

# SCHOOL INPUTS AND EDUCATIONAL OUTCOMES IN SOUTH AFRICA\*

ANNE CASE AND ANGUS DEATON

We examine the relationship between educational inputs—primarily pupil-teacher ratios—and school outcomes in South Africa immediately before the end of apartheid government. Black households were severely limited in their residential choice under apartheid and attended schools for which funding decisions were made centrally, by White-controlled entities over which they had no control. The allocations resulted in marked disparities in average class sizes. Controlling for household background variables, we find strong and significant effects of pupil-teacher ratios on enrollment, on educational achievement, and on test scores for numeracy.

## I. INTRODUCTION

That educational inputs should be important determinants of educational outcomes is a proposition that appeals to common sense, but is nevertheless controversial in the literature both for developed and less-developed countries. Surveys by Hanushek [1986], for developed countries, and [1996], for developing countries, argue that school facilities have at best tenuous effects on outcomes, particularly on test scores. Kremer [1996] emphasizes that such a negative overall assessment of the evidence rests on Hanushek's interpretation of statistically insignificant findings as evidence *against* an effect of school quality, but argues that there is a singular absence of evidence from developing countries that the pupil-teacher ratio is an important determinant of outcomes.

One of the difficulties in estimating the impact of class size on outcomes is the potential endogeneity of school inputs. Parents who care about education may move to be close to good schools and may be willing to pay higher housing prices to do so. Parents who

\* This work reported here was originally funded under a cooperative agreement between the Institute for Policy Reform and the Agency for International Development, and later by the John D. and Catherine T. MacArthur Foundation within their network on poverty and inequality in broader perspectives. We are grateful to Harold Alderman, Orley Ashenfelter, Samuel Bowles, David Card, Carlo del Ninno, Bruce Fuller, Robert Jensen, Lawrence Katz, Dulcie Krige, Alan Krueger, Francie Lund, Pieter LeRoux, Caroline Minter Hoxby, Richard Murnane, Ron Ridker, Cecilia Rouse, T. Paul Schultz, Charles Simkins, Geert Steyn, Servaas van der Berg, Pieter van der Wolf, and seminar participants at Boston College, Cornell University, Harvard University, Princeton University, the University of Cape Town, the University of Michigan, the University of Natal Durban, and the University of the Witwatersrand for useful discussions and comments during the preparation of this paper.

care about education will typically engage in political action to increase local school quality and funding. Such parents may also spend extra effort to ensure that their children are progressing well in school. In such cases, a positive relationship between school resources and outcomes for children may be due to unobserved parental tastes for education, and it is often difficult to disentangle the effects of such tastes from those of school inputs. In spite of the difficulties, two recent papers have made significant progress. Krueger [1999] analyzes data from the Tennessee Student/Teacher Achievement Ratio experiment (the only large randomized class size experiment ever conducted in the United States) to measure the impact of class size on student test scores, and finds significant effects of class size on the test scores of young children (grades K–3). Angrist and Lavy [1999] use Maimonides' rule governing maximum class sizes in Israel to examine the relationship between class size and test scores for fifth graders there. Maimonides' rule yields highly irregular patterns in class size that are precisely mirrored in student tests scores.

In this paper we examine the relationship between educational inputs and school outcomes in South Africa immediately before the end of apartheid government, and, in doing so, we add to what is known about the impact of school quality on child outcomes. There are three features of the South African system that are particularly salient. First, Black households were severely limited in their residential choice under apartheid. Second, funding decisions for most Black schools were made centrally, by White-controlled entities on which Blacks were not represented and over which they had no control. These two features limited the two most obvious ways in which Blacks could affect the conditions under which their children were educated. Third, the allocations resulted in marked disparities in average class sizes even across areas as large as magisterial districts, with some districts averaging 20 children per teacher in Black schools, and others upwards of 80 children per teacher. The unusually large variation in pupil-teacher ratios provides an excellent opportunity to examine their effects on outcomes.<sup>1</sup>

Beginning early in the century, the White South African

1. In this paper we have little option but to follow the apartheid racial classifications of Black, Coloured, Asian, and White; we use capitals and the Anglicized spelling throughout to register the specialized use. Magisterial districts correspond roughly to counties: there are 363 in the country; the average (median) population in each is 100,000 (38,000).

government pushed for a “Bantustan” system, in which Black families were assigned to “homelands” based upon their language, regardless of where the family had previously resided. This program gained momentum in 1970, with the passage of the “Black Homeland Citizenship Act.” The South African government forced millions of Blacks into homelands, and made it extremely difficult for family members in a homeland to join a migrant working in a city or in a mine. Lack of mobility was consciously built into the South African migratory labor system. As a by-product, the system severely limited the ability of households to move to areas with better schools.

Under the apartheid regime, resources for Black schools were, with the exception of the “independent” homelands, centrally controlled. Funding for Black schools in the provinces that the White government wanted ultimately to comprise South Africa (Cape, Orange Free State, Natal, and Transvaal) was controlled centrally by a Department of Education and Training (DET), and was set apart from the bodies responsible for the education of the White population. School funds for the homelands that were slated to become independent (the so-called Self-Governing Territories (SGTs) of KwaZulu, Lebowa, Gankankulu, KwaNdebele, KaNgwane, and Qwa-Qwa) were earmarked centrally by the Minister of National Education. The “independent” homelands (Ciskei, Transkei, Bophuthatswana, and Venda) had more control over their budgets. However, most of the money these governments had to spend came through the South African Department of Foreign Affairs, which played an important role in determining how that money would be spent [Republic of South Africa 1994]. This system generated marked discrepancies in educational funding per pupil across racial groups and place of residence. Taking Blacks in the DET schools as unity, funding levels for Whites, Asians, Coloureds, Blacks in SGTs, and Blacks in homelands were, respectively, 1.85, 1.61, 1.59, 0.74, and 0.67 [South African Institute of Race Relations 1997]. Given the very limited control that the Black population had over location and resource allocation, an unusually large fraction of the variation in school resources across districts was independent of the educational choices of Black parents and their children; see Section II below for further discussion.

We examine the effects of pupil-teacher ratios and school facilities on educational outcomes, including school attendance, educational attainment, and test scores. We use data from the

South African Living Standards Survey (SALSS) that was carried out jointly by the South African Labor and Development Research Unit (SALDRU) and the World Bank in the last five months of 1993—just prior to the change of government. The survey was supplemented by a series of community questionnaires on local facilities, and by a literacy and numeracy survey administered to a subset of individuals in the base survey. We make use of both of these supplements, and of administrative data on pupil-teacher ratios by race and by magisterial district. The largest part of our paper is devoted to a national analysis of the relationship between pupil-teacher ratios and educational outcomes. Because we are not controlling for other school-based inputs, and because in South Africa, other inputs follow the supply of teachers, our purpose is not to assess the specific role of class size among other competing uses of resources, but to measure the effects of resources in general. Except when stated to the contrary, all subsequent references to the effects of pupil-teacher ratios should be understood in this sense.

Our empirical analysis shows marked effects of school quality as measured by pupil-teacher ratios, on outcomes for Black children. Controlling for household background variables—which themselves have powerful effects on outcomes, but have no effect on pupil-teacher ratios—we find strong and significant effects of pupil-teacher ratios on enrollment, on educational achievement, and on test scores for numeracy.

The paper is laid out as follows. Section II discusses the data that are used in our empirical work and provides an overview of educational inputs and outputs in South Africa, with particular focus on the distribution by race of pupil-teacher ratios, of educational facilities, and of test scores. Section III contains our empirical analysis. We start with an analysis of enrollment and educational attainment in relation to race, age, family background variables, and pupil-teacher ratios, an analysis that can be carried out for the complete sample in the main SALSS. We then present a similar analysis of test scores using the smaller sample of individuals to whom the literacy survey was administered. We also look at the role of facilities other than pupil-teacher ratios, although this requires a restriction of the sample to those living in clusters covered by the community questionnaire. Section IV is a summary and conclusion which also contains a discussion of the relationship between our results and those in the literature on school resources in developing countries.

## II. EDUCATIONAL INPUTS AND OUTCOMES IN SOUTH AFRICA: DATA AND OVERVIEW

### *II.1. Data*

There are five data sources used in this study, three associated with the SALDRU-World Bank survey, one from administrative records, and one from the 1991 census. We summarize each of these here; a more detailed discussion of the data used is provided in the Appendix. The South African Living Standards Survey (SALSS) is our main source of household and individual data, and we refer to it as SALDRU1. The survey collected data from 8848 households in 360 clusters in a nationwide survey, including what were then described as “independent” homelands. The household roster of the survey collected data on 43,974 individuals, and each of these reported (among other things) age, relationship to the head of household, and highest educational standard achieved. An education section of the questionnaire asks questions of all individuals aged from 7 to 24. More broadly, the survey collected household information, including place of residence, magisterial district, urbanization (metropolitan, urban, and rural) and the detailed expenditure and income data that can be used to compile comprehensive (if noisy) measures of total monthly income and expenditure. These last were constructed by the survey team at SALDRU and are included in the public use versions of the data.

The main SALSS was supported by a community questionnaire that was administered to “knowledgeable” local people. In principle, the community questionnaire would allow us to match school facilities to each household in SALDRU1 but, in practice, the survey organizations did not treat the community questionnaire as seriously as they did the main survey, and the information is missing for a substantial number of clusters. Two of the most important questions, on the numbers of pupils and teachers in each school, were provided for any primary (secondary) school by only 137 (113) and 156 (110) clusters, respectively. We do not know why some clusters were covered and others not, but there is a serious risk of selection bias if we simply drop the households who live in clusters without the necessary information. For the pupil-teacher ratios, where the data are poorest, we have an alternative source of data, *The Education Atlas*, which we shall describe below and link with the community information.

Our third data source (SALDRU3) contains test results on literacy and numeracy for a subset of the individuals in the main

SALSS. The fullest description of the literacy module is given in Fuller, Pillay, and Sirur [1995, pp. 13–17]. One in every six households in the SALSS was asked to participate in the literacy study, and 1340 did so. In each selected household, the aim was to have two people take a test, one person to be between the ages of thirteen and eighteen (inclusive) and one person to be older, ideally a parent of the adolescent. We have scores from 2381 individuals in 1322 households. There are 1039 individuals in the thirteen to eighteen group, and 1330 who are older than eighteen. (A few have age missing or younger than thirteen.) Of these 2369 individuals, 190 are “new” in that they do not appear in the SALSS data, something that in theory should be impossible. Of these “new” people, 63 are in the thirteen to eighteen age group, and once these are removed, we have 976 adolescents who took the test and whose full information is also contained in the SALSS files. It is this subsample to which we pay most attention in the analysis of test scores. The 976 nonnew adolescent test-takers are split 53:47 in favor of girls, which is not significantly different from equality.

