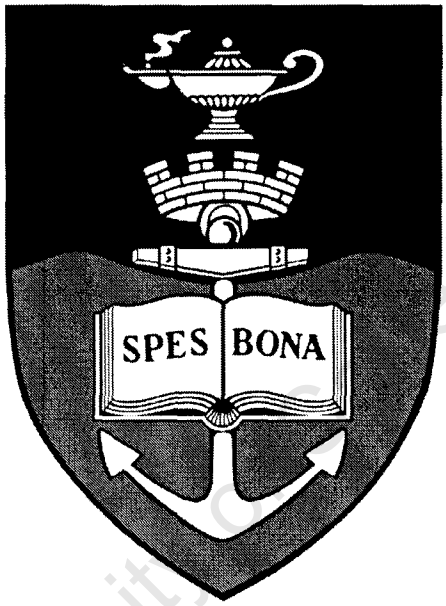
School Drop-Out in the Western Cape

The Role of Prior School Progress, Ability, Pregnancy & Family Background



Catherine Kannemeyer (KNNCAT003)

Masters Dissertation

November 2008

School of Economics

University of Cape Town

Supervisor: Professor Ingrid Woolard

ABSTRACT

School Drop-Out in the Western Cape: The Role of Prior School Progress, Ability, Pregnancy & Family Background

This paper examines the role of major factors that lead to school abandonment in the Western Cape area of South Africa. Family background, pregnancy, individual ability as well as prior-school progress are all factors which exert influence on the decision to abandon school, the effect of which can be examined using rich household data and individual event histories from the Cape Area Panel Study (2002 – 2005) and incorporating school-level data from the Schools' Register of Needs (2000). The effect of educational inputs is considered to vary by race, and pregnancy is a major push-out factor for girls, and, as such Black and Coloured males and females are considered separately.

Results show that prior school progress is strongly correlated with eventual dropout, and this is a concern in the light of prior research indicating that school progress in Black schools is a stochastic process. The effect of family income on dropout is only significant for the Coloured population, which likely reflects the extremely limited labour market options for Black dropouts. In terms of girls' attendance of school, results indicate that teenage pregnancy has a strong impact on school dropout, although the strength of this effect differs by race.

ACKNOWLEDGEMENTS

My special thanks go to Professors Ingrid Woolard and Murray Leibbrandt for their supportive supervision and valuable input.

I would like to acknowledge funding from the National Income Dynamics Study Scholarship Fund.

PLAGIARISM DECLARATION

1. I know that plagiarism is wrong. Plagiarism is to use another's work and pretend that it is one's own.

This paper entitled: *School Drop-Out in the Western Cape: The Role of Prior School Progress, Ability, Pregnancy & Family Background* is my own work.

2. I have used the Harvard convention for citation and referencing. Each significant contribution to, and quotation in, this paper from the work(s) of other people has been attributed, and has been cited and referenced.

3. I have not allowed, and will not allow, anyone to copy my work with the intention of passing it off as his or her own work.

Signature: Date: 25/11/2008

LIST OF TABLES

LIST OF FIGURES

1. INTRODUCTION

According to the South African Department of Education (cited in Flisher, 2007) 60% of children in South Africa who enrol in Grade 1 drop out before completing high school. Despite near universal primary school enrolment, high levels of grade repetition are common. Schooling in South Africa is compulsory up until the end of Grade 9 or until learners turn 15 years old, and despite high levels of enrolment beyond this point, these levels of enrolment are not matched by successful school completion.

Adverse circumstances associated with premature school leaving include reduced economic prospects in terms of the range of jobs available and earnings potential, poorer mental and physical health, as well as an increase in the risk of engaging in criminal activity. (Beauvais et al., 1996 cited in Flisher, 2007).

Discrimination in the Apartheid era was based on race, and extended to the quality and type of education made available to different ethnic groups. In the post Apartheid era, it is contended that discrimination in the labour market is perpetuated, but only indirectly, as it has been found that the market for skilled Blacks clears, and that wage differentials are based on skills rather than race, and so it is contended that wages are only indirectly discriminatory. (Fallon & Lucas, 1998) The importance of improving educational attainment and skills training of Black South Africans is hence paramount to ending wage discrimination in South Africa.

Employment outcomes are particularly dire for school dropouts (Beutel & Anderson, 2007; Flisher, 2007; South et al., 2007; Cardoso & Verner, 2006; Rosenthal, 1998), and this paper seeks to identify influential factors associated with dropouts, who are likely to be confined to the congested unskilled segment of the labour market. It is also important to note that their poor labour markets prospects may have led them to drop out, rather than dire labour market prospects solely being a consequence of dropping out of school.

In terms of the human capital framework (Becker, 1962), the amount of educational investment undertaken is a function of the expected returns to education. Poor economic prospects may be a causal factor in the decision to drop out of school, as well as a consequence of school dropout. If the schooling environment is characterised as one where grade progression has a large stochastic component and educational quality varies indeterminately[1], in the face of limited outside options, it is very likely that the result of the individual decision whether or not to remain in the schooling system is suboptimal.

[1] Incomplete schooling is not subject to standardised assessment

This study examines the odds that Black and Coloured youth in the Western Cape drop out of school. It examines the effect of poverty, parental input, prior performance in school, school quality as well as literacy and numeracy levels on the likelihood that individuals fail to complete school. The investigation is conducted using a sample of young people enrolled in school in the Western Cape Area in 2002. This is a sub-sample of the Cape Area Panel Study which originally sampled youth aged between 14 and 22, some of whom would have already finished school at the time of the original survey. The study is limited to the Black and Coloured populations as a result of the dropout rate for white youth in the original sample being close to zero.

The advantage of using CAPS (2002-2005) is that it provides detailed individual, household and school level (SRN 2000) information on the youth of interest.

Due to reproductive differences, the decision to drop out of school will differ for each of the sexes, and teenage mothers are less likely to finish school. As such, boys and girls are considered separately.

This paper is comprised of 5 sections including the introduction. Section 2 conducts an overview of the South African education system and literature regarding the causes of school dropout is reviewed thereafter. Section 3 provides an introduction to the data and preliminary analysis. Section 4 describes the econometric methods employed, presenting results as well as limitations of the analysis. Section 5 concludes.

2. LITERATURE REVIEW & OVERVIEW OF THE SOUTH AFRICAN EDUCATION SYSTEM

In this section, an overview of the South African Education system and school-level inputs into educational outcomes is provided. This is followed by a summary of the major economic theories of educational attainment, looking at the application of Human Capital Theory as first described by Becker (1962), as well as Signalling Theory introduced by Spence (1973). Thereafter discussion regarding the impact of family background and the process of girls' schooling decisions specifically follows.

Primary and secondary schooling in South Africa is composed of 12 grades, primary school spanning Grades 1-7 and secondary school Grades 8-12. Completion of Grade 12 is referred to as matriculation, and Grade 12 is often referred to as "matric." Matric examinations are conducted nationally, and set and marked externally, and are one of the only standardised scores available of educational attainment. South African legislation dictates that schooling attendance is compulsory from the year in which a learner turns 7, until they turn 15, or complete Grade 9, whichever occurs sooner (The South African Schools Act (84), 1996). Grade 9 is the only other instance of semi-standardised assessment, as national examinations are set externally, but marked internally, ostensibly following externally set memoranda (The Further Education and Training Act (98), 1998).

Thus, for the most part, grade promotion is not a standardised process, and to the extent that failure induces learners to drop out, it is not necessarily the case that dropouts are less capable learners.

Lam *et al.* (2007) analyze the large racial differences in progress through high school in South Africa, using the CAPS dataset. Characterising schooling as a lottery, they find that grade progression in Black schools is weakly related to actual ability and learning, in contrast to the patterns observed in former Coloured & white schools. They do, however, find that ability matters for performance in matric, across race groups. This is very likely a function of Grade 12 being externally examined and marked – as far as teachers in poor schools use inadequate measures of assessment to determine the progression of learners through school, stochastic grade progression is not surprising. This could mean that we are less likely to observe learners, Blacks especially, dropping out after being retained in a grade, especially in the situation of scarce employment opportunities. Additionally, Lam *et al.* (2007) find that enrolment of Blacks seems unaffected by income, even in the instance of failure in Grade 8.

Lam *et al.* (2007) depict that an equilibrium characterised by high levels of failure, high fractions of the eligible population enrolled, and high fractions of learners putting in low effort can be explained by an environment with a large stochastic component to grade advancement. This equilibrium resembles the

situation in most Black schools. Most importantly, both the probability of grade advancement and the probability of enrolment 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. Furthermore, the impact of grade retention on future enrolment will be lower when there is a high stochastic component.

Both grade retention and dropout play a role in the differences in the educational attainment differentials between race groups. According to calculations by Case & Deaton (1999) using national data, "Blacks gain only 0.61 years of attainment for each year of age compared with 0.76 for Coloureds, 0.88 for Whites, and 0.95 for Asians. They are unable to distinguish which part of the differential can be attributed to grade repetition or dropping out however.

With regard to the distribution of school-level educational inputs, prior to the current post-Apartheid era, the South African Education system was strictly segregated, with different departments responsible for the education of Blacks, Coloureds, Whites and Asians. In the Cape, funding for Black schools was administered by the Department of Education and Training (DET), funding for Coloured schools by the House of Representatives, funding for white schools by the House of Assembly, and funding for Asians by the House of Delegates. This system generated substantial disparities in educational funding per pupil across racial groups. Treating funding levels for Blacks in the DET schools as unity, relative funding levels for Whites, Asians, Coloureds were, respectively, 1.85, 1.61, 1.59 at the end of Apartheid (South African Institute of Race Relations 1997 cited in Case & Deaton, 1999)].

Differences in funding under Apartheid were related to the variation in teacher quality according to school type (Case & Deaton, 1999). Teacher quality matters immensely for educational outcomes, a fact which was highlighted in a special review of education which appeared in the Economist in October 2007. Using data collected by the OECD's Programme for International Learner Assessment (PISA) - it was found that the best systems do not have the highest expenditure per learner, nor were the teachers in some of the top performing countries necessarily highly remunerated (relative to their counterparts in other countries). McKinsey, quoted in the Economist's special education review, stated that schools need to do three things in order to ensure that learners meet their potential – "get the best teachers; get the best out of teachers; and step in when pupils start to lag behind." (How to be Top, 2007)

Markley (2004) cites Sanders' (1999) and Wenglinsky's (2000) work using longitudinal data from Tennessee where both found that teacher effectiveness is the single biggest contributor to student success, outweighing all other factors, such as class size, socioeconomic status, and gender in determining educational outcomes.

