	(0.48)	(0.51)	(0.47)	(0.53)	(0.53)
	-0.16	-0.11	-0.13	-0.17	-0.17
2011					
Black	2.36	2.46	2.81	2.53	2.45
	(0.44)	(0.40)	(0.44)	(0.46)	(0.49)
White	2.42	2.57	2.92	2.73	2.71
	(0.49)	(0.49)	(0.51)	(0.55)	(0.51)
	-0.07	-0.11	-0.11	-0.20	-0.25

Table 3: Wage Rates by Race and Education

Table 4 reports years of government aid and spousal support for both the employed and the nonemployed. These data were used by Neal (2004) to explain the pattern of selection in 1990. Government aid comprises AFDC, Food Stamps and SSI in 1990 and TANF, Food Stamps, and SSI in 2011. Years of government aid and of spousal income refer to the years 1986 - 1990 and 2007 - 2011, respectively. Among young black women, dependence on government aid went down among those who were not employed, while receipt of aid increased among the employed. This reflects the change in the welfare programs from AFDC to TANF and the increase of participation in Food Stamps[2]. This is further evidenced by the increase in the fraction of employed young black women with 5 years of government support. This rose from 9 percent to 35 percent. The fraction of nonemployed young black women with 5 years of government support fell from 50 percent to 30 percent from 1990 to 2011. Among young white women, receipt of government aid went up for both the employed (from 1 percent to 3 percent) and nonemployed (from 12 percent to 18 percent) with a larger increase for those who were not employed. Overall, the fraction of these women receiving 5 years of government support

---

[2]Participation in Food Stamps across the entire population more than doubled between 1990 and 2011 from about 20 million to over 40 million.

was considerably less than for young black women. Years of spousal income increased significantly for nonemployed young black women, but fell slightly among the employed. Similarly, the fraction of employed young black women with 5 years of spousal income fell from 12 percent to 9 percent, while the fraction of nonemployed young black women with 5 years of spousal income rose from 5 percent to 12 percent. For young white women, years of spousal income increased whether employed or not. The fraction of employed young white women with 5 years of spousal income was steady at 30 percent while the fraction of nonemployed young white women with 5 years of spousal income rose from 39 to 41 percent. Overall, the fraction of young white women with 5 years of spousal support was considerably higher than for young black women.

	Years of Gov't Aid	Years of Spousal Income	Fraction w/ 5 yrs of aid	Fraction w/ 5 yrs of Spousal Income	Years of School	No of Obs
Table 4: Receipt of Government Aid and Spousal Income by Race and Employment Status						
1990						
Black Women						
employed	0.99	1.42	0.09	0.12	13.06	810
	(1.64)	(1.86)			(1.93)	
nonemployed	3.58	0.63	0.50	0.05	11.52	232
	(1.99)	(1.38)			(1.97)	
White Women						
employed	0.30	2.51	0.01	0.30	13.42	1501
	(0.91)	(2.07)			(2.28)	
nonemployed	1.09	3.00	0.12	0.39	12.37	347
	(1.79)	(2.03)			(2.29)	
2011						
Black						
employed	1.76	1.35	0.35	0.09	13.92	663
	(1.90)	(1.74)			(2.61)	
nonemployed	3.02	1.33	0.30	0.12	12.20	253
	(1.97)	(1.78)			(2.52)	
White						
employed	0.48	2.76	0.03	0.30	15.12	1031
	(1.16)	(1.97)			(2.64)	
nonemployed	1.79	3.15	0.18	0.41	12.78	225
	(1.99)	(1.98)			(2.78)	

3 Median Regressions

Neal (2004) estimated four median regressions before imputing wages to adjust for sample selection. In this section, we replicate these median regressions using our restricted sample (26 - 31 year olds) for 1990 and 2011. In a simple median regression with only age held constant, Neal found a gap of 0.18 log wage points between the young white and black women who reported a wage in the 1990 interview. This regression included 79% of the black women and 83% of the white women in the NLSY79 sample. For women with no reported wage in the 1990 interview, Neal then expanded the sample size for the regression by using the wage reported in the interview between 1988 and 1992 that occurred closest to January 1990. This second median regression included 90% of the black women and 94% of the white women in the sample. Including these more marginally attached women increased the gap to 0.21 log wage points.

