



# On the sources of the black–white test score gap in Europe <sup>☆</sup>

Eleonora Patacchini <sup>a</sup>, Yves Zenou <sup>b,c,\*</sup>

<sup>a</sup> Università di Roma, “La Sapienza”-Piazzale Aldo Moro, 5-00185 Roma, Italy

<sup>b</sup> Department of Economics, Stockholm University, 106 91 Stockholm, Sweden

<sup>c</sup> Research Institute of Industrial Economics (IFN), Sweden

## ARTICLE INFO

### Article history:

Received 18 December 2007

Received in revised form 14 November 2008

Accepted 18 November 2008

Available online 24 November 2008

### Keywords:

Ethnic minorities

Education

Parental involvement

Oaxaca decomposition

### JEL classification:

I21

J15

Z10

## ABSTRACT

Differences between black and white students in Britain are investigated. If black parents would invest in education as much as white parents do, the racial test score gap in mathematics and reading would be reduced by 7 and 9%.

© 2008 Elsevier B.V. All rights reserved.

## 1. Introduction

While studies on racial gaps in outcomes are pervasive in the United States (see, Neal, 2006 for a review), relatively few are conducted elsewhere. In recent years, Europe has acquired the features of a multicultural and multiethnic society. Yet, empirical analyses of racial differences in socio-economic behavior in Europe are still quite limited. This dearth of research is partly motivated by the scarcity of data on these topics. The few (and recent) existing works on the racial gap in school achievements are snapshots on a sample of students (see, e.g. Bradley and Taylor, 2004; Gillborn and Mirza, 2000; Modood, 2005) documenting some sets of correlations. A notable exception is Wilson et al. (2005) where the focus is on the evolution in test scores between different ethnic groups and whites in England. Using the information on personal and school characteristics contained in a universe dataset of state school students in England, they investigate the roles of poverty, language, school quality and teacher influence on education. They find that none of these factors fully explains the empirical evidence. Their conclusions point to the fact that other forces like ethnic differences in attitudes, for which they have no measure, should be at work.

This paper aims precisely at uncovering the importance of these factors in shaping the racial gap in Britain, focussing in particular on the role of parental involvement in the child's education. We exploit the unique characteristics of the National Child Development Study, which is a longitudinal survey containing very detailed information on cohort's members residential neighborhood, family environment and background, parental behavior, attitudes, school and education. The richness of this information provides us with a set of “non standard” variables to account for the heterogeneity of our sample and the dataset longitudinal feature allows us to tackle endogeneity issues. We document the existence of strong *parenting differences between races* and we find that a non-negligible part of the test score racial gap can be explained by parental involvement in a child's education.

Parental interest, however, can proxy for a broad set of environmental and behavioral factors. Our findings have thus to be interpreted as suggestive evidence indicating that racial differences in behavioral factors, which are usually unobserved, might be responsible for at least a part of the racial differences in academic performance. The extent of the effects might be large.

## 2. Data

Our empirical analysis is based on data from the National Child Development Study (NCDS). It is a longitudinal survey that follows all British persons who were born between the 3rd and 9th of March 1958, with follow-up surveys in 1965 at age 7 (NCDS sweep one), in

<sup>☆</sup> We are grateful to an anonymous referee, Raquel Fernandez and the participants of the conference on “The Economics of Diversity, Migration, and Culture” in Bologna, September 2006, for very helpful comments and suggestions.

\* Corresponding author. Tel.: +46 8162880; fax: +46 16159482.

E-mail address: [yves.zenou@ne.su.se](mailto:yves.zenou@ne.su.se) (Y. Zenou).

