The Racial Wage Gap in the United States

Kai Liao, Yuchen Xin, Shuhan Zeng

April 2021

1 Introduction

Race differential is a significant issues in the United States. Nearly all surveys shows that the wage, compensation, unemployment rates, non-wage compensation, and occupations are very difference across different race groups. One of the most important differentials is the wage gaps across different race groups. Economists study these issues by building economic models and using empirical evidence. Economists keep trying to add more details and get more insights from economic model. They also tried to fix the endogenous problem of their empirical evidence. We will summarize the previous research in this field, and then introduce ours empirical strategies and findings towards this problem.

1.1 Economic Theories

Many economic theories explain the discrimination in the labor market. The most frequently used models are statistical discrimination model and taste-based model. As the evidence grows, economists also find that details in the application process can also have significant influence on labor market. An important detail leads to lexicographic search models which we will elaborate further in the empirical evidence section.

A fundamental class of models is the statistical discrimination models (Arrow, 1973; Coate and Loury, 1993). These models constructed a economic world of imperfect information. It says that for African-Americans, some observable signals is not as precise as the exact same signal for other race groups (Bertrand and Mullainathan, 2004). Part of these perception is due to stereotypes and racism. It is important to note that the stereotypes and racism would decrease the return on labor capital investment of African Americans, and will lead to bigger wage gaps. Besides this feedback effect, stereotypes and racism will result in lack of employer investment on African-Americans. This self-confirming stereotypes are studied not only by economists, but also studied by psychologists. They find that the external stereotypes would generate change how people think about themselves and people tend to develop toward the direction of stereotypes.

Another class of models is the Taste-based models (Becker, 1971). Taste-based models (Becker, 1971) decompose the discrimination in the labor market to the prejudiced preference of employers, customers, or coworkers. For example, employees would like to work with people in the same race group. This preference will lead to segregated work forces. However, theoretically, this "preference" will eventually eliminate wage gaps. Another example is the taste of customers. The "taste" of customers make hiring a specific race employer more profitable. Some custom surveys shows that the race of car dealers have significant effect on the decision of buyers. The most studied preference comes from the employer side. The discussion of employer's stereotypes in the taste-based models is similar to statistical discrimination models. Black (1995), Borjas and Bronars (1989), and Bowlus and Eckstein (1998) find that, if we consider a taste-based model with search cost, the prejudice of customer and employer will put minority workers at a more disadvantage position. This negative effects will not go away in the equilibrium. The Taste-based models also imply that the wage gaps in different occupations are different, since the taste of employers, customers, or coworkers vary across jobs.

1.2 Empirical Evidence

The wage gaps across wage groups have many alternative explanations. First, it could be due to systematic labor market discrimination. Second, when hiring an employee, employers might have more information of the applicants than researchers. This information became unobservables in the empirical study. Third, the wage gaps may due to the skill gaps across wage groups when keeping other observable constant. To provide robust evidence of labor market discrimination, one must incorporate well-designed stategies.

One of the most convincing empirical evidence is from the field experiment. To identify the discrimination in the labor market, Bertrand and Mullainathan (2004) use resumes that are randomly assigned with African-American- or White-sounding names. They find that African-American names receive 33 percent less callbacks for interviews. They propose the lexicographic model to explain the labor market discrimination against African-Americans. This discrimination in the recruitment process may due to the fact that employers don't want to read through every resume. Therefore, some employers stop reading further when they see an African-American name. A limitation of their study is the external validity. Although the evidence of discrimination in the "resume-scanning" process is robust, this result cannot show whether the discrimination is resistant in the workplace and eventually lead to wage gaps.

Instead of using data of all occupations, some papers take advantage of the transparency of wages and performance in some specific occupations, espcially professional sports (Kahn, 1991). The data of performance is very rich in professional sports. These data can be used to control the bias caused by unobserved skills and productivity. However, the results from different sport leagues are very inconsistent. A more important issue of this empirical evidence is that we cannot imply the status of labor market discrimination in other occupation by just observing the sport communities.

To avoid the external validity concern, economists also use observational data to study discrimination in the labor market. The observational data must have a good evaluation of the skills of labors. Altonji et al (2012) and Neal and Johnson (1996) use the National Longitudinal Surveys of Youth (NLSY) data to study discrimination in the labor market. This is a panel data set which contains the scores of Armed Forces Qualification Test (AFQT). They find that black and Hispanic women earn more than white women, but black men earn less than white men when AFQT is held constant. Base on the research of Johnson (1996) and Altonji et al (2012), our paper uses panel data method, Heckman selection model, and decomposition model to further explore the data and solve the identification issues.