Table 5 presents the results of simple median log wage regressions for women aged 26 - 31. The regressions presented in columns 1 and 2 include 78% of the black women and 81% of the white women in the 1990 sample and 72% of the black women and 81% of the white women in the 2011 sample. The first and third columns show the estimated coefficient on a black dummy from a median regression with no other covariates, while the second and fourth columns show results from a median regression that controls for race, potential experience, and education. Columns 1 and 3 differ slightly from those in Neal (2004) which also included a control for age.[3] Columns 2 and 4 exactly replicate the regressions reported in Neal (2004), which include potential experience and education. As in Neal (2004), we use indicator variables for potential experience. In column 1 with only a control for race, the median log wage gap rose from 0.174 to 0.254 between 1990 and 2011. The gaps reported in these simple median regressions are the raw log wage gaps at the median for the two years that were reported in Table 3. After controlling for potential experience and education in column 2, however, the estimated median log wage gap falls from 0.169 to 0.143. Note that controlling for potential experience and education has a small negative effect

---

[3]We ran these regressions with the age variable and the coefficients on the race dummy were very similar. This is not surprising as the age distributions across the two groups of women are essentially the same. We did not show these because we use potential experience in columns 2 and 4 (as Neal did) and felt it was unappealing to use age in one column and potential experience in the other.

on the median log wage gap in 1990, but it has a substantial negative effect in 2011. This latter effect reflects the large increase in the fraction of employed white women with college degrees.

Table 5: Median Log Wage Regressions				
	Working in year		Working in at least one of last 5 yrs.	
1990				
Black	-0.174	-0.169	-0.212	-0.163
	(0.026)	(0.024)	(0.029)	(0.026)
Schooling	–	0.107		0.113
		(0.005)		(0.006)
Number of observations	2311	2311	2694	2694
2011				
Black	-0.254	-0.143	-0.283	-0.161
	(0.029)	(0.028)	(0.034)	(0.023)
Schooling	–	0.081	–	0.077
		(0.005)		(0.004)
Number of observations	1694	1694	2053	2053

Since we are using data from the NLSY97 from 2011, we cannot assign wages to women who did not report a wage in 2011 by using wages from 2012 or 2013. Instead, we assign wages to these women based on the wage, if any, reported in the closest year prior to 2011 in the sample frame, i.e., in the period 2007 - 2010. For symmetry, we do the same for women in the NLSY79; i.e., we assign wages to women who did not report a wage in 1990 by using the wage, if any, earned in the closest year prior to 1990 in the sample frame, i.e., in the period 1986 - 1989. This second set of regressions includes 90% of the black women and 95% of the white women in the 1990 sample and 93% of the black women and 96% of the white women in the 2011 sample. As was found by Neal (2004), expanding the sample causes the

median log wage gap to rise. With this expanded sample, we find that the median log wage gap rises from 0.212 to 0.283 between 1990 and 2011 if we do not control for potential experience and education. After controlling for these human capital variables, the median log wage gap is 0.163 in 1990 and 0.161 in 2011, i.e., there is no significant change in the log wage gap among the expanded sample between 1990 and 2011.

