Estimating Counterfactual Densities: An Application to Black-White Wage Differentials in the US*

Donal O'Neill[†] Olive Sweetman[‡] and Dirk Van de gaer[§]

June 2000

Abstract

In this paper we use a semi-parametric procedure, developed by DiNardo et al (1996), to estimate the distribution of the racial wage gap in the US and to examine the extent to which the forces underlying this wage gap vary throughout the distribution. In carrying out our analysis, we focus on recent work by Neal and Johnson who argue that one test score explains much of the average racial wage gap for men. Our results show that the wage differential varies significantly throughout the distribution. Furthermore Neal and Johnson's conclusion appears to be driven by forces working at the upper end of the distribution

*We would like to thank John Kennan and participants at the Dublin Labour Studies Group Workshop (December 1999), the European Society of Population Economics Annual Conference (Bonn, June 2000) the EALE/SOLE Conference (Milan, June 2000) and the Irish Economic Association annual conference (Waterford, April 2000) for helpful comments regarding this paper.

[†] National University of Ireland Maynooth

[‡] National University of Ireland, Maynooth.

[§] Ghent University, Belgium.

1 Introduction

Over the last twenty years several studies have attempted to estimate the proportion of the racial wage gap which can be attributed to discrimination. The vast majority of these papers have used the procedure developed by Blinder (1973) and Oaxaca (1973). This procedure uses the fact that the OLS regression line passes through the mean of the data to decompose the average wage gap into a component due to differences in characteristics and a residual component. The residual is often interpreted as discrimination. While this decomposition is useful for looking at the mean wage gap it tells us nothing about what is happening at other parts of the distribution. For example an average wage differential of 10% is consistent not only with a situation where all blacks are underpaid by 10% but also with one in which half of the black workers are underpaid by 20% and the other half receive the same wage as white workers. Furthermore we might be interested in knowing if the same factors account for the differential at different parts of the distribution.

In a recent study DiNardo, Fortin and Lemieux (1996) developed a semi-parametric procedure for estimating counterfactual distributions. They used this procedure to account for changes in wage inequality over time. In this paper we use this approach to estimate the distribution of the racial wage gap in the U.S., and also to examine the extent to which the forces underlying this wage gap vary throughout the distribution. Jenkins (1994) developed an alternative approach for analyzing the distribution of the wage gap. However, this procedure restricts the estimated parameters of the wage regression to be the same throughout the distribution. Recent work by Bonjour and Gerfin (1997) used a procedure developed by Donald, Green and Paarsch (1995) to look at the gender gap in Switzerland, while Gardeazabal and Ugidos (2000) used quantile regressions to examine the gender wage gap in Spain. The advantage of the approach developed by DiNardo et al over these techniques is that it requires fewer parametric restrictions. In carrying out our analyses we pay particular attention to recent work by Neal and Johnson (1996) who argue that 'one test score explains .. much of the [racial] gap for men'. We show in our paper that the wage differential varies significantly throughout the distribution. Furthermore Neal and Johnson's conclusion appears to be driven by forces working at the upper end of the distribution.

The next section explains the procedure we use to estimate the counterfactual distribution of wages. Section 3 describes the data used in our analysis, while section 4 examines the role of skills in explaining the distribution of the black-white wage gap. In section 5 we examine possible explanations of our findings and see whether these are compatible with some traditional theories of discrimination. Section 6 concludes.

2 Estimating Counterfactual Distributions

A key component of the traditional Blinder-Oaxaca wage decomposition is the estimation of the wage of the average black-worker when paid according to the white wage structure. In this paper we estimate a similar counterfactual , for the entire distribution, not just at the mean. To estimate this counterfactual distribution one could simply use the white distribution of earnings. However since white workers may not have the same distribution of characteristics as blacks, this may not be appropriate. If for example blacks were lower skilled than whites, then we should reweight the white sample so that low(high) skilled whites receive a high(low) weight. This is the intuition behind the estimator proposed by DiNardo et al (1996).