The fourth data set comes from *The Education Atlas of South Africa* [Krige et al. 1994]. The Atlas presents a picture of education in South Africa in 1991; it is based on administrative data and contains a large number of maps showing educationally relevant data—fractions of pupils by race, numbers of teachers, numbers of children out of school, income levels—organized by magisterial district. The data that we use in this paper are the numbers of teachers and numbers of enrolled pupils of each racial group in each district, which we have matched to the magisterial district for each cluster in the SALSS survey. Of the 363 districts in the country, 188 show up in the survey. Of these, 104 contain a single cluster, 43 two clusters, 23 three clusters, up to one magisterial district (Johannesburg) that contains thirteen clusters. The fact that our pupil-teacher ratios are based on enrollment and not attendance is a potential problem if DET officials targeted attendance in appointing teachers; poor attendance relative to enrollment would thereby be associated with high pupil-teacher ratios. We can do no more than note the problem along with the fact that the bias runs in the opposite direction if the DET had targeted enrollment.

Because disaggregation by district is reasonably fine, and because there is a good deal of variation across districts, it is plausible that the Atlas data will give a good indication of the

pupil-teacher ratios that are relevant for the households in our survey. However, communities are often a good deal smaller than districts, and the Atlas data do not separate primary and secondary pupils and teachers, so that we have only one proxy for both. In order to assess the relevance of these data, we examine how well the magisterial district data explain the limited number of pupil-teacher ratios available from the community questionnaire (SALDRU2). Table I shows the results of regressing the SALDRU2 pupil-teacher ratios on the corresponding Atlas pupil-teacher ratios, first for primary schools, then for secondary schools, and finally for both, with the combined SALDRU2 pupil-teacher ratio computed by summing the secondary and primary teachers in the cluster's schools and dividing by the sum of the primary and secondary pupils. Because the Atlas publishes separate pupil-teacher ratios for each race, we use the community demographic

TABLE I  
RELATIONSHIP BETWEEN PUPIL-TEACHER RATIOS FROM COMMUNITY SURVEY AND  
PUPIL-TEACHER RATIOS BY MAGISTERIAL DISTRICT FROM THE EDUCATION ATLAS

	Primary pupil-teacher ratio (SALDRU2)			Secondary pupil-teacher ratio (SALDRU2)			Primary and secondary pupil-teacher ratio (SALDRU2)		
	All races	Blacks		All races	Blacks		All races	Blacks	
Constant	6.26 (2.0)	3.81 (0.9)	-3.95 (0.3)	-2.47 (0.8)	5.37 (1.1)	14.1 (1.0)	0.92 (0.4)	4.79 (1.1)	6.15 (0.5)
Atlas pupil-teacher ratio	0.87 (9.0)	1.07 (5.4)	1.11 (3.7)	1.01 (10.1)	0.61 (2.5)	0.60 (1.6)	0.95 (11.3)	0.78 (3.8)	0.83 (2.4)
Black	—	-6.17 (1.2)	—	—	8.30 (1.6)	—	—	3.21 (0.7)	—
Coloured	—	-2.30 (0.7)	—	—	-0.42 (0.2)	—	—	-0.97 (0.4)	—
Asian	—	-3.62 (0.8)	—	—	0.49 (0.1)	—	—	-1.18 (0.4)	—
$R^2$	0.39	0.40	0.19	0.52	0.52	0.06	0.60	0.60	0.15
$F$ and $p$ -value for									
constant 0,	3.47	3.33	0.09	3.10	2.08	0.73	0.43	0.65	0.16
slope 1	0.03	0.04	0.91	0.05	0.13	0.49	0.65	0.52	0.85
Number of clusters	130	130	58	97	97	39	88	88	33

$t$ -statistics are in parentheses. In regressions for all races, Black, Coloured, and Asian are dummy variables indicating the main racial group in the community. White is the omitted category.

Source: authors' calculations from SALDRU2 and *Education Atlas*. The dependent variable in columns 7 through 9 is the sum of reported mean number of primary and secondary pupils divided by the sum of the reported mean number of primary and secondary school teachers, by cluster.

questionnaire to identify the main racial group in the community and to select the appropriate ratio from the Atlas.

For all the regressions the Atlas pupil-teacher ratio is a strong predictor of the community level pupil-teacher ratio. For the separate primary and secondary regressions, we can only just reject the hypothesis that the slopes of the regressions are unity and their intercepts zero, so that the community level pupil-teacher ratio is equal to the magisterial district pupil-teacher ratio up to measurement error. In fact, this hypothesis cannot literally be correct, because the secondary school pupil-teacher ratios are lower than the primary school pupil-teacher ratios, but it seems that apart from this difference in levels, community pupil-teacher ratios move one for one with district pupil-teacher ratios. This interpretation is consistent with the results in the last three columns, where we use the combined pupil-teacher ratio from SALDRU2 and where we cannot reject the hypothesis that the intercept is zero and the slope is unity. Note also the relatively high  $R^2$  statistic, albeit on a sample of only 88 clusters. We also report regressions with race dummies to check that the effects of the district pupil-teacher ratios are not simply picking up differences by race, and separately for Black schools only. The significance of the district ratios remains in all of these specifications.

The regression results in Table I show that a substantial fraction of the intercluster variation in pupil-teacher ratios is explained by interdistrict variation. Note also that although the district ratios are noisy measures of the cluster ratios, their use as explanatory variables will not cause the usual attenuation bias because the “measurement errors” are here the difference between the cluster measure and the district average, and these are (by definition) uncorrelated with the district means that are being used as proxies. There will of course still be a loss of efficiency in using the less precise—and less variable—district level measures. Put differently, the use of district averages should not be thought of as using an error-ridden proxy, but as an instrumental variable procedure that uses district dummies as instruments. This interpretation also highlights a benefit associated with the use of magisterial district averages, which is a potential reduction in endogeneity bias from any parental control that exists. Parents may (or may not, in the case of Black parents) have some control over local school resources, but they have less over average resources in their magisterial district.

The Atlas data are merged into the SALSS data first by

matching clusters to districts and merging in the pupil-teacher ratios for each race, and then by selecting the appropriate pupil-teacher ratio on a *household* basis using the race of the household head. (The SALSS did not collect race data on all members of the household.) We are thus effectively assuming that children attend state schools according to their racial group.

Our final source is the 30 million individual records from the 1991 South African Census. With the exception of South Africans who lived in the so-called “independent” homelands of Transkei, Ciskei, Bophuthatswana, and Venda, the 1991 Census enumerated all South Africans. We will use the racial composition of magisterial districts from the census when looking at patterns of funding for Black schools.

*II.2. Inputs and Outputs*

Figures I and II show educational attainment in relation to age, where the former is measured by years of school completed. The data come from the main SALSS survey in response to a question about the highest educational level attained, which we have converted to years of education at the rate of one year per

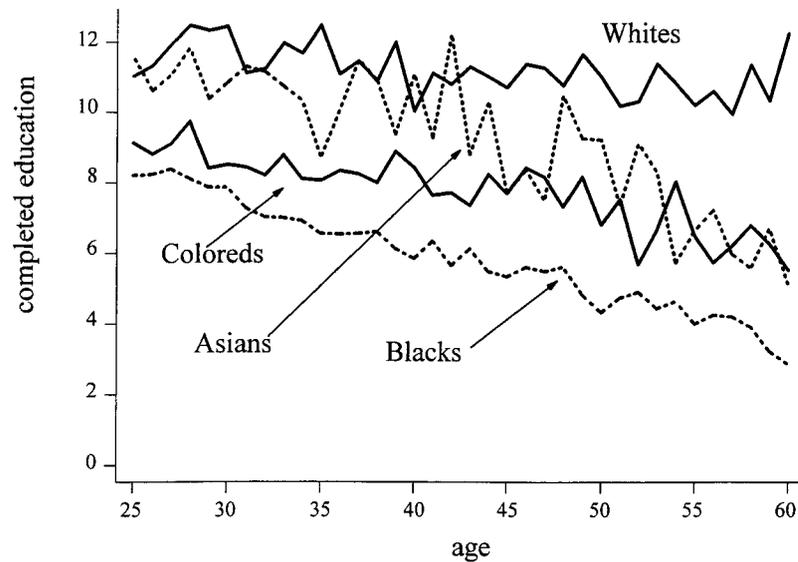


FIGURE I  
Number of Years of Education Completed by Age and Race for Adults

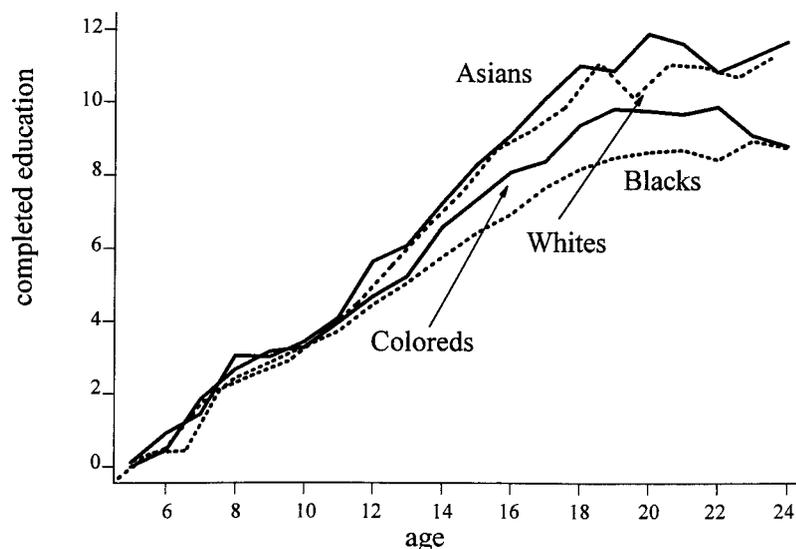


FIGURE II  
Educational Attainment by Age and Race, Ages 5 to 24

standard. While we shall often refer to this as years of education, it will be less than the number of years spent in education for students who repeat standards, or who are not in full-time education. Figure I shows years of education for adults by race and age, where adults are defined as those 25 and over. We adopt this high age cutoff because many young adults are still completing their education, particularly as a consequence of the disruptions to the high school system since 1976. White adults of all ages have on average eleven years of education. Except for young adult Asians, educational attainments are less for the other groups, but have been increasing over time, so that younger cohorts have more education than older cohorts. This is particularly true for Asians, the youngest of whom have educational attainments equal to or in excess of Whites. Increasing education is also apparent for Blacks, so that those born in 1968 have three years more of educational attainment than those born in 1933. The increase is slowest for Coloureds, where the average educational attainments of 60 year olds is comparable to that of Asians, while that of 25 year olds is comparable to that of Blacks.