**Table 1**  
Description of variables

		Whites	Blacks
Parental interest	First principal component of responses to survey questions on parental interest	0.79 (0.47)	0.26 (0.36)
Mathematics test score	Child's age-16 standardized arithmetic test scores, NCDS3	0.23 (0.76)	-0.69 (0.12)
Reading test score	Child's age-16 standardized reading test scores, NCDS3	0.21 (0.78)	-0.64 (0.13)
Female	Dummy taking value one if the child is female, NCDS1	0.48 (0.52)	0.54 (0.54)
Birth order	Child's birth position with respect to brothers and sisters, NCDS2	1.41 (0.88)	2.46 (0.99)
Birth weight	Weight of baby in ounces at birth, NCDS1	118.03 (31.25)	100.13 (35.95)
Free school meals	Dummy taking value one if the child receives free school meals, NCDS2	0.23 (0.34)	0.18 (0.46)
Parent age	Average of parents' age at the child's age 16, NCDS1	41.09 (8.26)	37.19 (10.02)
Parent education	Average of parents' completed years of schooling, NCDS3	14.25 (2.31)	10.13 (2.02)
Parent income	Weekly net wage of father (or mother if no father figure), NCDS3	24.00 (11.12)	15.71 (12.15)
Parent employed	Dummy taking value one if both parents are working, NCDS3	0.77 (0.38)	0.58 (0.26)
Single parent	Dummy taking value one if there is no regular father or mother figure, NCDS2	0.003 (0.054)	0.09 (0.17)
Parent health problems	Dummy taking value one if at least one parent suffered from serious ill-health, including death of mother or father, NCDS2	0.05 (0.22)	0.04 (0.31)
Financial problems	Dummy taking value one if the family experienced financial difficulties, NCDS2	0.08 (0.24)	0.15 (0.25)
Domestic tensions	Dummy taking value one if the family experienced domestic tensions, NCDS2	0.11 (0.10)	0.16 (0.17)
Family size	Number of people in household, NCDS3	4.22 (1.66)	5.12 (1.88)
English language	Dummy taking value one if English is the language spoken at home, NCDS2	0.98 (0.04)	0.58 (0.17)
Accommodation tenure	Dummy taking value one if it is a owner occupied house, NCDS1	0.44 (0.10)	0.26 (0.32)
Students quality	Proportion of pupils studying for GCE and SCE O-levels, NCDS3	29.99 (16.34)	12.84 (21.12)
School composition	Percentage of pupils with non-manual fathers, 9 bands, NCDS3	4.89 (2.34)	3.01 (1.13)
Teachers quality	Number of teachers with more than 6 years of experience, 9 bands, NCDS3	1.59 (2.35)	0.78 (2.25)
High-skilled population	Census ward proportion of persons with A-levels or above, Census 1971	13.11 (4.05)	10.97 (3.04)
Black population	Census ward black residents over total residents, Census 1971	0.13 (0.14)	0.31 (0.20)
Total population	Census ward total residing population (thousands), Census 1971	59.79 (73.29)	80.12 (90.33)
Unemployment rate	Census ward unemployed over active population, Census 1971	0.05 (0.03)	0.10 (0.04)

Notes. A parent is the mother or the father or a person acting as mother or father. GCE and SCE O-levels are exams taken mainly by pupils in grammar schools, nationally the top 20% by ability. Mean values and standard deviations (in parentheses) are reported. *T*-tests for differences in means are all significant at least at the 10% significance level.

1969 at age 11 (NCDS sweep two), in 1974 at age 16 (NCDS sweep three), in 1981 at age 23 (NCDS sweep four), and in 1991 at age 33 (NCDS sweep five). For the purpose of this paper, we exploit only the first three waves, which report children test scores at school and parents' involvement during these years. This dataset contains detailed parental and child information, as well as data on family background, home environment and school quality. It also reports area of residence identifiers for cohort's members residential addresses in 1974 (NCDS sweep three), which allows us to match NCDS data with the 1971 Census data, obtaining a detailed picture of the residential neighborhood community when cohort's members were teenagers (age 13).<sup>1</sup> We distinguish between white and black children, and we consider Black Africans only (hereafter blacks).<sup>2</sup> Excluding individuals with missing or incomplete information on our target variables, we are left with 1848 individuals followed from 7 to 16 years old. Blacks are roughly 15% of the total population in our final sample (1571 whites and 277 blacks).