Let R denote race, $R = B$ denotes black and $R = W$ denotes white and assume that wages depend only on a composite measure of skill, which we denote by t. The relationship between skill and wages can differ for black and white workers. To find the black wage distribution when black workers are paid according to the white wage structure, we can view each individual observation as a vector (w, t, R) consisting of a wage w, a level of skill, t, and an indicator for race. The density of wages for a particular race, $f_{RR}(w)$, can be written as the integral of the density of wages conditional on skill t, $f_R(w|t)$ over the distribution of skill $f_R(t)$:

$$f_{RR}(w) = \int f_R(w \mid t) f_R(t)\, dt \tag{1}$$

Since we want to estimate counterfactual wage distributions it is important to distinguish between the wage structure of a particular race $f_R(w \mid t)$ and the skill level associated with that race, $f_R(t)$. If we assume that paying blacks the same wages as whites for each level of skill does not alter the white wage structure $f_W(w \mid t)$, then the counterfactual density $f_{WB}(w)$ is given by:

$$f_{WB}(w) = \int f_W(w \mid t) f_B(t)\,dt = \int f_W(w \mid t) f_W(t)\,\psi(t)\,dt \qquad (2)$$

where $\psi(t)$ is the re-weighting function, defined as $dF_B(t)/dF_W(t)$. $F_R(t)$ is the cumulative distribution function corresponding to the density $f_R(t)$.

$f_{WB}(w)$ represents the counterfactual density, the density that would have prevailed if black individuals (or individuals with black skills) had been paid according to the white wage structure. $f_{WW}(w)$ represents the white density, and has been defined in (1). From equation (2) we can see that the counterfactual wage density is the white density adjusted by an appropriate weighting function. The weight $\psi(t)$ represents the ratio of the probability that you have skill level t given that you are black to the probability that you have skill level t given that you are white. If for instance more white people were highly skilled then this ratio would be less than one and these people would get a low weight in the density estimation. Once the weight is calculated we can use it to estimate the counterfactual density by weighted kernel methods:

$$f_{WB}(w) = \frac{1}{n_w} \sum_{i \varepsilon wh} \frac{1}{h} \psi(t_i) K \left(\frac{w - W_i}{h} \right) \qquad (3)$$

where N_W is the set of white individuals in our sample, n_W is the number elements in this set, $K()$ is the kernel function and h is the bandwidth. Throughout this paper we assume a Gaussian Kernel and the bandwidth is chosen using Scott's optimal bandwidth for the normal density.

3 Data

In this section of the paper we use the above procedure to analyze the distribution of the black-white wage gap in the U.S.. The data we use are taken from the National Longitudinal Surveys of Youth (NLSY). The NLSY is a panel data set that follows 12,686 youths born between 1957 and 1964.

These data have been used by others (Altonji and Pierret (1997) and Neal and Johnson (1996)) to examine differences in the average wages of black and white workers. We extend these papers by examining the entire distribution of the wage gap. We focus only on male workers. We use wage data from 1993 and drop observations for which the hourly wage is less than $1 per hour or greater than $75 per hour in 1993. In 1980, NLSY respondents were administered a battery of ten tests referred to as the Armed Services Vocational Aptitude Battery. These tests examined knowledge of general science, verbal and mathematical reasoning and job related issues such as auto and mechanical comprehension. A subset of these tests, namely Word Knowledge, Paragraph Comprehension, Arithmetic Reasoning and Mathematics Knowledge are combined to form the Armed Forces Qualification Test (AFQT).[1] Neal and Johnson(1996) argue that controlling for differences in the AFQT, can explain a large proportion of the average black-white wage gap. To allow us to compare our findings with those from Neal and Johnson we restrict our sample to individuals who were 18 or younger when they took the test. We adjust the test scores for age effects and standardize the results to have mean zero and variance 1. After imposing these restrictions we are left with a sample of 1370 individuals. Summary statistics for this sample are presented in Table 1.