Figure II shows attainment by race and age for people aged 5 to 24. Young Asians in the SALSS have more education than other

groups, although there is little difference between Asians and Whites. At age ten educational attainment does not differ much by race, but from ages ten to eighteen there is increasing dispersion between the races. Between ten and eighteen and looking across the age cross section, Blacks gain only 0.61 years of attainment for each year of age compared with 0.76 for Coloureds, 0.88 for Whites, and 0.95 for Asians. From the data to hand, we cannot tell how much of these differences can be attributed to the separate influences of dropout rates, repetition rates (which are only important after age eleven), and part-time versus full-time education.

That differences in pupil-teacher ratios may explain part of the fanning out in completed education is clear from Figure III, in which years of completed schooling are plotted against age at different pupil-teacher ratios for Black children aged ten to eighteen. Until roughly age thirteen we find little relationship in the raw data between the pupil-teacher ratio and completed education. Through age twelve, children in magisterial districts with average pupil-teacher ratios between 60 and 70 pupils per

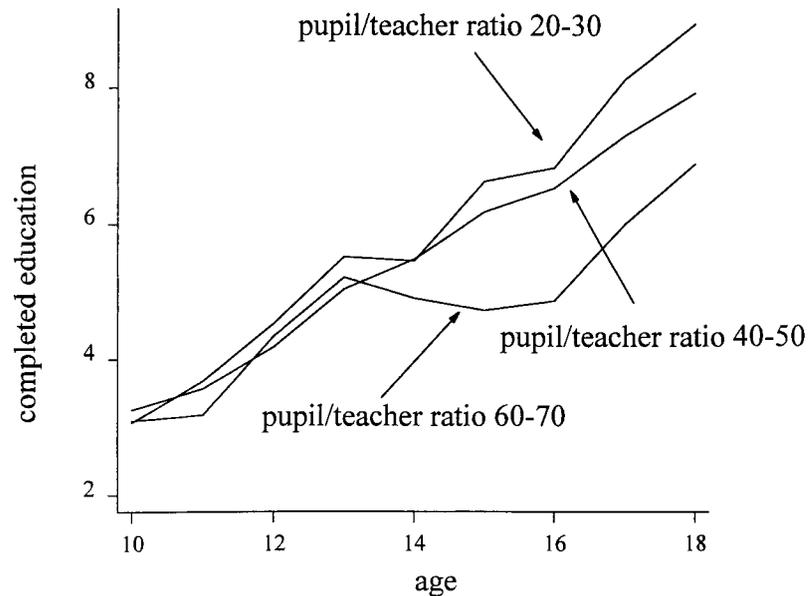


FIGURE III  
Completed Education by Age at Different Pupil-Teacher Ratios, Blacks 10-18

teacher advanced through school about as quickly as children in magisterial districts with average ratios between 20 and 30 pupils per teacher. Beyond that age, after which students must pass a provincewide examination in order to advance to the next grade, we see a pattern in which Black children in districts with fewer pupils per teacher advance more quickly. This fanning out of attainment for Black children between districts mirrors that in Figure II for educational attainment between races.

The SALSS questionnaire asks each person aged from 7 to 24 whether or not they are enrolled in formal education. The fractions reporting enrollment are tabulated by race and age—starting at age eight—in Table II. These data are not directly comparable to the educational attainment data in Figures II and III. Attainment comes from the integral of past enrollments, not the current enrollments reported here, and a year of enroll-

TABLE II  
SCHOOL ENROLLMENT BY RACE

Age	Black		Coloured		Asian		White	
	Percent enrolled	Pct with senior cert						
8	.907	.000	.918	.000	1.00	.000	1.00	.000
9	.927	.001	1.00	.000	1.00	.000	1.00	.000
10	.954	.000	.990	.000	1.00	.000	1.00	.000
11	.961	.000	1.00	.000	1.00	.000	1.00	.000
12	.966	.000	.989	.000	1.00	.000	1.00	.000
13	.963	.000	1.00	.000	1.00	.000	1.00	.000
14	.950	.000	.958	.000	1.00	.000	1.00	.000
15	.932	.001	.986	.000	1.00	.000	1.00	.000
16	.889	.002	.930	.000	.960	.000	1.00	.015
17	.855	.013	.857	.000	.912	.000	.954	.032
18	.790	.047	.795	.097	.917	.136	.970	.189
19	.677	.089	.439	.207	.455	.600	.769	.560
20	.607	.120	.288	.579	.438	.929	.500	.769
21	.477	.179	.264	.526	.105	1.00	.345	.950
22	.383	.252	.093	.600	.125	.667	.296	.809
23	.360	.220	.047	.667	.000	—	.273	.889
24	.228	.320	.020	1.00	.077	1.00	.266	.706

Percent with a senior certificate is the percent of students currently enrolled who hold a senior certificate (matric/form 5/senior certificate) or have completed (form2, form3, or form 4) and received a diploma. Source: SALDRU1. Files: m8\_hrost.dta and m4\_ed1.dta.

ment will not necessarily always lead to a year of completed education.<sup>2</sup>

Table II shows that all Asian and White children are in school up to age fifteen, and that substantial numbers continue beyond age fifteen, particularly among Whites. Enrollment rates among Black and Coloured children are lower. They start school later than Asians or Whites—which in spite of possible cohort and repetition effects appears to be inconsistent with the equality of achievement by age 10 in Figure II—and they stop going to school earlier and in greater numbers so that, by age 18, less than 80 percent of Blacks and Coloureds are enrolled, as opposed to 92 percent of Asians and 97 percent of Whites. Beyond age 18 there are substantial fractions of Blacks in education—more than a third at age 23—as they work to complete their education. At these ages, the Black and White data look similar, but for very different reasons; the Whites are in tertiary education and the Blacks are catching up on high school. Table II reports on the percentage of enrolled children of each age who have a senior certificate—equivalent to graduation from high school. For Blacks between 22 and 24 who are enrolled in the educational system, only between a quarter and a third have a senior certificate, as opposed to around 80 percent of enrolled Whites in the same age group.

The survey also asked, for those between the ages of 7 and 24 who were not enrolled, the most important reason for nonenrollment. Of Blacks 21.4 percent report “expense” as the main reason for nonenrollment; none of the Whites or Asians give this reason. “Illness,” “pregnancy,” and “could not cope with school work” are also much more important for Blacks than for Asians or Whites. These figures need to be interpreted with caution, particularly given the very different age profiles of enrollment in Table II. The high numbers of people listing “pregnancy” among the Blacks reflect the large numbers of Blacks attempting to complete high school in their early twenties.

Costs of education are reported by Blacks as an important reason for not being in school; it is possible using the SALSS to look at the size and structure of educational expenses. The average Black household with at least one member aged from 8 to 24 spends 40 Rand a month (on average 2.9 percent of total

2. Where appropriate, we note at the bottom of the table the name of the original World Bank file used in the creation of that table. These files are available on line from the World Bank at <http://www.worldbank.org/lsmis/guide/select.html>.

expenditure) on educational expenses, on fees, uniforms, transportation, school meals, and books. For Whites the corresponding figures are 248 Rand a month or 4.6 percent of total expenditure. Black households spend R13.8 a month per child in primary school and nearly twice as much, R25 per child, in secondary schools. For Whites the figures are ten and seven times as much, R129 for a primary school child, and R165 for a secondary school child, so that even though the average White household spends five times as much in total household expenditure as the average Black household, the share of educational expenses in the budget is larger for Whites. About a quarter of educational expenditures by Black households goes to school uniforms, and rather more than another quarter on transportation and school meals. These three items account for less than a quarter of the educational budget of White households.

Figure IV shows summary information for the pupil-teacher ratios from the Atlas data as merged into the SALSS sample. Each person in SALDRU aged 7 to 24 is assigned the pupil-teacher ratio for the appropriate racial group in the district in which he or she

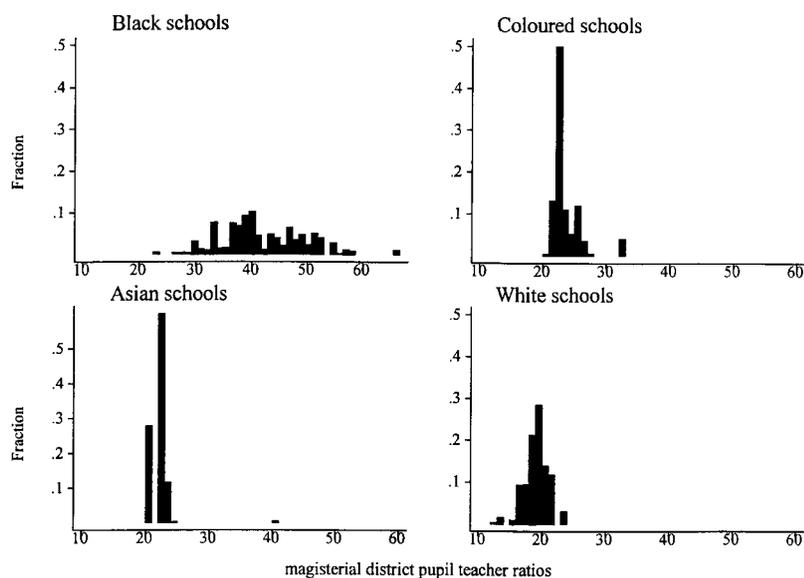


FIGURE IV  
Distributions of Pupil-Teacher Ratios for School-Age Children in SALSS, by Race, from District Data

lives. These are the data shown in Figure IV. While these are filtered through the SALSS, and are therefore not the same as the national figures given in *The Educational Atlas*, the means for each race are almost identical, as should be the case for a large national sample. Note that the data as shown understate the true variance of pupil-teacher ratios across individuals, because they ignore the dispersion across individuals in the same magisterial district.