As with most longitudinal surveys, attrition and incomplete information is an issue in the NCDS. We gain confidence on our sample as follows. Firstly, although the NCDS does not oversample ethnic minorities and sample sizes are not large initially, 92% of the black cohort's members remain in all the three waves. Concerning the white population, we lose almost 25% of our observations from the first to the third wave. However, comparing descriptive statistics of the whole sample and our selected one, it does not appear that we lose representativeness. Secondly, we compare our sample descriptive statistics with the ones obtained by Wilson et al. (2005) in their study on school attainments of England's ethnic minorities. As described in the introduction, they use a universe dataset of state school students in England (accounting for around 93% of all pupils). We find the same

qualitative results on the levels and evolution of the black–white gap across the school years.<sup>3</sup>

The NCDS provides rich qualitative information on parental involvement in the child's education at different ages of the child. We use two alternative proxies, one that is based on the teachers and headmasters perception of parental involvement and the other that is derived from parental self-reported information.<sup>4</sup> Specifically, the school questionnaire of different NCDS sweeps (1965, 1969, 1974) contains the following question: "With regard to the child's educational progress, does the mother/father appear: over concerned about the child's progress and/or expecting too high a standard? Very interested? To show some interest? To show little or no interest?" We use the NCDS sweep one (age 7), i.e. when the child has just started school. This choice should ensure that this indicator of parental interest in a child's education is not driven by the child's schooling performance, thus *avoiding possible endogeneity problems in our empirical exercise*. Indeed, the parents' interest in his/her child education at that early stage as perceived by headmasters (based on parent's telephone calls asking information about courses and textbooks, frequency of parental visits to the school, meetings with teachers, etc...) should closely proxy parental care independently from children future schooling achievements.<sup>5</sup>

<sup>3</sup> See Patacchini and Zenou (2007) for further details.

<sup>4</sup> This selection avoids problems in the measurement of parental inputs due to the fact that parental involvement at school might not correspond to parental investment at home. Indeed, there may be ample investment at home for parents too shy and too deferential to visit the school. And there may be relatively little investment at home from parents who are very assertive at school. We use both sources of information (i.e. school and home).

<sup>5</sup> Our analysis has however also been performed using the information provided by the NCDS data on the mothers' attitude towards pregnancy in the antenatal period (namely, whether the mother who declares being a smoker prior to pregnancy quits smoking after four months of pregnancy and the mother's total number of antenatal visits without abnormality during pregnancy) as instruments for parental involvement in the child's education. Indeed, these variables should be predictors of parental care and, according to our data, they are not directly correlated with children's future school performance. The qualitative results remain unchanged, providing further support to the validity of our findings. The results of this robustness check are not reported here for brevity. They can be found in Patacchini and Zenou (2007).

<sup>1</sup> All data can be obtained from the UK Data Archive. We acknowledge the original data creators and depositors. They bear no responsibility for the analyses and interpretations presented here.

<sup>2</sup> Children with Black Africans and Black Caribbeans background have very different education trajectories. We consider only Black Africans to obtain cultural homogeneity. Small sample sizes on the other nonwhite ethnic groups prevent us to consider other interracial differences in test scores. Non-White and Non-Black African children are thus excluded from our sample.

Our alternative proxy is based on parental self-reported information on the frequency they read to the child. It is also taken when the child is of age 7. Specifically, in the NCDS sweep one (parental questionnaire) it is asked: “Does the mother/father read to, or read with, the child?” with possible answers “Yes, at least every week; Yes occasionally; Never or hardly never”.<sup>6</sup>

Because of the obvious high level of correlation between the four observed measures of parental interest,<sup>7</sup> we apply a Principal Component Analysis (PCA) to reduce the number of dimensions without much loss of information. A basic analysis, which looks for orthogonal solutions, shows a first principal component that alone accounts for roughly 93% of the total variance. This implies that the information content of our four different variables can be appropriately summarized by one derived variable. We then use this artificial variable (first principal component) in our subsequent analysis as predictor of parental interest.<sup>8</sup>

The other variables used in our regression analysis are described in Table 1. Table 1 also reports our sample descriptive statistics separately for white and black pupils.

The individual test scores have been normalized to have zero mean and a standard deviation of one for the overall sample on each of the tests (mathematics and reading). The resulting black–white gap is slightly smaller but comparable to the one registered in the US. Indeed, Table 1 shows that blacks score on average 0.92 standard deviation worse than whites in mathematics and 0.85 standard deviation worse in reading (while in the US the average difference amounts to roughly one standard deviation). Looking at our target variable, i.e. parental involvement,<sup>9</sup> Table 1 documents that black parents appear to be much less interested in their children's education than white parents.<sup>10</sup> Not surprisingly, Table 1 also confirms large differences in family, neighborhood and school quality between whites and blacks. The relative contribution of these factors in shaping the racial test score gap is then assessed in our regression exercise.

