Southern Africa Labour and Development Research Unit

SALDRU

Schooling as a Lottery: Racial Differences in School Advancement in Urban South Africa

by
David Lam, Cally Ardington & Murray Leibbrandt

WORKING PAPER SERIES
Number 18

About the Authors and Acknowledgments

David Lam is Professor of Economics and Research Professor in the Population Studies Center at the University of Michigan. Murray Leibbrandt is Professor of Economics and Director of the Southern Africa Labour and Development Research Unit (SALDRU) at the University of Cape Town. Cally Ardington is a Senior Research Officer in SALDRU at the University of Cape Town. Support for this research was provided by the U.S. National Institute of Child Health and Human Development (Grants R01HD39788 and R01HD045581), the Fogarty International Center of the U.S. National Institutes of Health (D43TW000657), and the Andrew W. Mellon Foundation. An earlier version of this Working Paper was published as Economic Research Southern Africa (ERSA) Working Paper, No. 56, 2007. Funding from ERSA for the ERSA Working Paper and comments from the ERSA referees are both gratefully acknowledged.

Recommended citation

Lam D, Ardington C & Murray Leibbrandt (2008) Schooling as a Lottery: Racial Differences in School Advancement in Urban South Africa. A Southern Africa Labour and Development Research Unit Working Paper Number 18. Cape Town: SALDRU, University of Cape Town

---

ISBN: 978-0-9814031-8-2

Working Papers can be downloaded in Adobe Acrobat format from www.saldru.uct.ac.za.
Printed copies of Working Papers are available for R15.00 each plus vat and postage charges.

Contact Details
David Lam
Email Addy: davidl@umich.edu
Cally Ardington
Email Addy: cally.ardington@uct.ac.za
Murray Leibbrandt
Email Addy: murray.leibbrandt@uct.ac.za

Orders may be directed to:
The Administrative Officer, SALDRU, University of Cape Town, Private Bag, Rondebosch, 7701,
Tel: (021) 650 5696, Fax: (021) 650 5697, Email: brenda.adams@uct.ac.za

Schooling as a Lottery: Racial Differences in School Advancement in Urban South Africa

David Lam
University of Michigan

Cally Ardington
University of Cape Town

Murray Leibbrandt
University of Cape Town

Abstract

This paper develops a stochastic model of grade repetition to analyze the large racial differences in progress through secondary school in South Africa. The model predicts that a larger stochastic component in the link between learning and measured performance will generate higher enrollment, higher failure rates, and a weaker link between ability and grade progression. Using recently collected longitudinal data we find that progress through secondary school is strongly associated with scores on a baseline literacy and numeracy test. In grades 8-11 the effect of these scores on grade progression is much stronger for white and coloured students than for African students, while there is no racial difference in the impact of the scores on passing the nationally standardized grade 12 matriculation exam. The results provide strong support for our model, suggesting that grade progression in African schools is poorly linked to actual ability and learning. The results point to the importance of considering the stochastic component of grade repetition in analyzing school systems with high failure rates.

Introduction

Grade repetition is one of the most important problems in educational systems in many developing countries. In sub-Saharan Africa, where the problem is particularly severe, repetition rates are often 20% per grade (Lee, Zuze, and Ross, 2005), contributing both to low average levels of schooling and high schooling inequality. In spite of the wide recognition of the importance of grade repetition, research on the determinants of progress through school remains very limited. The goal of this paper is to advance our understanding of grade repetition by analyzing progress through secondary school in South Africa. More than a decade after the end of apartheid there continue to be large racial differences in schooling outcomes in South Africa. As we will show, grade repetition plays a key role in explaining these differences.

South Africa has almost universal primary school enrollment, with enrollment rates remaining high into the teenage years (Anderson, Case, and Lam, 2001). Ultimate schooling attainment is mostly determined between ages 14 and 22, the years when young people may drop out or fail out of secondary school, may pass or fail their grade 12 matriculation exam, and may or may not go on to post-secondary education. This paper looks at one of the most critical periods in this transition, the period following grades 8 and 9. Using a new panel study of youth collected in Cape Town, we are able to follow students through three years of secondary school. We find large racial differences in grade advancement – 84% of white students who were in grades 8 and 9 in 2002 successfully advanced three grades by 2005, compared to 44% of coloured students and only 32% of African students. While dropping out is one reason for these differences, we show that high rates of grade repetition play a fundamental role. While only 32% of African students in grades 8 and 9 in 2002 had advanced three grades by 2005, 74% were still enrolled in school.

The importance of grade repetition has been pointed out in a number of other developing countries. Gomes-Neto and Hanushek (1994) documented repetition rates of 20-54% per grade in primary school in Brazil. They found that lower student test scores were associated with increased probability of grade repetition, a result consistent with our results below. Jacoby (1994) found that 21% of 7-12 year-olds had repeated at least one grade in Peru. He found that household income and assets reduce grade repetition, concluding that borrowing constraints play an important role. As pointed out by Lee, Zuze, and Ross (2005), grade repetition is an even more serious problem in sub-Saharan Africa than in other regions, with primary school repetition rates of over 20% per grade in many countries. Although the importance of these high rates of grade repetition is widely recognized, research on grade repetition remains limited. This is due in part to data limitations, with few data sets providing direct information on grade repetition. We take advantage of new longitudinal data collected with a strong focus on grade repetition, allowing us to get a clearer picture of this important component of schooling inequality in South Africa.

As a framework for understanding the determinants of progress through school, we develop a stochastic model of grade advancement. From the perspective of a student trying to decide whether to enroll in a given grade and how much effort to invest in school, performance in school in a given year depends on systematic components such as prior learning, student effort, and inputs from home and school, as well as a stochastic component that reflects imperfect links between actual learning and measured performance. We show that high variance in this stochastic component can generate an equilibrium characterized by high enrollment, low effort, and high rates of grade repetition, features that are typical of predominantly black schools in South Africa. We also show that higher variance tends to reduce the impact of variables such as prior learning and household income on the probability of grade advancement.

After developing our theoretical model, we analyze the determinants of grade advancement and school enrollment using a rich set of variables from the Cape Area Panel Study, a recently-collected longitudinal survey of young people in Cape Town. These variables include previous school outcomes, scores on a baseline literacy and numeracy evaluation, and household variables such as income and parental schooling. Our empirical results are highly consistent with our theoretical model. While there is a strong impact of baseline test scores and household income on progress through grades 8-11, the effect is much weaker for African students than for coloured and white students. We interpret this as evidence that the African school environment does a poor job translating ability and resources into measured performance. Also, in line with our model, we find that African students are less likely to drop out of school than coloured students after failing a grade. As a strong test of our model, we show that our results change systematically when we look at pass rates on the nationally standardized grade 12 matriculation exam. The impact of baseline test scores and income are as large for African students as coloured students in predicting pass rates on the grade 12 exam. This suggests that the weaker impact of baseline test scores and income for Africans in grades 8-11 is due to a poor system of evaluation in those grades.

1. Historical Background and Empirical Regularities

A. South African schools and the legacy of apartheid

A series of cross-national standardized tests have shown that South African learners are not internationally competitive (Van der Berg, 2005; Crouch and Vinjevold, 2006). For example, in the comparative international testing program in mathematics and science - TIMMS - South African grade 6 students performed the lowest out of 50 countries (Reddy, 2006). South Africans perform poorly even within Africa (Van der Berg and Louw, 2006). Since the South African population is dominated by non-white groups that were disadvantaged under apartheid, an obvious explanation for this poor performance is that it reflects a lingering legacy of educational inequities from the apartheid era. Some support for this is found in the fact that a small pocket of white South Africans do very well on these international tests amid otherwise poor aggregate performance. Similar disparities emerge when analyzing grade 12 matriculation exam results, especially in mathematics and science (Van der Berg 2005, Bhorat and Oosthuisen 2006).

Moving beyond simple descriptions of these disparities to more detailed explanation has proven to be elusive. There is mounting evidence that the disparities are no longer simply a problem of school access or government budget allocations (Fiske and Ladd, 2004; Crouch and Vinjevold, 2006; Van der Berg, 2005). Indeed, in these dimensions the post-apartheid government has achieved major progress and substantial equalization by race. The literature suggests, however, that progress on enrollments and budget equalization has not led to equalization of educational outcomes. One reason for this is that budget allocations provide an aggregate view of educational equalization that masks remaining inequities at the school level. There has been some reduction in the inequality in pupil-teacher ratios that was shown by Case and Deaton (1999) to have an important impact on inequality in schooling outcomes in 1993. Due to large disparities in school fees, however, the equalization of government funding has not fully equalized pupil-teacher ratios and other school inputs (Fiske and Ladd, 2004; Yamauchi, 2005) Although there are greater possibilities to exercise school choice in the post-apartheid environment, constraints facing students are such that most black students are still in schools with poor educational infrastructure.[1]

[1] Van der Berg (2005) records the beginnings of an increase in within-race inequality in terms of both inputs and outputs as one would expect given increased options for all learners. However, Seloud and Zenou (2003) provide a useful model of this constrained optimization process that shows how hard it is for previously disadvantaged South Africans to improve their schooling through school choice.

A fair amount of research using an education production function approach has analyzed the role of these input inequities on educational performance (Case and Deaton, 1999; Crouch and Mabogoane, 1998, 2001; Van der Berg, 2005; and Bhorat and Oosthuisen, 2006). The overriding conclusion from these studies is that even after controlling for infrastructure differences a large part of student performance remains unexplained. This has led van der Berg (2005) and Bhorat and Oosthuizen (2006) to speculate about the role of less quantifiable aspects of school quality such as school management and teacher quality. Since these variables are not well measured in the school-level data that are used for these studies, it is difficult to asses the role of these factors.

Detailed classroom-level and school-level research by South African educationalists gives a sense of some of these realities. Hoadley (2007) concludes that South African schools are "struggling to meet their current educational mandates in their three core functions: teaching, learning and management." She documents high teacher absenteeism, especially in more poorly resourced schools. Even in schools in which there is not a culture of absenteeism, a recent teacher workload study shows that significant teaching time is spent in non-teaching functions. Chisholm *et al* (2005) find that teachers spend only 46% of formal school time on teaching and learning, with as little as 10% of the time spent on teaching in some schools. School governance problems are also serious, with a recent ministerial review of school governance showing that in most cases school management teams cannot fulfill the functions allocated to them (DOE, 2004). Finally, the culture in the schools is often far from a culture of learning with evidence of violence and sexual abuse in some schools (Brookes et al, 2004; Human Rights Watch, 2001; USAID, 2003).

It is widely recognized that school infrastructure and school-level variables are only part of the story. Another important long-run impact of apartheid (and apartheid education in particular) is that it leaves black parents and black communities without the resources to create a favorable home environment for learners. Some researchers attempt to control for this in school-level studies by merging community level socio-economic variables into the school-level analyses. However, as these variables are only loosely connected to any actual learner and to any school, this is not completely convincing. Given the limitations of school-based research, there would seem to be high returns to changing the perspective to that of the student and viewing post-apartheid education through their eyes in the context of their household, community and school. Such a view can be pursued with household surveys. We analyze a new panel study of youth that provides us with much more detailed information on young people, their households, and their communities, than is available in school-based studies. Another important strength over school-based data is that we are able to follow young people over time, whether or not they remain in school. This permits us to study, for example, whether failing a grade leads students to drop out of school, and allows us to link baseline characteristics with later school outcomes.

A major focus of this paper is the comparison of schooling outcomes for African, coloured, and white youths. These three population groups were subject to very different treatment under apartheid. Whites had advantages in a wide range of areas, including significantly higher expenditures on schooling, privileged access to the labor market, unrestricted residential mobility, and better access to social services. Africans had the least access to services and the most restrictions on work and migration, with a large gap in expenditures on schooling. The coloured population, which is heavily concentrated in the Western Cape (including Cape Town), occupied an intermediate status under apartheid, with higher expenditures on schooling, fewer restrictions on residential mobility, and better access to jobs than Africans. This history of racial inequity in education is more than a matter of historical interest. There are still strong inequalities in the current South African education system ranging from small groups of world class schools to a large set of poorly functioning schools. This spectrum continues to bear a strong racial footprint. The same is true of the home circumstances of the students in these schools. As we document below, there continue to be enormous racial differences in variables such as household income, differences that we will take account of in analyzing progress through school.

B. Data: The Cape Area Panel Study

This paper uses the Cape Area Panel Study (CAPS), a longitudinal study of youth and their families in metropolitan Cape Town. Details about CAPS, a collaborative project of the University of Cape Town and the University of Michigan, are available in Lam, Seekings, and Sparks (2006)[2]. Wave 1, collected in 2002, included 4,752 young people aged 14-22. Cape Town has three predominant population groups – the distribution in the 2001 census was 48% coloured, 32% African/black, and 19% white. CAPS oversampled areas classified as predominantly African and white. Cape Town is the only major city in South Africa to have substantial numbers of white, coloured, and African residents, providing unique opportunities to study the changing nature of inequality after the end of apartheid.