Apart from the obvious and very large difference in pupil-teacher ratios between Blacks and the other groups, the most notable features of Figure IV is that pupil-teacher ratios are much more dispersed within the Black population than is the case for the other groups. To some extent, this is a consequence of the fact that some groups—particularly Coloureds and Asians—are concentrated in a relatively small number of districts, but the differences are striking.

The relationship between pupil-teacher ratios and district socioeconomic characteristics are explored in Panel A of Table III. For Black schools (column 1) and White schools (column 2), we regress the log (pupil-teacher ratio) by magisterial district on mean characteristics (by race) of each district. For Blacks there is no relationship between the log (pupil-teacher ratio) and the mean total household expenditure, used here as a measure of household well-being. Neither is there a significant relationship between the log (pupil-teacher ratio) and the mean years of education completed by the household head, a measure that may proxy both for household well-being and a household's taste for education. These findings, which are robust to the use of pupil-teacher ratios in place of the log ratios, and to the replacement of total expenditure with (the noisier) total household income, are important because, as we shall see below, both household expenditure and education of the household head have large and significant effects on children's educational attainment. That these characteristics have no effect on pupil-teacher ratios in the schools is consistent with the view that, under apartheid, Black parents had limited control over the provision of school quality. The results for White schools, in the second column, are somewhat different. Although, once again, the socioeconomic variables are jointly insignificant predictors of pupil-teacher ratios, there is some evidence of an income effect through total household expenditure. This result needs to be seen in the light of the low mean and low variability of pupil-teacher ratios among Whites. The

TABLE III  
DETERMINANTS OF PUPIL-TEACHER RATIOS  
Panel A: Socioeconomic Determinants of Pupil-Teacher Ratios

Dependent variable: log(pupil-teacher ratio) in schools attended by:		
	Blacks	Whites
Mean total household expenditure (1000s)	-.003 (0.2)	-.020 (2.6)
Mean household size	.006 (1.2)	.034 (1.7)
Mean years of education of household head	.000 (0.0)	.006 (0.9)
Fraction of female heads in magisterial district	-.072 (1.9)	-.008 (0.1)
<i>F</i> -test: joint significance of SES variables ( <i>p</i> -value)	1.50 (0.20)	2.05 (0.10)

Robust *t*-statistics are in parentheses. Also included in regressions: indicators for metro/urban/rural area and province. All SES variables here are race-specific magisterial district means, calculated from SALDRU1. There are 299 (70) clusters in the Black (White) regressions.

Panel B: Racial Composition of the Magisterial District

Dependent variable: log(pupil-teacher ratio) in schools attended by:		
	Blacks	Whites
Population White/total population	-.130 (1.4)	.188 (1.5)
Population Asian/total population	.367 (2.5)	.138 (0.8)
Population Black/total population	.227 (5.0)	.062 (1.4)
Log (total population)	.008 (1.1)	.037 (2.9)
Number of observations	283	253
<i>F</i> -test: joint sig of pop fraction variables ( <i>p</i> -value)	18.21 (.0000)	1.04 (.3761)
<i>F</i> -test: fraction(pop white) = -fraction(pop black) ( <i>p</i> -value)	0.60 (.4404)	2.60 (.1082)
<i>R</i> <sup>2</sup>	.1799	.0849

*T*-statistics are in parentheses. Population data are from the 1991 South African Census.

uniformly generous funding and small class sizes in White schools is in itself a measure of White parents' control over the school system, so that the lack of a relationship between parental characteristics and pupil-teacher ratios for Whites has different causes than that for Blacks.

That parental socioeconomic characteristics have little or no

effect on pupil-teacher ratios for Blacks is consistent with the general presumption that Black families had limited control over school resources under apartheid. Nevertheless, we have seen that there is a great deal of variability in pupil-teacher ratios across Black schools, even when aggregated up to the district level, and we rely on this variability to identify the effects of interest. It is therefore important to know, not only what does not affect pupil-teacher ratios, but what does.

We have devoted a considerable amount of time to this topic, interviewing South African academics, as well as current and ex-officials of the Ministry of Education and the (old) Department of Education and Training (DET). Our hope was to discover the rules under which teachers were allocated to schools, and to match those rules to our data along the lines of Angrist and Lavy [1999]. While there were guidelines on class sizes at both primary and secondary levels, these were not closely observed. Much of the variation in pupil-teacher ratios appears to have been generated by a (largely Afrikaner) bureaucracy that exercised its own discretion, but was not very responsive to the variation in needs that followed outflows of teachers or inflows of pupils. In addition, some of the very high pupil-teacher ratios may come from areas where classes were organized on the "platoon system," with multiple shifts. More generally, requests for resources flowed from headmasters and inspectors to the regional offices of the DET, but were decided centrally in Pretoria. During the disruptions of schools during the early 1980s, inspectors from the DET were not permitted to enter many schools, so that it would have been difficult for the DET to allocate resources appropriately, even if it had wished to do so.

Even so, there are clearly some avenues through which Black parents could influence the organization of the schools. For example, there appears to have been some migration of children from urban to rural schools during the disruptions. We do not know how many children were involved, but the incoming children would have inflated rural pupil-teacher ratios and might have had poor outcomes because of the disruption at origin rather than the high pupil-teacher ratios at destination. (But the empirical results to follow are unaffected by adding controls for whether the child migrated in the last five years.) The DET was also to some extent responsive to headmasters' requests, which in turn were presumably affected by parental pressure. South African schools tend to use school fees to supplement resources, and there

are even accounts of parents constructing their own schools in the face of indifference from Pretoria. Nevertheless, there is a stark difference from the situation for White parents. In some areas, the state matched contributions of White parents on a Rand for a Rand basis, a scheme that was never extended to Black schools. Even when pressed, none of our informants identified political pressure—for example through local ANC strength—as playing any role in the allocation of resources, even in the early 1990s.

Panel B of Table III shows that pupil-teacher ratios among Blacks are significantly affected by the racial composition of the district as measured in the 1991 census. Results are reported separately for Black schools (column 1) and White schools (column 2) from regressions in which the  $\log(\text{pupil-teacher ratio})$  in each district is regressed on the proportion of the district population that is Black, White, and Asian. For Black schools the racial composition of the magisterial district is a significant determinant of the district's pupil-teacher ratio, and the larger the fraction Black, the more pupils for every teacher. In contrast, the racial composition of the district plays no role in the allocation to White schools.<sup>3</sup> For Black schools we cannot reject the hypothesis that a one-percentage point increase in the proportion of the population that is Black has the same effect on school resources as a one-percentage point decrease in the proportion of the population that is White, so that it is the ratio of Black to White that matters. In this respect, and although the mechanisms are certainly different, South African allocation decisions prior to the end of apartheid mirror those in the United States in the early part of the century [Bond 1934, p. 244, cited in Boozer, Krueger, and Wolkon 1992, p. 287]. Indeed, the South African data sit almost exactly on (an extension of) the regression line of relative pupil-teacher ratios for black and white schools against fraction of the population that is black that was fitted by Boozer et al. to southern American states for 1930. The account of resource allocation given above provides no immediate explanation for these results. They do not come from the superior funding of urban over rural schools, nor from the higher ratio of Africans in rural areas. Rather, we suspect either that it was easier for headmasters to pressure the DET in areas where the fraction of Whites was high, because the headmasters themselves were more

3. The  $F$ -statistic of the joint significance of the race variables is 18.21 ( $p$ -value = 0.0000) for Black schools, and 1.04 ( $p$ -value = 0.3761) for White schools.

likely to be White, because they had more access to White officials, or because Whites themselves would exert pressure on behalf of their employees or servants. It is also possible that the DET had its own prejudices in favor of areas with a large White presence.

Although the data on school facilities are compromised by the lack of completeness in SALDRU2, we report some information in Table IV. The table reports the fraction of schoolchildren (from age five to eighteen) who have access to each of the facilities shown in the rows of the table. These are computed by assigning to each school-age child a one or zero depending on whether or not the facility is reported as present in the cluster, and then averaging over all nonmissing observations. The results confirm that it is not only pupil-teacher ratios that are unequally allocated. Only 11 percent of Black children live in communities where the primary school has a library, compared with almost all Asian and White children. Only 39 percent of Black children live in communities where there is a secondary school with a science laboratory, again compared with nearly all Asian and White children. These figures are very similar to those reported in the Republic of South Africa [1997] for biology laboratories (only 23 percent of all schools); science laboratories (31 percent); and library facilities (17 percent). Sports facilities are also unequally distributed, and even if we believe that sports facilities are less important to educational achievement than are libraries or laboratories, their distribution is likely to mirror school facilities in general. It is also worth noting that, as claimed in the introduction, the existence of these

TABLE IV  
SCHOOL FACILITIES: FRACTION OF SCHOOL AGED CHILDREN (5 TO 18) WITH EACH FACILITY AVAILABLE IN THEIR CLUSTER

	Black	Coloureds	Asian	Whites	Number of clusters reporting
Primary school library	.110	.854	1.00	.943	279
Secondary school library	.380	1.00	1.00	1.00	227
Secondary school laboratory	.390	1.00	1.00	.958	227
Primary school sports facility	.602	.612	1.00	.971	281
Secondary school sports facility	.635	.779	.824	.999	228
Primary school swimming pool	.012	.059	.005	.397	273
Secondary school swimming pool	.034	.014	.174	.446	222

Source: SALDRU2 (Community Questionnaire).

facilities is positively correlated with pupil-teacher ratios. For example, for Black children aged ten to eighteen the correlation between the log(pupil-teacher ratio) in their magisterial district with the presence in their cluster of a primary school library is  $-0.69$ , of a secondary school library is  $-0.59$ , and of a secondary school laboratory is  $-0.53$ .