Examining the South African education system, Crouch and Mabogoane's (1998) paper on the role of learning resources, social advantage and education management in the performance of South African schools highlights critical factors within schools which influence learners' ability to progress adequately within the schooling system. The quality of teachers is considered the greatest predictor of educational outcomes, and Crouch and Mabogoane (1998) find that the qualifications of educators are found to have a more profound effect on learning outcomes than any other measure of a school's resource endowment/quality.

While the impact of teacher quality is clearly apparent, assessment of teacher quality is more difficult, and this type of data is not frequently available (and is not available for this study, but its importance must be acknowledged).

Case & Deaton (1999), suggest that despite contrary findings in other influential international studies in both a developed (Hanushek, 1986 cited in Case & Deaton, 1999) and developing world context (Hanushek, 1996 cited in Case & Deaton, 1999), in addition to teacher quality, class size has a significant impact on educational outcomes. In the only large randomized class size experiment ever conducted in the United States (Kreuger, 1999 cited in Case & Deaton, 1999) class size was found to affect performance on learner test scores, and Angrist & Lavy (1999 cited in Case & Deaton, 1999) also find similar results in an Israeli context.

Case & Deaton (1999) combine five data sources to create a national sample of rich household and school level data to examine the relationship between pupil-teacher ratios and educational outcomes in South Africa. They find a significant impact of pupil-teacher ratios on educational attainment for Black learners, and provide evidence that not only is the amount of educational attainment a function of class-size, but that reducing class-size improves the pace of educational attainment.

Human Capital Theory is the major economic paradigm regarding schooling decisions. The human capital perspective on educational attainment describes each individual's education and training decisions as a function of the expected payoffs from the accumulation of a stock of productive capital. Schooling is one major investment activity through which one can improve one's productive abilities and hence raise real income prospects. (Becker, 1962)

In terms of the human capital framework then, the optimal level of schooling is a function of the net return on educational investment. A key stylized fact regarding the educational optimising decision is that the relationship of log-earnings to schooling is essentially linear (Mincer, 1974 cited in Ashenfelter & Rouse, 1998).

The direct cost of schooling is comprised of fees and other miscellaneous costs including the cost of transport, books and stationery. While most schools charge fees, the South African Schools Act (84), (1996) upholds that no individual should be denied the right to education on the basis of being unable to pay school fees.

The direct cost of going to school is quite low for Blacks especially. Of the school-going sample in 2002 the annual amount spent on school fees at the household level was less than R1 per day up to the 90th percentile.[2] Hence, it is quite possible that the cost of schooling/limited availability of income will have no effect on school attendance.

The availability of income will have a particularly small effect on school attendance if we consider the opportunity cost of attending school. From the age of 16 – the legal working age – the opportunity cost of schooling would include foregone wages. However, the odds of being unemployed are high, and the quality of education of schooling obtained varied, being especially poor in Black schools – hence the opportunity cost of school attendance in terms of foregone wages is low.

The odds of being employed also affect the return to education. Van der Berg (2005) cites Bhorat and Leibbrandt (2001) who showed that education had a positive effect on the odds of being employed for Blacks, as well as improving their earnings, with the returns to secondary education being particularly high. However, the unemployment rate in South Africa is incredibly high, especially for youth with schooling less than Matric – exceeding 40% for this category, with Blacks worst affected, according to Bhorat (2007).

Other studies suggest that returns to tertiary education are much higher than returns to having only a matric (Keswell & Poswell, 2004 and Moll, 1998) which may lead to learners remaining in the school system in order to obtain a matric, especially if the cost of continuing education is low and outside options are limited.

Thus, the low cost of education in conjunction with the low odds of finding employment, could have two effects. It may lead to high levels of dropout, as the expected returns to education are tempered by the low odds of employment, or, as a result of the lack of outside opportunities, and the positive (albeit low) odds of education improving job prospects, Black youth may remain in school.

Another stylized fact is that the educational optimising decision is strongly influenced by the family background of learners, including ability and the home learning environment as well as socioeconomic circumstances. (Ashenfelter & Rouse, 1998).

[2] Author's own calculations from CAPS 2002

Higher ability individuals are considered less likely to drop out of school on average, as the expected returns from education are higher for individuals with greater ability, and in addition to education providing some general skills training, educational certification is also one way of signalling innate ability. Signalling theory is another economic theory describing the educational investment decision, where the decision to invest in education is a function of signalling one's innate ability rather than necessarily accumulating general or specific skills as is posited in human capital theory. (Spence, 1973)

The relationship between inherited genetic ability and educational attainment is subject to much debate. Some views consider inherent genetic ability to ultimately determine educational attainment, while others consider that education improves individual skills, compensating for, rather than simply augmenting differences in inherent ability.

Difficulties exist with the estimation of the effect of ability on educational outcomes, as innate ability is an unobservable individual characteristic. Some studies make use of test scores as a proxy for ability, but these scores are also a source of endogeneity. Studies by Herrnstein and Murray (1994) were based on the premise that ability is genetically determined, and used test scores to measure the effect of genetically endowed ability on education. Neal and Johnson (1996) argue that the test scores used by Herrnstein and Murray (1994) do not measure inherent ability; but rather reflect differences in "acquired skill" or manifest ability.

Ashenfelter and Rouse (1998) study identical twins, and suggest that returns to schooling may be slightly lower for higher ability individuals suggesting that "schooling compensates for genetic differences in ability". This runs counter to the notion that genetics determines educational attainment; rather it suggests education mediates the differences in latent ability.

Further evidence comes from experiments in Dallas where pupils of average ability were taught by teachers deemed in the top fifth of the profession and ranked amongst the top 10% of learners as a result; while, where they were allocated to teachers from the bottom fifth of the profession, they end up at the bottom. (Rivkin, Hanushek & Kain, 1998)

Other main factors in home background influencing cognitive and educational outcomes referred to in the literature include family income, parents' education, household structure, as well as broader influences of community and socio-economic context.

Belley and Lochner (2007) examine two cohorts (1979 & 1997) of New York City learners from the National Longitudinal Survey of Youth, determining the changing role of family income and ability on educational attainment. They find that cognitive ability is an important predictor of educational outcomes in both cohorts, while family income does not play a big role in high school completion. They

do, however, find that family income matters for tertiary education, as does cognitive ability – this might be explained by the jump in cost of schooling between secondary and tertiary schooling, and their finding that the 1997 cohort are more credit-constrained than the 1979 cohort.

Household poverty is one of the factors found to be associated with lower levels of educational attainment. (Cardoso & Verner, 2006; Rosenthal, 1998, Scott-Jones, 1984, Brooks-Gunn *et al.*, 1997) Long-term family income matters in terms of influencing learner ability as well as earlier grade progression. (Cameron & Heckman, 2001)

One way in which household poverty has an indirect effect on educational attainment is through its association with poor physical health (Scott-Jones, 1984; Brooks-Gunn *et al.,* 1997). Malnutrition affects children's capacity to concentrate and motivation and which in some cases leads to limited cognitive development and learning disabilities which might lead to grade repetition, and lower test achievement scores.

Poverty is correlated with other household characteristics such as family structure, lower levels of parental education, fewer educational resources and poorer neighbourhoods, which make it difficult to tease out the separate effects of these variables, referred to as "poverty co-factors" by Brooks-Gunn *et al.,* (1996), on educational attainment.

With regards to family structure, children from two-parent families generally have lower probabilities of dropping out of high school than do children in other family structures. (Björklund *et al.*, 2007; Rosenthal, 1998). Which parent one lives with, i.e. one's mother or father, also affects one's enrolment decision and progress through school according to Al-Samarrai & Peasgood's (1998) and Anderson & Lam's (2003) findings in Tanzania and South Africa respectively.

Al-Samarrai & Peasgood (1998) and Anderson & Lam (2003) also find that parental education has a strong impact on enrolment and grade repetition. Cameron & Heckman (2001) conclude that long-term family background (measured by parental education) is more important in explaining school attendance than income shortages.

In South Africa, parents of the current generation of youth also suffered from disparities in educational access, and their capacity to assist their children is a function of the education they themselves obtained under Apartheid. In addition to resource disparities, apartheid policies constrained Blacks (who comprise the ethnic majority in South Africa) to low levels of poor quality education (Beutel & Anderson, 2007). According to van der Berg (2006) "A less obvious but perhaps even greater impact of the apartheid

education system was on educational quality differentials which were not reflected in attainment figures and skills levels."

Thus, in addition to the large amount of unemployment in Black and Coloured households causing a financial constraint on access to better quality education – we have the problem of parents being unable to provide educational assistance and compensate for the shortfalls in the public schooling system. Indian and Coloured individuals were also confined to inferior education, relative to the privileged white minority, but were at less of a disadvantage than Blacks under the Apartheid regime.

Parents' educational deficiencies combined with the enduring legacy of schools being under-resourced (both in terms of physical assets and skilled teachers), imply that there is the daunting prospect of the Apartheid system perpetuating itself. That said, while parents differ in their capacity to provide direct educational assistance, simply taking an interest in their children's education, and demonstrating that they value education through homework supervision, for example, could improve their offspring's educational outcomes.

The effect of parent's education and parental presence may vary according to the gender of the parent and child concerned. Thomas (1996) notes that, in many countries, the impact of maternal education on the human capital of children is bigger than that of the father and observes that in South Africa, the parental-education effects are not significantly different.

Gender has also been found to affect the odds of dropout, but is not as strong a predictor as ethnicity or socioeconomic status. Studies exist in which males have been found to be at slightly higher risk of dropping out than females (Combs & Cooley, 1968; Ekstrom *et al*, 1986; Peng & Takai, 1983; Rumberger, 1987 cited in Rosenthal, 1998), but studies have also found that the risk of dropout is independent of gender. (Rumberger, 1990 cited in Rosenthal, 1998).