4 Controlling for Sample Selection

We next control for sample selection using Neal's imputation method. Since he was using median regression, Neal (2004) assigned arbitrarily low wages to women who did not report a wage in any year during the period 1988 - 1992 and who would likely have earned less than the median (Group A) and arbitrarily high wages to women who did not report a wage for this period but would have likely earned more than the median (Group B). In particular, he used the following criteria to define Group A: (i) no wage reported for the period 1988 - 1992, (ii) no post-secondary education, (iii) government aid in every year from 1988 - 1992, and (iv) no spousal support over the period 1988 - 1992. Group B consisted of women who met the following criteria: (i) no wage reported for the period 1988 - 1992, (ii) at least a high school education, and (iii) substantial support from their spouse. Neal (2004) used two definitions of "substantial spousal support." The first required a woman to have a spouse with *average* earnings over the period 1988 - 1992 at the 90th percentile or above in the earnings distribution of men 25 - 35 of the race of the woman; the second definition was the same, except that the 75th percentile was used as the cutoff. Neal (2004) considered three different imputation rules. The first assigned $1 to all women in Group A. The second assigned $1 to all women in Group A and $30 to all women in Group B using the 90th percentile definition of substantial spousal support. The third assigned $1 to all women in Group A and $30 to all women in Group B using the 75th percentile definition of substantial spousal support. The median regression holding age constant was then run including these imputed wages. This regression included observed or imputed wages for 97% of the black women and 95% of the white women in the sample. The estimated coefficient on the black dummy in the median regression rose (in absolute value) from 0.21 to 0.26. Most of the increase (from 0.21 to 0.25) came from imputing wages to the women in Group A, mostly black women. A second set of median regressions, this time controlling for education and potential experience, was also run. The basic message from this second set of results is similar, i.e., the median log wage gap increases.

We estimate median log wage regressions for 1990 and 2011 for women aged 26 - 31 using essentially the same three imputations. Since we do not have data from 2012 or 2013, we use the years 1986 - 1990 and 2007 - 2011 to implement the criteria used to define Groups A and B. When we add all the imputed wages, 96% of the black women and 97% of the white women

are included in the 1990 regressions and 95% of the black women and 97% of the white women are included in the 2011 regressions.

The results are shown in Table 6. The median regressions are quite similar across the three imputations. This reflects the fact that when adjusting for selection, almost all of the action comes from adding Group A since there are relatively few women in Group B. Including the wage imputations increases the raw median log wage gaps significantly in both years. Comparing column 1 in Table 5 with column 5 in Table 6, the median log wage gap rose from 0.174 to 0.256 in 1990 and from 0.254 to 0.315 in 2011. That is, after we account for the nonparticipants with imputed wages, there is a larger gap between young white and young black women in 2011 than there was in 1990. The major factor responsible for this differential selection bias is education. Young black women with a high school degree or less were less likely to be employed than were their white counterparts in both 1990 and 2011. Young black women with a college degree or more were more likely to be employed than were their white counterparts in 1990 but less likely to be employed than college-educated white women in 2011. When we control for potential experience and education, adjusting for selection leads to a small increase in the estimated median log wage gap in 1990 but has virtually no effect on the gap in 2011. Once we control for potential experience and education, the median log wage gap in 1990 goes from 0.169 (Table 5, column 2) to 0.194 (Table 6, column 6) when we impute wages to the women in Groups A and B. That is, once we control for potential experience and education, there appears to be almost no effect of correcting for selection in 2011, whereas in 1990 correcting for selection results in an increase in the median log wage gap. In short, when we do not control for potential experience and education, we find that the median log wage gap among those who actually worked *increased* substantially between 1990 to 2011 (Table 5) as did the median log wage gap that we would have expected to see if all women had been working (Table 6). Once we control for potential experience and education, we find a lower median log wage gap (as did Neal (2004)) and we find that this decreased between 1990 and 2011. When we control for selection, the gap increases for 1990, as Neal (2004) found, but is unaffected in 2011.

Table 6: Median Log Wage Regressions after Wage Imputations						
	Rule 1		Rule 2		Rule 3	
1990						
Black	-0.248	-0.187	-0.249	-0.189	-0.256	-0.194
	(0.031)	(0.025)	(0.034)	(0.025)	(0.034)	(0.022)
Schooling	–	0.123	–	0.124	–	0.125
		(0.006)		(0.009)		(0.008)
Number of observations	2772	2772	2782	2782	2792	2791
2011						
Black	-0.312	-0.136	-0.312	-0.141	-0.315	-0.144
	(0.036)	(0.021)	(0.036)	(0.022)	(0.035)	(0.020)
Schooling	–	0.104	–	0.106	–	0.105
		(0.008)		(0.008)		(0.008)
Number of observations	2082	2082	2085	2085	2092	2092