2 Model Specification

Firstly, we use linear regression and random effect model to examine the data as follow:

$$\ln wage = \alpha + \beta_1 black + \beta_2 hispanic + \gamma \mathbf{X} + \varepsilon \tag{1}$$

where $\mathbf{X}$ is the control variables. *black* and *hispanic* are racial indicators.

Nevertheless, without selection correction, the above specifications only refer to workers and don't take the people who choose not to work into account. The above wage equations may introduce a selection bias when people don't make working decisions randomly. For example, due to discrimination, African Americans may be less likely to receive callback for interviews and successfully get a job. In this case, discrimination may influence both black people's working choice and their labor market performance, and consequently the estimated wage gap with simple OLS is biased.

In order to correct this sample selection bias, we plan to use a two-stage Heckman selection model (1979) to investigate the wage gap between the while and the African-American in the United States. First, we use probit models to estimate the parameters in the selection equation. With the predicted value of the selection equation, we gain a set of Inverse Mills' ratio. Second, we plug the Inverse Mill's ratio obtained from the first step into the wage equation and finally get the wage gap between the African-American and the white. $\hat{\beta}_1$ is the estimated parameter of our interest. At the same time, for the purpose of controlling more individual-specific attributes related to both wages and race, we add more control variables in the Heckman models, such as marital status, parents' education background, and children.

In the model the equation of primary interest is defined as:

$$lnw_i = black_i\beta_1^g + hispanic_i\beta_2^g + \mathbf{X}_i\beta^g + \epsilon_i \tag{2}$$

where lnw is the logarithm of wage rate, and X_i includes all kinds of control variables. The coefficient of race dummy (β_1^g) captures the wage gap between the white and the African-American for each gender.

The selection equation as:

$$d_i^* = black_i\beta_1^g + hispanic_i\beta_2^g + \mathbf{Z}_i\gamma^g + u_i \tag{3}$$

$d_{ig} = 1$, if $d*_{ig} > 0$, otherwise $d_{ig} = 0$

where lnw_i is only observed, if $d_i = 1$. The selection equation can be interpreted as people's employment rate equation. If the offered wage rate is higher than the wage rate asked, people would choose to work and $d_i = 1$. We could only observe the wage rates of workers and the shadow prices are unobserved.

2.1 Decomposition of the race gap

In this part, we mainly decompose the race gap of different wage levels into three factors: coefficients-related gap, covariates-related gap and residuals-related gap. First of all, we let y be the outcome of our interests and x be the independent variable of our interest. a linear relationship is assumed between the quantiles of y and x similarly to OLS that assumes a linear relationship between the mean of y and x. In our research, the dependent variable is the logged wage rate and the covariates including race dummies and other control covariates. Thus, the quantile regression coefficients can be interpreted as rates of return to the different characteristics at the specified quantile of the conditional distribution.

Besides, we also assume that $q^{(\hat{\beta},x)}$ is a consistent and asymptotically normally distributed estimator of q^{θ}, which is the population's θth quantile of y.

We then use the same framework as that in paper "Decomposition of differences in distribution using quantile regression" by Blaise Melly to decompose the differences in wage distributions between the black and the white.

According to Melly(2005), the wage gap between the African-American and the white can be decomposed as follows:

We first estimate the counterfactual distribution of wages that would have prevailed among black people if the distribution of individual attributes had been as it is among white people. Thus, the difference between $q^{(\hat{\beta}^{black}, x_{black})}$ and $q^{(\hat{\beta}^{black}, x_{white})}$ is explained by changes in characteristics. Secondly, in order to separate the effects of coefficients from the effects of residuals, we estimate the distribution that would have prevailed if the median return to characteristics had been the same as among the white but the residuals had been distributed as among the black. Note that the τ th quantile of the residuals distribution conditionally on x is consistently estimated by $x(\hat{\beta}(\tau) – \hat{\beta}(0.5))$, and we define the J $\times$ 1 vector $\hat{\beta}^{(m_{black}, r_{white})}$ where its jth element is given by:

$$\hat{\beta}^{(m_{black}, r_{white})}(\tau_j) = (\hat{\beta}^{black}(0.5) + \hat{\beta}^{white}(\tau_j) - \hat{\beta}^{white}(0.5)) \tag{4}$$

.