Table V reports on the test scores. These results are restricted to the test-takers who are also present in the full SALSS survey, and we show them separately for the group aged thirteen to eighteen. The test itself consists of fourteen questions that cover comprehension, practical mathematics, and literacy. These are designed to cover material in the school syllabus up to Standard 5, the end of primary school, normally reached by age twelve. The questions are posed in English or in the respondent's mother tongue and are read aloud by the interviewer. They pose simple practical problems, so that, for example, question 1 (in English) asks how long a bus trip would take to cover a given distance at a specified speed. The respondent then reads a two-paragraph passage in his or her mother tongue, and is asked six questions to test comprehension of the written material. There are then four computational problems (e.g., 25 percent of 228 Rand equals how many Rand?) and two practical math problems. Each question was graded correct or incorrect, and total scores computed. For the purpose of this paper, we calculated "mathematics" scores and "literacy" (or comprehension) scores. The former is the total number of correct answers on the first two questions, the four computational questions, and the two practical math questions, so that scores range from 0 to 8. The first two questions are also

TABLE V  
LITERACY AND NUMERACY TEST SCORES

Means for age:	Literacy scores		Numeracy scores	
	Blacks	Whites	Blacks	Whites
13	2.64	5.32	1.77	4.89
14	2.87	4.57	2.02	3.86
15	2.92	5.76	2.25	4.15
16	3.35	6.00	2.44	5.21
17	3.14	5.81	2.22	5.38
18	3.46	5.88	2.59	5.65
Number of observations, ages 13–18	756	108	756	108

added into the literacy scores (they are as much about comprehension as computation) together with the six answers from the comprehension question, so that as with the math score, the literacy score ranges from 0 to 8.

We report only on Blacks and on Whites, since the disaggregation by age leaves too few Coloureds and Asians to give useful averages. Both groups did better on the comprehension tests than on the math tests, and, for Blacks, test scores improve with age; we shall see in Section III that this improvement comes from years of education, not age. For Whites the evidence of improvement with age is more limited or absent. Since the test is set at around the Standard 5 level, which most Whites (and Asians) have completed by age twelve, it is plausible that they do as well as they are ever going to do by age thirteen.

We conclude with a first look at the relationship between educational attainment and various measures of economic success at both the household and individual level. Figures V and VI show plots of the logarithm of wage on educational attainment. These graphs present information for only those individuals who report formal sector earnings, which in South Africa is only about a

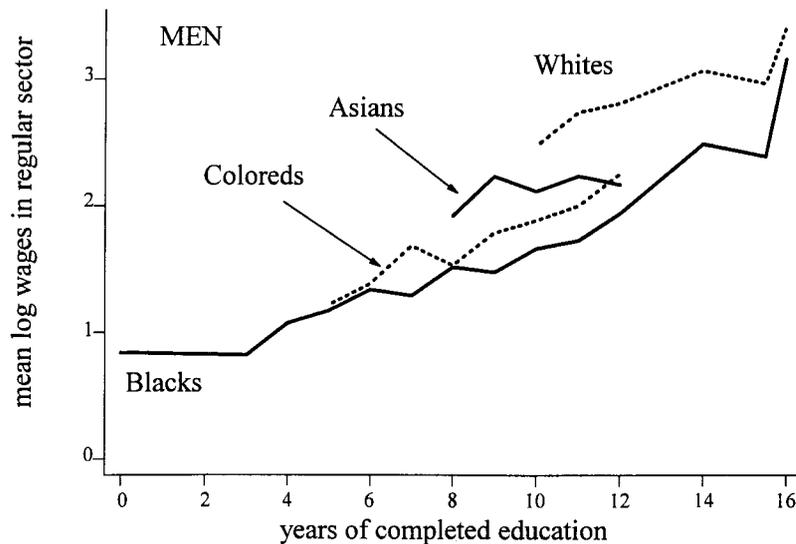


FIGURE V  
Mean Log Formal Sector Wage Rates for Men by Race and Years of Completed Education

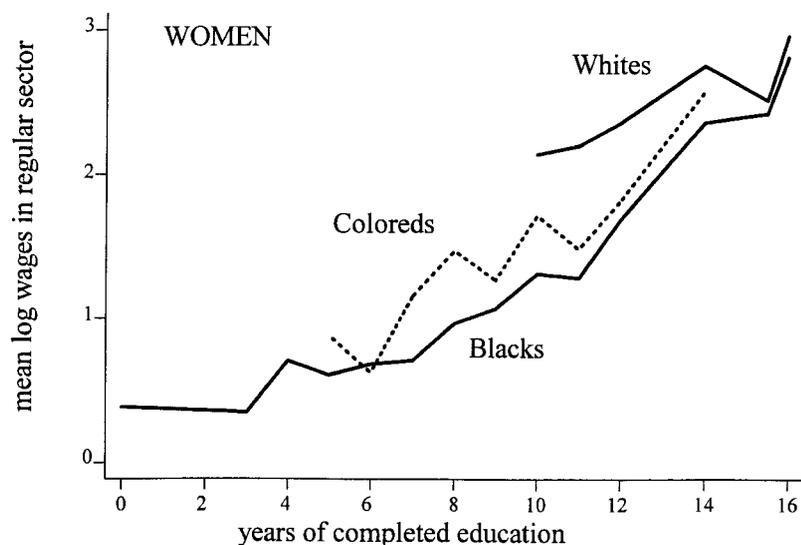


FIGURE VI  
Mean Log Formal Sector Wage Rates for Women by Race and Years of Completed Education

quarter of the potential labor force. The graphs do not cover all groups for all education levels because we do not show the plot when there are insufficient observations, usually ten or fewer at a given level. Except at the lowest levels, higher education is associated with higher formal-sector wage rates for both men and women. The graphs also show that, at higher levels of education, the *slopes* of the education-wage profiles are steeper for Blacks than for Whites. These results are consistent with findings in Moll [1995, 1996] and Mwabu and Schultz [1996]. However, it should be remembered that most Whites have more than ten years of education, and most Blacks less than ten years, so that an additional year of education from the mean is associated with about the same rate of increase of wages for both groups.<sup>4</sup>

4. Similar relationships are seen when one graphs the education of household heads and two other economic measures, total household income and total household expenditure. As in the figures shown here, the slope for better educated Blacks is steeper than that for Whites. The relationships are also replicated in regressions of wages on a set of education dummies with controls for age, age squared, and age cubed. Such regressions allow for the negative correlation of completed education with age; see Figure I. Controlling for age, however, does equalize the estimated rates of return to tertiary education for Blacks and Whites.

### III. EDUCATIONAL INPUTS AND EDUCATIONAL OUTCOMES: ECONOMETRIC ANALYSIS

In this section we examine the relationship between school inputs, particularly the pupil-teacher ratio, and various measures of educational outcomes, including educational attainment, enrollment rates, the reasons for not being in school, educational expenditures, and test scores. We present the results of a series of regressions in which the pupil-teacher ratio (or the presence of other facilities) is an explanatory variable. Among the other controls are age, urbanization, sex, and various measures of family background, such as whether the household is headed by a woman, household size, the educational attainment of the head, and the logarithm of total household expenditure per capita. We think of head's education as both a direct input into the educational process and a measure of household resources. We also proxy household well-being with total household expenditure per household member, a quantity that is both better measured than income and a better indicator of longer-term living standards.

In the regressions presented below, we have chosen to focus on the log of the pupil-teacher ratio, instead of the ratio itself. There is no theoretical reason to prefer the ratio over its logarithm, and the latter has the advantage of making the regressions invariant to whether we work with the pupil-teacher ratio or the teacher-pupil ratio. The results are robust to replacing the logarithm with the pupil-teacher ratio, which is seen in at least one specification in each table that follows.

#### *III.1. Pupil-Teacher Ratios and Educational Attainment of Adolescents*

Table VI presents the results of an analysis of years of completed education for children aged ten to eighteen. The dependent variable in the regression in column (1) is educational attainment measured as years completed; among the explanatory variables, age is entered as a series of dummies, one for each year, with age ten omitted. Age is the immediate correlate of educational attainment for young people who are still in the educational system, and it is clearly the primary determinant of years of education attained. We also include a dummy variable for sex, various household socioeconomic characteristics, including the education of the head, household size, and a measure of household resources, as well as dummies for urbanization and provinces. In

TABLE VI  
DETERMINANTS OF YEARS OF COMPLETED EDUCATION FOR CHILDREN AGED 10 TO 18

	Blacks						TSLS (7)
	Whites (1)	(2)	(3)	(4)	(5)	P/T Ratio (6)	
Indicator: age = 11	.851 (10.3)	.393 (7.0)	.365 (6.4)	—	—	—	.380 (5.3)
12	1.66 (14.8)	1.08 (16.4)	1.08 (16.3)	—	—	—	1.09 (14.5)
13	2.65 (17.8)	1.70 (23.8)	1.69 (22.5)	—	—	—	1.80 (21.7)
14	3.71 (24.6)	2.40 (29.2)	2.42 (29.2)	—	—	—	2.49 (25.2)
15	4.70 (28.5)	3.08 (35.1)	3.08 (34.0)	—	—	—	3.16 (31.2)
16	5.74 (32.8)	3.56 (35.9)	3.56 (35.3)	—	—	—	3.66 (31.8)
17	6.40 (25.3)	4.26 (39.5)	4.27 (40.0)	—	—	—	4.28 (32.8)
18	7.06 (26.2)	4.77 (50.1)	4.78 (47.6)	—	—	—	4.79 (41.1)
Age in years	—	—	—	.617 (63.6)	1.04 (5.1)	.478 (5.5)	—
Log (pupil-teacher ratio)	-.265 (0.4)	-1.82 (4.3)	-2.05 (4.5)	-1.80 (4.3)	1.01 (1.5)	.030 (1.8)	-3.43 (1.7)
Col (6): pupil-teacher ratio	—	—	—	—	—	—	—
Female	-.007 (0.1)	.496 (11.7)	.490 (11.3)	.496 (11.6)	.496 (11.8)	.496 (11.8)	.460 (8.7)
Head of household's completed educ	.061 (2.7)	.076 (9.8)	—	.076 (9.8)	-.051 (1.6)	-.051 (1.6)	.078 (8.1)
Log (household size)	-.669 (1.6)	.188 (2.6)	—	1.85 (2.5)	.178 (2.4)	.181 (2.4)	.170 (1.8)
Indicator: female head	.136 (0.8)	.120 (2.1)	—	.121 (2.1)	.116 (2.1)	.120 (2.1)	.149 (1.9)
Log (expenditure per household member)	-.086 (0.5)	.506 (9.5)	—	.508 (9.5)	-.398 (2.5)	-.390 (2.4)	.424 (6.5)
Age*head's completed education	—	—	—	—	.009 (3.6)	.009 (3.7)	—
Age*log (P/T ratio)	—	—	—	—	-.208 (4.3)	-.005 (4.3)	—
Col (6): age*P/T ratio	—	—	—	—	—	—	—
Age*log (expenditure per household member)	—	—	—	—	.064 (5.3)	.064 (5.2)	—
F (head ed, age*head ed)	—	—	—	—	48.8	48.9	—
F (P-T, age*pup-teacher)	—	—	—	—	15.7	12.9	—
F (log exp, age*log exp)	—	—	—	—	47.2	46.7	—
Number of obs	629	7103	7103	7103	7103	7103	4965

*T*-statistics are in parentheses. Regressions are estimated with robust standard errors, allowing for correlation between observations from the same cluster. Metro and province indicators are included in all regressions. *Source*: SALDRU1 and *Education Atlas*. One magisterial district is dropped because of an outlier in White population (47 observations in Umlazi, KwaZulu removed). Seven observations are removed for individuals identified as head of household. In column (6) pupil-teacher ratio and the interaction of age with the pupil-teacher ratio used in place of log (pupil-teacher ratio) and its interaction with age.