CAPS Wave 1 included a household questionnaire providing data on all household members. A young adult questionnaire was administered to up to three residents aged 14-22, covering a wide range of variables including schooling, employment, and fertility. It also included a literacy and numeracy evaluation (LNE) which will feature prominently in our analysis. Youth respondents were interviewed a second time in either 2003 or 2004 and a third time in 2005. Table 1 shows sample sizes in Wave 1 and Wave 3 for respondents who were enrolled in grades 8 to 12 in 2002, the sample used in the analysis below. As seen in Table 1, there were 2,344 respondents in grades 8-12 in Wave 1, 47% of whom were African. The "weighted percent" column shows that when we adjust for oversampling the African group is 30% of those enrolled in grades 8-12. The white sample is considerably smaller, a result in part of a lower white response rate.[3] The overall attrition rate between Wave 1 and Wave 3 was 19%, with significant differences across population groups. The African attrition rate is 22%, with most attrition due to migration back to the rural Eastern Cape province that is the main sending region for Africans living in Cape Town. The coloured population has its roots primarily in Cape Town, a factor contributing to its lower 11% attrition rate. The 32% attrition rate for whites includes both migration out of Cape Town (including out of South Africa) and a significant number of refusals.

C. School enrollment, grade repetition, and work

This section provides an overview of key patterns in school enrollment, grade repetition, and labor force activity that form a backdrop for the school transitions we analyze below. Figure 1 shows three indicators of schooling at each age from 6 to 20 based on retrospective reports of CAPS respondents aged 20-22 in 2002. The top panel shows the proportion who were enrolled in primary or secondary school (through grade 12) at each age. Enrollment rates for all groups are close to or above 90% for all ages between 9 and 15, with female enrollment slightly higher than male enrollment for all three population groups until around age 18. Africans lag behind in starting school. Only 80% of Africans were in school at age

---

[2] Additional detail and technical documentation is available on the CAPS web site, www.caps.uct.ac.za.

[3] As in most South African household surveys, CAPS response rates were high in African and coloured areas and low in white areas. Household response rates were 89% in African areas, 83% in coloured areas, and 46% in white areas. Young adult response rates, conditional on participation of the household, were quite high, even in white areas. Given household participation, response rates for young adults were 93% in African areas, 88% in coloured areas, and 86% in white areas (Lam, Seekings, and Sparks, 2006).

8, compared to 99% for coloured and white 8-year-olds. Above age 9 Africans have enrollment rates of 95% to 99%, similar to those of coloureds and whites. Coloured enrollment rates begin to fall above age 15, with Africans having higher enrollment rates than coloureds at all ages above 15. Enrollment rates for whites drop rapidly at age 18, a reflection of the fact that most whites complete grade 12 by that age.[4]

The second panel of Figure 1 shows the number of primary and secondary grades completed at each age for 20-22 year-old Wave 1 respondents. Whites advance almost one grade per year, reaching a mean of about 8 grades by age 14. Although coloureds start school at a similar age as whites, and have almost identical enrollment rates, they lag behind whites in grade advancement from an early age. By age 14 coloured females were about 0.5 grades behind white females, with a similar gap for males. Africans start school later and advance through grades more slowly. By age 14 grade attainment was 5.8 grades for African males, two full grades behind white males. Because of high enrollment rates for Africans in the late teens, Africans almost catch up with coloured grade attainment by age 20. The figure also shows a female advantage in grade attainment in all three groups. As pointed out by Anderson, Case, and Lam (2001), girls move through school faster than boys, with female schooling exceeding male schooling by about one full grade among recent African cohorts who have finished schooling.

One valuable feature of the CAPS data is that it provides direct measures of grade repetition. For each grade of schooling respondents were asked whether they passed the grade, failed the grade, or dropped out before completing the grade. The bottom panel of Figure 1 shows the cumulative number of primary and secondary grades failed at each age, as reported by respondents age 20-22. Coloured and African students both fail grades at a much higher rate than whites, with higher failure rates for males. African and coloured males have failed an average of one grade by age 17. Taken together, the three panels in Figure 1 document a school environment characterized by almost universal primary education, high enrollment rates up to at least age 16, with grade repetition playing a large role in explaining the racial gap in schooling. Africans have particularly high rates of grade repetition, combined with high enrollment rates into the late teenage years.

While this paper focuses on schooling, it is important to keep in mind the labor market environment faced by youth during and after the school-age years. Decisions by young

[4] A significant fraction of whites continue to post-secondary education. We focus on enrollment through secondary school to demonstrate the continued enrollment in secondary school of Africans above age 18.

people about whether to stay in school and how much effort to apply to school will be affected by the opportunity cost of their time and by the expected impact of schooling on wages and employment. Table 2 shows the percentage of young people who did any work for pay or family gain during the 12 months prior to the CAPS Wave 1 survey in 2002. Work is defined broadly, including any work done during the year. This includes work during school vacations, so work is not necessarily directly competing with school. As the table shows, there are enormous differences in the work experience across racial groups. At age 17, over half of white males and females report having worked in the last year, compared to only 1% of African females and 7% of African males. Coloured youth are in between, with 26% of both males and females having worked in the last year at age 17. At age 22 only 24% of African female and 35% of African males report having worked in the last year, compared to over 75% of the other four gender/race groups.

Summarizing the patterns in Figure 1 and Table 2, African teenagers in Cape Town have high rates of school enrollment, high rates of grade repetition, and low rates of employment. These patterns are similar to those for African youth in all of South Africa (Anderson et al., 2002). Limited labor market opportunities, driven in part by spatial segregation that is a legacy of apartheid, is presumably important in explaining both low employment and high school enrollment. Coloured youth have significantly higher employment rates than African youth, a reflection of both closer geographic proximity to jobs and the legacy of coloured labor preferences that existed in the Western Cape under apartheid. There appears to be more of a tradeoff between school enrollment and work among coloured youth, especially for males. Whites have the highest rates of employment along with the highest levels of school enrollment and schooling attainment, an indication that work and school in the teenage years are not entirely incompatible.

2. A Stochastic Model of Grade Repetition

In this section we develop a theoretical model that provides a framework for understanding school enrollment and progress through school in South Africa. We focus in particular on the combination of high enrollment rates and high rates of grade repetition documented in Figure 1. Making schooling decisions in an environment with high failure rates may lead to a number of important outcomes. One is that students and their parents may find it difficult to predict their probability of success in a given school year. Given large crowded schools with limited resources, it is hard for students and parents to control or predict key inputs such as the quality of their teachers, the ability of their classmates, or the financial status of the school. In addition to the fact that learning is compromised in such an environment, it is also likely that the evaluation of learning is imperfect. While students everywhere tend to rationalize failure to be the result of bad luck, there may be more truth to these perceptions in poor schools with high failure rates.

To model this environment, consider a stochastic model of grade advancement. Suppose that students are evaluated at the end of the school year based on a final score *S*. One component of this score is the students' actual knowledge at year's end, which we characterize by a learning production function $K=F(\mathbf{X})$, where **X** is a vector of inputs that affect knowledge at the end of the year, including prior knowledge, effort, school inputs, and family background characteristics such as parental schooling. The score for student *i* also includes a stochastic component u_i reflecting a wide variety of reasons for discrepancies between knowledge and scores. In the most literal sense, these include errors in marking exams and recording grades. More broadly they include problems in the school environment that causes learning to be unrewarded in the evaluation of students to determine advancement. For example, weak teachers in bad schools may teach and test in such a disorganized way that mastery of course material has little impact on final evaluations. Our review above drew attention to some frequently cited problems in South African schools such as teacher absenteeism and disorganized school administration. These may all contribute to an environment in which there is a weak link between actual learning and measured performance.

Assume that we can summarize this environment with a linear model

$$S_i = \boldsymbol{\beta}'\mathbf{X}_i + u_i \tag{1}$$

where **X** is a vector representing the systematic determinants of student performance and *u* is a stochastic component that is uncorrelated with the variables in **X**. We assume that there are a large number of independent components in *u*, making it reasonable to assume that it is normally distributed, $u \sim N(0,\sigma)$. We assume that perceptions of students and their parents about the process driving Equation (1) are consistent with reality, but for purposes of understanding decisions about enrollment and effort it is only necessary to assume that Equation (1) describes those perceptions, whether or not they are correct. Students pass the current grade if $S_i > T$, where *T* is a threshold established for all students at the same grade. The probability of passing is

$$P(S_i > T) = P[u_i > T - \boldsymbol{\beta}'\mathbf{X}_i] = 1 - \Phi\left[\frac{T - \boldsymbol{\beta}'\mathbf{X}_i}{\sigma}\right], \tag{2}$$

where Φ is the cumulative of the standard normal distribution.

A. The effect of characteristics on passing

We can use Equation (2) to analyze the impact of characteristics on the probability of passing. Consider some characteristic which is one component of **X**, such as previously acquired human capital, parental education, or a measure of school quality. Denote this variable by X_1, and its corresponding coefficient in Equation (1) by β_1. To be concrete, consider the impact of mother's schooling on the probability of passing, assuming that one year of mother's schooling increases a student's score by β_1 points. We differentiate Equation (2) to get:

$$\frac{\partial P(S_i > T)}{\partial X_{1i}} = \frac{\beta_1}{\sigma}\phi\left(\frac{T - \boldsymbol{\beta}'\mathbf{X}_i}{\sigma}\right) = \beta_1 f\left(T - \boldsymbol{\beta}'\mathbf{X}_i\right), \tag{3}$$

where ϕ is the density of the standard normal distribution and f is the density of the normal distribution with mean zero and standard deviation σ. Equation (3) has a number of interesting implications. First, since the density of u is highest at the mean (assumed to be zero), the effect of an increase in X_1 will be largest for students who would be close to the passing threshold independent of their draw from u. Students in either the high end or the low end of the distribution will have little effect on their probability of passing if they raise their score by one more point, while students close to the threshold will have a large effect.

It is also clear from Equation (3) that the marginal effect of characteristics depends on the standard deviation σ. Evaluated near the mean, the effect of X_1 is a negative function of the standard deviation σ. For those near the passing threshold, a higher variance in the random component of the score will reduce the marginal payoff to an extra point, and thus reduce the impact of characteristics. The intuition is straightforward. The only time an increase in the score will affect whether the student passes or fails is when the student is just below the threshold. For students who would have been just above the threshold based on the deterministic component alone, their probability of being at the threshold decreases as the variance of the stochastic component increases. Defining $g_i = T - \boldsymbol{\beta}'\mathbf{X}_i$ as the gap between the deterministic component of the score and the threshold, and taking the derivative of Equation (3) with respect to σ,

$$\frac{\partial^2 P}{\partial X_{1i}\partial\sigma} = \left(\frac{\beta_1}{\sigma}\right) f\left(g_i\right)\left[\frac{{g_i}^2}{\sigma^2} - 1\right]. \tag{4}$$

The cross-partial derivative in Equation (4) is negative when $|g_i| < \sigma$ and positive when $|g_i| > \sigma$.

Figure 2 illustrates several features of the model. The top panel shows the density for three normal distributions with mean zero, with standard deviations equal to 1, 1.1, and 2. The bottom panel shows the corresponding cumulative distributions. Point A in the top panel, which sits at one standard deviation for the distribution with σ=1, is one of two fixed points of the density when σ is increased (the other being at -σ). An increase in the standard deviation causes a decrease in the density in the range within one standard deviation of the mean and an increase in the density outside that range. This implies that when the deterministic component is within one standard deviation of the passing threshold, an increase in the variance of the stochastic component will reduce the impact of characteristics on the probability of passing. Consider two students with identical characteristics who are in schooling systems that differ only in the variance of the stochastic component, with $\sigma_2 = 2\sigma_1$, and with the same β_1 coefficient in both regimes. Suppose both students are exactly at the passing threshold based on the deterministic component, $T - \boldsymbol{\beta}'\mathbf{X}_2 = T - \boldsymbol{\beta}'\mathbf{X}_1 = 0$, implying that they will pass if the draw is positive and fail if the draw is negative. Looking at the densities around zero and at Equation (4), we see that the marginal effect of one additional point on the probability of passing is twice as high in the low-variance regime. Put another way, an increase in the value of some characteristic that causes a β_1 point increase in the deterministic component of the grade will have twice as large an impact on the probability of passing in the low-variance environment, evaluated for a student near the passing threshold.

When the deterministic component is in either tail outside one standard deviation of the threshold, an increase in the stochastic variance will increase the impact of characteristics. Since 68% of the draws from the stochastic component will be within one standard deviation, most draws will lie in a range in which increased stochastic variance leads to a decrease in the impact of characteristics on passing.[5] Comparing two identical students in one of the tails of the distribution, the student in the high-variance regime will have a *higher* impact of characteristics on the probability of passing. The intuition is straightforward. A student at the top of the distribution in terms of the systematic component would pass with certainty if the variance of the stochastic component were zero. An increase in the stochastic variance increases the probability that she will get a large negative draw from the stochastic component, and increases the payoff to raising her score by one point. For a student at the bottom of the distribution, a one point increase would have almost no chance of pushing her over the passing threshold in a low-variance regime. But in a high-variance regime there is a greater chance that her one additional point will be matched by a high positive draw from the stochastic component, pushing her over the threshold.