Therefore, the difference between $q^{(\hat{\beta}^{(m_{white}, r_{black})}, x_{white})}$ and $q^{(\hat{\beta}^{black}, x_{white})}$ is due to changes in coefficients since characteristics and residuals are kept at the same level. Finally, the difference between $q^{(\hat{\beta}^{white}, x_{white})}$ and $q^{(\hat{\beta}^{(m_{white}, r_{black})}, x_{white})}$ is due to residuals. The final decomposition is the following:

$$q^{(\hat{\beta}^{wh}, x_{wh})} - q^{(\hat{\beta}^{bl}, x_{bl})} = (q^{(\hat{\beta}^{wh}, x_{wh})} - q^{(\hat{\beta}^{(m_{wh}, r_{bl})}, x_{wh})}) + (q^{(\hat{\beta}^{(m_{wh}, r_{bl})}, x_{wh})} - q^{(\hat{\beta}^{bl}, x_{wh})}) + (q^{(\hat{\beta}^{bl}, x_{wh})} - q^{(\hat{\beta}^{bl}, x_{bl})}), \tag{5}$$

where the first bracket represents the effect of changes in residuals, the second the effects of changes in (median) coefficients and the third the effects of changes in the distribution of the covariates.

3 Data and Results

3.1 Data and Descriptive Statistics

The NLSY79 Cohort is a longitudinal project that follows the lives of a sample of American youth born between 1957-64. The cohort originally included 12,686 respondents ages 14-22 when first interviewed in 1979; after two subsamples were dropped, 9,964 respondents remain in the eligible samples. Data are now available from Round 1 (1979 survey year) to Round 28 (2018 survey year).

This dataset includes the scores of Armed Forces Qualification Test (AFQT). AFQT is considered to be a good measurement of skills. It a screening test developed in 1950 by the Department of Defense to determine a person's eligibility for acceptance into U.S. military service by assessing his or her mental ability qualification. Originally consisting of 100 multiple-choice items measuring vocabulary, arithmetic, spatial relations, and mechanical ability. It is generally a good control of labor production in their future jobs.

The summary statistics of the main variables we used in this paper is given by table 1.

3.2 Is There a Discrimination in the Labor Market?

The result of pooled OLS estimation is given by table 2 and table 3. Column (1) does not include any control of skills or education. We can see that African-American and Hispanic men and women do have wage gaps when compared with other race groups. Column (2) uses education year and its square term to control skills level and productivity of labors. The results are vary similar to column (1). Column (3) is one of our main result. We use AFQT score and its quadratic term to control skills level and productivity of labors. Although we still find that African-American men has lower wages, the effect of discrimination is less than column (1) and column (2). And

African-American women actually receive higher wages when skills level is controlled. Their wages is approximately 6% higher than the baseline race group.

To explore more from this dataset, we use panel data method to estimate the effects job market discrimination on wage. Specifically, we use two panel data methods. One is year fixed effect model in table 5. And the other is random effect model in table 6. Both methods further confirm that when AFQT is controlled, the wage of African-American men is less than the baseline race group. The differences is 0.8% by fixed effect model and 11.5% by random effect model. However, the wage of African-American women is higher than the baseline race group. The differences is 22.8% by fixed effect model and 5.6% by random effect model.

3.3 Heckman Model: Is There a Discrimination in the Labor Market?

In order to take unemployed people into account, we implement two-step Heckman models to estimate the wage gap between the black and the white.

3.3.1 First step: estimate selection equation

According to the two-step procedure introduced by Heckman (1979), the first step is to estimate the selection equation with Probit models.

As shown in Table 8 and Table 9, there is a significant labor participation gap between African Americans and white Americans. Specifically, African-American men are 15% less likely to work than white males, and African-American women are 9% less likely to work than white females. As mentioned in the theory section, the difference of labor participation between African Americans and white Americans can be caused by both taste-based discriminations and statistical discriminations: African-American have less opportunities to receive callbacks for interviews (Bertrand

and Mullainathan, 2004), and finally they are more likely to be unemployed.

When we estimate the selection equations, we add much more control variables. Most of the estimated results are not counterintuitive (full results can be seen in Table 8 and Table 9). Both men and women with higher educational attainments are more likely to work. Married people are less likely to work than unmarried people. Specifically, marriage has a much larger negative effect on women than men, because married women always need to burden heavy housework, which forces them to spend less time in the labor market. High income gained by spouses tends to encourage people to work. The reason might be that a person with high income usually has a large social network which can help their spouse find a good job. Children would encourage fathers to work and discourage mother to work, implying a traditional social norm that fathers need to take the family financial burden and mothers should spend much more time with children. People with high-level stills are more likely to work than people with low-level skills.

3.3.2 Second Step: estimate wage equation

The estimated coefficients of Probit model are used to compute the inverse mills ratio. Then we run regressions on independent variables, control variables, and inverse mills ratio. Due to discriminations, African Americans and White Americans are not evenly distributed across occupations and industries. In order to control this difference, we include occupation dummies and industry dummies in this model. Table 10 and 11 include the full results.