A number of hypotheses might exist to explain females being at greater risk of dropping out than men, particularly in societies where teen pregnancy is prevalent, or where women's education is not highly valued. Even if females are not at a greater risk of dropping out than men, it seems likely that different factors influence school girl dropout.

In many developing countries, girls' educational attainment is significantly lower than that of their male peers. (Filmer, 2000) Lower educational attainment by girls is borne out in India, for example, where much of the gender gap in educational attainment can be explained by the intra-household allocation of education. (Kingdon, 2001) In rural China it is found that girls' school enrolment is more income elastic than boys' enrolment. (Song, Appleton & Knight, 2006) In terms of levels of attainment, a number of South African studies have concluded that the gap in educational attainment between genders, within

race groups, is virtually non-existent. (Case & Deaton, 1999) In fact, for recent cohorts, educational attainment of females has outstripped that of their male peers. (Anderson & Lam, 2003)

Thomas' (1996) findings suggest that investments in human capital in South Africa are not allocated differentially to sons relative to daughters. Thomas (1996) also notes that human capital investment cannot solely be determined by returns in the labour market, since women of all races are less likely to participate in the labour market than men and typically earn lower wages when they do work. It has also been found that in white and Asian households, girls benefit more from maternal income than their male counterparts, when parental income is not pooled. Where income is pooled, however, paternal income tends to have a higher impact on white and coloured females than males, while the reverse is true for Asian households (Thomas, 1996).

Meekers and Ahmed (1999) examine pregnancy related school dropout in Botswana, from the viewpoint that schoolgirl pregnancy limits the capacity of teen mothers to complete their education, stunting socioeconomic advancement which results in adverse consequences for their offspring. Their results indicate that the problem of schoolgirl pregnancy is more severe than is commonly assumed, and they recommend that programmes be established to facilitate the return of young mothers to school, as well as improving pregnancy prevention programmes.

Meekers and Ahmed (1999) highlight that many pregnant schoolgirls are compelled to drop out of school because of school-level policy on expulsion, and that schoolgirls often have difficulty combining motherhood with formal education. They report that more than half of schoolgirl dropouts occur as a result of pregnancy, and that the percentage of schoolgirls who drop out due to pregnancy is as high as 10% in some African regions. (Gyepi-Garbrah, 1985; Division of Family Health, 1988; Ferguson, Gitonga and Kabira, 1988; Mwateba, Paxman and Weil, 1988; Ferguson 1989; Lesetedi *et al.*, 1989 cited in Meekers and Ahmed, 1999).

While teenage pregnancy does lead many of these mothers to drop out of school, Hill & Jepsen (2006) study a cohort of dropouts in the U.S. and show that many of these individuals have completed secondary school by the time they reach their mid-twenties. Case & Deaton (1999) also suggest that the large numbers of Blacks attempting to complete high school in their early twenties cited prior pregnancy as a reason for interrupted schooling.

In summary, this section has demonstrated that the historical legacy of education in South Africa and the disparities associated with it are related to differences in the process of advancement through school between races. Different theories regarding the relationship between schooling and ability have been presented, and it is noted that while school progress is a function of inherent ability, it also serves to

mediate the effect of differences in inherent ability. In addition, other empirical findings suggest that investment in human capital is not solely a function of the expected returns from schooling in South Africa where females, despite being expected more likely to remain jobless and to earn less than their male counterparts, invest in similar levels of education. While no evidence has been provided to consider that human capital investment is subject to discrimination on the basis of gender, because of teenage pregnancy and its association with girls dropping out, it is advisable to examine males and females independently.

3. DATA DESCRIPTION

SURVEY DESIGN & METHODOLOGY

Rosenthal (1998) reports that the notion of dropout is not consistently defined in the literature and varying definitions lead to several different measures of its incidence. In most circumstances, to be considered a dropout the individual must have missed an entire year of school – but rarely can researchers definitively decide that this is a state of permanent dropout. Most authors acknowledge this issue and then proceed to treat all individuals who have not completed school and are not currently attending as school dropouts. (See Flisher, 2007; Udry & Chantala, 2003; and South *et al.*, 2007)

CAPS is a very rich data source, from the perspective that while it is a household survey, it incorporates schooling data from a national dataset provided by the department of education – the Schools Register of Needs (2000). While the school data is losing currency, and are classroom level averages of aggregate school data, it is contended that this data does much to provide a broader view of the factors affecting educational attainment. Advantages of using CAPS is that the data it contains was collected at the time where youth are most at risk of dropping out, however, some youth remain in school beyond the period for which data is available (and are still at risk of dropping out), and for some of the youth who do drop out in the observation period, this may not be a permanent state of drop out.

Rosenthal (1998) observes that most studies use retrospective rather than longitudinal designs. Some complications arise with the use of retrospective histories rather than longitudinal data. Retrospective histories tend to rely on recording historic events, and may not capture information on all relevant variables at the time that the event occurs. CAPS Wave I collects retrospective data for the youth sample, from their time of birth up to the year 2002. Data was recorded in annual intervals relating to schooling enrolment and performance, living arrangements, and the incidence of marriage and pregnancy as well as whether respondents worked in a given year.

Where retrospective histories are more comprehensive a further source of difficulty creeps in – limited recall accuracy, a phenomenon which is a function of the type of information being recalled. (See Russell Bernard *et al.,* 1984) To illustrate, certain types of events are easier to recall – e.g. marriage, pregnancy etc. but providing retrospective data on income, for example, is far harder, especially for tax-exempt households.

A major difficulty arises with this type of data because (i) information may be inaccurately recalled, (ii) it often relates primarily to outcomes, and (iii) rarely extends to causal factors associated with the event of interest. In order to deal with the lack of concurrent explanatory data, researchers often use the most recent measure on the variable of interest. Using the most recent measure of a given explanatory variable

may introduce bias into the estimation process. For example, using current socioeconomic status to explain dropout 10 years ago potentially suffers from simultaneity bias. Lower socioeconomic status may lead to school dropout, but school dropout may lead to lower socioeconomic status as well. If data is not available around the time that the event occurred, results regarding the effect of socioeconomic status may be misleading.

Ideally, assessing the risk of dropout and its co-factors would be conducted on longitudinal or panel data arranged by school entry cohort – with both events and explanatory data being collected simultaneously. This type of data is rarely found because of the expense of conducting panel surveys. Longitudinal design allows one to employ techniques for analyzing panel data, or using survival analysis to model the risk of dropout, thus taking full advantage of the effect of time-varying co-factors.

In the situation where studying school dropout is not only interesting from the vantage point of whether the event occurs or not, but also the time taken for the event to occur, survival analysis is a useful technique to employ. Learners are at risk of experiencing an event over a given duration of time and the time spent in school may be of interest for a number of reasons. Survival analysis can be used to model educational attainment using time-to-event data. Another advantage it provides is that, unlike simple cross-sectional studies, survival analysis can accommodate censoring. Censoring occurs when individuals are still at risk for an event, but analysis of data occurs before risk period is over, for whatever reason. In terms of this paper, survival analysis can accommodate individuals who are still in school, and at risk, to a greater extent than cross-sectional OLS or standard binary dependent variable Maximum Likelihood Estimation (Jenkins, 2005).

It is now useful to examine why standard binary dependent variable models are problematic when it comes to estimation of duration dependent risks. It seems intuitive that one might simply wish to evaluate whether a given individual has made a transition from one state to another using a binary dependent variable to contrast states, and apply logit or probit regression. In this way, one could get around the censoring issue – observations that made the transition would be denoted by a 1, and censored observations would be denoted by a 0.

This strategy is suboptimal for a number of reasons. It fails to account for the differences in time periods for which each individual is at risk of experiencing the event of interest. One could get around this by considering whether a transition occurred within some predetermined time interval, but this still obscures a vast amount of information; especially if we have a particular interest in when a transition occurs, rather than focusing solely on whether a transition occurs or not (Jenkins, 2005).

The CAPS sample is poorly suited to survival or panel analysis, as the original CAPS sample examined youth aged between 14 and 22, many of whom had already completed or dropped out of school before being surveyed. While retrospective data was collected, retrospective histories collected do not extend to the explanatory variables of interest. Hence the chosen method of modelling was to examine simple binary transitions, from being enrolled to dropping out, with school incomplete. In order to minimize simultaneity bias, the sample was restricted to individuals who were currently enrolled in primary or secondary education in 2002, who did not complete school in 2002, and who were re-interviewed in at least one subsequent wave.

DESCRIPTION OF THE DATA

As previously discussed the data used for analysis is sourced from the Cape Area Panel Study (CAPS) (2002-2005) as well as the Schools Register of Needs Survey (2000). The CAPS survey is a longitudinal study of youth in the Cape Town metropolitan area and, presently comprised of three waves, designed with the intention of tracking the transition of youth in the Western Cape from school into the labour market. Data in CAPS provides demographic information; including rich measures of home background variables. Literacy and numeracy evaluations were administered in the first wave of the survey, and are used as a proxy for ability. In addition, as indicated in the previous section, the first wave of CAPS collected a retrospective annual history for each youth, and this is used to characterise progress through school.

The Schools' Register of Needs (SRN) was compiled in 2000 and is a national survey of South African schools, comprised of information on school management and resources. This study makes use of two variables from the SRN, pupil-teacher ratios and former education department of the school under Apartheid. School name information (not available in the public release) in the CAPS data allows one to link these school-level variables to each respondent, and the merged data are available from the CAPS website. These two data sources in conjunction are used to provide individual, school level and household level data.

This study utilises data from all three waves of CAPS, (2002), (2004a & b) and (2005). The respondents were aged 14-22 years when they were interviewed in 2002. A combination of wave I, II and III data is used, as households were only interviewed in waves I and III. Wave I household and demographic data is used for home-background and individual-level variables respectively, while data from waves I-III are used to determine enrolment status, and progress patterns.

The regression analysis is limited to the sample of individuals enrolled in school in 2002 who do not complete matric in that year. 2002 is used as the baseline, and transitions from being in school to not

being enrolled can occur in 2003, 2004 or 2005. The event of interest is missing one year or more of school. The study is limited to the first instance of missing a year of school in the period post 2002. Hence, it is possible that individuals may return to school (as they are still of a school going age) but in this study one qualifies as a dropout if one misses a year of school or more. This follows from the definition of dropout used in Udry and Chantala (2003) and South *et al.* (2007).