In column (7) log (pupil-teacher ratio) is instrumented on the fractions of the population in the magisterial district that are Black, Asian, and White. (*Source*: 1991 South African Census.) Transkei, Ciskei, Bophuthatswana, and Venda (TBVC) were not enumerated in the census and are missing from the TSLs estimation. If we attribute a 100 percent Black population to the magisterial districts in TBVC, the TSLs estimate of the log (pupil-teacher ratio) coefficient is -3.19 (robust *t*-statistic = 1.6).

In regressions estimated using log (expenditure per adult) and the log (number of adults in the household) in place of log (expenditure per household member) and the log (number of household members), but with a specification identical to that in column (2), the coefficient on log (pupil-teacher) ratio is -1.82, with a *t*-statistic of 4.2, and the coefficient on log (expenditure per adult) is .467 with a *t*-statistic of 5.5.

column (1) we show the regression for Whites and, in subsequent columns, regressions for Blacks only. There are 7103 Black children aged ten to eighteen in the SALSS with complete information, but only 629 Whites, and correspondingly fewer Coloureds and Asians, the results for whom we do not present.

The first two columns show how educational status rises with age, and the coefficients on the age dummies rise as we move down the columns. The differences between successive dummies would be unity if each child moved with certainty from one grade to the next, and this is essentially what happens in column (1) for Whites, where the differences are 0.81, 0.99, 1.06, 0.99, 1.04, 0.96, and 0.66, with the last figure presumably dropping off as some White children leave school at seventeen. Among Black children, however, the annual increments vary from 0.48 to 0.70, and average only a little over 0.6, so that the average Black child obtains only two-thirds of a year of education for each additional year of age.

In both columns the pupil-teacher ratio has a negative impact on attainment for age, although the effect is small and insignificant for Whites; for Blacks it is  $-1.82$  with a  $t$ -value of 4.3. According to this estimate, reducing the average Black class size from 40 to 30 pupils per teacher, or a quarter, for example, would raise average educational attainment by 0.52 years, equivalent to adding about ten months to age. Reducing it by 20, from the average Black to White ratio, would raise attainment by 1.26 years, about quarter of the actual years of education for a fourteen year old Black child.<sup>5</sup> One obvious interpretation of the difference in the effects between Whites and Blacks is that the effect of additional resources is nonlinear, and that decreases in the pupil-teacher ratio have a much larger effect when there are 40 (or 60) pupils per class than when there are 20. Rerunning the Black regression with the square of the log(pupil-teacher ratio) as well as its level (not shown in the table) provides no support for this supposition. The square of the log has the wrong sign (positive) but is not significantly different from zero.

What turns out to be more important is to allow for interactions between age and the pupil-teacher ratio. School resources act not only on educational attainment itself, but also on the process of acquiring education, so that students make more rapid

5. If the pupil-teacher ratio is used in place of its logarithm, the coefficient is  $-0.04$  with a  $t$ -ratio of  $-3.6$ ; this gives somewhat larger effects than those calculated above.

progress through the educational system when there are more teachers for each student. Low pupil-teacher ratios may also cause pupils to start education earlier so that the pupil-teacher ratio can then affect both the intercept of the relationship between attainment and age, and the slope, which is the rate at which age turns into educational attainment. The effect of a low pupil-teacher ratio on attainment should be larger for older children who have had more exposure. It is clumsy to allow these interactions with a list of dummies for age, so we show in column (4) a regression in which age is restricted to appear linearly, and in column (5) we build on this specification by adding an interaction term between age and the pupil-teacher ratio, whose coefficient of  $-0.208$  attracts a  $t$ -value of  $-4.3$ . The  $F$ -test for the joint inclusion of the log(pupil-teacher) ratio and its interaction with age is  $15.7$ . The combination of intercept and interaction terms show that if the pupil-teacher ratio were halved, from 40 to 20, it would have little effect on attainment at age six, but would add three-quarters of a year to educational attainment by age ten, and a year and a half by age fifteen. This interactive specification makes most sense to us, and these are our preferred estimates.<sup>6</sup> (The corresponding results using the pupil-teacher ratio itself are shown in column (6); the calculated effects are now somewhat smaller than in the logarithmic specification, 0.4 at age ten and 0.9 at age fifteen.)

Gender and household characteristics have important effects in the regressions. Black female students have on average about half a year of educational attainment more than Black male students, and among Black students there are the expected positive effects of household resources and of the education of the household head. Head's education is a strong predictor of educational attainment among both Blacks and Whites; a head completing twelve years of education as opposed to eight years—the difference between completing primary school and completing secondary school—is predicted to raise the educational attainment of members of the household by a quarter of a standard. For Blacks the effect is about the same as that of reducing the pupil-teacher ratio by ten. This intergenerational influence will enhance the long-run benefits of lowering pupil-teacher ratios,

6. For both columns (2) and (5) we tested for differences between boys and girls by interacting a dummy for female with the log pupil-teacher ratio and with log of expenditure per household member. In neither case were these interactions jointly significant. For column (2) the  $F$ -test is 0.35; for column (5) 0.60.

since a better-educated current generation will have better-educated children.

For the purposes of this paper, the two most interesting contrasts are in the effects of resources, private resources through expenditure per head, and public resources through the pupil-teacher ratio. Household resources have no effect on the educational attainment of White children, but a marked positive effect on the education of Blacks. According to this, we might interpret the expenditure term as a resource effect; education is cheap enough so that funding is not a constraint for Whites but can be a serious constraint for the much poorer Blacks. The same appears to be true for the pupil-teacher ratio, whose strong negative effect is confined to Black pupils. The pupil-teacher ratio has no obvious effect on educational attainments of White children, among whom the pupil-teacher ratios are around nineteen. For the Black students, where the pupil-teacher ratios in primary and secondary schools are more than twice as big, higher pupil-teacher ratios have a strong and significant negative effect on attainment.

Column (3) explores the consequences of excluding the family background variables for Black students. Although these are significant in the regression in column (2), their removal does not affect the other coefficients, and in particular does not alter either the size or significance of the pupil-teacher ratio. These results are what we would expect from Table III, which showed negligible effects of family background variables on the pupil-teacher ratios, and suggest that family variables are important in their own right. They provide no support for the proposition that pupil-teacher ratios are simply picking up the effect of excluded background variables.

Column (5) allows not only for interactions between age and the pupil-teacher ratio, but also for interactions between age and head's education and the logarithm of per capita household expenditure. All three interaction terms are significant in this regression, so that the evidence is consistent with the idea that money, family education, and pupil-teacher ratios all work by helping people progress more rapidly through the educational system, presumably through some combination of increasing the probability of enrollment and of completing each standard more rapidly conditional on enrollment.

Consistent with the institutions of apartheid, we have found no evidence that pupil-teacher ratios for Blacks are affected by measured household characteristics, from which it seems implau-

sible that the estimated effects of the pupil-teacher ratio are coming from the influence of *unmeasured* household characteristics. Nevertheless, as a check, we consider possible instrumentation using the racial composition of magisterial districts. As we saw in Table III, pupil-teacher ratios can be predicted by racial composition, and the  $F$ -test of the joint significance of population fractions Black, Asian, and White in the first stage regression for column (7) is 2.85. The consistency of the TSLS estimates also requires us to assume that the racial composition of the district has no *direct* effect on educational attainment which, although plausible, is not necessarily the case. The TSLS estimates should therefore not be regarded as definitive, but as another strand of evidence. Even so, an overidentification test of these instruments yields a chi-square value of 3.58 ( $p$ -value = 0.17), which is acceptable. Given all this, the TSLS results are very similar to the OLS results in column (2); indeed, the effect of the  $\log(\text{pupil-ratio})$  is, if anything, more pronounced. In all specifications for Black children, the  $\log(\text{pupil-teacher ratio})$  appears to have a negative and significant effect on the children's years of completed schooling.

Another concern is that these results might be generated in part by reverse causation in the following way. Suppose that an unobservable characteristic causes children in some magisterial districts to go through the system more quickly, and to remain in school longer. In such districts, there will be a high fraction of children in secondary schools relative to primary schools and, because teacher-pupil ratios are higher in secondary schools, the overall teacher-pupil ratio will also be higher. As a result, there is a spurious correlation running from educational achievement to the teacher-pupil ratio. The ideal way to deal with this issue would be to include in the regressions not the overall pupil-teacher ratio, but pupil-teacher ratios for primary and secondary schools separately. But *The Educational Atlas* does not give such information nor, given that there is often only one cluster in a magisterial district, can we construct adequate information on primary and secondary enrollment for magisterial districts from SALDRU1. Nevertheless, note the following evidence. First, using the data available from SALDRU2 on primary and secondary school enrollments and teachers (33 African clusters out of 360 total clusters report on both), the average pupil-teacher ratio for African primary and secondary schools do not differ by much, 37.4 students per teacher for secondary and 39.6 for primary schools. Since the two class sizes are also highly correlated within clusters

(0.42), the differential weighting across clusters can have only limited impact on the pattern of the overall pupil-teacher ratios. Second, we used the primary and secondary pupil-teacher ratios from SALDRU2 to instrument the overall pupil-teacher ratio in those clusters where the data exist. These IV estimates continue to show a negative and significant effect of the pupil-teacher ratio on outcomes.<sup>7</sup> Third, we used the SALDRU1 data to estimate, for each magisterial district, the fraction of schoolchildren in secondary school. Together with the average difference in teacher-pupil ratios across secondary and primary schools, these ratios are used to “correct” the Atlas teacher-pupil ratios for the aggregation effect. When these corrected figures are used in the regressions, now for all the clusters, the results are exactly as shown in the table, which would appear to show that the reverse causation has a negligible effect on the results.