As shown in Table 10 and Table 11, there exists an significant wage gap between African-American men and White-American men, and an insignificant wage gap between African-American women and White-American women. With other characteristics including human capital related characteristics, region characteristics, job related characteristics, and individual characteristics the same, African-American men tend to gain 11% less than White-American men. According to

Table 11, women suffer less racial discrimination but more children penalties than men in the labor market. With other characteristics the same, women with children tend to gain 6% less than women without children. Women's wages are much more affected than family-based characteristics, such as children, spouses and marriage, rather than racial characteristics.

3.4 Decomposition of wage gap between the African-American and the white

As can be seen in table 12 and table 13, there is significant wage gap between the African-American and the white. Specifically, with other characteristics the same, the African-American men of all earning levels tend to gain 20% less than white men. Among all African-American men, median-income African-American suffer the most racial discrimination, and high-income African-American suffer the least, but the difference is not significant. Compared with African-American men, African-American women are less discriminated against. With other characteristics the same, African-American women of all earning levels tend to gain 10% less than white women. Among all African Americans, low-wage African-American women are less discriminated against. Compared with other groups, the discrimination against African-American women is not even significant.

By decomposing changes in distribution into three factors: changes in regression coefficients, changes in the distribution of covariates and residuals changes, we easily find out that most of the wage inequality between the white and the African-American can be explained by race differences in the distribution of characteristics and returns to the characteristics. The negative effect of characteristics on all Quantiles among both men and women indicates that if African-American workers' attributes had been rewarded the same with white workers, wages should have fallen. This result is consistent with discrimination theories that the external stereotypes would result in less employer investment on African Americans and then African Americans tend to receive less tenure-related training. At the same time, after African Americans recognize the discrimination against them, they may choose to invest less on human capital, which enlarges the wage gap.

The lower level of wages is also explained by the differences in coefficients, that is how workers' characteristics are rewarded. The result on coefficients suggests that African Americans tend to receive less returns on labor capital investment. Finally, the effect of the residuals is very small. It is even insignificant for men.

In sum, the observed race gap is around 20% to 25% for men and around 8% to 11% for women. The race difference in the distribution of characteristics and the race difference in coefficients make nearly equal contributions to the race gap of wages. The part due to the residuals is small.

4 Conclusion

Labor discrimination among different races has always been an important issue in the United States, in order to study the wage gaps between black and white people, we implemented panel data method, Heckman selection model, and decomposition model to explore this. OLS estimation gives us the result that African-American and Hispanic men and women do have wage gaps when compared with other race groups, yet African-American women actually receive higher wages when skills level is controlled, and the panel data method gives out a similar result as the wage for African-American women is higher than the baseline race group.

Besides, the Heckman model mainly states that African-American men have less opportunities to work and gained lower wages compared to white-American men. The race gap is much smaller for women. Marriage and children tend to have a larger influence on women's labor market performance.

Finally, the decomposition result shows that the observed race gap is around 20% to 25% for men and around 8% to 11% for women.The race difference in the distribution of characteristics and returns to characteristics are main contributors to the race gap of wages, and their contributions are nearly equal. Therefore, all of these results point to the fact that discrimination among races is still existing in the US, waiting for us to solve this problem, yet the inequality between genders should also be paid more attention. Only in this way can we really eliminate the wage gap among people of different races and genders.

References

[1] Joseph Altonji, Prashant Bharadwaj, and Fabian Lange. Changes in the characteristics of american youth: Implications for adult outcomes. *Journal of Labor Economics*, 30(4):783 – 828, 2012.

[2] Joseph Altonji and Rebecca Blank. Race and gender in the labor market. In O. Ashenfelter and D. Card, editors, *Handbook of Labor Economics*, volume 3, Part C, chapter 48, pages 3143–3259. Elsevier, 1 edition, 1999.

[3] George Borjas and Stephen Bronars. Consumer discrimination and self-employment. *Journal of Political Economy*, 97(3):581–605, 1989.

[4] Audra J . Bowlus and Zvi Eckstein. Discrimination and skill differences in an equilibrium search model*. *International Economic Review*, 43(4):1309–1345, 2002.

[5] Stephen Coate and Glenn C Loury. Will Affirmative-Action Policies Eliminate Negative Stereotypes? *American Economic Review*, 83(5):1220–1240, December 1993.

[6] James J. Heckman. Sample selection bias as a specification error. *Econometrica*, 47(1):153–161, 1979.

[7] LAWRENCE M. KAHN. Customer discrimination and affirmative action. *Economic Inquiry*, 29(3):555–571, 1991.

[8] Blaise Melly. Decomposition of differences in distribution using quantile regression. *Labour Economics*, 12(4):577–590, August 2005.

[9] Derek Neal and William Johnson. The role of premarket factors in black-white wage differences. *Journal of Political Economy*, 104(5):869–95, 1996.