The original wave I sample provided information for 4752 respondents, of whom 4078 were interviewed in Wave II, and only 3531 were interviewed in Wave III.

TABLE 1: Attrition in the CAPS dataset

Attrition in the CAPS dataset			
	No.	%	% attrition
Wave I	4752	100%	
Wave II	4078	86%	14%
Wave III	3531	74%	26%

Reasons for incomplete interviews can be found in the Appendix, Item 1. More than 30% of attrition is due to moving around the country, the most popular destination being the Eastern Cape, movements to the Eastern Cape being slightly greater than movements around the Western Cape. The overwhelming majority of those that move within the country are Black. Those who move to the Eastern Cape are most likely to be Black, have incomplete schooling (in 2002), and have attended either ex-DET or ex-homeland schools. Two thirds of this group are female. This implies that the incidence of dropout is likely to be higher than shown in this study due to the nature of attrition.

Table 2 compares the mean characteristics of the Wave I sample to Wave III. The population group proportion has changed, with Coloured representation increasing relative to that of Blacks and whites. Changes in the population indicate that the incidence of dropout is likely to be underestimated for the Black sample, but other changes are fairly marginal, and the results of the analyses are not likely to be hugely distorted.

Given the slight change in age group, it is also possible that the true incidence of dropout is higher than the subsequent analysis might suggest.

TABLE 2: Variation in Characteristics between waves

Attrition

Variation in characteristics (proportions/means)

	I	III
Black	45%	43%
Coloured	42%	48%
White	13%	10%
Male	45%	46%
Highest grade completed '02	9.17	9.08
Age '02	17.86	17.65
exDET	20%	21%
exhort	21%	24%
exHOA	9%	8%
New Education Dept.	3%	3%
Homeland	1%	1%
Department Missing	48%	43%
Log of Income (p.c)	6.3	6.26
Sample Size	4752	3531

Source: CAPS (2002-2005)

For the entire sample, it can be seen that the majority of both Black and Coloured youth have yet to complete school in 2002.

TABLE 3: Schooling Status (2002)

Entire Sample (2002)						
Schooling Status						
	Black		Coloured		Total	
Incomplete	1767	88%	1490	80%	3257	84%
Complete	241	12%	370	20%	611	16%
Total by Race	2008		1860		3868	

Examining the enrolment status of the 2002 sample, contingent on schooling being incomplete we find Blacks are more likely to be enrolled in school than Coloureds. The dropout rates above are lower than the figure cited by Flisher in 2007, which indicated that 60% of children in South Africa who enrol in Grade 1 drop out before completing high school, according to the South African Department of Education in 2003. This was a national figure however, and variation between provinces might explain some of the discrepancy in dropout rates, in addition to the fact that the level at which most dropouts occur is not made clear.

TABLE 4: Enrolment Status (2002)

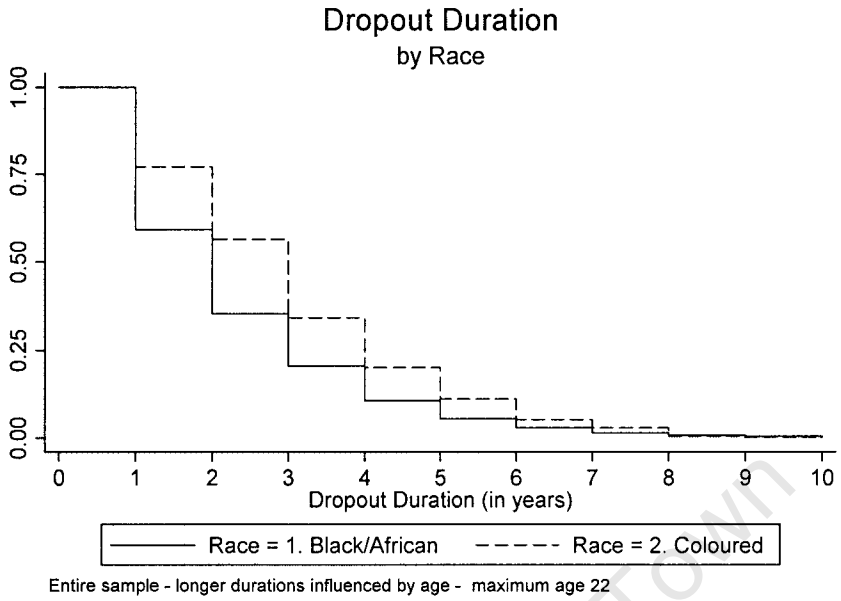
Enrolment Status (2002)						
School Incomplete						
	Black		Coloured		Total	
In School	1230	70%	922	62%	2152	69%
Not Enrolled	536	30%	567	38%	1103	31%
Total by Race	1766		1489		3255	

Using data from the retrospective histories provided by the respondents, the Table 5 summarises the number of years of schooling missed, according to population group. It is evident that the majority of learners have not missed a single year of school across population groups, yet similar levels of both Blacks and Coloureds have missed a year of school or more by 2002. While the bulk of the sample that dropped out have only missed one or two years, there is still a significant proportion of learners who have missed more. This is confirmed in the graph that follows, which plots dropout durations for Blacks and Coloureds who dropped out of school at any point.

TABLE 5: School Years Missed (2002)

Years of School Missed (2002)						
	Black		Coloured		Total	
None	1478	74%	1356	73%	2834	73%
1 Year	216	10.76%	115	6.18%	331	8.56%
2 Years	127	6.32%	105	5.65%	232	6.00%
3 Years	79	3.93%	112	6.02%	191	4.94%
4 Years	52	2.59%	71	3.82%	123	3.18%
5 Years	27	1.34%	45	2.42%	72	1.86%
6 Years	14	0.70%	30	1.61%	44	1.14%
7 Years	7	0.35%	12	0.65%	19	0.49%
8 Years	4	0.20%	12	0.65%	16	0.41%
9 Years	2	0.10%	1	0.05%	3	0.08%
10 Years	2	0.10%	1	0.05%	3	0.08%
Total	2008		1860		3868	

Figure 1: Dropout Duration (2002)



Short dropout periods occur for two reasons, (i) dropout is not always permanent & (ii) dropout duration is influenced by age. Nonetheless, given that the maximum age of the sample is 22, some learners have dropped out of school before the legal exit age, which is 15. (The South African Schools Act (84), 1996)

TABLE 6: Highest Level of Education completed prior to first instance of dropout

Level of Education Completed *prior to first year missed*		
Grade 1 or less	17	1.6%
Grade 2	11	1.0%
Grade 3	18	1.7%
Grade 4	33	3.1%
Grade 5	34	3.2%
Grade 6	82	7.8%
Grade 7	171	16.3%
Grade 8	170	16.2%
Grade 9	194	18.5%
Grade 10	185	17.6%
Grade 11	135	12.9%
Total	1050	

The prior table shows the highest level of education completed before the first instance of dropout. It is evident that the bulk of the sample, i.e. 65.2% of dropouts exit after having completed primary school. In addition, a substantial number of learners exit just prior to school completion.

TABLE 7: Years of schooling prior to first instance of dropout

Time to Dropout			
		% Dropouts	Cumulative %
1 Year	9	0.9%	0.9%
2 Years	6	0.6%	1.4%
3 Years	10	1.0%	2.4%
4 Years	11	1.0%	3.4%
5 Years	20	1.9%	5.3%
6 Years	33	3.1%	8.5%
7 Years	78	7.4%	15.9%
8 Years	100	9.5%	25.5%
9 Years	131	12.5%	37.9%
10 Years	216	20.6%	58.5%
11 Years	196	18.7%	77.2%
12 Years	154	14.7%	91.9%
13 Years	61	5.8%	97.7%
14 Years	20	1.9%	99.6%
15 Years	4	0.4%	100.0%
Total	1049		

Figure 2: Time in school before dropping out

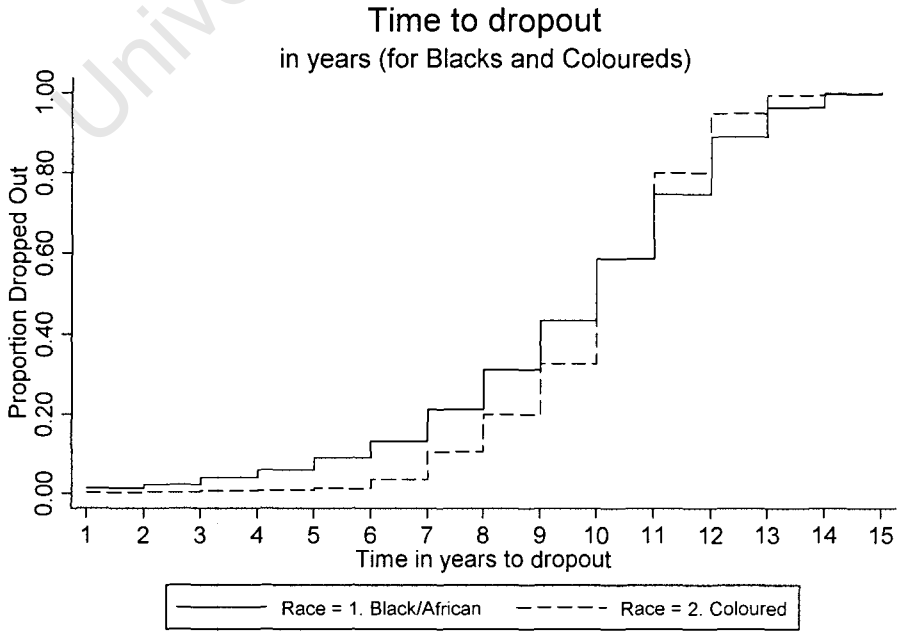


Figure 2 shows time (in years) spent in school before the first instance of dropout for Coloureds and Blacks. What is evident from this illustration is that while a bigger proportion of Blacks drop out early, this rate equalises after 10 years in school, whereafter a greater proportion of Coloureds drop out than Blacks.

What the preceding summary data has shown is that completion and enrolment patterns vary by race. It is also evident that most individuals spend a significant number of years in school before dropping out, i.e. most individuals observe the legal obligations regarding school attendance.