Equation (3) also reminds us of an econometric point that will be important in our empirical analysis below. If we estimate a standard probit regression of the probability of passing on some characteristic, the regression gives us an estimate of β/σ. If we estimate different regression coefficients for two different groups, in general we will not be able to distinguish between differences in the marginal impact of the characteristic on human capital accumulation (differences in β) and differences in the variance in the process that determines grade promotion (differences in σ). However, if the educational environment provides a situation in which we would expect the differences in σ between groups to decrease, we can use this to make some headway in distinguishing between differences in β and σ. As explained in part F below, a nationally standardized exam at the end of secondary schooling provides such a situation.

B. How hard do students work?

One set of variables in the learning production function is inputs of student time and effort – variables such as the amount of time spent on homework and the number of days attending school. The results derived above for the impact of characteristics on the probability of passing can be applied to the impact of effort. The results imply, for example, that an increase in the variance of the stochastic component will decrease returns to effort for

[5] This does not necessarily mean that the majority of students are in this range, since that depends on the distribution of the deterministic component. It is possible, for example, that most students are well above the passing threshold and are relatively unaffected by the stochastic component. This is discussed below.

students who are within one standard deviation of the passing threshold. Effort will presumably have returns in addition to its impact on the probability of passing, but this will be one important component of those returns.

An important difference between effort and other characteristics is that the level of effort is endogenous. Changes in the marginal returns to effort will presumably lead to adjustments in the amount of effort supplied. Assume that students equate the marginal return to their time spent on school to the opportunity cost of that time in other uses. This opportunity cost could be the wage in labor market work or the marginal utility of leisure. If there is a decline in the marginal return to time spent on school, the student can be expected to reduce that time to re-equilibrate the marginal returns across all uses of time. We will thus expect less effort from students who are within one standard deviation of the passing threshold if we increase the stochastic variance.

C. Who goes to school?

Assuming that school enrollment is a voluntary decision by children and/or their parents, those who enroll in a given school year will be those for whom the expected benefits exceed the expected costs. As an extreme simplification, suppose that attending school in any one particular year has zero payoff if the student does not pass that grade. If school enrollment requires no out-of-pocket expenses and has no opportunity cost, then every student should enroll since every student has some probability of passing the grade. Even for those with deterministic components of their final score below the threshold *T*, there is some probability that they will get a lucky draw from *u* and end up with a passing score. This might be thought of as the chance that the few things they learn happen to be the ones that get asked on the final exam, or that the teacher will incorrectly mark the exam. More realistically, there are in fact both direct costs and opportunity costs to being in school. There will therefore be some threshold probability of passing required for students to enroll in a given year. To see how this probability is affected by the variance of the stochastic component, we take the derivative of Equation (1) with respect to σ:

$$\frac{\partial P(S_i > T)}{\partial \sigma} = \left(\frac{T - \boldsymbol{\beta}'\mathbf{X}_i}{\sigma^2}\right)\phi\left(\frac{T - \boldsymbol{\beta}'\mathbf{X}_i}{\sigma}\right), \qquad (5)$$

Equation (5) tells us that the probability of passing increases with σ for those who would have failed based on the deterministic component, and decreases for those who would otherwise have passed. As σ increases, the expected probability of passing the grade is increasingly determined by the stochastic component. Consider two groups of students, a low-skilled group for whom $T > \boldsymbol{\beta}'\mathbf{X}_1$ and a high-skilled group for whom $T < \boldsymbol{\beta}'\mathbf{X}_2$. If the distribution of opportunity costs were the same for the two groups, then we would expect a higher fraction of the high-skilled group to be in school. But an interesting implication of the model is that an increase in the variance of *u* would tend to decrease the enrollment of high-skilled students at the same time that it increases the enrollment of low-skilled students. The reason is that the probability of high-skilled students passing goes down because of the increased chance of getting draws from the bottom of the distribution. The probability of low-skilled students passing goes up because of an increased probability of getting a draw large enough to push them over the passing threshold. An increase in the variance would therefore have the potential to diminish the difference in enrollments between low-skilled and high-skilled students, *ceteris paribus*.

The point is illustrated in Figure 2. The CDF in the bottom panel shows the probability of passing for different values of $\boldsymbol{\beta}'\mathbf{X}_i - T$. Suppose that given the opportunity cost of attending school, students only enroll if they have a 30% chance of passing. If σ=1, only those to the right of the line at Point B enroll. Following the line to the top panel, we see the marginal impact of characteristics for the lowest scoring students. If we increase the standard deviation to σ=2, the line dividing those with 30% probability of passing shifts to the left to Point C. As the stochastic variance increases, weaker students enroll from the bottom of the distribution. Looking at the distribution of the two densities to the right of Points B and C, we see that the average return to characteristics will tend to be lower when σ=2, although a precise statement depends on the distribution of the deterministic component.

It is easy to simulate examples that illustrate the model's predictions. Suppose, for example, that the passing threshold is 50 points. There are two groups of students, each of which would have a mean of 55 and a standard deviation of 10 in the absence of the stochastic component. The stochastic component has mean zero for both groups, with σ=10 for Group A and σ=20 for Group B. If students enroll who have an expected probability of passing of at least 30%, then we will observe the following: the enrollment rate will be 85% in Group A and 94% in Group B; the passing rate among those enrolled will be 72% in Group A and 61% in Group B; the average impact of one additional point on the probability of passing will be 54% larger in Group A than Group B. In other words, we generate differences that are similar to those observed between African and coloured students in Cape Town – the group with the larger stochastic component has higher enrollment, lower pass rates, and a lower impact of characteristics on passing.

D. Effects over multiple years

Following students over multiple grades, we can generalize (1) by adding a subscript t and making assumptions about the correlation of stochastic terms across years. The simplest case is to assume that u_{t+k} is uncorrelated with u_t for all k, an assumption that fits our characterization of the idiosyncratic nature of the stochastic term. The probability of passing all years from year 1 to year n is $P_{i,1n} = \prod_{t=1}^{n} P_{i,t} = \prod_{t=1}^{n} P(S_{i,t} > T_t)$. Consider the simple case in which $\mathbf{X}$, $\boldsymbol{\beta}$, σ, and T are the same every year, so that students would get the same score each year in the absence of the stochastic term. If the stochastic terms are uncorrelated across years then the probability of passing is identical every year, P_1=P_t= P_n. To analyze the impact of characteristics on the probability of passing n grades, it is helpful to take logs and look at the proportional impact. Taking the derivative of ln$(P_{i,1n})$ with respect to a characteristic X_1, and recalling the result from (3), we get

$$\frac{\partial \ln P_{i,1n}}{\partial X_{1i}} = \frac{n\beta_1}{\sigma P_t}\phi\left(\frac{T - \boldsymbol{\beta}'\mathbf{X}_i}{\sigma}\right), \tag{6}$$

where P_t is the probability of passing in any single year. Note that this is equivalent to (3) in the case where n=1 and we take the derivative of the log.

Equation (6) shows that the proportional impact of characteristics on the probability of passing increases as we track progress over more grades. If one additional IQ point gives a student a 2% higher probability of passing a single grade, then it will give the student a 10% higher probability of passing five consecutive grades.[6] While (6) is derived for the simple case in which every grade is identical, the result is quite general as long as some component of the stochastic component is uncorrelated across years. The stochastic component introduces noise into each year's results, causing some weak students to pass over better students and weakening the link between ability (for example) and scores. Over multiple years the better students pull ahead, as the systematic component dominates the uncorrelated stochastic component. The uncorrelated components, which in our model represent noise in the link between learning and evaluation, become less important when we look across more years. This implies that the proportional impact of characteristics on pass rates will be larger when we look at passing over multiple years.

E. Impact of failing on enrollment and future success

Another implication of the model is that the impact of failing grades on future enrollment will depend on the magnitude of the stochastic component. If students are uncertain about their own ability and likelihood of future success in school, then each year's scores (and decisions about promotion) are important signals about that ability. A larger stochastic component in scoring implies that grade promotion is a noisier signal about the student's ability and future probability of success. We expect, then, that past failure will be a weaker predictor of future enrollment in a regime with higher variance. Past failures will also be a weaker predictor of future probabilities of passing in the high-variance environment, since high variance weakens the link between failure and actual learning. While it is perceived variance that affects the link between past failure and future enrollment decisions, the actual variance will be important for the link between past failure and future promotion. The perceived variance may also play an additional role by affecting the impact of past failure on future effort, which in turn will affect future probabilities of passing.

F. Externally evaluated standardized exams

An important feature of the South African school system is the nationally standardized, externally evaluated matriculation exam given to all students at the end of grade 12. These exams are set in a process run by the Education Department. Exam papers are taken from the schools, pooled, and graded by external evaluators, with only the results sent back to the school. Performance on this exam has important consequences for both students and schools, with extensive media coverage of matric pass rates when they are announced each December. Preparation for the matric exam is a major focus of student effort during grade 12. The matric exam provides an interesting test of our model, since it implies that there are important differences between passing grade 12 and passing grades 8-11. Since the standardization and external evaluation should lead to both a reduction in variance and smaller differences in the variance of the stochastic component across racial groups, we would expect there to be a

---

[6]More precisely, a .02 log probability difference in 1 year implies a .10 log probability difference in 5 years.

larger impact of characteristics on pass rates and smaller racial differences in the impact of characteristics.

While the cases discussed above are highly stylized, some important realism is captured by the model. The most important is that it is easy to generate an equilibrium which has the following features – relatively high fractions of students fail each grade, high fractions of the eligible population are enrolled, and high fractions of students expend low effort on school. These are broadly the features observed in predominantly black schools in South Africa. The somewhat puzzling combination of high failure rates and high enrollment rates can be explained by an environment with a large stochastic component to grade advancement. The model has a number of empirical implications that can be analyzed using survey data such as CAPS. Most importantly, both the probability of grade advancement and the probability of enrollment will tend to be less affected by characteristics such as ability and family background in an environment with a larger stochastic component to measured performance. Second, the impact of failing grades on future enrollment and grade advancement will be lower when there is a high stochastic component. Third, the impact of characteristics on passing will be larger when we look at passing over multiple grades than when we look at passing a single grade. Fourth, the impact of characteristics will be larger and the differences between population groups will be smaller for passing the standardized grade 12 matriculation exam than for passing earlier grades.

3. Empirical Evidence

In this section we analyze the extent to which empirical evidence on progress through school is consistent with our stochastic model of grade repetition. We begin with a descriptive overview of grade progression for our sample of 8^{th} and 9^{th} graders in 2002. We then estimate regressions to test some of the specific predictions of the model as described above.

A. Grade Progression between 2002 and 2005

The 2005 CAPS Wave 3 data make it possible to follow the progress of young people who were in school in 2002. We begin by focusing on the experience of respondents who were in grades 8 and 9 in 2002. If they remained in school and passed all grades, these students would have reached grades 11 and 12 by 2005. Table 3 shows the activities in 2005 of those where in grade 8 and 9 in 2002. About 93% of whites who were in grade 8 in 2002 had advanced to grade 11 or 12 by 2005. The experience of African and coloured youth is very different. Among Africans who were in grade 8 in 2002, only 36% had advanced to either grade 11 or grade 12. About the same percentage, 37%, were in grade 10, implying that they had progressed two grades in three years. About 18% of Africans who had been in grade 8 in 2002 were not enrolled in 2005, with only 3% not enrolled and working. Coloured youth who were in grade 8 in 2002 were less likely than Africans to be enrolled in 2005, but those who were enrolled were more likely to have maintained normal grade progression. About 47% were in grade 11 or 12, with 13% in grade 10. A higher percentage of coloured youth appear to have dropped out to work, with 12% in the "not enrolled/working" category. African youth are much more likely to stay in school than coloured youth, in spite of their higher rates of grade repetition. About 82% of African students who were in grade 8 in 2002 were still enrolled in school in 2005, compared to 64% of coloured students.

Results for those who were in the 9th grade in 2002 show similar patterns. Among white students, 85% had reached grade 12 or had graduated to post-secondary schooling by 2005. This compares to 29% for African students and 43% for coloured students. About 30% of African students are in grade 11, and another 11% are only in grade 10. About 70% of African students who were in grade 9 in 2002 are still enrolled in school in 2005, compared to 62% of coloured students. The patterns in Table 3 begin to illustrate some of the predictions of our model. Coloured students have easy access to traditionally coloured schools, schools that are generally of higher quality than traditionally African schools. Coloured students are more likely than African students to make normal progress and to drop out, with African students having both higher enrollment and higher failure rates. These issues will be explored in much greater detail below.

B. Characteristics affecting progress through school

In this section we provide an overview of some of the individual, household, and community characteristics we will use in our regressions. One interesting feature of CAPS is the literacy and numeracy evaluation (LNE) that was administered to all youth respondents in Wave 1. This was a self-administered written test taken after completion of the young adult questionnaire. The test had 45 questions and took about 20 minutes to complete. Respondents could choose to take the test in English or Afrikaans. There was no version in Xhosa, the home language of most African respondents. The English language test was taken by 99% of African respondents, 43% of coloured respondents, and 64% of white respondents. In interpreting the results it is important to keep in mind that most white and coloured students took the test in their first language, while Africans took the test in a second language. It must also be noted, however, that English is the official language of instruction in African schools and is used for many tests such as the grade 12 matriculation exam. We use the LNE scores as a measure of cumulative learning at the time of the 2002 interview. Performance on the test reflects a combination of many factors, including innate ability, home environment, and the quantity and quality of schooling to that point.