Table 2 reproduces the central findings from the Neal and Johnson study.[2] The table reports the results from regressions of the log hourly wage on race, age and AFQT or education . In keeping with Neal and Johnson we observe that the AFQT test score explains a large proportion of the average black-white wage gap. The average differential falls from .26 to .06 when we include a measure of test scores. The test scores themselves are also highly significant. In contrast we see that differences in education do not explain much of the black-white wage gap even though education is significant in determining earnings.

[1] For a more detailed discussion of the AFQT measure see Cawley et al (1997).

[2] The results we present here differ slightly from those reported in Table 1 of Neal and Johnson. They use pooled wage data from 1990 and 1991, whereas the wage data we use refer to 1993. Nevertheless the main conclusion of their analysis is still evident with these data.

4 Distributional Analysis of the Black-White Wage Gap

In this section of the paper we extend the analysis of Neal and Johnson by examining the entire distribution of wages. We begin by estimating the log wage gap at each centile of the wage distribution. The results are presented in Figure 1. We can see from this that there is significant variation around the mean wage gap, which is indicated by the horizontal line at .26. At the low end of the distribution, the wage gap is as low as .05. This may be attributable to the compression of the wage distribution due to the minimum wage. The gap then rises to a high of about .36 at the 30th percentile and then falls thereafter. There is no strong evidence of a 'glass ceiling', in that blacks do not appear to be disproportionately excluded from high paying jobs.[3] To test whether these differentials are significantly different, we estimated quantile regressions of wages on race for each of the deciles. The estimation procedure used provides an estimate of the entire variance-covariance matrix of the system by bootstrapping which can be used to test restrictions across equations. The null hypothesis of equal differentials at all deciles was rejected with a p-value close to zero.

In the remainder of this paper we consider the role of test scores in explaining the wage differential at different points in the distribution. To do this we use the procedure outlined in Section 2 to estimate the distribution of wages for a group of workers who are paid according to the white wage structure and like whites in all other ways except they have black test scores. To calculate the weights we discretize the test score distribution by dividing it into quintiles.[4] Since whites tend to be overrepresented in the upper portions of the test distribution this needs to be taken into account when carrying out the counterfactual. To do this we attach a lower weight to whites in this part of the distribution. The weight used is the proportion of blacks in the each quintile of the test distribution relative to the proportion of whites.

[3]For a discussion of this topic in relation to gender wage differentials see Albrecht et al (2000).

[4]In practice we can use more than one variable and these do not have to be discrete variables.

The results from this analysis are presented in Figure 2. Here bb denotes the cumulative distribution of wages for black workers and ww denotes the white distribution. wb denotes the counterfactual distribution, namely the distribution of wages of workers with the black distribution of test scores but paid according to the white wage structure. Differences between the ww and wb distributions reflect differences in test scores. Differences between wb and bb may reflect discrimination or differences in other factors. We return to the interpretation of these differences in the next section of the paper.

Following Bonjour and Gerfin(1997) we note that the horizontal difference between bb and ww measures the difference between male and female earnings at different percentiles. This can be decomposed into a wage-structure and a characteristics effect. We see from Figure 2 that the difference between the black and the white wage at the upper end of the distribution is eliminated when we take account of differences in test scores. At the lower end of the wage distribution, however, the test score component and the residual component are both important in explaining the wage gap. Figure 3 shows the proportion of the wage gap due to test scores throughout the distribution. This rises from about 45% at the 30th percentile to over 100% at the 90th percentile.[5] Thus while the work of Neal and Johnson provides evidence of the importance of test scores in accounting for the average black-white wage gap, the results in our paper highlight a significant role for the residual component at the lower end of the distribution.[6]

5 Explaining the Findings

One interpretation of the above findings is that wage discrimination (defined as unequal expected wages for equal skills - measured here by AFQT scores)

[5]We have also examined this issue parametrically using quantile regressions of wages on race and AFQT. The results of this exercise are consistent with the above findings in that the conditional racial wage differential is insignficantly different from zero at the upper end of the distribution. However, the differential is significantly different from zero at the lower end of the distribution. Furthermore the null hypothesis of equality of the the conditional wage gap throughout the distribution is rejected.