For the purposes of regression analysis the sample is restricted to individuals who are enrolled in the traditional school system in 2002, who are re-interviewed in one or more subsequent waves, for whom sufficient schooling data is available. Analysis is limited to Blacks and Coloureds, because a negligible number of Whites drop out of school. In addition, none of the Coloured girls or Black males who attend white schools drop out either.

It can also be seen that the proportion of grades successfully completed relative to years enrolled is lowest for the Black sample – leading them to comprise the biggest proportion of learners still enrolled in school. This is a preliminary indicator that grade repetition does not automatically lead Black learners to drop out permanently.

TABLE 8: Summary Data for those enrolled in school in 2002

Summary Statistics for those enrolled in school (2002)		
	Black	Coloured
Enrolled	1232	912
Mean Years Enrolled	10.04	10.23
Mean Enrolment level (Grade)	9.18	9.68
Proportion of years successfully completed	90.49%	93.93%

Explanatory variables used in the analysis are at the individual, school and household level. Individual-level variables include gender, population-group, highest level of education completed in 2002, the sum of school years missed up to and including 2002, the sum of years failed before 2002, failure prior to the last year in the sample, as well as z-scores from the literacy and numeracy evaluation administered in wave I. Household level variables relate to household composition: the proportion of life lived with either parent, as well as household "knowledge resources": biological parents' education (which is not strictly a household variable, as the data relates to biological parents, not all of whom are co-resident), parents' help with homework as well as financial

resources: per capita household income. School quality controls relate to pupil teacher ratios, and former school department – this data is taken from SRN (2000) in relation to the schools respondents attended in 2002, and it is assumed that learners do not frequently swap between school types.

Separate summary data is provided for male and female youth, according to race and dropout status under items 3 and 4 in the appendix.

Examining the summary data for Black males: it is evident that relative to Black males who do not drop out, Black male dropouts have lower LNE scores, but do not appear to be from lower income households. The biological parents' years of education relating to individuals for whom we have data do not appear to vary greatly between those who drop out and those who do not drop out and it appears that similar amounts of education data are missing for dropouts and non-dropouts. Proportions of life lived with parents also do not vary between Black males who drop out and those who do not, but parental assistance is higher among individuals who did not drop out.

The pupil teacher ratio for Black male dropouts appears lower than for those who did not drop out, but there is also a higher level of missing pupil-teacher ratio data missing for Black male dropouts. The lower pupil-teacher ratio for dropouts suggests that this data is problematic. Black males attend predominantly exDET schools, and it is worth noting that none of the Black males who attend former white schools dropped out.

Black male dropouts have missed more years of school, have failed more years of school, and are also more likely to have failed in the year preceding their exit from school.

Examining the summary data for Coloured males, dropouts performed substantially worse on literacy and numeracy evaluation, and appeared to come from slightly more impoverished households. The biological parents' years of education relating to individuals for whom we have data does is substantially lower for those who drop out and given that greater proportions of education data are missing for dropouts and non-dropouts, it is probable that parents' years of education are even lower for dropouts than indicated by the data available. The proportion of life lived with one's father is significantly lower for those Coloured males who dropped out, and these individuals are less likely to report that they receive help with homework from their parents.

Pupil teacher ratios for Coloured male dropouts appear lower than for those who did not drop out, but as with Black males, there is also a higher level of missing pupil-teacher ratio data missing for dropouts. The lower pupil-teacher ratio for dropouts once again suggests that this data is problematic, and appears biased by missing data, thus results relating to class size are unlikely to be

meaningful. Coloured males predominantly attend former Coloured schools, and very few of those who attend former White schools dropped out.

In terms of their school progress history, Coloured male dropouts have missed significantly more years of school, have failed more years of school, and are also more likely to have failed in the year preceding their exit from school.

Given the association of failure with eventual dropout, it is worthwhile to do a preliminary examination of the association of failure with performance on the literacy and numeracy evaluation (kernel density plot provided in the Appendix – Item 5). From the plot, it is apparent that Coloured males who dropped out performed far worse on the literacy and numeracy evaluation, whereas there appears to be very little difference in the performance of Black males who dropped out, and those who did not. To the extent that the LNE scores are a reasonable proxy for the potential to succeed at school, this indicates that Black males dropping out appears far more stochastic than Coloured males dropping out.

Looking at the summary data for the female youth (see Appendix - Item 4, both Black and Coloured dropouts are far more likely to have given birth than their counterparts who did not drop out. In addition dropouts performed worse on the literacy and numeracy evaluation, and the Coloured girls appear to be from poorer households although the difference between Black girls' home income is not remarkable.

Both Black girls and Coloured girls who drop out appear have biological parents who are less educated than those who do not drop out, but the difference is greatest when comparing Coloured females. The amount of missing education data for parents is similar between Black females who dropped out, and those who did not, but this data is more likely to be missing for Coloured girls who dropped out. The proportion of life lived with one's parents does not differ between Black girls who dropped out and those who did not, whereas Coloured girls who did not drop out have spent a greater proportion of their lives living with their parents.

For Coloured girls who dropped out both parents are less likely to have helped with homework, whereas only Black dropouts report having less assistance from their mothers only, and no difference in the lower level of assistance from their fathers.

Pupil-teacher ratios seem to be higher for the Coloured girls who dropped out, but counter-intuitively appear slightly higher for Black girls who did not drop out. Once again, this is probably a function of there being more missing pupil teacher ratio data for dropouts. The majority of Black

girls attend exDET schools, and the majority of Coloured girls attend exHOR schools, and it is worth noting that none of Coloured girls who attend former White schools dropped out.

Both Black and Coloured girls who dropped out are more likely to have missed a year of school, be further behind in school due to grade retention, and have failed in the year prior to dropping out.

Additional information is provided about the association of failure with performance on the literacy and numeracy evaluation for Black and Coloured girls. (See the kernel density plot provided in the Appendix – Item 6). Here it seems that failure is associated with poorer performance on the evaluation for both Coloured and Black girls.

This preliminary description of the data as well as the differences in school quality and assessment capabilities has provided motivation for independent analysis of the population groups. Given that the precise effect of contributing factors to school dropout were considered to vary by gender as well as race, these populations are analysed independently. The primary results presented in the next section relate to school dropout – but a primary determinant of school dropout was found to be failure prior to dropping out, hence some analysis of the co-factors of failure (assumed to be similar to dropout) was undertaken.

4. RESULTS

Regression output can be found in the appendix. Item 7 contains the failure regressions, and Item 8 contains the dropout regressions. First, a brief discussion of the failure regression results is provided, after which discussion of the main dropout regression results follows. Please note that all regression coefficients are in exponentiated form, i.e. they are odds-ratios. The exponentiated coefficient describes the increase in the odds of failure for a one unit change in the explanatory variable. This form of presentation is recommended by Peng *et al.* (2002) who compare various methods of reporting logit results in educational research. In addition, Gould (2000) also advocates that results can most easily be interpreted using odds-ratios.

Failure Regressions

Many of the covariates included in the failure regressions for males are not found to be significant, but it is worth noting that performance on the literacy and numeracy evaluation matters for Coloured males, while it is not significant for Black males which suggests that grade retention for Coloured males is very likely a function of educational competency, whereas the same is not true for Black males. This is not out of line with Lam *et al.'s* (2007) finding that school progression is stochastic in Black schools, but not in Coloured schools. In addition, attendance of former Whites only schools lowers the odds of failing substantially – this is very possibly a function of better teaching.

For Coloured males, their biological mother's education appears to reduce the odds of failure slightly for each additional year of education.

When examining the odds of failure for Black and Coloured girls, performance on the Literacy and Numeracy evaluation was significant in reducing the odds of failure for both groups, although improved performance on the Literacy and Numeracy evaluation reduced the odds of failure by a greater magnitude for Coloured girls than Black girls. The effect of increased income was also insignificant for Black and Coloured girls as well as boys. Neither biological parents' education nor the proportion of life spent living with their parents seems to matter for either group of girls. There might be some concern that the effect of parents' education is being obscured, but the same proportion of education data was missing for both individuals who failed and those that did not, so unless the missing parents were substantially different to those for whom data exists, this concern is limited.

Pupil-teacher ratios are not significant, but this is unsurprising, given that the means for individuals who had failed were slightly lower than those who did not fail. Further difficulties associated with the use of the pupil-teacher ratios provided in the data is that they refer to the departmental allocation of teachers, which is unlikely to reflect the actual attendance of teachers, especially in poorer ill-functioning schools.

For both groups of girls, attendance of former white schools is associated with lower odds of failure than attending exDET schools (for Black girls) and exHOR schools (for Coloured girls). Once again, this suggests that the individuals who have managed to attend better schools are different to (have better access to educational resources and may be of higher ability) those who remained confined to schools historically associated with their race.

Dropout Regressions

With these regressions, it is interesting to observe that performance on the literacy and numeracy evaluation mattered for each group under consideration, except for Black males. While this is unsurprising, given the summary data presented in the previous section, this result is interesting because to the extent that the LNE z-scores serve as a reliable proxy for the ability to succeed in school, it suggests that different forces may be at play in the decision to drop out for Black males in comparison to Black females.

Simple attitudinal differences might explain the variation in the dropout decision making process, which may both be consistent with the Human Capital Framework. Black males may recognise that independent of their ability, education within the exDET system will not significantly improve their odds of employment/increase their wages and hence choose to terminate the schooling prematurely, while Black females may be more optimistic about the value of education in relation to their own abilities. As Black females are even less likely to be employed than Black males (Thomas, 1996) this result may simply reflect a lower opportunity cost of remaining in school.

For the Coloured population, it is evident that better performance in the LNE is associated with a bigger lowering of the odds of dropping out, and, to the extent that the schools Coloureds attend function better than Black schools in terms of evaluating educational performance and rewarding ability this result makes intuitive sense.

Higher family incomes lower the odds of Coloured youth dropping out, but the same does not apply to Black youth. This could be because of better outside options available to Coloured youth than Black youth either in the job market or in terms of support from relatively better off families.

In terms of the highest level of education completed, the reaching of the legally obligated level of schooling is strongly significant in increasing the odds of dropping out for Black males only, and has no impact on the other groups being considered. If we interpret performance on the literacy and numeracy evaluation as an indicator of the potential to complete school, given that there is no difference in average performance (or income) between the Black males who drop out and those who do not there is scope for further analysis to identify other push-out factors e.g. household shocks.