As a robustness check, we construct bounds for the median log wage gaps. We construct a lower bound on the median log wage gap by assuming that all the young black women who were not included in the regressions reported in Table 6 had wages above the median and all the young white women who were not included in those regressions had wages below the median. This, of course, reduces the raw median log wage gaps, but, as shown in column 1 of Table 7, the median log wage gap in 2011 is still greater than that of 1990, as was the case in Tables 5 and 6. When we control for potential experience and education, the median log wage gap in 2011 is still less than that of 1990. This is also consistent with the pattern observed in Tables 5 and 6. We construct an upper bound by assuming that all the young black women who were not included in the Table 6 regressions had wages below the median and all young white women who were not included in those regressions had wages above the median. As shown in column 3 of Table 7, this naturally leads to higher raw median log wage gaps in both years, but the raw gap in 2011 is again greater than that in 1990. Finally, in column 4 of Table 7, we show the upper bounds on the median log wage gaps after adjusting for the education and potential experience. This reduces the gap in both years, but the effect is larger in 2011, leading to a lower median log wage gap in 2011.

	Lower Bound		Upper Bound	
Table 7: Lower and Upper Bounds on Estimated Gaps				
1990				
Black	-0.196	-0.138	-0.316	-0.233
	(0.025)	(0.028)	(0.034)	(0.021)
Schooling	–	0.134		0.123
		(0.008)		(0.010)
Number of observations	2890	2890	2890	2890
2011				
Black	-0.249	-0.105	-0.354	-0.197
	(0.031)	(0.024)	(0.026)	(0.025)
Schooling	–	0.104	–	0.103
		(0.009)		(0.009)
Number of observations	2172	2172	2172	2172

5   Conclusion

In sum, we replicate the analysis of Neal (2004) for 1990. We find, as he did, that when we do not control for human capital variables, the observed median log wage gap between young white women and young black women underestimates the true, selection corrected gap, i.e., the gap had all these women been working. This selection effect reflects the fact that relatively more young black women than young white women who would be expected to be at the low end of the wage distribution were nonparticipants in 1990. This is likely due to the fact that a disproportionate number of young black women participated in AFDC, which discouraged labor force participation. Young white women who did not work in 1990 were less skewed toward the bottom of the wage distribution. We reestimated the basic median regression using the NLSY97 data for 2011. Given that the welfare program after 1990, TANF, encouraged recipients to work and given that a disproportionate number of recipients were black, we might have expected that the gap between young white women and young black women would have increased and that the selection correction would have made less difference in 2011. We find that the median log wage gap did in fact increase between 1990 and 2011, but we also find that the selection correction is still large and positive in 2011. That

is, we find that the observed median log wage gap in 2011 still underestimates the gap that would occur if all the women had been working. The reason for this is that there was a significant increase in the education levels of young women, which was concentrated among young white women.

When we adjust for potential experience and education, the results are quite different. Looking first at the women who reported a wage in the survey year, in 1990, controlling for these basic human capital variables has a neglible effect on the median log wage gap. However, in 2011, adding controls for education and potential experience substantially reduces the median log wage gap. Indeed, after controlling for these variables, the gap in 2011 is less than in 1990. When we adjust for selection, the median log wage gap increases in 1990, but, in 2011, the effect of the selection adjustment is negligible. The bottom line is that, even after adjusting for selection, the median log wage gap between young black and young white women increased between 1990 and 2011, but once we control for basic human capital variables (education and potential experience), this result is reversed.

References

[1] Altonji, Joseph, Prashant Bharadwaj and Fabian Lange (2012), "Changes in the Distribution of Characteristics for American Youth: Implications for Adult Outcomes," *Journal of Labor Economics*.

[2] Moffitt, Robert and Karl Scholz (2010) "Trends in the Level and Distribution of Income Support," in Jeffrey Brown, ed., *Tax Policy and the Economy*, NBER 2010.

[3] Neal, Derek (2004) "The Measured Black-White Wage Gap among Women is Too Small," *Journal of Political Economy*, 112(1, pt.2), S1-S28.