### *III.2. Enrollment Rates and Reasons for Nonenrollment*

Table VII shows estimates of linear probability models for enrollment status. These are run for Blacks aged from eight through eighteen inclusive. A dummy variable equal to one for enrollment and zero otherwise is regressed on the pupil-teacher ratio and the family background variables, together with (not shown) a series of age indicators and urbanization dummies.

The results are qualitatively consistent with our findings on educational attainment in Table VI, particularly as concerns the effects of public and private resources on education. The log of the pupil-teacher ratio has a significant negative effect on the probability of being in school. Using the same example as before, a cut in the pupil-teacher ratio of a quarter would increase the probability of enrollment for Black students by 0.02 in a single year. The same cut is estimated in Table VI to increase average educational attainment for ten to eighteen year olds by 0.52, which is larger than the prediction from the enrollment effects (between four and twelve years of risk times 0.02 per year). As we have already noted, enrollment and attainment are not mechanically linked;

7. For the whole sample, the effect of the magisterial district pupil-teacher ratio (the level, not the logarithm) on years of completed education (comparable to results in Table VI) is  $-0.039$  with a  $t$ -statistic of 3.6. In the 33 clusters for which we have information on pupil-teacher ratios for both secondary and primary schools from SALDRU2, the corresponding OLS parameter estimate is  $-0.069$  with a  $t$ -statistic of 3.1. In these same clusters, instrumenting the magisterial district pupil-teacher ratio with the pupil-teacher ratios in primary and secondary schools from SALDRU2, the two-stage least squares estimate of the effect of pupil-teacher ratio is  $-0.075$  with a  $t$ -statistic of 2.2.

TABLE VII  
SCHOOL ENROLLMENT: BLACKS AGED 8 TO 18

Dependent variable: currently enrolled in school = 1					
Explanatory variables:					
Log (pupil-teacher ratio)	-.070 (2.1)	1.71 (1.8)	1.72 (1.8)	—	—
Log (pupil-teacher ratio) <sup>2</sup>	—	-.240 (1.8)	-.241 (1.8)	—	—
Pupil-teacher ratio	—	—	—	-.002 (2.2)	.007 (1.1)
Pupil-teacher ratio <sup>2</sup>	—	—	—	—	-.0001 (1.3)
Indicator: female	.010 (1.6)	.010 (1.6)	.088 (3.7)	.010 (1.6)	.088 (3.6)
Female*age interaction	—	—	-.006 (3.1)	—	-.006 (3.1)
Household head's educ	.006 (4.8)	.006 (4.9)	.006 (4.9)	.006 (4.8)	.006 (4.9)
Indicator: female head	.014 (1.7)	.014 (1.8)	.015 (1.8)	.014 (1.7)	.014 (1.8)
Log (expend per hh mem)	.037 (4.9)	.037 (4.9)	.037 (4.9)	.037 (4.9)	.036 (4.8)
Log (household size)	.029 (2.6)	.030 (2.8)	.030 (2.7)	.030 (2.7)	.029 (2.6)
Number of obs	8958	8958	8958	8958	8958
R <sup>2</sup>	.061	.063	.064	.062	.064

Linear probability model. Robust *t*-statistics are in parentheses, allowing for correlation between observations from the same cluster. Also included in the regressions are age indicators, urban, and rural indicators. *F*-test of joint significance of the pupil-teacher ratio and its square in column 5 = 2.6, (*p*-value .07).

Source: SALDRU1. Files: m4\_ed1.dta, m8\_hrost.dta, hhexpl.dta and the *Education Atlas*.

enrollment does not imply attendance, let alone passing on to the next grade, and dropouts can reenroll at various ages, so that it is hard to use the comparison as a cross-check. As with attainment, the effect of a four-year increase in head's education is about the same as a ten-pupil drop in the pupil-teacher ratio, and, in both cases, household resources enhance education.

Column 2 allows for nonlinearity in the effect of log(pupil-teacher ratio) on the probability of enrollment. The probability of Black enrollment peaks at a log pupil-teacher ratio of log(36), which is at the low end of pupil-teacher ratios observed for Blacks in our sample; only 17 percent of Blacks aged eight to eighteen live in magisterial districts with average pupil-teacher ratios of 36 or below. Moreover, the marginal impact of an increase in the log(pupil-teacher) ratio increases in absolute value as the pupil-

teacher ratio increases. An additional student per teacher at a pupil-teacher ratio of 40 is associated with a reduction in the probability of enrollment of 0.15 percent while at a pupil-teacher ratio of 60 is associated with a reduction of 0.43 percent.

Column 3 allows for different attendance patterns for males and females and shows that girls are more likely to be enrolled until age fourteen and less likely thereafter, a finding that is consistent with the superior levels of attainment of girls shown in Table VI.

### *III.3. Determinants and Effects of Test Scores*

We now turn to the much smaller sample of individuals who took the comprehension and numeracy tests appended to the main survey. Since the selection of the adult sample of test-takers was problematic, and since the pupil-teacher ratio can only be reliably linked with those now in school, we confine our attention to those between thirteen and eighteen years of age, of which there are 763 Blacks, and 89 Whites. The samples of Asians and Coloureds are too small for separate analysis, and we do not feel justified in pooling them with either of the other two groups.

The results for the White subsample are presented in the first two columns of Table VIII; there are no significant effects of any of the variables on either of the two scores. While this is consistent with the general lack of such findings in the literature, the small sample size must be kept in mind; the estimated effects of the pupil-teacher ratio are insignificantly different from zero but are also consistent with the existence of a large negative influence of class size on test scores. For the larger sample of Black adolescents, the results are more precise. Columns 3 and 4 show the regressions with education omitted, columns 5 and 6 with years of completed education as a conditioning variable, and columns 7 and 8 with education included but without the family background variables. Since pupil-teacher ratios affect the amount of education, which in turn is likely to affect test scores, we are interested in regressions both unconditional and conditional on education. The conditional regression measures the direct effect of educational quality on test scores with quantity held constant, while the unconditional regression gives the total (reduced-form) effect of quality, both direct and indirect. In the absence of education, age has a positive effect on test scores, as does the education of the household head, household resources, and female headship. Higher pupil-teacher ratios negatively affect both scores, but the size of

TABLE VIII  
DETERMINANTS OF LITERACY AND NUMERACY TEST SCORES

	Whites		Blacks		Blacks		Blacks		Blacks			
	Literacy	Numeracy	Literacy	Numeracy	Literacy	Numeracy	Literacy	Numeracy	Literacy	Numeracy		
Years of completed education	.225 (1.5)	-.052 (0.3)	—	—	.225 (7.7)	.270 (8.7)	.273 (9.8)	.300 (10.2)	.262 (6.4)	.333 (8.4)	.252 (6.4)	.333 (8.3)
Age	.091 (0.5)	.355 (1.8)	.161 (4.6)	.146 (3.9)	.024 (0.6)	-.019 (0.5)	-.006 (0.2)	-.037 (0.9)	-.034 (0.7)	-.079 (1.5)	-.035 (0.7)	-.078 (1.5)
Log (pupil-teacher ratio)	1.15 (0.6)	3.24 (1.4)	-.156 (0.5)	-.778 (2.2)	.006 (0.0)	-.584 (1.7)	-.123 (0.4)	-.825 (2.5)	.228 (0.5)	-.946 (1.8)	—	—
Pupil-teacher ratio	—	—	—	—	—	—	—	—	—	—	.008 (0.7)	-.022 (1.8)
Primary school library	—	—	—	—	—	—	—	—	-.352 (1.1)	-.301 (0.7)	-.343 (1.0)	-.302 (0.7)
Secondary school library	—	—	—	—	—	—	—	—	.863 (3.1)	.399 (1.2)	.877 (3.2)	.403 (1.3)
Secondary school laboratory	—	—	—	—	—	—	—	—	-.452 (1.7)	-.365 (1.2)	-.450 (1.7)	-.361 (1.2)
Head of household's completed educ	.078 (1.4)	.055 (0.9)	.072 (4.3)	.078 (4.3)	.053 (3.2)	.054 (3.1)	—	—	.053 (2.1)	.034 (1.2)	.054 (2.1)	.035 (1.3)
Log (expenditure per household member)	.569 (1.3)	.824 (1.7)	.392 (3.7)	.318 (2.8)	.231 (2.2)	.125 (1.1)	—	—	.172 (0.9)	.226 (1.1)	.168 (0.9)	.226 (1.1)
R <sup>2</sup>	.200	.251	.132	.122	.125	.203	.175	.183	.241	.229	.242	.228
Number of obs	89	89	763	763	763	763	777	777	383	383	383	383

T-statistics are reported in parentheses. Source: SALDRU1 and SALDRU3. All regressions include metro indicators, log of household size, indicator for female-headed household, and an indicator that the respondent is female. When age is interacted with log (pupil-teacher) ratio in regressions otherwise identical to those in columns 3 and 4, the coefficients on the interaction terms are positive and insignificantly different from zero.

the effect is five times as large for the mathematics test scores as it is on the comprehension scores. The standard errors are much the same for both, so that only the effect on the math score is significantly different from zero.<sup>8</sup>

Columns 5 and 6 include educational attainment as a regressor, which has a strong positive effect on test scores—four additional years generates one additional correct answer on the tests—and which removes the effect of age. The estimated coefficients of pupil-teacher ratios are diminished (absolutely) by the exclusion of the indirect effect on educational attainment; there is no estimated direct effect of the pupil-teacher ratio on the comprehension score, and the reduction in the size of the effect on the math score leaves it significant only at the 10 percent level. The comparison between these results and those in columns 7 and 8 shows that the results are robust to the inclusion or exclusion of the background variables. In their absence, the pupil-teacher ratio has somewhat larger effects, and is more precisely estimated, but the difference is not very marked.

The last four columns look at the effect of facilities other than pupil-teacher ratios and experiments with the pupil-teacher ratio instead of its logarithm. We work here with Black adolescents only, and the samples are further reduced (and likely selected) by the problems with SALDRU2. We show the regressions with years of education included—the results are similar with the expected alterations if education is excluded—as well as the family background variables. We focus on three measures of facilities, whether the local primary school has a library, whether the secondary school has a library, and whether the secondary school has a laboratory. We selected these because they are the most likely to have a direct effect on the test scores, although we recognize that they are also likely to serve as indicators of the general quality and endowments of schools.