### 3. Understanding racial school performance differences

Do racial differences in rearing children really matter in explaining the interracial school achievement gap?

We start by performing a standard analysis of the racial test score gap, i.e. the estimation of a linear regression model where the individual test scores are expressed as a function of a set of regressors, including parental involvement in the child's education.<sup>11</sup> We consider the exams taken at age 16, which give the crucial test scores both for further education and drop-out teenagers. We exploit the longitudinal information contained in the NCDS to account for parental, school and neighborhood-level inputs during the school years and to avoid the endogeneity issues arising when using a contemporaneous measure of parental interest (as explained in Section 2).

<sup>6</sup> Non-response corresponds to all of these parenting variables. Missing values are reported in less than 5 % of the observations.

<sup>7</sup> Namely, the mother's and father's involvement in their child's education as reported by the teachers and headmasters (school questionnaire) and the self-reported frequency the mother and the father read to their children (parental questionnaire).

<sup>8</sup> The details of this PCA are available upon request.

<sup>9</sup> Our (derived) indicator of parental interest has been normalized to be between 0 and 1.

<sup>10</sup> Because one may argue that black parents are not very interested in education because they are not educated themselves, we also computed the inter-race differences in parental interest for high- and low-educated parents separately. Our (derived) predictor of parental interest has a lower mean value for children of low-educated parents. However, there are still large (statistically different) inter-race differences within each type of parents. This suggests that the different priorities given to education by black and white parents might in fact be related to cultural differences.

<sup>11</sup> Different variations of the test score production function (as in Todd and Wolpin, 2003; Hansen et al., 2004) have also been used. The main results remain qualitatively unchanged.

**Table 2**  
Responsiveness of test scores to covariates

Variable	(1) Dep. var.: Mathematics test scores		(2) Dep. var.: Reading test scores	
	Whites	Blacks	Whites	Blacks
Parental interest	0.1867*** (0.0035)	0.0678*** (0.0201)	0.2118*** (0.0030)	0.0715*** (0.0169)
Female	0.0122** (0.0055)	0.0430** (0.0214)	0.0152** (0.0070)	0.0689** (0.0340)
Birth order	-0.0049* (0.0028)	-0.0030 (0.0049)	-0.0086 (0.0055)	-0.0051 (0.0061)
Birth weight	0.0008** (0.0004)	0.0013*** (0.0004)	0.0002** (0.0001)	0.0007*** (0.0002)
Free school meals	-0.0022** (0.0010)	-0.0103 (0.0099)	-0.0006** (0.0003)	0.0091 (0.0090)
Parent age	-0.0215 (0.0389)	-0.0117 (0.0501)	-0.0186 (0.0190)	-0.0105 (0.0217)
Parent education	0.0780*** (0.0155)	0.0087*** (0.0026)	0.0895*** (0.0196)	0.0092*** (0.0030)
Parent income	0.0611*** (0.0102)	0.0075** (0.0033)	0.0771*** (0.0167)	0.0080*** (0.0002)
Parent employed	0.0318*** (0.0070)	0.0245** (0.0117)	0.0199** (0.0090)	0.0171** (0.0082)
Single parent	-0.1470** (0.0704)	-0.1010** (0.0425)	-0.1444** (0.0709)	-0.0998** (0.0468)
Parent health problems	-0.0565** (0.0286)	-0.0163** (0.0082)	-0.0545** (0.0239)	-0.0120** (0.0060)
Financial problems	-0.0150 (0.0162)	-0.0053 (0.0113)	-0.0174* (0.0103)	-0.0082 (0.0085)
Domestic tensions	-0.0761*** (0.0155)	-0.0029 (0.0108)	-0.0804*** (0.0095)	-0.0051* (0.0029)
Family size	-0.0029** (0.0014)	-0.0106 (0.0105)	-0.0051** (0.0026)	-0.0152 (0.0150)
English language	0.0116 (0.0131)	0.1701*** (0.0447)	0.0105 (0.0101)	0.1572*** (0.0352)
Accommodation tenure	0.0075 (0.0072)	0.0052 (0.0205)	0.0104 (0.0102)	0.0097 (0.0239)
Students quality	0.0029** (0.0013)	0.0127*** (0.0005)	0.0065** (0.0032)	0.0139*** (0.0006)
School composition	0.0056** (0.0024)	0.0327*** (0.0108)	0.0104*** (0.0029)	0.0421*** (0.0118)
Teachers quality	0.0111** (0.0051)	0.0335*** (0.0101)	0.0129** (0.0063)	0.0339** (0.0154)
High-skilled population	0.0021** (0.0010)	0.0240*** (0.0060)	0.0029** (0.0013)	0.0272*** (0.0065)
Black population	-0.0159** (0.0073)	-0.0252** (0.0122)	-0.0102** (0.0045)	-0.0197** (0.0098)
Total population	0.0004** (0.0002)	0.0007** (0.0003)	0.0002** (0.0001)	0.0006** (0.0003)
Unemployment rate	-0.1839** (0.0900)	-0.3865*** (0.1166)	-0.1409** (0.0698)	-0.3446*** (0.1051)
R <sup>2</sup>	0.4815	0.5412	0.4923	0.5500