Both Coloured males and females have significantly lower odds of dropping out the closer they are to completing their schooling. However, this does not imply that they successfully complete school; as failure in matric is not considered to be the same as dropping out.

Missing a year or more of school before 2002 increases the odds of dropping out for Black males and Coloured females, but not for the two other groups. Across population groups for both genders grade retention prior to 2002 raises the odds of dropping out by broadly similar amounts for each year failed. Failure between 2002 and 2005, preceding the final year of observation has a universally significant and strong impact on the odds of dropping out in the following year. This suggests that learners take into account recent failure when making the decision to drop out, but whether this is because of perceived increased costs of remaining in school, lower benefits from remaining in school, or disillusionment is uncertain. To the extent that Lam *et al.*'s (2007) suggestion that Black schools on average are characterised by a stochastic grade promotion process is true, accompanied by the earlier finding that literacy and numeracy levels are not significant in terms of which Black males fail, the resulting decision to drop out may not be based on any real difference in the capacity to successfully complete school.

Biological parents' education was included in the model to capture the correlation between genetically inherited traits and educational attainment, and does not appear to impact on the odds of dropping out, except in the case of the fathers of Coloured males, where more educated fathers reduce the odds of dropping out. The significant coefficient on mother's education being missing for Black males suggests that this group of offspring are different, but it is unclear why this is the case.

For Coloured girls, the proportion of life lived with their mothers significantly reduces the odds of dropping out, while living with their fathers raises the odds of dropping out significantly. While the two are correlated, there are reasons to expect that living with one's mother has a different impact on a girl's wellbeing to living with one's father. Thus while the results found may initially appear peculiar, they may make more sense if it is poorer and less capable parents who tend to stay together, and more capable mothers leave their husbands.

Assistance with homework from mothers mattered for both Black boys and girls, but not for Coloureds, and assistance from their fathers mattered for both Coloured and Black males. This is interesting as Blacks received less assistance than their Coloured counterparts, but the assistance those who receive assistance appear to benefit from it. Homework assistance from fathers only matters for the boys lowering the odds of dropping out significantly for both Coloured & Black males. More males reported receiving assistance from their fathers than the girls, and evidently their contribution was valuable. Mothers assisting lowered the odds of both Black girls and boys dropping out, but had no effect on Coloured boys and girls. Whether competent assistance is provided is questionable, but it appears that

taking an interest and being willing to assist provides some motivation for the boys, as well as Black girls; although it is interesting to observe that the gender of the parent providing assistance seems to be significant.

With respect to the school quality variables, it is unsurprising that the pupil-teacher ratios did not turn out to be significant; they appeared peculiar when examined under the summary data section, and it is thought that the theoretical pupil-teacher allocation is not the same as the actual attendance or input of teachers in poorer schools.

Before examining whether the school-type coefficients are significant - it is first worth noting that none of the few Black males in former White schools, and none of the Coloured girls in former White schools dropped out (and these observations were not included in the model). For Black males attending a former Coloured school lowered the risk of dropping out, while this effect was not significant for Black females. However, attending former White schools, or schools established after the end of Apartheid lowered the odds of Black girls dropping out. Coloured males fared worse if they attended a school which fell under the category other, which refers to schools which were unclassifiable.

The impact of child-bearing for Black and Coloured females raised the odds of dropping out immensely. It is worth noting however that the magnitude of the size of this coefficient may be upwardly biased to the extent that unobservable household characteristics influence the odds of becoming pregnant and dropping out as well. In addition, given Case & Deaton's (1999) observation that many Black child mothers return to school in their twenties, the incidence of permanent drop out may not be as high as the CAPS dataset indicates, due to the age bracket of the sample.

5. CONCLUSION

This study has shown that prior progress through school is a big predictor of school dropout. It has demonstrated that there are differences in the process of advancement through school between races. Notably, in the Coloured population failure is associated with literacy and numeracy levels, suggesting that advancement through school is a meritocratic process. For the Black population, the association between the capacity to succeed at school and the actual incidence of dropping out is far less clear-cut, especially for Black males. School environments which are characterised by a lack of adequate resources and disruptive learning environments imply that the individual capacity to succeed in these schools is limited, and exiting school prematurely may be entirely rational.

For the most part, parental interest and input is shown to be associated with lowering the odds of dropping out, whilst biological parents' education has an insignificant effect on the odds of dropping out of school.

Pregnancy has a significant association with school dropout for girls, and while it is possible that this effect is overstated, because of the association between pregnancy and other adverse home circumstances. While teenage pregnancy is concern, especially when considering the well-being of the offspring of teen mothers, this result may simply reflect the lower opportunity cost of pregnancy for the girls who become pregnant, who may have had poor prospects before becoming pregnant anyway.

In terms of labour market or returns-to-schooling literature, the results suggest that educational investment for the Black population relies on unreliable signals of an individual's capacity to complete school. The significance of the effect of income for the Coloured population on the decision to drop out is far more likely to be a rational response in relation to their labour market prospects and the limited gains to be had from remaining in school for the least capable and poorest Coloured individuals.

APPENDIX

Item 1

Final Result of Wave III interview

	Freq.	Percent
1. Completed	3531	74.62%
2. Not Available	210	4.44%
3. Refused	191	4.04%
4. Deceased	29	0.61%
6. Moved within Cape Town	134	2.83%
7. Moved within South Africa	299	6.32%
8. Moved Abroad	67	1.42%
9. Moved No details	232	4.90%
10. Temporary Living Situation	29	0.61%
11. Mentally Unfit/Disabled	2	0.04%
14. No contact	8	0.17%
Total	4732	

Source: CAPS (2002-2005)

Item 2

Destination
Wave III interview unsuccessful

	Freq.	Percent
Eastern Cape	160	35.6%
Free State	4	0.9%
Gauteng	52	11.6%
Kwa-Zulu Natal	9	2.0%
Limpopo	2	0.4%
Mpumalanga	2	0.4%
Northern Cape	7	1.6%
North West	1	0.2%
Western Cape	156	34.7%
Outside S.A.	56	12.5%
Total	449	

Source: CAPS (2002-2005)

Item 3

BM = Black males CM = Coloured males

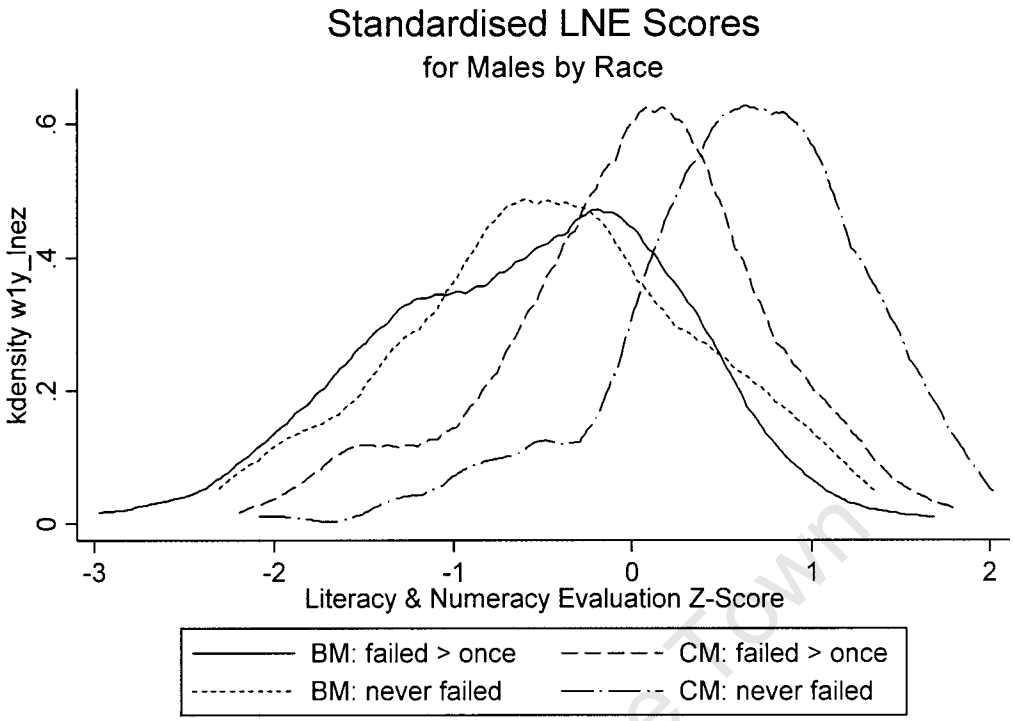
Boys	*Did not drop out*		*Dropped Out*	
	BM	CM	BM	CM
Variable	Mean			
Standardised LNE Score	-0.5211126	0.4227288	-0.6562501	-0.1464277
Log of Income (p.c.)	5.577541	6.779575	5.55191	6.166093
Mother's Education	7.72807	9.040201	7.542169	7.685393
Father's Education	7.54386	9.708543	7.072289	7.505618
Father's Education missing	0.3952255	0.2226563	0.4195804	0.3732394
Mother's Education missing	0.331565	0.1953125	0.2937063	0.3450704
Proportion of life lived with father	0.4898456	0.7013608	0.486894	0.5970554
Proportion of life lived with mother	0.8337478	0.9164163	0.8049045	0.896656
Mum helps with homework	0.2175066	0.3164063	0.1118881	0.2253521
Dad helps with homework	0.061008	0.2148438	0.020979	0.056338
Pupil-teacher ratio	29.8435	29.09375	27.03497	28.76056
Pupil-teacher ratio missing	0.1299735	0.03125	0.1888112	0.084507
exDET	0.7135279	0.0117188	0.7132867	0.028169
exHOA	0.0344828	0.125	0	0.0140845
exHOR	0.0875332	0.8476563	0.0699301	0.8873239
New Education Department	0.0848806	0	0.0699301	0.0070423
Former Dept. Missing	0.0583554	0.015625	0.0979021	0.0633803
School years missed	0.0954907	0.0117188	1.433566	1.338028
Prior Failure	0.0954907	0.1289063	0.3076923	0.4014085
Sum of failures before 2002	0.8435013	0.3945313	1.20979	0.9295775
Sum of Failures (2002 - 2005)	0.3076923	0.234375	0.4405594	0.5704225
Grade 7 (2002): marginal	0.7241379	0.953125	0.8461538	0.8732394
Grade 8 (2002): marginal	0.5676393	0.796875	0.7832168	0.6549296
Grade 9 (2002): marginal	0.3687003	0.5234375	0.4475524	0.3873239
Grade 10 (2002): marginal	0.2175066	0.359375	0.2167832	0.1619718
Grade 11 (2002): marginal	0.0848806	0.1835938	0.048951	0.028169
Observations	374	256	142	141