The effects of the pupil-teacher ratios on test scores are not significantly different from the results in the columns where facilities are excluded. Higher pupil-teacher ratios reduce the math score but have an insignificant positive effect on the comprehension scores. We find no effect of a primary school library, but the presence of a secondary school library has a

8. As was the case in Table VI, we found no evidence of differential effects by sex; for example, in columns 3 and 4, the *F*-tests for interactions of a female dummy with log expenditure per household member and log pupil-teacher ratio were 1.29 (literacy) and 0.77 (numeracy), respectively.

significant and large—more than 85 percent of a correct answer— influence on the literacy test score. Secondary school laboratories attract an insignificant negative coefficient.

Although the evidence here cannot be given much weight—the sample sizes are small and many clusters are excluded for reasons that we do not understand—the results are consistent with the commonsense view that the provision of facilities—in this case a library—enhances the related skill, and does so in addition to the effects of class size.

#### IV. SUMMARY AND CONCLUSIONS

On the eve of democratic elections in South Africa, as remains the case today, educational resources were (and are) sharply different by race, with pupil-teacher ratios in Black schools more than twice as high as those in White schools. In this environment we find that poorly resourced schools, defined as those with high pupil-teacher ratios, discourage educational attainment conditional on age, lower test scores, and lower the probability of being enrolled in education. These results for educational attainment and enrollment are based on a large sample of pupils and schools; those for test scores on a much smaller subsample. The effects of the pupil-teacher ratio on attainment and enrollment are confined to Blacks, consistent with the view that at the small class sizes that characterize education for the other racial groups, reductions in class size have little or no effect. Pupils in better-off Black households do better in their education, and we find no parallel for Whites. That the education of Blacks but not Whites is constrained by financial resources is further supported by the fact that many Blacks who are not in school—but not Whites—report lack of resources as the reason. Although not reported here, the results on educational attainment and enrollment can be very closely replicated using data from the much larger (30,000 versus 8,800 households) 1995 October Household Survey conducted by the South African Central Statistical Office two years after the survey used here.

These results differ sharply from what is often thought to be a consensus, that school resources do not matter very much. In his much cited review and meta-analysis, Hanushek [1995] writes “research demonstrates that the traditional approach to providing quality—simply providing more inputs—is frequently ineffective,” and Hanushek’s conclusion is cited by officials in South Africa in

support of the current policy of aiming for class sizes that are closer to apartheid levels for Blacks than those for Whites. Even so, our results are less exceptional than suggested by Hanushek's meta-analyses or those of Harbison and Hanushek [1992] and Fuller [1986].

There are two main points. First, and as emphasized by Harbison and Hanushek, the studies included in the meta-analyses are of very mixed quality. Many have never been peer reviewed, some are no longer available, and when they are, it is often difficult to discover exactly what data were used and how the analysis was done. The descriptions of econometric procedures—including the Gauss-Markov theorem—are sometimes so exotic as to raise serious doubts about the validity of the results. Nor do the meta-analyses tell us what other variables are being controlled for in any given study. For example, the insignificance of class size on outcomes means something different depending on whether or not expenditure per pupil is also included in the regression. We would argue that the only reasonable inference from a meta-study is that the mass of this literature permits no conclusion whatever.

Second, if we disregard the meta-analyses and move instead to individual studies, the conclusions are often far from "resources do not matter." Once again, a key issue is conditioning. Many studies are concerned with the estimation of detailed educational production functions that try to sort out the effects of different educational inputs, including class size, but also such factors as classroom aids, blackboards, textbooks, uniforms, and so on. In such analyses, class size is sometimes important and sometimes not. For example, in the analyses of seven countries reported in Lockheed and Komenan [1989] and Jimenez and Lockheed [1995], regressions for achievement scores include such variables as aptitude test scores, whether or not the class follows an enriched curriculum, the level of child expectations about the extent of future education, as well as time spent by students at the blackboard or listening to lectures, and time spent by teachers maintaining order, doing administrative duties, or setting tests. It is hard to know how pupil-teacher ratios could be reduced while holding such variables constant, or why the insignificance of class size in such regressions—and in several of these studies it shows a significant negative effect on outcomes—has any relevance for a national policy on class size.

By the same token, it should be reemphasized that our results

apply to changes in the pupil-teacher ratio in South Africa that would hold constant the mix of teachers and other resources, and perhaps most importantly, that would hold constant the quality of teachers. We have little doubt that the positive effects of reducing class size would be reduced if the quality of teachers were to be reduced in an attempt to expand the number of teachers.

#### DATA APPENDIX

##### *A.1. The Main SALSS (SALDRU1)*

The main South African Living Standards Survey (SALSS) was in the field during the last five months of 1993. This is our main source of household and individual data, and we refer to it as SALDRU1. The survey collected data from 8848 households in 360 clusters in a nationwide survey, including what were then described as “independent” homelands. The sample was stratified by province, and used a two-stage self-weighting design in which clusters were selected with probability proportional to size, and an equal number of households selected from each. The clusters are sometimes well-defined communities, particularly in rural areas, but often are not. In urban areas the clusters may be no more than enumeration areas (census tracts), while in many rural areas in South Africa there are no well-defined villages that would make natural clusters or communities. In spite of the self-weighting design, weights had to be introduced *ex post* to compensate for various practical difficulties, such as a few clusters that were too dangerous to be visited, and higher rates of refusal by Whites. However, the weights are not very variable, particularly within race and, for most calculations, it makes little difference whether or not they are incorporated.

One problem with SALDRU1 should be noted. Interviewers were asked to weigh and measure all children aged six and younger. In order to shorten the interviews, some younger children were reported as being seven years old. This misreporting is likely to cause problems with other survey responses for those reported as being seven year olds. There are too many of them in the survey, and if there appear to be too few seven year olds in school, it may be because some of them are actually younger. To

avoid the problem, we often report results only for children aged eight and above.

*A.2. The SALSS Community Questionnaire (SALDRU2)*

The main SALSS was supported by a community questionnaire which was administered to “knowledgeable” local people. The sections of the questionnaire cover the demographic composition of the cluster (the main population groups by race and religion), as well as economic, educational, and health infrastructure. The section dealing with education asks whether there is a primary school in the community, if so, how many, and, if not, how far it is to the nearest one. For each of up to three primary schools attended by children in the cluster, information was obtained on whether it was government or private, whether a school for boys or girls or both, the number of classrooms, whether or not it had a library, a sports ground, and a swimming pool, and the number of students and teachers. The same information was collected for up to three secondary schools used by the community, with the addition of questions on whether the secondary school had a science laboratory, and whether it taught academic subjects only, technical subjects only, or a mixture of both.

In principle, the community questionnaire would allow us to match school facilities to each household in SALDRU1 but, in practice, there are serious problems. The survey organizations collecting the data did not treat the community questionnaire as seriously as they did the main survey, and the information is missing for a substantial number of clusters. For primary schools the bulk of the questions were answered for between 271 and 281 out of 360 clusters, but two of the most important questions, on the numbers of pupils and teachers in each school, were provided for any school by only 137 and 156 clusters, respectively. For secondary schools the situation is even worse. For the questions other than those about numbers of pupils and teachers, there are answers for between 202 and 230 clusters, but we have data on pupils and teachers for even one school for only 113 and 110 clusters, respectively. We have no direct information on why some clusters were covered and others not, but it is clear that there is serious risk of selection bias if we simply drop the households who live in clusters without the necessary information.

## REFERENCES

- Angrist, Joshua D., and Victor Lavy, "Using Maimonides' Rule to Estimate the Effect of Class Size on Scholastic Achievement," *Quarterly Journal of Economics*, CXIV (1999), 533-575.
- Boozer, Michael, Alan Krueger, and Shari Wolkon, "Race and School Quality Since *Brown v. Board of Education*," *Brookings Papers on Economic Activity: Microeconomics* (1992), 269-338.
- Bond, Horace Mann, *The Education of the Negro in the American Social Order* (New York: Prentice-Hall, 1934).
- Fuller, Bruce, "Raising School Quality in Developing Countries: What Investments Boost Learning?" World Bank Discussion Paper No. 2, 1986.
- Fuller, Bruce, Pundy Pillay, and Neeta Sirur, "Literacy Trends in South Africa: Expanding Education While Reinforcing Unequal Achievement," South African Labor and Development Research Unit, mimeo, 1995.
- Hanushek, Eric, "The Economics of Schooling: Production and Efficiency in Public Schools," *Journal of Economic Literature*, XXIV (1986), 1141-1177.
- \_\_\_\_\_, "Interpreting Recent Research on Schooling in Developing Countries," *World Bank Research Observer*, X (1996), 227-246.
- Harbison, Ralph W., and Eric A. Hanushek, *Educational Performance of the Poor: Lessons From North Eastern Brazil* (Oxford: Oxford University Press for the World Bank, 1992).
- Jimenez, Emmanuel, and Marlaine E. Lockheed, "Public and Private Secondary Education in Developing Countries: A Comparative Study," World Bank Discussion Paper 309, 1995.
- Kremer, Michael, "Research on Schooling: What We Know and What We Don't. A Comment on Hanushek," *World Bank Research Observer*, X (1996), 247-254.
- Krige, Dulcie, Sandy Cairns, Bulelwa Makalima, and Di Scott, *The Education Atlas of South Africa*, (Durban: Education Foundation, 1994).
- Krueger, Alan B., "Experimental Estimates of Education Production Functions," *Quarterly Journal of Economics*, CXIV (1999), 497-532.
- Lockheed, Marlaine E., and Andre Komenan, "Teaching Quality and Student Achievement in Africa: The Case of Nigeria and Swaziland," *Teaching and Teacher Education*, V (1989), 93-113.
- \_\_\_\_\_, "Human Capital, Cognitive Skill and Schooling in South Africa," Chicago, mimeo, February 1995.
- \_\_\_\_\_, "The Collapse of Primary Schooling Returns in South Africa, 1960-90," *Oxford Bulletin of Economics and Statistics*, LVIII (1996), 185-209.
- Mwabu, Germano, and T. Paul Schultz, "Wage Premia for Education and Location by Gender and Race in South Africa," Yale University mimeo, January 1996.
- Republic of South Africa, *School Registry of Needs 1997* (1997).
- Republic of South Africa, *The Financing of Education in Terms of the Constitution of The RSA, 1983 (Act No. 110 of 1983)*, National Education Policy Branch, Department of National Education, Report: NATED 02-329 (94/05), May 1994.
- South African Institute of Race Relations, *South Africa Survey 1996/97* (Johannesburg: South African Institute of Race Relations, 1997).