[6]We have repeated this exercise using education to measure the distribution of skills. The results are shown in Figure 4. In keeping with the finding of Neal and Johnson, we see that education has little effect on the wage gap throughout the distribution.

is lower at higher levels of the test scores. To examine this issue we reestimate the model in Table 2, allowing both the intercept and the slope to vary by race. The results are presented in Table 3. These show that black workers receive a lower wage than white workers with the same test score. However the interaction term on the slope coefficient is not significantly different from zero, which implies that the level of discrimination is constant for different values of the test score. Our data suggest, therefore, that the wage schedules for blacks and whites are parallel yet the counterfactual wage distribution converges to the actual black distribution at high wage levels. These findings can be reconciled by considering a situation where each level of skill is associated with a non-degenerate distribution of wages. If the black wage distribution conditional on test scores has fatter tails than the white distribution then the counterfactual wage distribution will cross the black wage distribution even if the average wage schedules are parallel.

To see this, consider the case where the probability distribution function of the white wage schedule is degenerate over its domain: $w^W(t) : \left[\underline{t}^W, \bar{t}^W\right] \rightarrow \Re$. Suppose that the support of the black distribution of skills is given by $\left[\underline{t}^B, \bar{t}^B\right]$ where $\bar{t}^B < \bar{t}^W$. Let $F_{WB}(w)$ denote the counterfactual distribution of white wages if white workers had black skills. Equivalently, it is the distribution of black wages if they were rewarded on the basis of the white wage structure. The highest wage a white person will obtain in the counterfactual distribution is $w^W\left(\bar{t}^B\right)$, which implies that $F_{WB}\left(w^W\left(\bar{t}^B\right)\right) = 1$. The lowest wage that a white worker will receive equals $w^W\left(\underline{t}^B\right)$, and $F_{WB}\left(w^W\left(\underline{t}^B\right)\right) = 0$. The actual cumulative distribution function of black wages is $F_{BB}(w)$: it is the result of the interaction between the black distribution of skills and the black wage structure. Assume that the black wage schedule is stochastic at the top, and has a large enough tail, such that there is some positive probability that a black worker with skills $\bar{t}^B$ will get a wage higher than $w^W\left(\bar{t}^B\right)$. As a consequence, $F_{BB}\left(w^W\left(\bar{t}^B\right)\right) < 1$. Since the expected black wage for skill level $\underline{t}^B$ is below the wage for whites, we must have that $F_{BB}\left(w^W\left(\underline{t}^B\right)\right) > 0$. The result is that the black cumulative distribution of wages and the counterfactual cumulative distribution of wages will cross. For this result to hold, it is necessary that the black probability density function has a fatter tail than the white density function at high levels of t. Figure 5 provides some evidence that this is the case. This figure provides Kernel Density estimates of the wage distributions for black and white workers at the 90th percentile of the black distribution of test scores. These graphs suggest that variation in wages about the mean may be higher for blacks than whites at least at the upper end of the test score distribution.

A question arises as to how we should view this variation around the mean. Suppose an individual with a level of skill t receives an income y, which is a random drawing from a density $f(y|t)$. Differences in stochastic elements means that people with different characteristics, such as race,are confronted with different distributions. One could argue that the differences in the distributions reflect random processes and that we should only concern ourselves with expected outcomes. However, this does not seem entirely satisfactory. An alternative would be to try and model the processes underlying our analysis and incorporate these results into the evaluation of the black-white wage gap. We need a theoretical model which incorporates the fact that blacks have on average lower test scores than whites, the fact that the slopes of the wage schedules are equal for blacks and whites but the intercept is lower for black workers and the possibility that the variance around the wage schedules is larger for blacks than whites. If we consider the traditional statistical discrimination models Lundberg and Startz(1983) and Coate and Loury(1993), we note that a steeper wage schedule induces whites to invest more in skills than blacks. While our data indicate that whites obtain more skills than blacks, there is no evidence that this is due to steeper wage profiles. A recent extension of these models by Lundberg(1991) allows for a situation in which the slopes of the average profiles are the same for blacks and whites and blacks do better at the upper end of the distribution relative to whites. However in this case the intercepts of the profiles are equal which is not consistent with our findings.