[10] David Neumark. Detecting Discrimination in Audit and Correspondence Studies. *Journal of Human Resources*, 47(4):1128–1157, 2012.

Appendix

Table 1. Descriptive Statistics

Variable	Obs	Mean	Std. Dev.	Min	Max
hwage	138370	203.799	1338.618	0	100878.99
afqt	138370	44.535	28.589	1	99
black	138370	.225	.418	0	1
hispanic	138370	.161	.367	0	1
age	138370	29.738	8.192	14	51
highschool	138370	1	.015	0	1
yearschool	138199	12.987	2.332	0	20
marital	138370	.461	.498	0	1
child	138370	.872	1.127	0	10
educmother	138370	11.109	3.116	0	20
educfather	138370	11.08	3.903	0	20

Table 2. Log Wage Linear Regression, Men

	(1) lwage	(2) lwage	(3) lwage
black	-0.272***	-0.271***	-0.110***
	(-31.15)	(-31.13)	(-11.17)
hispanic	-0.135***	-0.134***	-0.00966
	(-13.57)	(-13.50)	(-0.91)
age	0.0764***	0.0764***	0.0757***
	(170.09)	(170.10)	(165.52)
highschool		0.520*	
		(2.53)	
afqt			0.00774***
			(15.81)
afqt_2			-0.0000287***
			(-5.83)
_cons	0.693***	0.172	0.410***
	(49.52)	(0.83)	(24.10)
N	88810	88810	84395

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 3. Log Wage Linear Regression, Women

	(1) lwage	(2) lwage	(3) lwage
black	-0.140***	-0.140***	0.0604***
	(-14.77)	(-14.78)	(5.68)
hispanic	-0.0454***	-0.0433***	0.112***
	(-4.12)	(-3.92)	(9.47)
age	0.0698***	0.0699***	0.0698***
	(143.97)	(144.01)	(142.91)
highschool		0.707***	
		(4.24)	
afqt			0.0118***
			(20.15)
afqt_2			-0.0000526***
			(-8.93)
_cons	0.610***	-0.0966	0.175***
	(40.34)	(-0.58)	(8.77)
N	82504	82504	79778

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 4. Log Wage Year Fixed Effects, Men

	(1) lwage	(2) lwage	(3) lwage
black	-0.0513	-0.0517	-0.008***
	(-1.62)	(-1.63)	(4.03)
hispanic	-0.120***	-0.119***	0.0174**
	(-21.08)	(-20.96)	(2.95)
age	0.0404***	0.0404***	0.0298***
	(43.47)	(43.53)	(32.01)
1979.year	0	0	0
	(.)	(.)	(.)
1980.year	-0.00692	-0.00742	0.00130
	(-0.36)	(-0.39)	(0.07)
1981.year	-0.00373	-0.00410	0.0154
	(-0.20)	(-0.22)	(0.83)
1982.year	-0.0412*	-0.0416*	-0.0137
	(-2.34)	(-2.36)	(-0.78)
1983.year	-0.0473**	-0.0479**	-0.00354
	(-2.68)	(-2.71)	(-0.20)
1984.year	-0.0471**	-0.0477**	0.0116
	(-2.64)	(-2.67)	(0.65)
1985.year	-0.0151	-0.0157	0.0564**
	(-0.83)	(-0.86)	(3.09)
1986.year	0.0151	0.0144	0.0964***
	(0.81)	(0.78)	(5.20)
1987.year	0.0740***	0.0732***	0.165***
	(3.90)	(3.86)	(8.71)
1988.year	0.0850***	0.0842***	0.197***
	(4.38)	(4.34)	(10.20)
1989.year	0.0422*	0.0414*	0.164***
	(2.14)	(2.10)	(8.31)

1990.year	0.0314	0.0305	0.164***
	(1.55)	(1.51)	(8.14)
1991.year	0.0165	0.0154	0.148***
	(0.78)	(0.73)	(7.02)
1992.year	-0.0103	-0.0113	0.127***
	(-0.47)	(-0.52)	(5.86)
1993.year	-0.00543	-0.00650	0.142***
	(-0.24)	(-0.29)	(6.43)
1994.year	-0.0248	-0.0259	0.135***
	(-1.08)	(-1.13)	(5.90)
1996.year	-0.0621**	-0.0634**	0.118***
	(-2.60)	(-2.65)	(4.94)
1998.year	-0.0881***	-0.0894***	0.115***
	(-3.50)	(-3.55)	(4.59)
2000.year	-0.121***	-0.123***	0.104***
	(-4.55)	(-4.60)	(3.91)
2002.year	-0.188***	-0.189***	0.0535
	(-6.66)	(-6.71)	(1.90)
2004.year	-0.319***	-0.320***	-0.0604*
	(-10.75)	(-10.81)	(-2.04)
2006.year	-0.387***	-0.389***	-0.105***
	(-12.49)	(-12.54)	(-3.39)
2008.year	4.969***	4.967***	5.274***
	(153.24)	(153.18)	(162.79)
highschool		0.446***	
		(3.80)	
afqt			0.00819***
			(30.17)
afqt_2			-0.0000265***
			(-9.69)
_cons	1.613***	1.166***	1.461***