Item 4

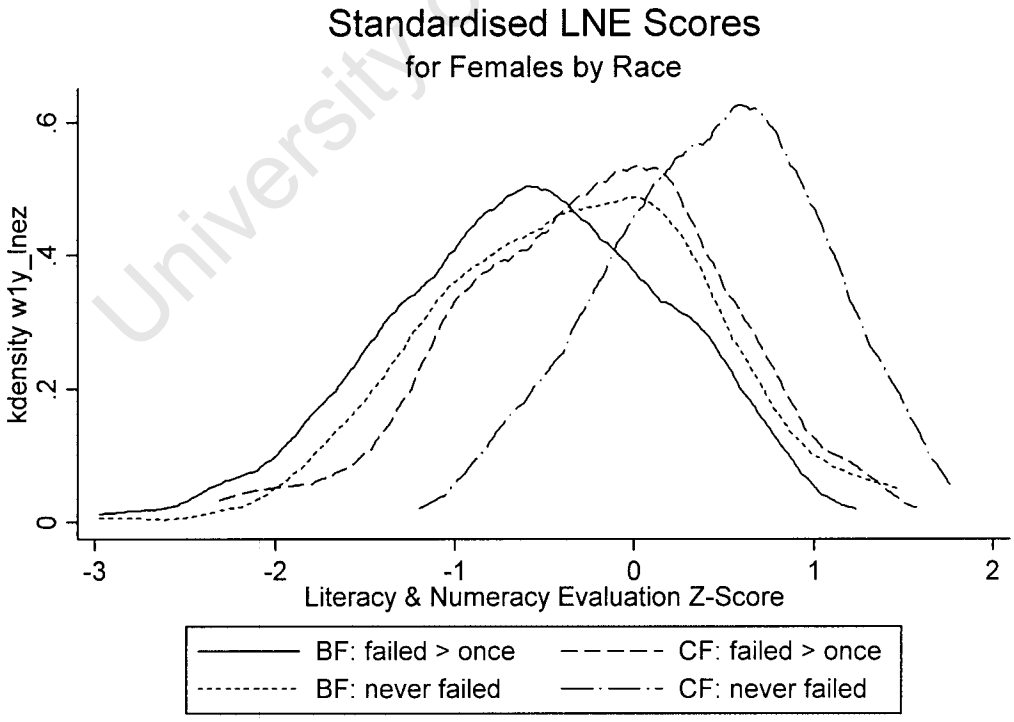
BF = Black females CF = Coloured females

Girls	***Did not drop out***		***Dropped Out***	
	BF	CF	BF	CF
Variable	Mean			
Ever given birth?	0.117773	0.081571	0.425	0.2484076
Standardised LNE Score	-0.4589148	0.286584	-0.6468052	-0.3173471
Log of Income (p.c.)	5.507992	6.6948	5.398527	6.090565
Mother's Education	7.538732	8.946502	7.120879	7.104651
Father's Education	7.25	9.641975	6.692308	7.651163
Father's Education missing	0.391863	0.265861	0.43125	0.4522293
Mother's Education missing	0.3126338	0.223565	0.3	0.3566879
Proportion of life lived with father	0.4696041	0.6582062	0.4616095	0.5315371
Proportion of life lived with mother	0.7878908	0.90461	0.7647277	0.8399486
Mum helps with homework	0.1541756	0.2628399	0.08125	0.1974522
Dad helps with homework	0.0492505	0.1661631	0.05	0.0828025
Pupil-teacher ratio	29.73019	28.24471	26.625	29.95541
Pupil-teacher ratio missing	0.1134904	0.0574018	0.19375	0.0318471
exDET	0.6959315	0.0151057	0.675	0.0191083
exHOA	0.0235546	0.1570997	0.00625	0
exHOR	0.111349	0.8096677	0.0875	0.9681529
New Education Department	0.0792291	0.0120846	0.05	0
Former Dept. Missing	0.0728051	0.0060423	0.14375	0.0127389
School years missed	0.0706638	0.0151057	1.45625	1.33758
Prior Failure	0.1584582	0.081571	0.35	0.4076433
Sum of failures before 2002	0.5888651	0.3413897	0.9	0.7898089
Sum of Failures (2002 - 2005)	0.3768737	0.1601208	0.5	0.5350318
Grade 7 (2002): marginal	0.8586724	0.9637462	0.8875	0.9363057
Grade 8 (2002): marginal	0.6980728	0.8308157	0.79375	0.7961783
Grade 9 (2002): marginal	0.4304069	0.6163142	0.55	0.4267516
Grade 10 (2002): marginal	0.2269807	0.4048338	0.3	0.1656051
Grade 11 (2002): marginal	0.0985011	0.1691843	0.1125	0.044586
Observations	461	329	159	155

Item 5



Item 6



Item 7

BM = Black males CM = Coloured males BF = Black females CF = Coloured females

Failure Regression

	Failure (2002 - 2005)			
	BM	*CM*	*BF*	*CF*
Standardized LNE total score	0.949	0.615	0.652	0.549
	[0.115]	[0.103]***	[0.076]***	[0.090]***
Log of income (p.c.)	0.943	0.893	0.913	0.891
	[0.102]	[0.154]	[0.095]	[0.133]
Mother's Education Missing	0.784	1.341	0.978	0.751
	[0.290]	[0.943]	[0.318]	[0.371]
Father's Education Missing	1.491	0.583	1.128	1.314
	[0.737]	[0.487]	[0.447]	[0.891]
Father's education	1.02	1.005	1.017	1.015
	[0.039]	[0.047]	[0.032]	[0.052]
Mother's education	0.997	0.924	1.038	0.985
	[0.027]	[0.033]**	[0.026]	[0.035]
Proportion of life lived with father '02	0.714	0.724	1.023	0.609
	[0.200]	[0.255]	[0.243]	[0.219]
Proportion of life lived with mother '02	0.87	2.324	1.467	1.112
	[0.352]	[1.368]	[0.474]	[0.562]
Pupil-Teacher Ratio (SRN 2000)	0.975	0.995	0.973	1.018
	[0.022]	[0.044]	[0.020]	[0.037]
Pupil-Teacher Ratio Missing	0.125	0.08	0.203	1.928
	[0.119]**	[0.137]	[0.197]*	[2.470]
exDET School	0.501		1.134	
	[0.427]		[1.046]	
exHOA School	0.063	0.135	0.107	0.181
	[0.088]**	[0.104]***	[0.141]*	[0.145]**
exHOR/exHOD School	0.522		1.189	
	[0.460]		[1.101]	
New Education Department	0.458		0.844	
	[0.418]		[0.809]	
Former School Department Missing	1.332		1.869	
	[0.944]		[1.247]	
Other department				0.509
				[0.463]
Observations	516	394	623	484

Coefficients presented in exponentiated form: i.e. they are odds-ratios

Robust standard errors in brackets

* significant at 10%; ** significant at 5%; *** significant at 1%

Item 8

BM = Black males CM = Coloured males BF = Black females CF = Coloured females

Dropout Regression

	Dropout Status (2003-2005)			
	BM	*CM*	*BF*	*CF*
Standardized LNE total score	0.817	0.635	0.75	0.511
	[0.122]	[0.132]**	[0.107]**	[0.115]***
Log of income (p.c.)	1.084	0.562	0.999	0.575
	[0.142]	[0.106]***	[0.127]	[0.116]***
Grade 7 in 2002	0.685	0.939	0.58	1.256
	[0.358]	[0.614]	[0.272]	[0.854]
Grade 8 in 2002	3.765	1.306	1.535	1.257
	[1.689]***	[0.530]	[0.602]	[0.529]
Grade 9 in 2002	0.789	1.501	1.399	0.991
	[0.250]	[0.647]	[0.423]	[0.359]
Grade 10 in 2002	0.944	0.595	1.344	0.471
	[0.336]	[0.275]	[0.441]	[0.192]*
Grade 11 in 2002	0.49	0.194	0.74	0.231
	[0.257]	[0.114]***	[0.310]	[0.137]**
Sum of Years Missed '02	1.432	3.374	1.192	2.209
	[0.267]*	[2.691]	[0.390]	[1.006]*
Sum of Years Failed '02	1.426	1.782	1.274	1.897
	[0.158]***	[0.353]***	[0.160]*	[0.335]***
Failed Prior	3.929	5.039	3.304	6.968
	[1.088]***	[1.729]***	[0.811]***	[2.510]***
Mother's Education Missing	0.475	0.957	0.622	0.9
	[0.206]*	[0.676]	[0.248]	[0.609]
Father's Education Missing	1.498	0.451	1.874	0.992
	[0.836]	[0.416]	[0.994]	[0.890]
Father's education	0.976	0.859	0.995	0.927
	[0.042]	[0.053]**	[0.042]	[0.057]
Mother's education	0.98	0.981	1	1.063
	[0.030]	[0.040]	[0.028]	[0.051]
Proportion of life lived with father (2002)	1.182	1.276	1.25	2.276
	[0.371]	[0.576]	[0.375]	[1.088]*
Proportion of life lived with mother (2002)	0.9	0.767	0.934	0.303
	[0.416]	[0.563]	[0.355]	[0.206]*
Mum helped with homework	0.543	0.638	0.522	0.869
	[0.196]*	[0.207]	[0.198]*	[0.286]
Dad helped with homework	0.308	0.307	1.152	1.013
	[0.191]*	[0.171]**	[0.609]	[0.416]