We could extend these models to allow employers to learn the true productivity of their workers over time. In the simplest possible extension we could have a two period model where a worker's productivity is completely known in the second period. This results in two different sets of wage schedules, one for each period. The first period wage schedule looks exactly like the one in the standard model. The second period wage schedule looks like the wage schedule that we found empirically. The wage schedule for blacks and whites are parallel to each other, the black wage schedule is the lowest one, and there is more variance around the mean for black workers. Hence, if, at the time our sample was taken, the productivity of the workers is known completely, then we have an explanation compatible with our findings within the theory of statistical discrimination. Since our workers are between 29 and 36 years old, this might not be entirely implausible. Models of statistical discrimination and learning by employers have recently been tested by Oettinger (1996) and Altonji and Pierret (1997). They do not find convincing evidence in favor of the hypothesis that firms use race as a source to discriminate statistically within a model of learning. There are other models that are potentially consistent with our findings, however. For instance models where social factors affect human capital accumulation can generate differences in the distribution of skills without requiring differences in the slopes of the wage profiles -see, e.g., Aigner and Cain (1977) or Lundberg and Startz (2000). Whether these models can be modified in a plausible fashion to explain the evidence on the variability of wages has not yet been examined.

6 Conclusion

This paper set out to extend the work of Neal and Johnson by looking at the black-white wage differential throughout the distribution and examining the forces behind this differential at each point of the distribution. In keeping with the findings of Neal and Johnson for the average worker, we find that the AFQT score is important for determining earnings differentials throughout the distribution. However the estimated effect is particularly pronounced at the top of the wage distribution where it accounts for all of the differential. Furthermore we argue that this result is not driven by less discrimination (defined as a situation of unequal wages for equal skills) at the upper end of the wage distribution. Rather the finding reflects greater variation in black wages around the conditional mean. This raises the question as to whether the greater risk due to the larger variability of wages around the mean should count as an additional disadvantage for black workers. To fully understand the implications of our findings it is important to develop a theoretical model which is consistent not only with average wage behavior but also the variation in wages around the mean. We view this as an important source of future work.

References

Aigner, D.J. and G.C. Cain (1977), Statistical Theories of Discrimination in Labor Markets, Industrial and Labor Relations Review, Vol.30, pp.175-187.

Albrecht, J, A. Broklund and S.Vroman (2000), "Is there a Glass Ceiling in Sweden ?," mimeo SOFI, Stockholm.

Altonji, J. and C. Pierret (1997), "Employer Learning and Statistical Discrimination," NBER Working paper #6279.

Altonji, J. and R. Blank (1999), "Race and Gender in the Labour Market," in O. Ashenfelter and D. Card (eds) Handbook of Labour Economics, Volume 3C, Elsevier Press, Amsterdam, pp. 3143-3261

Arrow, K (1973), "The Theory of Discrimination," in O. Ashenfelter and A. Rees (eds) Discrimination in Labour Markets, Princeton, N.J, Princeton University Press, pp. 3-33.

Blinder, A. (1973), "Wage Discrimination: Reduced Form and Structural variables, "Journal of Human Resources, col. 8, pp. 436-455.

Bonjour,D. and M.Gerfin (1997), "The Unequal Distribution of Unequal Pay An Empirical Analysis of the Gender Wage Gap in Switzerland," mimeo Universitat Bern

Cawley, J, K. Conneely, J. Heckman and E. Vytlacil (1997), "Cognitive Ability, Wages and Mertiocracy", in B. Devlin, S. Fienberg, D. Resnick and K.Roeder (eds) Intelligence and Success: Is it all in the Genes? Scientists Respond to the Bell Curve, New York, Springer Verlag.