Notes. Regional dummies are included. Standard errors in parentheses.

\*, \*\*, \*\*\* significant at 10, (5), [1] percent level respectively.

Because different covariates show a different impact for blacks and whites, we perform the analysis separately for blacks and whites.<sup>12</sup> The model has been estimated using alternative (increasing) sets of explanatory variables. The interesting result is that, although we use a large amount of information on individuals' socio-economic status, school and neighborhood quality, we are not able to completely explain the racial gap on the basis of racial differences in traditional observable variables. However, as soon as we include covariates that can be considered as proxies for a broad set of environmental and behavioral factors, such as domestic tensions, English language spoken at home, health problems of the parents or parental effort in investing in their children's education, the unexplained portion of the racial gap substantially shrinks.

<sup>12</sup> Conventional statistical tests always reject the use of cross-equation restrictions with our data. Neglecting this issue might lead to a biased estimation of the racial gap, which will be overstated (or understated) depending on whether the effect of a regressor is higher (or lower) for blacks with respect to whites.

**Table 3**  
Decomposition of racial differences in test scores

Test score gap relative to blacks		Table 3a	
		Part of gap explained by observables	Part of gap left unexplained
Whites			
Age 16	Mathematics	67.1%	32.9%
	Reading	62.2%	37.8%
		Table 3b	
		Additional contribution by group differences in parental involvement	Part of gap left unexplained
Parental interest	Mathematics	6.9%	26.0%
	Reading	9.2%	28.6%

Notes. See text for more details on calculation of the decomposition.

Table 2 shows the complete list of estimation results for the model specification which includes the more extensive set of covariates when using test scores in mathematics and reading as (alternative) dependent variables. Observe that the variables associated with statistical significant estimated coefficients are roughly the same for both races. This means that differences in school performance among blacks are driven by the same factors that drive differences among whites, thus suggesting that the black and white education production functions are not substantially different. However, the responsiveness to the different factors varies across races. For instance, black children are less responsive to changes in variables related to the family socio-economic status (e.g., parental education, income, employment) and in variables indicating difficulties at home (e.g., domestic tension, parental health problems, household financial problems) than white children. On the other hand, they are more responsive to changes in variables capturing the quality of school and residential neighborhood (e.g., students and teachers quality, school composition, percentage of neighborhood high skill population and local unemployment rate) than whites.

Let us now focus our attention on the main message of this paper, which is the finding of a possible important role played by interracial parental interest differences in explaining the racial test score gap. The estimated effect of parental interest is statistically significant both for blacks and whites and for both subjects.<sup>13</sup> In order to better quantify the contribution of parental involvement in closing the racial gap, we use the standard Oaxaca decomposition technique (Oaxaca, 1973). It allows to decompose any difference in a dependent variable between two groups into the part that is explained by different observable characteristics across groups and the proportion that is due to the same characteristics having a different effect between the two groups. In particular, the racial test score gap can be expressed as:

$$\bar{Y}^W - \bar{Y}^B = (\bar{X}^W - \bar{X}^B) \hat{\beta}^B + \bar{X}^W (\hat{\beta}^W - \hat{\beta}^B) \quad (1)$$