Pupil-Teacher Ratio (SRN 2000)	0.965	1.047	0.958	0.943
	[0.031]	[0.060]	[0.032]	[0.049]
Pupil-Teacher Ratio Missing	0.299	17.818	0.168	0.061
	[0.379]	[32.667]	[0.229]	[0.111]
exDET School	0.352		0.251	
	[0.277]		[0.253]	
exHOR/exHOD School	0.23		0.264	
	[0.197]*		[0.268]	
New Education Department	0.335		0.139	
	[0.290]		[0.160]*	
Former School Department Missing	0.573		0.509	
	[0.405]		[0.346]	
exHOA School		0.367	0.045	
		[0.324]	[0.074]*	
Other Department		4.556		3.075
		[4.114]*		[3.058]
Ever given birth?			5.191	5.453
			[1.370]***	[1.942]***
Observations	503	397	620	432

Coefficients presented in exponentiated form: i.e. they are odds-ratios
Robust standard errors in brackets
* significant at 10%; ** significant at 5%; *** significant at 1%

REFERENCE LIST

***The Employment of Educators Act, 1998* (Act No. 76 of 1998)**
http://wced.wcape.gov.za/documents/legislative_acts/a76-98.pdf

***The Further Education and Training Act, 1998* (Act No 98 of 1998)**
http://www.info.gov.za/gazette/acts/1998/a98-98.pdf

***The South African Schools Act, 1996* (Act No 84 of 1996)**
http://www.info.gov.za/acts/1996/a84-96.pdf

No Author (2007) "How to be Top" *The Economist (Print Edition)*
Available online: <http://www.economist.com/world/international/displaystory.cfm?story_id=9989914>

Al-Samarrai, S. and Peasgood, T. (1998), "Educational Attainments and Household Characteristics in Tanzania", *Economics of Education Review,* 17:4, pp. 395-417.

Anderson, K.G. and Lam, D. (2003), "Dynamics of Family Structure and Progress Through School in South Africa: Evidence from Retrospective Histories", Prepared for presentation at the Annual Meeting of the Population Association of America, Minneapolis, Minnesota, May 2003.

Ashenfelter, O. and Rouse, C. (1998), "Income, Schooling, and Ability: Evidence from a New Sample of Identical Twins", *The Quarterly Journal of Economics*, 113:1, pp. 253-284.

Becker, G. (1962), "Investment in Human Capital: A Theoretical Analysis", *The Journal of Political Economy*, 70:5, Part 2: Investment in Human Beings, (Oct., 1962), pp. 9-49. JSTOR.
Available online: <http://www.jstor.org/stable/1829103>

Belley, P. and Lochner, L. (2007), "The Changing Role Of Family Income And Ability In Determining Educational Achievement", *NBER Working Paper 13527,* National Bureau of Economic Research.
Available online: <http://www.nber.org/papers/w13527>

Beutel, A. and Anderson, G. K. (2007), "The Educational Expectations of South African Youth", *Sociological Focus,* 40:4. Available online: <http://faculty-staff.ou.edu/A/Kermyt.G.Anderson-1/papers/adol_edexp.pdf>

Bhorat, H. (2007), "Unemployment in South Africa: Descriptors & Determinants", Prepared for presentation to the Commission on Growth and Development, World Bank, Washington DC, October, 2007.

Björklund, A., Jäntti, M. and Solon, G. (2007), "Nature and Nurture in the Intergenerational Transmission of Socioeconomic Status: Evidence from Swedish Children and their Biological and Rearing Parents", *NBER Working Paper 12985*, National Bureau of Economic Research. Available online: < http://www.nber.org/papers/w12985>

Brooks-Gunn, J., Klebanov, P. and Duncan, G. (1996), "Ethnic Differences in Children's Intelligence Test Scores: Role of Economic Deprivation, Home Environment, and Maternal Characteristics", *Child Development*, 67:2, pp. 396-408.

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

Cardoso, A. and Verner, D. (2006), "School Drop-Out and Push-Out Factors in Brazil: The Role of Early Parenthood, Child Labour, and Poverty", *IZA Discussion Paper 2515.*Institute for the Study of Labour.
Available online: <http://ssrn.com/abstract=955862>

Case, A. and Deaton, A. (1999), "School Inputs and Educational Outcomes in South Africa", *The Quarterly Journal of Economics*, CXIV(3), pp. 1047-1084.

Crouch, L. and Mabogoane, T. (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, pp. 60-78. Available online: <http://www.rti.org/ddsp/documents/no_magic.pdf>

Fallon, P. and Lucas, R. (1998), "South African Labour Markets: Adjustment and Inequalities." *Informal discussion papers on aspects of the economy of South Africa No. 12.* Washington: The World Bank Southern Africa Department, 1998.

Filmer, D. (2000), "The Structure of Social Disparities in Education: Gender & Wealth", *Policy Research Working Paper 2268*, The World Bank Development & Research Group, Poverty & Human Resources, January 2000. Available online: <http://www-wds.worldbank.org/servlet/WDSContentServer WDSP/IB/2000/02/09/000094946_00012505525066/ Rendered/PDF/multi_page.pdf>

Flisher, A., Townsend, L., Chikobvu, P., Lombard, C. and King, G. (2007), "*Substance Use and High School Attrition: A Longitudinal Analysis of Substance Use and High School Attrition in Cape Town, South Africa."* Unpublished.

Gould, W. (2000), "Interpreting Logistic Regression in all its Forms", *Stata Technical Bulletin,* STB-53, January 2000, STATA CORP. Available online: <http://www.ats.ucla.edu/stat/Stata/library/sg124.pdf>

Hill, L. and Jepsen, C. (2007), "Positive outcomes from poor starts: Predictors of dropping back in" *Economics of Education Review*, 26, pp. 588–603. Science Direct.

Jenkins, Stephen. (2005), Survival Analysis. Lecture notes for Survival Analysis with Stata: Course EC968. Offered by the Institute for Social & Economic Research, University of Essex. Available online: <http://www.iser.essex.ac.uk/teaching/degree/ stephenj/ec968/pdfs/ec968lnotesv6.pdf>

Keswell, M. and Poswell, L. (2004), "Returns to Education in South Africa: A Retrospective Sensitivity Analysis of the Available Evidence", *South African Journal of Economics,* 72:4, pp. 834-360.

Kingdon, G. (2001), "The Gender Gap in Educational Attainment in India: How Much Can Be Explained?", *Journal of Development Studies*, 39:2, pp 25-53. Available online: <http://www.economics.ox.ac.uk/Members/geeta.kingdon/ PublishedPapers/jds2002gendergap.pdf.>

Lam, D., Ardington, C. and Leibbrandt, M. (2007), "Schooling as a Lottery: Racial Differences in School Advancement in Urban South Africa". Presented at the Annual Meeting of the Population Association of America. New York, New York. Available online: < http://caps.psc.isr.umich.edu/papers/Lottery.pdf>

Markley, T. (2004), "Defining the Effective Teacher: Current Arguments in Education", *Essays in Education*, Fall, 2004, Volume 11. Available online: <http://www.usca.edu/essays/vol112004/markey.pdf>

Meekers, D. and Ahmed, G. (1999), "Pregnancy-Related School Dropouts in Botswana", *Population Studies*, 53:2, pp195-209. JSTOR Available online: <http://links.jstor.org/sici?sici=0032-4728%28199907%2953%3A2%3C195%3APSDIB%3E2.0.CO%3B2-1>

Moll, P. (1998), "Primary Schooling, Cognitive Skills and Wages in South Africa", *Economica,* 65, pp. 263-84.

Neal, D. and Johnson, W. (1996), "The Role of Premarket Factors in Black-White Wage Differences", *The Journal of Political Economy*, 104:5, pp. 869-895.

Peng, C., Lee, K. and Ingersoll, G. (2002), "An Introduction to Logistic Regression Analysis and Reporting", *The Journal of Educational Research.* Available online: < http:www.episnu.net/zb41/download.php?id=board&page=1&sn1=&divpage=1&sn=off&ss=on&sc=on&select_arrange=hit&desc=desc&no=57&filenum=2>

Rivkin, S., Hanushek, E. and Kain, J. (1998), "Teachers, Schools, and Academic Achievement", *National Bureau of Economic Research Working Paper No. 6691*, NBER. Available online:<http://www.econ.ucdavis.edu/faculty/mepage/econ250b/teachers%20schools%20and%20achievement.pdf>

Rosenthal, B., (1998), "Non-School Correlates of Dropout: An Integrative Review of the Literature", *Children and Youth Services Review*, 20: 5, pp. 413-l33.

Russell Bernard, H., Killworth, P., Kronenfeld, D. and Sailer, L. (1984), "The Problem of Informant Accuracy: The Validity of Retrospective Data", *Annual Review of Anthropology.* 13, pp. 495-517. Available online: <http://social.cs.uiuc.edu/class/cs598kgk/papers/informant_accuracy.pdf>

Scott-Jones, D. (1984), "Family Influences on Cognitive Development and School Achievement", *Review of Research in Education*, 11, pp. 259-304.

Scott Long, J. (1997), *Regression Models for Categorical and Limited Dependent Variables.* Thousand Oaks: Sage Publications.

Song, L., Appleton, S. And Knight, J. (2006), "Why Do Girls in Rural China Have Lower School Enrollment?", *World Development*, 34:9, pp. 1639–1653. Elsevier Ltd.

South, S., Haynie, D. and Bose, S. (2007), "Learner mobility and School Dropout", *Social Science Research*, 36, pp.68–94.

Spence, M. (1973), "Job Market Signaling", *The Quarterly Journal of Economics,* 87: 3, pp. 355-374. JSTOR Available online:< http://www.jstor.org/stable/1882010 >

Thomas, D. (1996), 'Education Across Generations in South Africa', *The American Economic Review*, 86:2, Papers & Proceedings of the Hundredth and Eighth Annual Meeting of the American Economic Association San Francisco, CA, January 5-7 (May, 1996), 330-334. JSTOR. Available online: <http://www.jstor.org/stable/2118147>

Udry, J.R. and Chantala, K. (2003), "Missing school dropouts in surveys does not bias risk estimates", *Social Science Research*, 32, pp294–311.

Van der Berg, S. (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, 27-28 October 2005, Stellenbosch.

Van der Berg, S. (2006), How effective are poor schools? Poverty and educational outcomes in South Africa." Stellenbosch Economic Working Papers: 06/06, The Department of Economics and The Bureau for Economic Research at the University of Stellenbosch.