Coate, S. and G. Loury (1993), Will affirmative action eliminate negative stereotypes?, American Economic Review, Vol.83, 1220-1240.

DiNardo,J, N.Fortin and T.Lemieux (1996), "Labor Market Institutions and the Distribution of Wages, 1973-1992: A Semi-parametric Approach" Econometrica, Vol 64: 5, pp.1001-1044

Donald,S, D.Green and H.Paarsch (1999), 'Differences in Earnings and Wage Distributions between Canada and the United States: an Application of a Semiparametric Estimator of Distribution functions with Covariates, forthcoming Review of Economics and Statistics.

Gardeazabal, J and A. Ugidos (2000) "Measuring the Wage Distribution gender gap at quantiles," paper presented at the ESPE2000 conference, Bonn.

Jenkins,S.P. (1994) Earnings Discrimination Measurement - A distributional approach Journal of Econometrics, 61 pp81-102

Lundberg, S. (1991), The enforcement of equal opportunity laws under imperfect information: affirmative action and alternatives, Quarterly Journal of Economics, Vol.106:1, pp.309-326.

Lundberg, S and R. Startz (1983)"Private Discrimination and social intervention in competitive labour Markets," American Economic Review, vol. 73, pp. 340-347.

Lundberg, S. and R. Startz (1998), On the persistence of racial inequality, Journal of Labor Economics, Vol.16, 292-323.

Lundberg, S. and R. Startz (2000), Inequality and race, in K. Arrow, S. Bowles and S. Durlauf (eds.) Meritocracy and Economic Inequality, Princeton N.J., Princeton University Press, pp.269-295.

Neal,D. and W. Johnson: 'The Role of Premarket Factors in Black-White Wage Differences' Journal of Political Economy Vol 104:5 pp869-895

Oaxaca,R. (1973), Male-Female Wage Differentials in Urban Labor Markets, International Economic Review 14 693-709.

Oettinger,G. (1996), 'Statistical discrimination and the early career evolution of the Black-White wage gap Journal of Labor Economics, Volume 14, 52-78.

Phelps, E (1971) "The statistical theory of racism and sexism," American Economic Review, vol. 62, pp. 659-661

Scott, D (1992) Multivariate Density Estimation Theory Practice and Visualization, Wiley, New York

Table 1: Summary Statistics.

Variable Name	Blacks	Whites
Log Wage	2.18	2.44
Standardised AFQT	-.58	.51
Highest Grade Completed	12.7	13.5
N	467	903

Table 2: Wage Regression with Race Dummy for the Intercept.

Dependent Variable is Log Weekly Wage 1993 n=1370			
Black	-.26(.03)	-.06(.03)	-.204(.025)
Age	.036(.012)	.028(.011)	.030(.011)
AFQT		.182(.013)	
$AFQT^2$		.013(.010)	
education93			.074(.001)
R^2	.07	.189	.203

Table 3:Wage Regressions with Race dummy and Race interactions for slope.

Dependent Variable is Log Weekly Wage 1993 n=1370	
Black	-.07(.03)
Age	.03(.011)
AFQT	.18 (.017)
$AFQT^2$	.01 (.014)
AFQT*Race	-.005 (.034)
$AFQT^2 * Race$	.003 (.026)