	(72.92)	(9.74)	(65.43)
N	88810	88810	84395

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 5. Log Wage Year Fixed Effects, Women

	(1) lwage	(2) lwage	(3) lwage
black	0.0425	0.0424	0.228***
	(1.12)	(1.12)	(6.15)
hispanic	-0.0345***	-0.0326***	0.127***
	(-5.45)	(-5.14)	(19.33)
age	0.0249***	0.0251***	0.0139***
	(24.53)	(24.64)	(13.80)
1979.year	0	0	0
	(.)	(.)	(.)
1980.year	0.00995	0.0101	0.0225
	(0.48)	(0.49)	(1.10)
1981.year	0.0478*	0.0475*	0.0641**
	(2.34)	(2.33)	(3.19)
1982.year	0.00634	0.00607	0.0429*
	(0.33)	(0.32)	(2.27)
1983.year	0.0244	0.0238	0.0740***
	(1.27)	(1.24)	(3.90)
1984.year	0.0315	0.0312	0.0887***
	(1.62)	(1.61)	(4.63)
1985.year	0.0560**	0.0555**	0.127***
	(2.83)	(2.80)	(6.48)
1986.year	0.105***	0.104***	0.188***
	(5.21)	(5.17)	(9.46)
1987.year	0.141***	0.141***	0.241***
	(6.88)	(6.84)	(11.90)
1988.year	0.175***	0.174***	0.294***
	(8.27)	(8.24)	(14.10)
1989.year	0.152***	0.151***	0.283***
	(7.08)	(7.04)	(13.31)

1990.year	0.133***	0.132***	0.282***
	(6.05)	(6.01)	(12.93)
1991.year	0.166***	0.164***	0.301***
	(7.12)	(7.07)	(13.12)
1992.year	0.145***	0.143***	0.301***
	(6.08)	(6.02)	(12.77)
1993.year	0.121***	0.119***	0.284***
	(4.96)	(4.89)	(11.80)
1994.year	0.128***	0.126***	0.306***
	(5.10)	(5.04)	(12.38)
1996.year	0.112***	0.111***	0.313***
	(4.30)	(4.25)	(12.13)
1998.year	0.0858**	0.0844**	0.313***
	(3.13)	(3.08)	(11.55)
2000.year	0.100***	0.0979***	0.351***
	(3.45)	(3.36)	(12.17)
2002.year	0.0234	0.0211	0.298***
	(0.76)	(0.69)	(9.82)
2004.year	-0.233***	-0.236***	0.0529
	(-7.23)	(-7.31)	(1.66)
2006.year	-0.307***	-0.310***	-0.000316
	(-9.09)	(-9.17)	(-0.01)
2008.year	5.269***	5.266***	5.604***
	(149.09)	(149.03)	(160.30)
highschool		0.625***	
		(6.55)	
afqt			0.0115***
			(35.09)
afqt_2			-0.0000440***
			(-13.44)
_cons	1.684***	1.057***	1.424***

	(69.65)	(10.70)	(58.04)
N	82504	82504	79778

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 6. Log Wage Year Random Effects, Men

	(1) lwage	(2) lwage	(3) lwage
black	-0.276***	-0.276***	-0.115***
	(-23.99)	(-24.00)	(-9.72)
hispanic	-0.141***	-0.140***	-0.0136
	(-10.62)	(-10.58)	(-1.06)
age	0.0778***	0.0778***	0.0770***
	(172.90)	(172.90)	(167.96)
highschool		0.492	
		(1.86)	
afqt			0.00781***
			(13.27)
afqt_2			-0.0000298***
			(-5.02)
_cons	0.649***	0.157	0.372***
	(44.89)	(0.59)	(20.13)
N	88810	88810	84395

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 7. Log Wage Random Effects, Women

	(1) lwage	(2) lwage	(3) lwage
black	-0.151***	-0.151***	0.0564***
	(-11.97)	(-11.99)	(4.51)
hispanic	-0.0542***	-0.0518***	0.109***
	(-3.69)	(-3.53)	(7.83)
age	0.0724***	0.0724***	0.0714***
	(148.65)	(148.66)	(145.63)
highschool		0.742***	
		(3.51)	
afqt			0.0120***
			(17.57)
afqt_2			-0.0000543***
			(-7.87)
_cons	0.523***	-0.218	0.117***
	(33.33)	(-1.03)	(5.43)
N	82504	82504	79778