Figure 3 presents kernel density estimates of the distribution of the combined literacy and numeracy scores for each population group, using the sample of those enrolled in grade 8-12 in 2002 (the score is standardized to zero mean and unit variance for the full sample of 14-22 year-olds). Racial differences in test scores are striking. There is only a small area of overlap between the test scores of Africans and whites, with a much higher variance among Africans. The distribution of scores for coloureds sits between, with considerable overlap with both the white and African distributions. The mean standardized score is -0.6 for Africans, 0.01 for coloureds, and 1.14 for whites, implying a 1.7 standard deviation gap between whites and Africans. The standard deviation of African scores is 60% larger than the standard deviation of white scores.

Another key variable in our regressions is the log of per capita household income in 2002, as reported by an adult respondent in the Wave 1 household questionnaire. Figure 4 plots kernel densities for the income distribution of each population group, standardized to the mean of the combined population. Once again we see large racial differences, with a difference in mean log income between whites and Africans of almost 2.5. Exponentiated, this implies that white youth in 2002 lived in homes with over 10 times higher per capita household income than Africans. As was the case with test scores, a striking feature of Figure 4 is the very small range in which the African and white income distributions overlap, with the coloured distribution sitting in between.

An additional factor to consider in explaining school progress for 8th and 9th graders is the extent to which students were already behind in school in 2002. As shown in Figure 1, grade repetition is an important feature of the school experience of both African and coloured youth, and by grades 8 and 9 there is considerable variation in the age of students. Figure 5 shows the age distribution for 9th graders in 2002. White 9th graders are concentrated at age 15, with only about 15% at age 16. By contrast, the modal age of African 8th graders is 16, with a wide distribution ranging between ages 14 and 22. Roughly 25% of African 9th graders are age 18 or older.

Table 4 presents descriptive statistics for the sample of students who were enrolled in grades 8 or 9 in 2002, the sample that will be used for our first set of regressions. The first row shows the dependent variable in our first regressions, an indicator of whether the student advanced three grades by 2005 – 8th graders reached at least grade 11 and 9th graders reached at least grade 12. This variable equals 0 for any other outcome, including dropping out before reaching the target grade or being in a grade below the target grade in 2005. As seen in Table 4, the percentage of students advancing three grades varies enormously by race: 32% for Africans, 44% for coloureds, and 84% for whites. Table 4 also shows the percentage that was still enrolled in school in 2003, 2004, and 2005. Looking at enrollment in 2004, which we will use as an outcome in our second set of regressions, there is considerable variation across racial groups. About 94% of whites were enrolled in 2004, compared to 87% of Africans and 78% of coloureds. Table 4 presents three measures of grade failure. The number of grades failed by 2002, which we will use as an independent variable in our first regressions, varies from 0.7 for Africans to 0.5 for coloureds and 0.2 for whites. As shown in the next row, about 50% of Africans failed at least one grade by 2002. The percentage who failed their grade in 2002, which we will use in our regressions analyzing enrollment in 2004, varies from 13% for Africans to 1% for whites.

Table 4 presents means of several household characteristics that will be included in the regression. The large differences in the log of per capita household income were already noted. We also use an indicator of whether the household experienced a negative household shock between 2002 and 2005, as indicated by death of household member, job loss, marital disruption, or loss of a grant or remittance. About 26% of African youth lived in households experiencing a shock, compared to 16% of coloured youth and 4% of white youth. The mothers and fathers of African youth have 4-5 years less schooling than the parents of white youth, with schooling missing for about 40% of Africans[7]. We also include in our regressions the age-sex-specific unemployment rate for individuals with less than 12 years of schooling in the census sub-place. This varies from 81% for Africans to 27% for whites.

Table 4 also presents useful background information about school characteristics, school fees and school mobility. The variables for former department indicate that 77% of African

[7] Parental schooling comes from the household questionnaire when the parent is co-resident, and is collected from the young adult directly when the parent is not co-resident.

youth attend schools that were classified as African schools (Department of Education and Training) under apartheid. About 11% attend formerly coloured (House of Representatives) schools, 3% attend formerly white (House of Assembly) schools, and 9% attend schools that were created since 1994 and hence have no "former department" classification.[8] Note that 87% of coloured students are in formerly coloured schools and 91% of white students are in formerly white schools. The annual school expenditure variable shows the enormous differences in school fees. African students paid an average of 323 rands per year (roughly 32 dollars), coloured students paid 744 rands, and white students paid 5,890 rands. Since these fees are often used to hire extra teachers, the differences in fees translate into differences in pupil-teacher ratios. This is seen in the next row of Table 4, which shows a mean pupil-teacher ratio of 32.6 for Africans, compared to 24.1 for whites. There is some movement between schools, with 19% of Africans changing schools between 2002 and 2003. For all races respondents are most likely to change schools after grade 8. Although a fair number of respondents change schools, very few move to a school that fell under a different department before 1994. Between 2002 and 2003, when most school changes occurred, only 3.8% of Africans, 2.1% of coloureds and no whites changed former department.

C. Probit regression for progress through school

This section presents results of probit regressions in which our dependent variable indicates progress through school between 2002 and 2005. One of our key empirical questions is whether there are racial differences in the impact of individual and household characteristics on grade advancement. Given the differences in school environment discussed above, we hypothesize that Africans have lower β coefficients in the learning production function and/or a higher σ for the stochastic component, both implying that Africans will have lower estimated probit coefficients (β/σ) in a regression of school advancement on characteristics. As in Cameron and Heckman (2001), who estimate separate models for whites, blacks, and Hispanics in U.S. data, we assume from the outset that we should estimate separate regressions for Africans, coloureds, and whites. For each coefficient and each pairwise combination of races we test for equality of the probit coefficients. While it is impossible to distinguish between differences in β and σ from the probit regressions alone,

[8] Due to slow residential de-segregation, new schools are no less racially distinct than schools existing prior to 1994. For example, all new schools attended by African respondents were located in African

we will argue below that restrictions imposed by the standardized grade 12 matriculation exam help us identify the separate contributions of these two components.

Table 5 presents the first set of probits, which analyze the probability that those enrolled in grade 8 and 9 in 2002 advanced three grades by 2005. Columns 1-3 present the coefficients, columns 4-6 present tests of equality of coefficients for each pairwise comparison, and columns 7-9 present marginal effects evaluated at a common set of characteristics across samples. We estimate large effects of the LNE score and the number of previous grades failed, demonstrating the importance of prior learning and school performance. This is consistent with the results of Gomes-Neto and Hanushek (1994), who found that test scores were an important predictor of grade repetition in Brazil. The magnitude of the effects can be seen in columns 7-9, where we present marginal effects evaluated at a constant set of characteristics – a female who was in grade 8 in 2002, one previous failed grade, LNE score and log income at zero (the sample means), parents' schooling equal to 8 years, no household shock between waves, and a local age-sex-specific unemployment rate of 50%. Looking at the coefficients and the marginal effects, we see that previous grades failed has a much less negative effect on grade advancement for Africans than for coloureds and whites. At our assumed baseline characteristics, having failed one additional grade by 2002 is associated with a 7 percentage point decline in the probability of advancing 3 grades for Africans, compared to a 27 percentage point decline for coloureds.

We also see that the LNE score has a smaller positive effect for Africans. A one standard deviation increase in the LNE score is associated with a 12 percentage point increase in the probability of advancing three grades for Africans, compared to 33 percentage point increase for coloureds. The impact of log per capita household income is not statistically significantly for Africans, but is strongly positive for coloureds and whites. In Columns 4-6 of Table 5 we test for the equality of coefficients between pairs of racial groups across regressions. As shown in Column 4, we can reject the hypothesis that Africans and coloureds have equal coefficients on previous grades failed, the LNE test, and income. The small white sample leads to large standard errors on the white coefficients, making it impossible to reject equality of the African and white coefficients on these same variables, in spite of large differences in the point estimates.

The predicted impact of the LNE scores and household income can be seen in the predicted probabilities graphed in Figure 6. The top left panel shows the predicted

---

townships.

probability of advancing three grades as a function of LNE scores, based on the coefficients in the three separate probits by race. The predicted values are calculated for a female in grade 8 in 2002 with parents' education equal to 8 years and all other variables at their means. There is a very strong effect of the LNE scores on grade advancement, with the steepest slope around an LNE score of zero for coloured students (in addition to zero being the overall sample mean, it is close to the mean for the coloured subsample). The slope for Africans is flatter over most of the range between -2 and +2 standard deviations. Referring back to our theoretical model, this flatter slope could result from either a smaller coefficient on LNE scores in the underlying learning production function or from a higher variance in the stochastic component of measured performance (or some combination of the two).

The upper right panel of Figure 6 shows a similar pattern for the impact of income on grade advancement. The impact of income is much smaller for Africans than for white or coloured students. This result may seem surprising, since we might expect to find large effects of income over the range of income covered by the African sample. The poorest part of the African sample is in deep poverty, while the upper tail has income levels that should make it much easier to keep children in school and provide inputs to support their progress through school. It also differs from Jacoby's (1994) results for Peru, where income was an important predictor of grade repetition. Our interpretation of this low impact of income on African grade advancement is that it is a symptom of the inefficient and chaotic school environment, which is ineffective in translating either higher ability or better resources into measurable improvements in school performance.

Figure 6 demonstrates another important implication of our results. The predicted probability of advancing three grades is higher for Africans than for coloureds over a large range of test scores and income at the bottom of the distribution. This includes comparisons at zero, the mean of both the LNE score and income. Our results imply that the gap between coloured and African students in grade progression is more than fully explained by differences in characteristics, with test scores and income explaining most of the difference. Even the enormous gap between Africans and whites can be mostly explained by differences in the variables included in the regressions in Table 5. Recalling Figures 3 and 4, an important caveat to this result is the lack of overlap in the distribution of test scores and incomes. This is most serious in the African-white comparisons, where there is virtually no overlap in the distributions of these two key variables. This means that the predicted values for whites at low levels of income and test scores are largely out-of-sample projections. There is nothing we can do about this lack of common support, since it is simply a manifestation of the enormous racial inequality that continues to exist in South Africa. In the case of the African and coloured comparisons, the problem is less severe. The African and coloured distributions of both test scores and income have considerable overlap, including the range in which Africans have higher predicted probabilities of advancement than coloureds.

Looking at other variables in our probit in Table 5, we find no significant differences in grade advancement of males and females. This is consistent with the patterns shown in Figure 1 and in other research showing that there is no female disadvantage in schooling outcomes in South Africa, at least through secondary school. Parental schooling has surprisingly weak effects on grade advancement, with none of the coefficients statistically significant at conventional levels. This is surprising given the high variance in parental schooling in our sample and the wide range of research that finds strong effects of parental schooling on children's schooling outcomes. For Africans these coefficients continue to be insignificant even when the LNE scores and number of grades failed are omitted (results not shown). For coloured students we estimate a significant positive effect of father's schooling when the previous performance outcomes are omitted. Negative household shocks such as the death or job loss of a household member have significant negative effects on grade advancement for both coloureds and whites, though not for Africans.

The neighborhood unemployment rate is not significant for any group. We include it to capture two possible effects. On the one hand, the opportunity cost of time may affect effort in school or the probability of dropping out. On the other hand, better employment prospects might stimulate young people to stay in school and work harder in school. These effects may be cancelling out, although it is also possible that census subplaces do not capture the appropriate labor market. While white and coloured youth appear to have much better job opportunities than African youth due to geographical proximity, family networks, and language skills, there may not be sufficient geographical variation in job opportunities within racial groups to identify an effect.

D. Regressions for school enrollment

Table 6 presents regressions in which the dependent variable is school enrollment in 2004, continuing to use the sample of respondents who were enrolled in grade 8 or 9 in 2002. We include a dummy variable for whether the respondent failed their grade in 2002 in order to see whether students drop out or return to school after failing. Other variables are the same as those in Table 5. We exclude whites from these regressions because over 95% of whites are enrolled in 2004, making it difficult to estimate meaningful regressions. Columns 1-2 present the probit coefficients, Column 3 shows the test of equality of African and coloured coefficients, and Columns 4-5 show marginal effects evaluated at the same set of characteristics for each sample.

As in Table 5, we estimate significant negative effects of the number of previous grades failed and significant positive effects of LNE scores. The point estimates are larger in magnitude for coloureds than for Africans, though in neither case can we reject equality of the coefficients. Looking at marginal effects, a one standard deviation increase in the LNE score is associated with a 2.5 percentage point increase in the enrollment probability for Africans and a 6.2 percentage point increase for coloureds. The estimated effect of household income is statistically insignificant for Africans but strongly positive for coloureds. While it is surprising that income does not affect African enrollment, we interpret it as indicating that the combination of low opportunity cost, high returns to schooling, and imperfect evaluation make the benefits of being enrolled sufficient to overcome the direct costs such as fees and uniforms. Failing the grade in 2002 has a negative effect on 2004 enrollment for both Africans and coloureds, but the effect is much greater for coloureds and we reject equality of the coefficients. The marginal effects imply that failing in 2002 reduces the probability of enrollment in 2004 by 48 percentage points for coloureds, compared to 2 percentage points for Africans. This is consistent with our interpretation of the response of Africans and coloureds to differences in the school environment. Failing a grade prior to 2002 is a weaker predictor of future success in school for Africans than for coloureds. Consistent with this, Africans are less likely to drop out if they fail their grade in 2002.