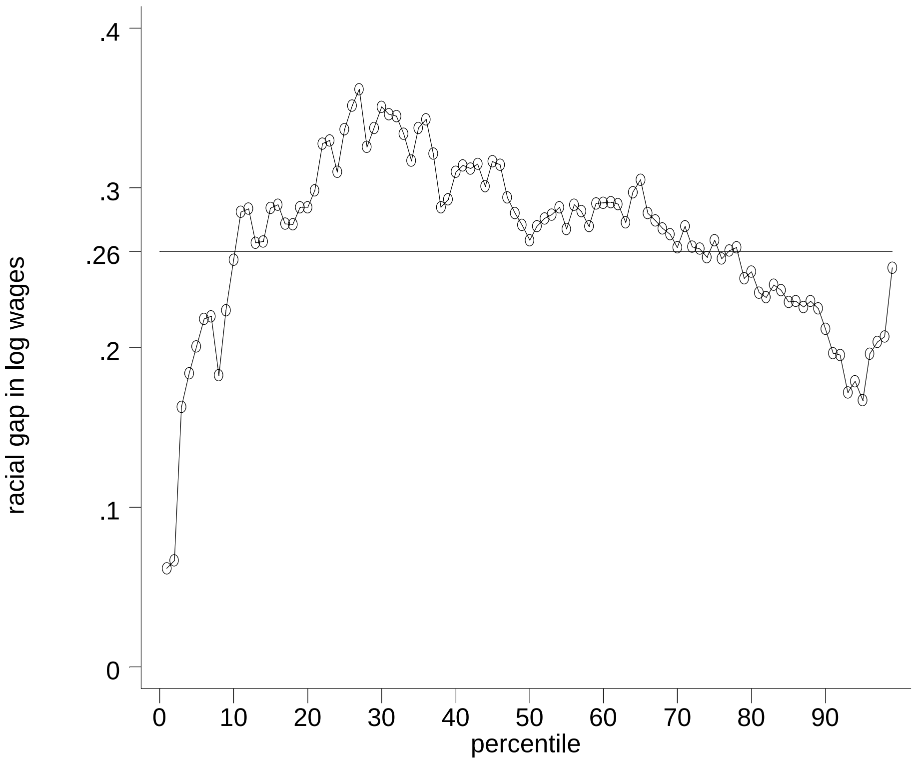

Figure 1: Racial Wage Gap at each Centile of the Distribution.

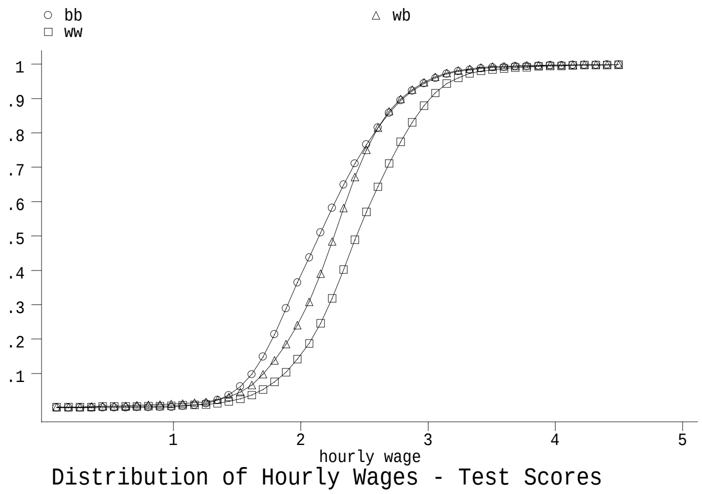

Figure 2: Counterfactual Wage Distributions controlling for AFQT scores.