where  $\bar{Y}^W$  ( $\bar{Y}^B$ ) is the average test score (in mathematics or reading) of whites (blacks),  $\bar{X}^W$  ( $\bar{X}^B$ ) is a row vector of average values for the

observable characteristics of whites (blacks) and  $\hat{\beta}^W$  ( $\hat{\beta}^B$ ) is the vector of coefficient estimates for whites (blacks) from a regression analysis run separately for each group (results given in Table 2). The first term in the decomposition represents the part of the gap that is due to intergroup differences in average values of the observables, and the second term captures the portion that is “unexplained”.

We report in Table 3a the percentage of racial gap explained by observable variables and the one left unexplained when the model specification includes the more extensive set of controls (described in Table 1), with the exclusion of parental involvement. Table 3b then displays the added contribution (in terms of percentage of explained gap) stemming from the inclusion of parental involvement among the explanatory variables.

Table 3a shows a substantial proportion of the gap left unexplained, which amounts to roughly 33% and 38% for mathematics and reading test scores, respectively. According to Table 3b the additional contribution of parental involvement to explain the racial test score gap is roughly 6.9% and 9.2% for mathematics and reading respectively.<sup>14,15</sup> Such evidence thus reveals a potential important role of interracial parental involvement differences in explaining the racial test score gap, which is larger for reading than arithmetic tests. This is not unreasonable, given that math skills might be strongly related to ability and genetic characteristics whereas children's reading performance is more likely to be the result of living habits, role models and parental investment in their children education.

## References

- Bradley, S., Taylor, J., 2004. Ethnicity, educational attainment and the transition from school. *The Manchester School* 72, 317–346.
- Gillborn, D., Mirza, S., 2000. Educational inequality: Mapping race, class and gender. A synthesis of research evidence. OFSTED, London.
- Hansen, K., Heckman, J.J., Mullen, K.J., 2004. The effect of schooling and ability on achievement test scores. *Journal of Econometrics* 121, 39–98.
- Modood, T., 2005. The educational attainments of ethnic minorities in Britain. In: Loury, G.C., Modood, T., Seles, S.M. (Eds.), *Ethnicity, Social Mobility and Public Policy*. Cambridge University Press, Cambridge.
- Neal, D., 2006. Why has black–white skill convergence stopped? In: Hanushek, E., Welch, F. (Eds.), *Handbook of Economics of Education*. Elsevier, Amsterdam, pp. 511–576.
- Oaxaca, R., 1973. Male–female wage differential in urban labor markets. *International Economic Review* 14, 693–709.
- Todd, P.E., Wolpin, K.I., 2003. On the specification and estimation of the production function for cognitive achievement. *Economic Journal* 113, F3–F33.
- Patacchini, E., Zenou, Y., 2007. The racial test score gap and parental involvement in Britain. IZA Discussion Paper No. 2613. Bonn.
- Wilson, D., Burgess, S., Briggs, A., 2005. The dynamics of school attainment of England's ethnic minorities. CMPO Working Paper No. 05/130. University of Bristol.

<sup>14</sup> The use of the coefficient estimates for whites as weights in estimating the contribution from group differences in observed variables (i.e.  $\hat{\beta}^W$  instead of  $\hat{\beta}^B$  in expression (1)) produces only slightly lower estimates of the proportion of the gap explained by our predictor of parental interest (6.6 and 8.9 percent for mathematics and reading, respectively).

<sup>15</sup> If we include only our parenting variables as explanatory variables of the racial test score gap in the Oaxaca decomposition, we find that they alone account for a large portion (roughly the 40%) of the gap. This indicates, not surprisingly, that they are highly correlated with the other family background variables. Our exercise here aims at uncovering the existence and the extent of the additional contribution of these parenting variables, once the influence of traditional individual and family background observed variables has been accounted for. The motivation is that these parenting variables might also capture the effects of a broad set of environmental and behavioral factors, which are usually unobserved, that may be responsible for at least a part of the racial differences in academic performance.

<sup>13</sup> T-tests between those estimated race-specific coefficients reject the null hypothesis of equal effect, thus providing a formal test for racial differences in the importance of parental interest on children's education.