The bottom panels of Figure 6 plot the predicted enrollment as a function of LNE scores and per capita household income for Africans and coloureds, with separate predictions for those who failed and passed in 2002. Several features of the graphs are worth noting. First, the lines for Africans are much flatter than the lines for coloureds, showing the much weaker responsiveness of enrollment to prior learning or income for Africans. Second, predicted enrollment is higher for Africans over a broad range of LNE scores and income. Even Africans who failed in 2002 have a higher predicted probability of being enrolled in 2004 than coloured students who passed in 2002 over much of the low range of LNE scores. Finally, we see that the gap in predicted enrollment between those who passed in 2002 and those who failed in 2002 is much larger for coloured students. Taken together, we see an equilibrium for Africans that is characterized by high enrollments that are only weakly related to previous performance and household income.

E. Grade 12 matriculation exam

The nationally standardized and externally evaluated grade 12 matriculation exam provides an interesting comparison to grade advancement from grades 8 to 11. Prior to the grade 12 exam the decision about whether to pass a student is made at the school level, based on a combination of graded material during the school year, end-of-year exams, and subjective evaluation by teachers. Given the structure of the matriculation exam, matric pass rates should have a smaller stochastic component than the school-level pass decisions for grades 8-11, and the stochastic component of matric pass rates should be similar across racial groups. We should therefore find that the impact of prior learning on the probability of passing the matric exam is larger and more equal across racial groups than was the impact of prior learning on grade 8-11 pass rates. In this section we present regressions in which the outcome is passing the grade 12 matriculation exam. The sample is all CAPS respondents who were enrolled in grade 12 in 2002, 2003, or 2004 and reported matriculation exam results. We also present separate regressions for passing grade 9, 10, and 11, using the sample of students who were enrolled in these grades in 2002, 2003, or 2004.

Table 7 presents pass rates and mean characteristics for the samples for each grade. The first row of the top panel gives the mean pass rate for students the first time they took the grade 12 matric exam. About 78% of African 12th graders passed the exam on their first attempt, compared to 89% of coloureds and 99.5% of whites. Given the almost universal pass rate for white students, we will not include white students in our regressions. The second row shows the percentage who passed "with exemption," a higher pass that qualifies students for admission to university. Only 18% of African students passed with exemption, compared to 22% of coloured and 60% of white students. The third row shows that about 11% of African students took the exam more than once between 2002 and 2004. We will include these multiple attempts in our regression, correcting the standard errors for clustering at the individual level. Looking at other characteristics in Table 7, we see that the sample of grade 12 students is, not surprisingly, a selective sample of the students observed at lower grades. Comparing 12th graders in the top panel with 9th graders in the bottom panel, the mean LNE score of 12th graders is about half a standard deviation above the mean for 9th graders for African and coloured students.

Table 7 also documents the large differences in pass rates across grades. Pass rates are low in grade 11, with only 69% of African students and 80% of coloured students passing. This is consistent with the widely held view that teachers and school administrators hold back

11th grade students who they feel are not ready to pass the matric exam, motivated in part by a desire to increase the school's pass rate. Interestingly, however, pass rates for Africans are equally low at grade 10 as grade 11, and for coloureds the 70% pass rate in grade 10 is the lowest of any grade. Grade 9 pass rates are higher, at about 85% for both African and coloured students.

Table 8 presents regressions in which the dependent variable is equal to 1 if the student passes a given grade. We are particularly interested in the results using the standardized grade 12 matriculation exam in the top panel. The most striking result of this regression is that the coefficient on the LNE score is now slightly higher for Africans than it is for coloureds. This is in contrast to the lower impact of LNE scores for Africans that we saw on grade progression at grades 8-11 in Table 5. We also estimate a lower impact of LNE scores for Africans than coloureds in the separate regressions for grade 9, 10, and 11 in Table 8, although only the grade 10 estimates are statistically different at conventional levels.

Looking at marginal effects evaluated at a constant set of characteristics in Columns 4-5, the point estimates imply that a one standard deviation increase in the LNE score is associated with a 15 percentage point higher probability of passing matric for Africans, compared to a 6 percentage point increase for coloureds. The effect of income for Africans is also larger and closer to the effect for coloureds in the matric regressions than it was in the regressions for advancing three grades between 2002 and 2005 or in the separate regressions for grades 9, 10, and 11. In our grade 12 results we continue to find a smaller impact of previous grades failed for Africans than for coloureds. This is expected, since our model implies that failing grades is less of a signal about prior learning for Africans. Indeed, the fact that previous failed grades have no significant effect on the probability of passing the matric exam for Africans is entirely consistent with our interpretation of the weak connection between learning and evaluation in African schools.

Comparing the results from the grade 12 regression with the results for grades 9-11, it is striking how much larger the impact of income and LNE scores are in grade 12 for Africans. The marginal impact of LNE scores on the probability of passing is 3 to 6 times larger for Africans in grade 12 than it is in grades 9, 10, or 11. The marginal impact of income for Africans is small and often not significantly different from zero at grades 9, 10, and 11, but becomes strong and significant in grade 12. For coloureds the impact of LNE scores and income are much more similar across grades. All these patterns are consistent with a regime in which there is a weak link between learning and evaluation for Africans in grades prior to grade 12, with the situation suddenly changing at the standardized grade 12 exam. For coloured students the evaluation of performance in grades 9, 10, and 11 is not substantially different from the evaluation in grade 12.

Given our probits for advancing three grades in Table 5, the results for passing the matric exam in Table 8 are quite remarkable. While an extra point on the LNE exam has less than half the impact on the probability of advancing three grades for African students compared to coloured students, an extra LNE point has roughly equal impact for Africans and coloureds on the probability of passing the matriculation exam. This provides strong support for our interpretation that the racial difference in the impact of LNE scores on grade advancement is due to a weaker link between learning and evaluation in African schools. Put another way, this suggests that it is a larger σ rather than smaller βs that cause Africans to have smaller probit coefficents in Table 5. If racial differences in the grade advancement regressions were due to an interaction between prior learning and school quality (causing lower β coefficients in the learning production function), then we should see the same kind of differences showing up in the matric regressions. The results suggest that initial human capital does translate into higher learning in both the African and coloured schools, but that this learning does not translate equally into grade advancement.

Table 8 also allows us to evaluate another prediction of our model – that the impact of characteristics on grade advancement increases as we look over a larger number of grades. As shown in Equation 6, the prediction is most straightforward for the proportional impact, since the absolute passing rate declines as we look across multiple grades. Evaluated at the X_1 vector of characteristics used for the marginal effects in Table 8, the predicted probability of passing for Africans is 0.78 in grade 12, 0.82 in grade 11, 0.85 in grade 9, and 0.95 in grade 8. Combining these with the marginal effect of LNE scores, a one standard deviation increase in the score implies a 19% increase in the probability of passing grade 12, a 5% increase at grade 11, and a 3% increase at grades 9 and 10. The predicted probability of passing three grades for Africans in Table 5 is 0.46. A one standard deviation increase in the LNE score implies a 27% increase in the probability of passing three grades. In other words, the proportional impact of the LNE score on passing three consecutive secondary grades is five times greater than the impact on passing grade 11 alone and nine times greater than the impact on passing grade 9 or 10 alone.

Results for coloureds are similar: While a one standard deviation increase in the LNE score implies a 18% increase in the probability of passing grade 10 and a 9% increase in the probability of passing grade 11, it implies a 33% increase in the probability of passing from grade 8 or 9 to grade 11 or 12. Similar results apply to the impact of household income. These results are entirely consistent with our stochastic model of grade repetition, suggesting that there is a random component of passing that is uncorrelated with ability and uncorrelated across years. The importance of this component declines when we look at passing over multiple grades, increasing the impact of characteristics such as the LNE score.

4. Conclusions

Our theoretical model demonstrates that a large stochastic component to grade advancement can have important effects on who attends school, how much effort they invest in school, and how individual and household characteristics affect the probability of grade advancement. We show that by increasing the variance in the stochastic component of grade advancement we can generate an equilibrium that looks very much like African schools in Cape Town – high rates of enrollment, low levels of effort, and high rates of grade repetition. This theoretical model also implies that characteristics such as parental income and previous school performance will have a lower impact in African schools than in coloured or white schools, assuming that African schools have a larger stochastic component in grade advancement.

The results of our probit regressions are highly consistent with our theoretical model. We find a strong effect of test scores and household income on the probability of grade advancement for all races. However, the effect of these variables is significantly weaker for African students. While this could indicate that there is an interaction between school quality and other inputs, it is also consistent with a higher variance in the random components of grade advancement in African schools. This high variance helps explain the high school enrollment among African students, even in the face of high failure rates. For these students, high school has elements of a lottery, with even low-ability students having an incentive to be enrolled. Strong evidence in support of our interpretation is provided by the fact that the impact of our baseline test score on the probability of passing the nationally standardized grade 12 matriculation examination is slightly higher for African than for coloured students, suggesting that the payoff to ability is equalized when the stochastic component of evaluation is equalized.

From a policy perspective the strong impact of household income and indicators of previous achievement such as test scores and the number of grades behind in 2002 remind us of the importance of quality primary schooling and the disadvantages of growing up in poor households. The results highlight persistent racial differences in the schooling environment and the signals that this sends to learners. In drawing attention to the differential translation of learning into grade advancement our model and empirical results highlight one particular serious policy challenge; namely, the need to strengthen the link between assessment and actual learning. The fact that this link is so much stronger for African students when assessment is nationally standardized and external to the school points to serious weaknesses in the ability of these schools to adequately evaluate student ability and learning.

Our analysis has important implications beyond South Africa. Our results suggest that a school's ability to accurately assess performance and determine which students advance to higher grades is a critical and understudied dimension of school quality. This is obviously of greatest importance when significant fractions of students are held back, but it may be important in any system in which some students fail. With high rates of grade repetition throughout Latin American and sub-Saharan Africa, this dimension of schooling would seem to deserve closer study. The strong empirical support for our theoretical model suggests that there are high returns to thinking systematically about the impact of imperfect evaluation on enrollment, effort, and grade advancement in systems with high levels of grade repetition.

References

Anderson, Kermyt G., Anne Case, and David Lam (2001) "Causes and Consequences of Schooling Outcomes in South Africa: Evidence from Survey Data." *Social Dynamics* 27(1): 1-23.

Bhorat, Haroon and Oosthuizen, Morne (2006) "Determinants of Grade 12 Pass Rates in the post-*Apartheid* South African Schooling System, DPRU Working Paper, University of Cape Town.

Brookes H., Shisana, O. & Richter, L. (2004). The national household HIV prevention and risk survey of South African children. Pretoria: Human Sciences Research Council

Cameron, Stephen V., and James J. Heckman (2001) "The Dynamics of Educational Attainment for Black, Hispanic, and White Males," *Journal of Political Economy* 109(3): 455-499.

Case, Anne, and Angus Deaton (1999) "School Inputs and Education Outcomes in South Africa," *Quarterly Journal of Economics* August, 114(3): 1047-84.

Chisholm, L., Hoadley, U., and wa Kivilu, M. (2005). Educator Workload Report. Pretoria: Human Sciences Research Council.

Crouch, Luis and T. Mabogoane (2001) "No magic bullets, just tracer bullets: The role of learning resources, social advantage and education management in improving the performance of South African schools," *Social Dynamics* 27(1):60-78.

Crouch, Luis and T. Mabogoane (1998) "When the Residuals Matter More than the Coefficients: An Educational Perspective," *Studies in Economics and Econometrics* 22(2)

Crouch, Luis and Vinjevold, Penny (2006) South Africa: Access before quality, and what to do now? Unpublished Paper, RTI, North Carolina.

Department of Education (2004). Review of school governance in South African public schools. Report of the Ministerial Review Committee on School Governance. Pretoria: Department of Education.

Fiske, Edward and Helen Ladd, H (2004) *Elusive equity: Education reform in post-apartheid*, Brookings Institution, Washington, D.C.

Gomes-Neto, João Batista, and Eric A. Hanushek (1994) "Causes and Consequences of Grade Repetition: Evidence from Brazil," *Economic Development and Cultural Change*, 43 (1): 117-148.

Hoadley, Ursula. (2007) The boundaries of care: education policy interventions for vulnerable children. Paper presented at Education & Poverty Reduction Strategies: Issues of Policy Coherence Conference. 21 - 23 February. Pretoria.

Human Rights Watch (2001). Scared at school: Sexual violence against girls in South African schools. New York: Human Rights Watch.

Jacoby, Hanan (1994) "Borrowing Constraints and Progress Through School: Evidence from Peru" *The Review of Economics and Statistics*, 76(1): 151-160.

Lam, David, Jeremy Seekings and Meredith Sparks (2006) The Cape Area Panel Study: Overview and Technical Documentation for Waves 1-2-3. University of Cape Town, December 2006.