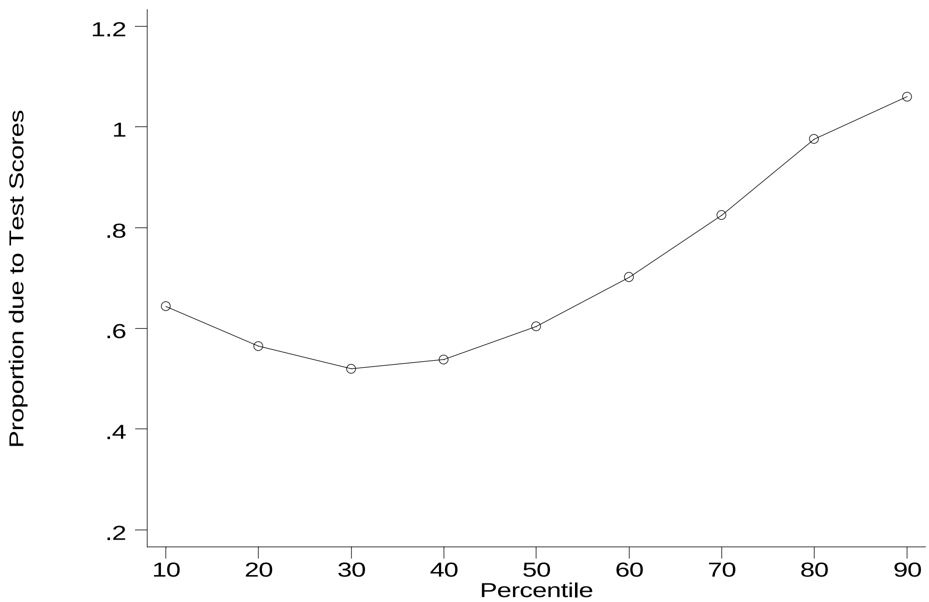

Figure 3: Proportion of the Wage Differential due to Test Scores

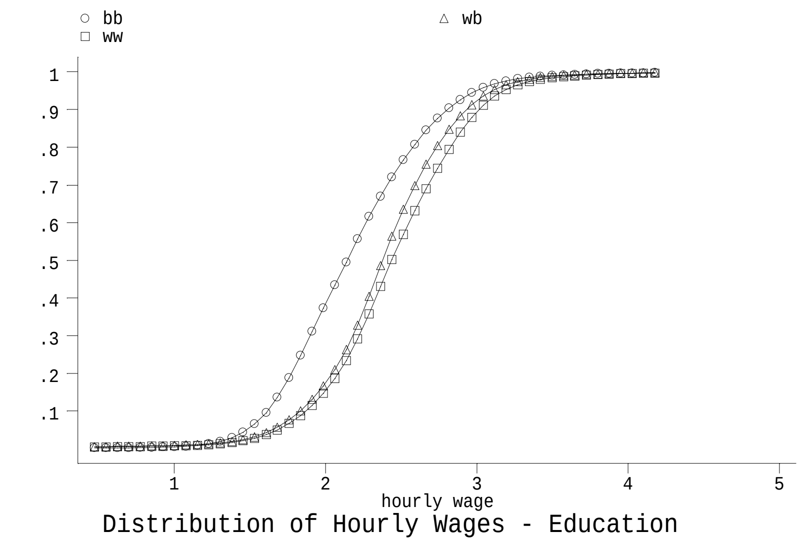

Figure 4: Counterfactual Wage Distributions controlling for Education.

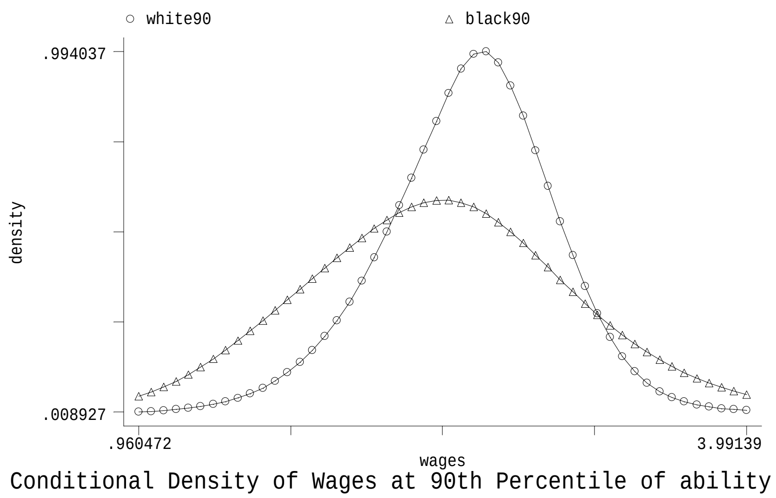

Figure 5: Conditional Wage Distributions at 90th percentile of the black test score distribution.