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 8: Probit Model Results, Men

	b	se	t	p
age	.1939397***	.005997	32.33934	0
age2	-.0030753***	.0000911	-33.77189	0
black	-.1464718***	.0189522	-7.728487	1.09e-14
hisp	-.0002181	.0226948	-.009612	.9923309
yearschool	.0189644***	.0043962	4.3138	.000016
child	.0702587***	.0174238	4.032344	.0000552
child2	-.0084547	.0044573	-1.896838	.0578494
marital	-.1184487***	.0242814	-4.878157	1.07e-06
logspousal	.0508076***	.0024495	20.74177	0
afqt	.0102511***	.0009815	10.44395	0
afqt2	-.0000713***	9.86e-06	-7.225465	4.99e-13
educmother	-.0090246**	.0033148	-2.722508	.0064788
educfather	-.0034375	.0026334	-1.305365	.1917685
urban	-.0137217	.0178276	-.7696873	.4414854
_cons	-1.770142***	.0921834	-19.20239	0

Table 9: Probit Model Results, Women

	b	se	t	p
age	.219393***	.0059912	36.61916	0
age2	-.0035437***	.0000892	-39.71329	0
black	-.0933354***	.0188574	-4.949537	7.44e-07
hisp	.038421	.0226062	1.699579	.0892101
yearschool	.0499664***	.0041627	12.00344	0
child	-.0664167***	.0140629	-4.722836	2.33e-06
child2	-.0092849**	.003491	-2.659647	.0078223
marital	-.3695536***	.0223946	-16.50192	0
logspousal	.0451647***	.0020963	21.54524	0
afqt	.0210456***	.0010256	20.51997	0
afqt2	-.0001823***	.00001	-18.15635	0
educmother	-.0038132	.0031562	-1.20818	.2269779
educfather	-.0136054***	.0023934	-5.684458	1.31e-08
urban	.0329875	.0168426	1.958574	.0501627
_cons	-2.599113***	.0922241	-28.18259	0

Table 10: Heckman Model Results, Men

	b	se	t	p
age	-.8034725***	.0136224	-58.98186	0
age2	.0124694***	.0002023	61.62466	0
black	-.1177805***	.0180317	-6.531873	6.50e-11
hisp	.0028686	.0205371	.1396765	.8889156
yearschool	.0497587***	.0041303	12.04729	0
marital	-.1628317***	.0220681	-7.378608	1.60e-13
child	.0828204***	.0152964	5.414358	6.15e-08
child2	-.0146623***	.0038983	-3.761159	.0001691
fulljob	.0191813	.0199264	.9626061	.3357452
2.occup	.1428706***	.0261023	5.473487	4.41e-08
3.occup	.0977385**	.0375675	2.601678	.0092769
4.occup	-.1567296***	.0296615	-5.283943	1.26e-07
5.occup	-.0188788	.0265666	-.7106223	.4773183
6.occup	-.1656691	.1773466	-.934154	.3502244
7.occup	-.1101545***	.0265104	-4.155149	.0000325
10.occup	-.1394837***	.0269137	-5.18263	2.19e-07
11.occup	-.1097528***	.0299694	-3.662159	.0002501
12.occup	-.4176581	.3258849	-1.281613	.1999786
2.ind	.4151795***	.0674047	6.159506	7.30e-10
3.ind	.3796358***	.0393353	9.651283	0
4.ind	.272592***	.0375338	7.262581	3.80e-13
5.ind	.3480922***	.0407943	8.532856	0
6.ind	.0582874	.0382048	1.525658	.1270951
7.ind	.2954054***	.0488316	6.049472	1.45e-09
8.ind	.2043764***	.0402301	5.080182	3.77e-07
9.ind	-.2574086***	.0688685	-3.737682	.0001857
10.ind	.0037228	.0712457	.0522524	.9583276
11.ind	.1645512***	.0417966	3.936953	.0000825
12.ind	.2475396***	.0460648	5.373731	7.71e-08
penplan	.2035729***	.0146632	13.88325	0
logspousal	.0315384***	.0021136	14.92158	0
afqt	.0011317***	.0003348	3.38046	.0007236
educmother	.0025657	.0029083	.882211	.3776627
educfather	.0105316***	.0023467	4.487772	7.20e-06
tenure	.0003123***	.0000344	9.074676	0
urban	.1679078***	.0168149	9.985628	0
_cons	14.0873***	.2310395	60.97357	0