Lee, Valerie E., Tia Linda Zuze, and Kenneth N. Ross (2005) “School Effectiveness in 14 Sub-Saharan African Countries: Links with 6th Graders’ Reaching Achievement,” *Studies in Educational Evaluation*, 31: 207-246.

Reddy, V. (2006). *Mathematics and science achievement at South African schools in TIMSS 2003*. Cape Town: HSRC Press.

Selod, H and Zenou, Y (2003) Private versus public schools in post-apartheid South African cities: Theory and policy implications. *Journal of Development Economics* 71(2): 351-394

USAID (2003). Unsafe schools: A literature review of school-related gender based violence in developing countries.Washington: USAID

van der Berg, Servaas (2005) “Apartheid’s enduring legacy: Inequalities in education,” Paper delivered to the Oxford University/University of Stellenbosch conference on *The South African Economic Policy under Democracy: A 10 year review*, Stellenbosch, October 2005.

van der Berg, S. & Louw, M. (2006). Lessons learnt from SACMEQ II: South African student performance in regional context. Paper presented at the conference on Investment Choices for Education in Africa, 19-21 September, Johannesburg.

Yamauchi, Futoshi (2005). Race, equity and public schools in post-apartheid South Africa: Equal opportunity for all kids. *Economics of Education Review* (24): 213-33.

Table 1. Sample size by population group and attrition between waves, respondents in grades 8 to 12 in 2002, Cape Area Panel Study

	CAPS Wave1, 2002			CAPS Wave 3 2005			
Population Group	Sample size	Unweighted percent	Weighted percent	Sample size	Unweighted percent	Weighted percent	Rate of attrition
Black/African	1,090	46.5	29.7	845	44.6	28.2	22%
Coloured	910	38.8	48.7	814	43.0	53.6	11%
White	344	14.7	21.6	235	12.4	18.1	32%
Total	2,344	100.0	100.0	1,894	100.0	100.0	19%

Table 2. Percentage who worked in last 12 months, CAPS respondents in Wave 1, 2002

	African		Coloured		White	
Age	Female	Male	Female	Male	Female	Male
14	0.0	0.7	7.4	19.7	9.0	30.3
15	0.0	0.8	12.7	10.5	27.1	33.3
16	1.6	5.3	14.9	27.2	44.8	32.0
17	1.3	6.6	26.4	26.6	53.9	51.0
18	1.9	9.5	32.0	47.0	53.3	73.6
19	6.9	10.8	52.3	62.7	70.2	72.6
20	16.7	24.7	63.9	83.5	82.9	80.5
21	19.8	26.9	65.1	82.4	78.8	89.1
22	23.9	35.3	77.4	78.1	75.7	87.9
Sample Size	1,219	927	1,077	925	313	284

Table 3. Percentage in each grade or non-enrollment status in 2005, CAPS respondents in grades 8 and 9 in 2002

	8th grade in 2002			9th grade in 2002		
Status in 2005	African	Coloured	White	African	Coloured	White
Enrolled in grade 8	0.7	0.0	0.0	--	--	--
Enrolled in grade 9	6.3	2.2	0.0	0.5	0.0	0.0
Enrolled in grade 10	37.4	13.2	0.0	10.6	3.5	0.0
Enrolled in grade 11	33.4	45.7	85.3	29.5	15.5	8.0
Enrolled in grade 12	2.5	0.7	7.3	27.5	40.7	80.8
Post-secondary	1.6	1.9	0.0	1.3	1.8	3.7
Total enrolled	81.9	63.6	92.5	69.3	61.5	92.4
Not enrolled/not working	15.0	24.0	3.0	22.4	22.8	1.1
Not enrolled/working	3.1	12.5	4.5	8.3	15.7	6.5
Sample size	141	132	41	248	228	58

Table 4. Descriptive statistics, Cape Area Panel Study Waves 1-3, CAPS respondents enrolled in Grades 8 or 9 in 2002 and observed again in 2005

	African (N=386)		Coloured (N=350)		White (N=78)	
Variable	**Mean**	**Std. Dev.**	**Mean**	**Std. Dev.**	**Mean**	**Std. Dev.**
Advance 3 grades by 2005	0.322	0.468	0.438	0.497	0.835	0.374
Enrolled in 2003	0.954	0.209	0.917	0.276	0.973	0.163
Enrolled in 2004	0.872	0.335	0.779	0.416	0.943	0.234
Enrolled in 2005	0.738	0.440	0.622	0.486	0.906	0.294
Grade 9 in 2002	0.644	0.480	0.625	0.485	0.604	0.492
Female	0.549	0.498	0.479	0.500	0.555	0.500
Number of grades failed by 2002	0.715	0.883	0.470	0.677	0.190	0.482
Failed at least one grade by 2002	0.497	0.501	0.372	0.484	0.152	0.361
Failed grade enrolled in 2002	0.128	0.335	0.108	0.310	0.008	0.090
Standardized LNE total score	-0.601	0.787	0.010	0.696	1.137	0.481
Log per cap hh income (mean zero)	-0.712	0.871	0.209	0.870	1.832	0.828
Mother's education (grades completed)	8.38	2.82	8.73	2.55	12.23	1.52
Father's education (grades completed)	7.36	3.54	8.83	2.93	12.94	2.05
Mother's education missing	0.104	0.305	0.095	0.294	0.000	0.000
Father's education missing	0.404	0.491	0.321	0.468	0.079	0.272
Household shock 2002-2005	0.255	0.436	0.158	0.365	0.035	0.185
Local unemployment rate for age & sex	0.815	0.124	0.637	0.182	0.266	0.268
Former DET (African) school	0.768	0.422	0.019	0.136	0.000	0.000
Former HOA (White) school	0.033	0.178	0.083	0.276	0.914	0.283
Former HOR (Coloured) school	0.107	0.310	0.867	0.341	0.011	0.105
New school since 1994	0.091	0.289	0.009	0.093	0.075	0.266
Annual school expenses in 2002 (rands)	323.2	1140.7	744.4	1356.3	5890.1	6058.8
Pupil-teacher ratio in 2002	32.6	3.9	30.5	3.2	24.1	2.8
Moved school between 2002 and 2003	0.194	0.396	0.100	0.300	0.049	0.218
Moved school between 2003 and 2004	0.104	0.306	0.050	0.219	0.082	0.277
Moved school between 2004 and 2005	0.119	0.325	0.040	0.196	0.058	0.235
Moved former department 2002-2003	0.038	0.191	0.021	0.143	0.000	0.000
Moved former department 2003-2004	0.022	0.148	0.000	0.000	0.000	0.000
Moved former department 2004-2005	0.009	0.093	0.003	0.056	0.000	0.000

Table 5. Probit regressions for probability of advancing 3 grades between 2002 and 2005, CAPS respondents in grades 8 or 9 in 2002

	Probit coefficients			Tests for equality of coefficients			Marginal effects at X_1		
Variable	**African**	**Coloured**	**White**	**African-Coloured**	**African-White**	**White-Coloured**	**African**	**Coloured**	**White**
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**	**(9)**
Grade 9 in 2002	-0.22	-0.39	-1.023	0.629	3.681	2.135	-0.086	-0.149	-0.338
	[0.128]*	[0.172]**	[0.401]**	(0.429)	(0.056)*	(0.145)	[0.051]*	[0.066]**	[0.126]***
Female	-0.002	0.171	0.027	0.604	0.002	0.050	-0.001	0.067	0.011
	[0.149]	[0.166]	[0.626]	(0.438)	(0.965)	(0.823)	[0.059]	[0.066]	[0.249]
Number of grades failed,	-0.18	-0.681	-0.745	7.336	1.526	0.018	-0.071	-0.271	-0.297
Wave 1	[0.100]*	[0.156]***	[0.451]*	(0.007)***	(0.218)	(0.893)	[0.040]*	[0.060]***	[0.177]*
Standardized LNE total	0.308	0.824	0.538	8.126	0.137	0.205	0.122	0.328	0.214
score	[0.096]***	[0.154]***	[0.620]	(0.005)***	(0.712)	(0.651)	[0.038]***	[0.063]***	[0.246]
Log hh income per capita	0.106	0.354	0.524	2.976	1.797	0.286	0.042	0.141	0.209
	[0.092]	[0.112]***	[0.300]*	(0.086)*	(0.181)	(0.593)	[0.036]	[0.044]***	[0.118]*
Mother's schooling	0.014	0.061	0.196	0.844	1.667	0.896	0.005	0.024	0.078
	[0.034]	[0.039]	[0.138]	(0.359)	(0.198)	(0.345)	[0.014]	[0.016]	[0.054]
Mother's schooling	0.102	0.399		0.249			0.041	0.157	
missing	[0.398]	[0.445]		(0.618)			[0.159]	[0.168]	
Father's schooling	0.004	0.05	-0.178	1.283	1.410	2.164	0.002	0.02	-0.071
	[0.023]	[0.034]	[0.153]	(0.258)	(0.236)	(0.143)	[0.009]	[0.014]	[0.060]
Father's schooling	-0.088	0.389	-2.619	1.395	1.505	2.103	-0.035	0.153	-0.474
missing	[0.205]	[0.349]	[2.063]	(0.239)	(0.221)	(0.148)	[0.080]	[0.133]	[0.191]**
Household shock 2002-	-0.18	-0.466	-1.51	0.922	4.712	2.684	-0.07	-0.176	-0.419
2005	[0.173]	[0.243]*	[0.594]**	(0.338)	(0.031)**	(0.103)	[0.066]	[0.086]**	[0.160]***
Unemployment rate	0.299	-0.194	-0.086	0.392	0.119	0.010	0.119	-0.077	-0.034
	[0.652]	[0.445]	[0.973]	(0.532)	(0.731)	(0.919)	[0.255]	[0.177]	[0.387]
Constant	-0.255	-0.594	0.525	0.159	0.250	0.490			
	[0.629]	[0.576]	[1.505]	(0.691)	(0.618)	(0.485)			
Observations	386	350	78						

Robust standard errors in brackets in columns 1-3 and 7-9; p-value of F tests in parentheses in columns 4-6; * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 6. Probit regressions for probability of enrollment in 2004, CAPS respondents enrolled in grades 8 or 9 in 2002

Variable	Probits for enrollment		Test of equality of coefficients	Marginal effects at X_1	
	African	Coloured		African	Coloured
	(1)	(2)	(3)	(1)	(2)
Grade 9 in 2002	-0.194	-0.283	0.132	-0.021	-0.048
	[0.171]	[0.178]	(0.716)	[0.018]	[0.030]
Female	0.42	0.006	2.723	0.054	0.001
	[0.187]**	[0.170]	(0.1)	[0.024]**	[0.024]
Number of grades failed, Wave 1	-0.272	-0.428	1.389	-0.025	-0.059
	[0.087]***	[0.101]***	(0.24)	[0.015]*	[0.026]**
Standardized LNE total score	0.238	0.322	0.254	0.022	0.044
	[0.101]**	[0.130]**	(0.615)	[0.012]*	[0.021]**
Log hh income per cap.	0.002	0.27	2.558	0	0.037
	[0.120]	[0.117]**	(0.111)	[0.011]	[0.024]
Mother's schooling	0.073	-0.015	2.965	0.007	-0.002
	[0.031]**	[0.041]	(0.086)*	[0.004]	[0.006]
Mother's schooling missing	0.744	-0.551	5.426	0.036	-0.11
	[0.343]**	[0.433]	(0.021)**	[0.023]	[0.126]
Father's schooling	-0.025	0.143	7.991	-0.002	0.02
	[0.040]	[0.044]***	(0.005)***	[0.004]	[0.007]***
Father's schooling missing	-0.451	1.006	7.024	-0.059	0.066
	[0.382]	[0.396]**	(0.009)***	[0.069]	[0.029]**
Household shock between 2002 and 2005	-0.03	-0.167	0.198	-0.003	-0.026
	[0.194]	[0.241]	(0.657)	[0.018]	[0.042]
Local unemployment rate for age and sex	0.643	-0.388	1.758	0.058	-0.054
	[0.566]	[0.531]	(0.186)	[0.073]	[0.077]
Failed in 2002	-0.133	-1.429	15.824	-0.013	-0.417
	[0.230]	[0.231]***	(0)***	[0.026]	[0.107]***
Constant	0.766	0.795	0.001		
	[0.534]	[0.604]	(0.971)		
Observations	438	369			

Robust standard errors in brackets in columns 1, 2, 4, and 5; p-value of F tests in parentheses in column 3; * significant at 10%; ** significant at 5%; *** significant at 1%.
Marginal effects evaluated for female in grade 8 in 2002, 1 failed grade, LNE=0, log income=0, parental schooling=8, no shock, unemployment rate=0.6, passed grade in 2002.

Table 7. Pass rates and descriptive statistics,
CAPS respondents enrolled in grades 9 to 12 in 2002, 2003 or 2004

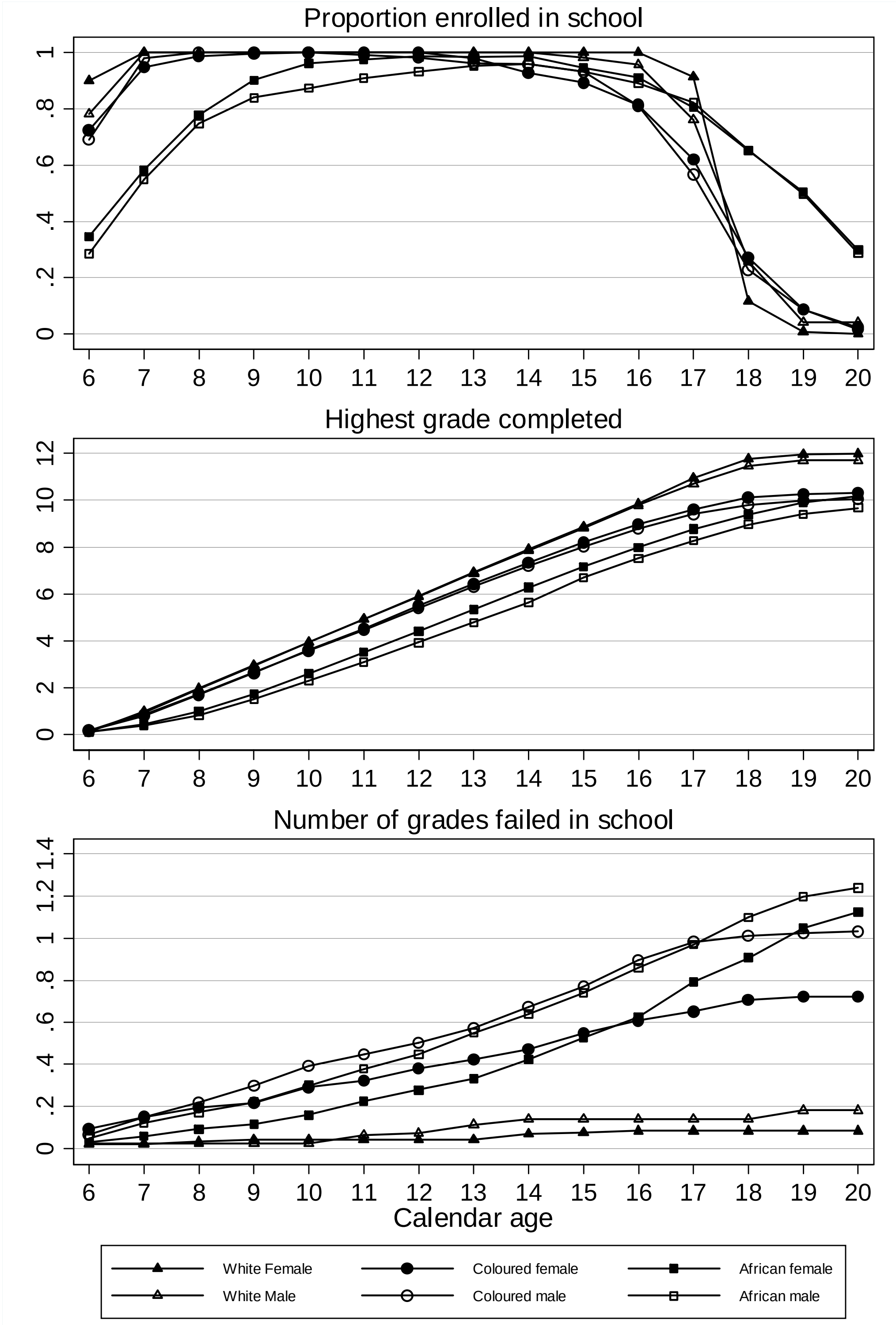
	African		Coloured		White	
Variable	**Mean**	**Std. Dev.**	**Mean**	**Std. Dev.**	**Mean**	**Std. Dev.**
Grade 12:						
Pass grade	0.780	0.415	0.893	0.310	0.995	0.068
Pass with exemption	0.186	0.390	0.220	0.415	0.599	0.492
Number of attempts at grade	1.108	0.337	1.018	0.133	1.000	0.000
Number of grades failed, Wave 1	0.703	0.856	0.395	0.679	0.156	0.379
Standardized LNE total score	-0.105	0.731	0.555	0.634	1.306	0.458
Log hh income per cap.	-0.448	0.948	0.652	0.787	1.782	0.741
Observations	279		349		125	
Grade 11:						
Pass grade	0.689	0.464	0.799	0.401	0.939	0.240
Number of attempts at grade	1.204	0.437	1.085	0.305	1.022	0.148
Number of grades failed, Wave 1	0.737	0.961	0.374	0.648	0.145	0.418
Standardized LNE total score	-0.326	0.776	0.394	0.643	1.231	0.535
Log hh income per cap.	-0.527	0.896	0.565	0.769	1.834	0.774
Observations	475		411		141	
Grade 10:						
Pass grade	0.697	0.460	0.707	0.456	0.933	0.251
Number of attempts at grade	1.221	0.457	1.171	0.396	1.057	0.232
Number of grades failed, Wave 1	0.738	0.900	0.463	0.733	0.154	0.420
Standardized LNE total score	-0.488	0.796	0.164	0.688	1.131	0.567
Log hh income per cap.	-0.648	0.882	0.353	0.847	1.806	0.768
Observations	560		474		118	
Grade 9:						
Pass grade	0.845	0.362	0.857	0.351	0.966	0.182
Number of attempts at grade	1.134	0.370	1.066	0.260	1.039	0.194
Number of grades failed, Wave 1	0.731	0.866	0.480	0.667	0.160	0.431
Standardized LNE total score	-0.619	0.767	-0.002	0.733	1.069	0.539
Log hh income per cap.	-0.723	0.890	0.239	0.865	1.838	0.727
Observations	498		380		88	

Table 8. Probit regressions for probability of passing Grade 12 matriculation exam and probability of passing Grades 9, 10, and 11

	Probit coefficient		Test of equality of coefficients	Marginal effects at X_1	
Variable	African	Coloured		African	Coloured
	(1)	(2)	(3)	(4)	(5)
Grade 12:					
Number of grades failed, Wave 1	-0.1	-0.324	1.424	-0.03	-0.043
	[0.107]	[0.118]***	(0.233)	[0.030]	[0.016]***
Standardized LNE total score	0.502	0.465	0.026	0.148	0.062
	[0.109]***	[0.169]***	(0.871)	[0.053]***	[0.031]**
Log hh income per cap.	0.266	0.289	0.019	0.078	0.039
	[0.104]**	[0.127]**	(0.89)	[0.041]*	[0.021]*
Observations	311	353			
Grade 11:					
Number of grades failed, Wave 1	-0.19	-0.193	0.001	-0.049	-0.042
	[0.051]***	[0.113]*	(0.981)	[0.014]***	[0.024]*
Standardized LNE total score	0.17	0.356	0.877	0.044	0.078
	[0.082]**	[0.133]***	(0.349)	[0.023]**	[0.037]**
Log hh income per cap.	-0.028	0.16	1.53	-0.007	0.035
	[0.080]	[0.106]	(0.216)	[0.021]	[0.027]
Observations	556	445			
Grade 10:					
Number of grades failed, Wave 1	-0.1	-0.225	0.888	-0.023	-0.069
	[0.058]*	[0.097]**	(0.346)	[0.013]*	[0.028]**
Standardized LNE total score	0.117	0.451	5.707	0.027	0.138
	[0.069]*	[0.093]***	(0.017)**	[0.016]*	[0.031]***
Log hh income per cap.	0.006	0.236	3.299	0.001	0.072
	[0.072]	[0.082]***	(0.069)*	[0.017]	[0.029]**
Observations	680	554			
Grade 9:					
Number of grades failed, Wave 1	-0.165	-0.397	1.779	-0.018	-0.041
	[0.071]**	[0.137]***	(0.182)	[0.009]**	[0.014]***
Standardized LNE total score	0.222	0.327	0.339	0.025	0.033
	[0.097]**	[0.128]**	(0.561)	[0.013]*	[0.020]*
Log hh income per cap.	0.173	0.116	0.116	0.019	0.012
	[0.095]*	[0.112]	(0.733)	[0.010]*	[0.013]
Observations	562	403			

Robust standard errors in brackets in columns 1, 2, 4, and 5; p-value of F tests in parentheses in column 3; * significant at 10%; ** significant at 5%; *** significant at 1%. Sample for each regression is the pooled sample of respondents enrolled in a given grade in 2002, 2003, or 2004; regressions also include variables used in Table 5; Marginal effects evaluated for female, 1 failed grade, LNE=0, log income=0, parental schooling=8, no shock, unemployment rate=0.6.

Figure 1. Schooling experience from retrospective histories
CAPS respondents age 21-22, 2002



Cape Area Panel Study Wave 1, 2002

Figure 2. Impact of higher variance on probability of passing, enrollment, and effort

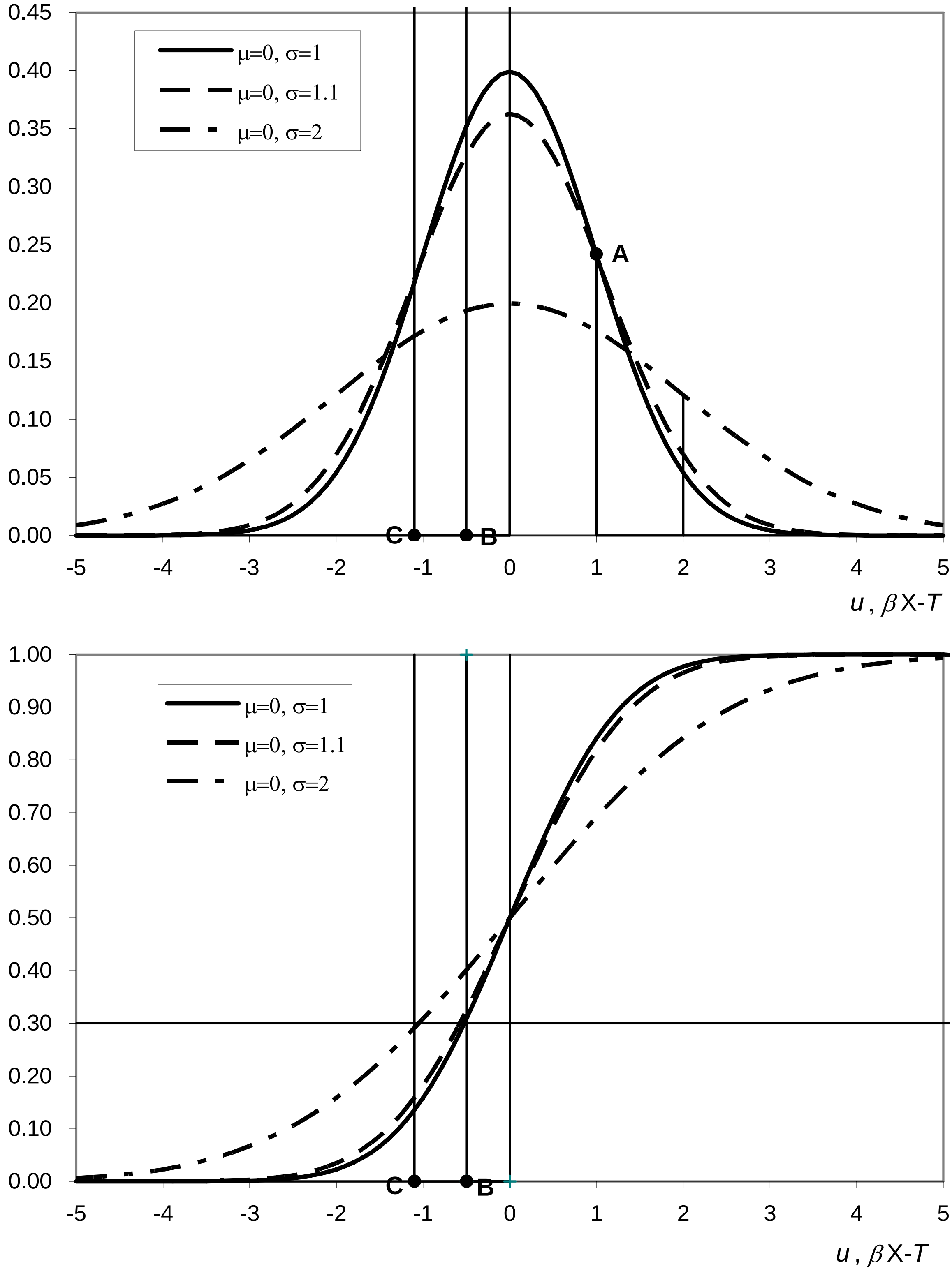


Figure 3. Kernel densities for scores on CAPS Wave 1 Literacy and Numeracy Evaluation