Table 11: Heckman Model Results, Women

	b	se	t	p
age	-.8846293***	.0151891	-58.24125	0
age2	.0136006***	.0002239	60.75475	0
black	-.0111241	.0211901	-.5249663	.5996066
hisp	.061463*	.0242764	2.531805	.0113477
yearschool	.0477364***	.00507	9.415389	0
marital	-.4493688***	.0284468	-15.79681	0
child	-.0680055***	.01625	-4.184945	.0000285
child2	.0108751*	.0042343	2.568358	.0102181
fulljob	-.0058383	.0165662	-.3524215	.7245222
2.occup	.2109008***	.0279616	7.542518	4.62e-14
3.occup	.0061972	.0395232	.1567985	.8754037
4.occup	-.0325234	.0230684	-1.409868	.1585785
5.occup	-.0555772	.0519055	-1.070739	.284287
6.occup	.0046643	.1045375	.0446181	.9644117
7.occup	-.0674795*	.0334256	-2.018798	.0435082
10.occup	-.1126385**	.0419506	-2.68503	.0072523
11.occup	-.0662892*	.028445	-2.330436	.0197831
12.occup	-.2344537**	.0898866	-2.608329	.0090985
2.ind	.3203678**	.1205667	2.657183	.0078797
3.ind	.3338624***	.0978221	3.412954	.0006426
4.ind	.1864887**	.0721998	2.582954	.0097958
5.ind	.2020118**	.0765566	2.638725	.0083219
6.ind	-.0366424	.0710117	-.5160051	.6058508
7.ind	.2081788**	.0736928	2.824957	.0047287
8.ind	.2198245**	.07481	2.938436	.0032987
9.ind	-.299405***	.0783724	-3.820285	.0001333
10.ind	-.2547868**	.0922371	-2.762302	.0057395
11.ind	.1589106*	.0715552	2.220812	.0263637
12.ind	.2799406***	.0754765	3.708976	.0002081
penplan	.1958243***	.0163265	11.99429	0
logspousal	.0540061***	.0029012	18.61531	0
afqt	.0010151*	.0004009	2.532165	.0113361
educmother	.0029425	.0033948	.8667523	.3860778
educfather	.0056689*	.0026467	2.141902	.0322013
tenure	.0003182***	.0000408	7.806718	5.77e-15
urban	.1480695***	.0200613	7.380868	1.57e-13
_cons	15.5451***	.2646278	58.74327	0

Table 12: Decomposition of Gap, Men

Component	Effects	Std. Err.	t	P	[95% Conf. Interval]	
Quantile .25						
Raw difference	-.236172	.006189	-38.16	0.000	-.248301	-.224043
Residuals	.013624	.007034	1.94	0.127	-.000162	.027411
Median	-.132062	.008931	-14.79	0.000	-.149566	-.114558
Characteristics	-.117735	.003747	-31.42	0.000	-.125079	-.11039
Quantile .5						
Raw difference	-.247797	.007175	-34.54	0.000	-.261859	-.233734
Residuals	-.00212	.005348	-0.40	0.837	-.012601	.008361
Median	-.112961	.010301	-10.97	0.000	-.133151	-.092771
Characteristics	-.132715	.004567	-29.06	0.000	-.141666	-.123765
Quantile .75						
Raw difference	-.202242	.018929	-10.68	0.000	-.239342	-.165143
Residuals	-.017773	.007676	-2.32	0.251	-.032817	-.002729
Median	-.091321	.015498	-5.89	0.000	-.121697	-.060945
Characteristics	-.093148	.007697	-12.10	0.000	-.108235	-.078062

Table 13: Decomposition of Gap, Women

Component	Effects	Std. Err.	t	P	[95% Conf. Interval]	
Quantile .25						
Raw difference	-.081015	.007692	-10.53 0.000	-.096091	-.065938	
Residuals	.016597	.007677	2.16	0.023	.001549	.031644
Median	-.068438	.00728	-9.40	0.000	-.082706	-.054169
Characteristics	-.029174	.005921	-4.93	0.000	-.040778	-.01757
Quantile .5						
Raw difference	-.107424	.007354	-14.61	0.000	-.121839	-.09301
Residuals	-.005888	.007094	-0.83	0.419	-.019792	.008015
Median	-.057825	.007292	-7.93	0.000	-.072117	-.043533
Characteristics	-.043712	.006566	-6.66	0.000	-.056581	-.030842
Quantile .75						
Raw difference	-.116216	.017913	-6.49	0.000	-.151325	-.081107
Residuals	-.017455	.011262	-1.55	0.043	-.039528	.004619
Median	-.061429	.008624	-7.12	0.000	-.078331	-.044527
Characteristics	-.037332	.011089	-3.37	0.001	-.059066	-.015599