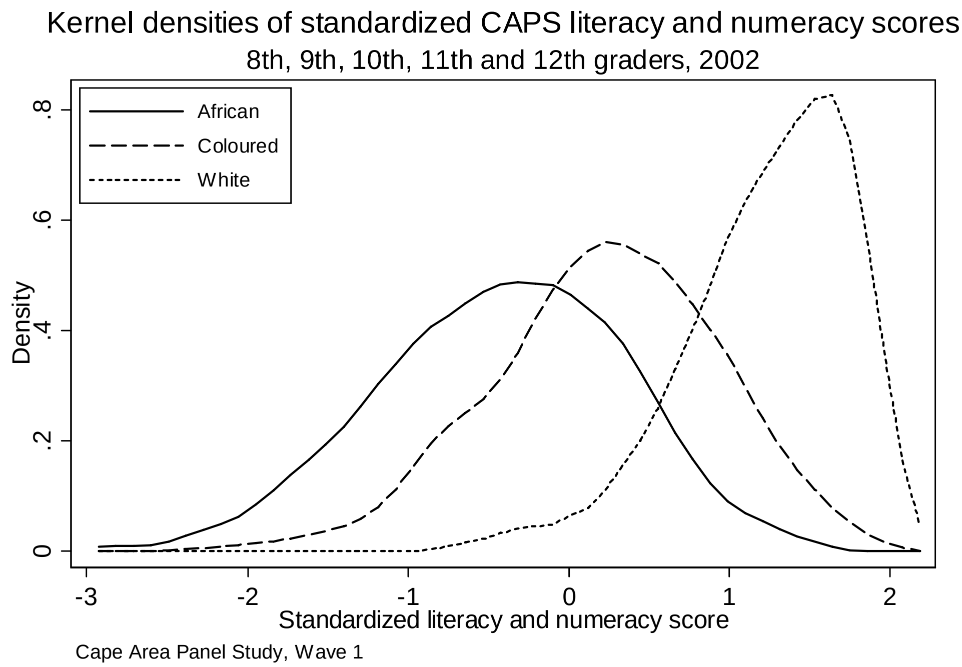


Figure 4. Kernel densities of log per capita household income, CAPS Wave 1

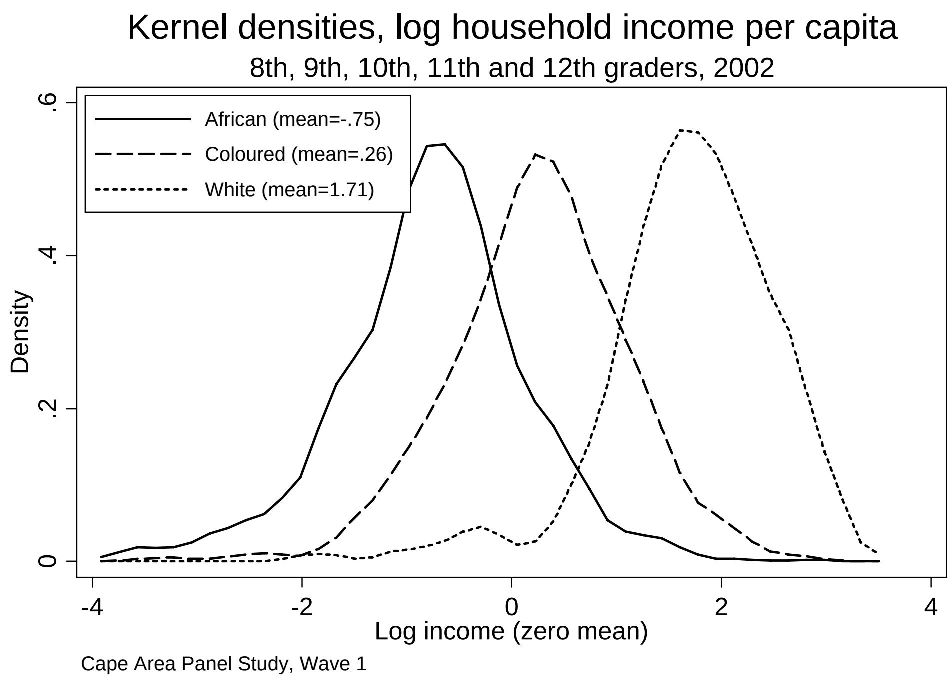


Figure 5. Age distribution of 9th graders, CAPS Wave 1, 2002

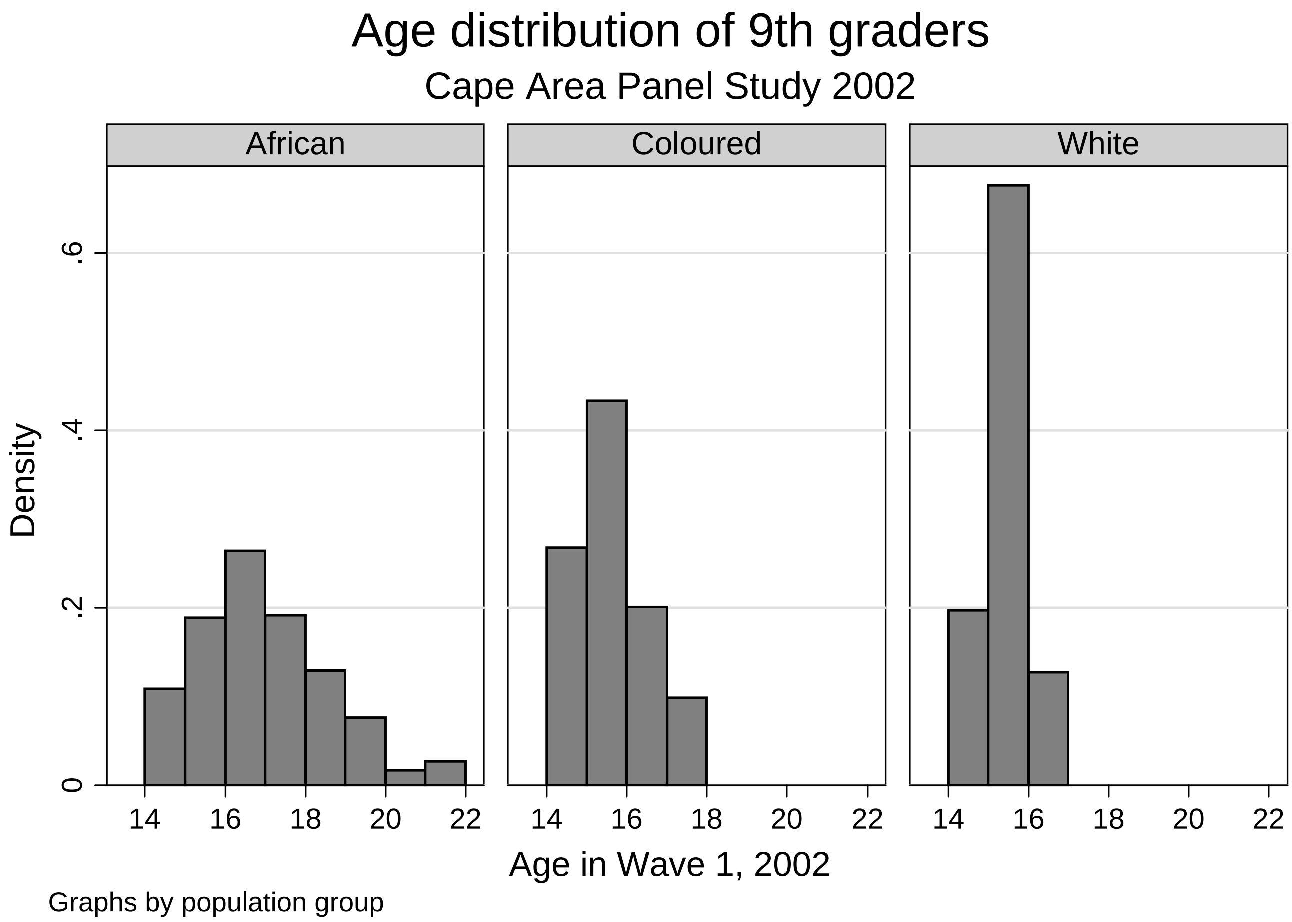
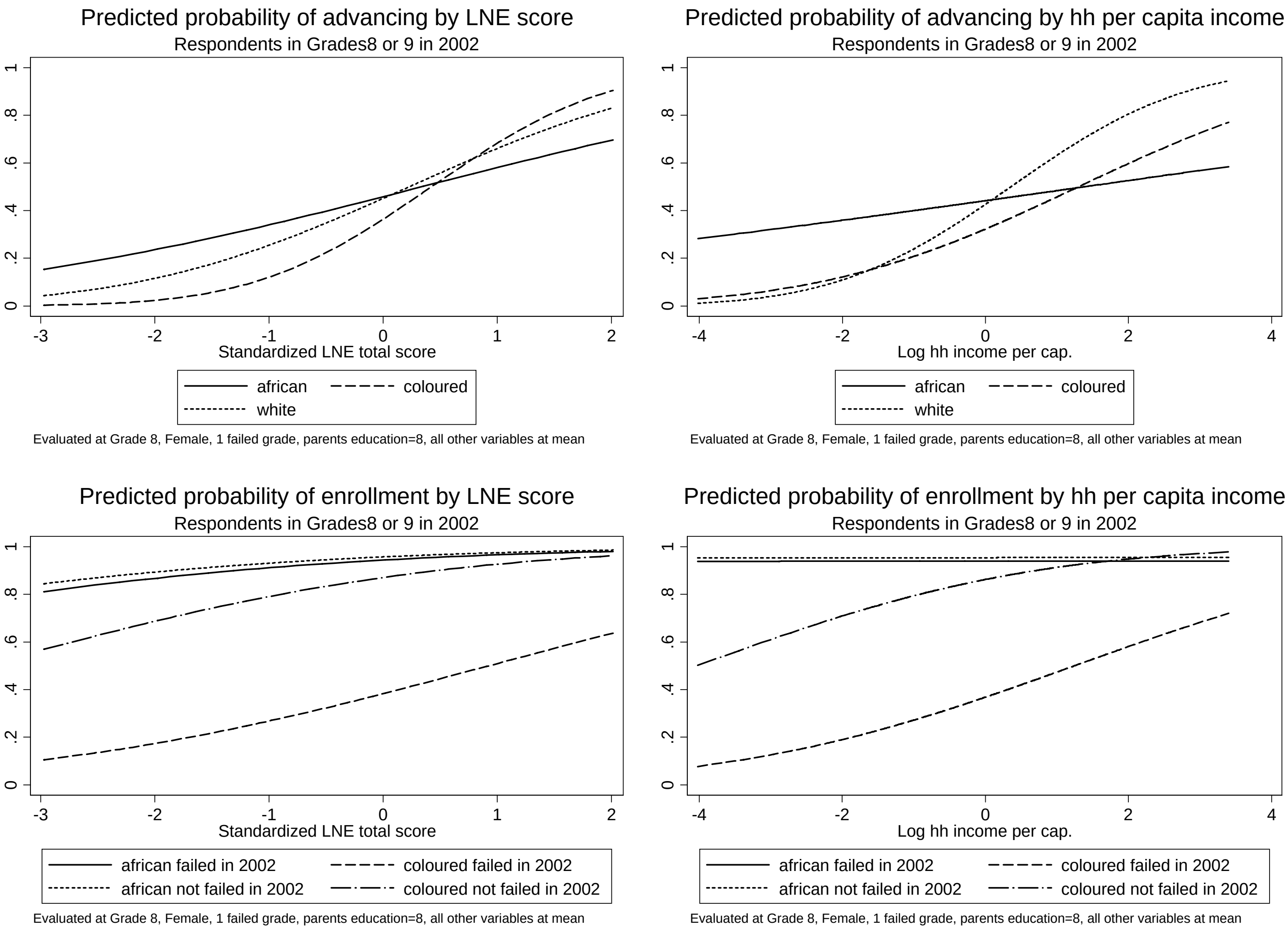


Figure 6. Predicted probability of advancing three grades by 2005 and being enrolled in 2004, by LNE score and per capita household income



The Southern Africa Labour and Development Research Unit

The Southern Africa Labour and Development Research Unit (SALDRU) conducts research directed at improving the well-being of South Africa's poor. It was established in 1975. Over the next two decades the unit's research played a central role in documenting the human costs of apartheid. Key projects from this period included the Farm Labour Conference (1976), the Economics of Health Care Conference (1978), and the Second Carnegie Enquiry into Poverty and Development in South Africa (1983-86). At the urging of the African National Congress, from 1992-1994 SALDRU and the World Bank coordinated the Project for Statistics on Living Standards and Development (PSLSD). This project provide baseline data for the implementation of post-apartheid socio-economic policies through South Africa's first non-racial national sample survey.

In the post-apartheid period, SALDRU has continued to gather data and conduct research directed at informing and assessing anti-poverty policy. In line with its historical contribution, SALDRU's researchers continue to conduct research detailing changing patterns of well-being in South Africa and assessing the impact of government policy on the poor. Current research work falls into the following research themes: post-apartheid poverty; employment and migration dynamics; family support structures in an era of rapid social change; public works and public infrastructure programmes, financial strategies of the poor; common property resources and the poor. Key survey projects include the Langeberg Integrated Family Survey (1999), the Khayelitsha/Mitchell's Plain Survey (2000), the ongoing Cape Area Panel Study (2001-) and the Financial Diaries Project.

Southern Africa Labour and Development Research Unit
School of Economics
University of Cape Town
Private Bag, Rondebosch, 7701
Cape Town, South Africa

Tel: +27 (0) 21 650 5696
Fax: +27 (0) 21 650 5697

Email: brenda.adams@.uct.ac.za
Web: www.saldru.uct.ac.za

SALDRU
School of Economics
UNIVERSITY OF CAPE TOWN