Does Individual Ability Play a Role in Educational Attainment over and above Household, School and other Socio-economic Circumstances?

Amy Kahn (KHNAMY001)

Masters Mini Dissertation

May 2007

School of Economics

University of Cape Town

Supervisor: Professor Murray Leibbrandt

ABSTRACT

This study looks at the impact of latent individual ability on educational attainment, specifically the decision to acquire tertiary education. The analysis aims to determine whether ability is significant over and above socio-economic status in determining educational attainment. The Cape Area Panel Study (CAPS) (2002-2005) and the School Register of Needs Survey (2000) provide data at the individual, household and school level, which are used to model the decision to apply for tertiary education for a sample of respondents who have completed matric. The CAPS data-set provides a set of literacy and numeracy test scores, which are regarded as measures of manifest ability. After purging out the effects of age, education level, as well as household and school-level characteristics from these test scores, a latent measure of ability is obtained, which is relatively independent of socio-economic status. Probit regressions are then used to determine the impact of this ability measure on the decision to apply for tertiary education. Population groups are analysed separately due to their wide disparities in household and school resources in the South African context. Furthermore, a household fixed effect analysis is used to control for any other unobservable household characteristics that may affect educational decision-making. The results indicate that overall, individual ability is significant in explaining the decision to pursue tertiary education over and above contextual factors. However, for the Black sample, ability is not significant in explaining this decision, and is significantly less important in explaining the attainment of a matric exemption compared to the Coloured and White groups. This suggests that a lack of resources and unfavourable learning environments, which are characteristic of the Black population, lead to a 'crowding out' of individual ability.

ACKNOWLEDGEMENTS

My special thanks go to Professor Murray Leibbrandt and Cally Ardington for their supportive supervision and valuable input.

I acknowledge funding from the National Institute of Child Health and Development and the National Institute of Aging (Grant R01 HD045581-01).

PLAGIARISM DECLARATION

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

2. This paper entitled: *Does Individual Ability Play a Role in Educational Attainment over and above Household, School and other Socio-economic Circumstances?* is my own work.

3. 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.

4. 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:

TABLE OF CONTENTS

Appendices

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

The South African population is characterised by a high level of inequality. This is reflected in disparities in unemployment and wage rate levels between different groups of the population. The prevailing view is that to a large extent these disparities stem from different levels of educational attainment as well as educational quality. More specifically, different groups of the population are born into vastly different circumstances and are faced with different opportunities, which in turn translate into different educational and labour market outcomes. Thus, inequities, or unequal opportunities, often result in enduring inequalities.

This study aims to determine whether individual ability plays a significant role in educational attainment over and above socio-economic background. More importantly, in the case of the most disadvantaged segments of the population, it evaluates whether or not ability is able to transcend these inequities and result in different outcomes from what initial circumstances may predict.

The analysis looks at the impact of ability as well as various other individual, household and school level characteristics on the decision to pursue tertiary education. The Cape Area Panel Study (2002-2005) provides a rich set of household and demographic data for a sample of young adults in the Cape Town metropolitan area, where population group is a strong indicator of home background and school quality factors. This data is combined with

the School Register of Needs survey (2000) which provides information on the respondents' schools.

Ability as a personal trait is difficult to identify; for example, it is often masked by disadvantaged circumstances, or can be enhanced by an environment that is highly conducive to learning. The analysis deals with this issue explicitly. The CAPS data-set provides literacy and numeracy test scores which are assumed to reflect socio-economic status, including household circumstance and school quality, to a large extent. After purging these factors (along with age and education level) from the test scores, a more latent ability measure is obtained. This measure is used to analyse the effect of individual ability on the decision to apply for tertiary education, controlling for household and school level characteristics. A fixed-effects model which is able to control for unobservable household characteristics is also used in order to confirm that the effect of ability is indeed independent from socio-economic status. In addition, population groups are analysed separately; thus, the opportunity to advance to higher levels of education on the basis of individual ability can be assessed within each segment of the population.

LITERATURE REVIEW

The aim of this study is determine the significance of the impact of individual ability on educational attainment, specifically the decision to acquire tertiary education. This particular decision variable has been chosen as the dependent variable in the analysis because in South Africa, returns to tertiary education have been shown to be much higher than returns to having only a matric (Keswell & Poswell, 2004 and Moll, 1998).[1] Thus, it is important to determine to what extent these returns are a reward for ability or effort as opposed to simply a 'reward' for having an advantaged background. In addition, Keswell (2004) finds that the rate of return to education for Whites has become increasingly larger than the rate of return for Blacks since the end of Apartheid. He suggests that this may lead to a reduction in incentives to acquire more education for Blacks. Thus, it is possible that the role of ability in the decision to pursue tertiary education may differ between population groups.

The following review takes a look at the literature surrounding the relationship between socio-economic status (including home background and school quality), educational outcomes and ability. It also reviews studies which analyse how socio-economic status translates into labour market outcomes either directly or via its effects on educational outcomes and ability. However, the first step is to evaluate the various definitions or conceptualisations of 'ability' as this has implications for how ability is seen to interact

[1] Moll (1998) argues that this reflects the low supply of tertiary-level skills in the South African economy as well as the high demand for these skills.

with other socio-economic indicators and therefore is highly relevant for the analyses which follow later on.

What is cognitive ability?

This study is concerned with latent or individual ability. The most important assumption in this respect is that achievement test scores are largely affected by home background and school quality factors and as a result they are measures of manifest ability rather than latent ability. Thus, in order to create a measure of latent ability, these socio-economic factors need to be purged out of the test scores. This measure can then be used to analyse the direct effect of ability on educational attainment, independent of school and home background factors. Therefore, in this study, ability is seen to reflect individual capability, which remains fairly constant with educational attainment and age and is also unaffected by socio-economic circumstance. In this way, individuals who have been exposed to the same life opportunities can be compared and the effect of individual ability on the decision to pursue tertiary education can be evaluated more precisely.

This conceptualisation of ability is in contrast to the view put forward by Herrnstein and Murray's book, *The Bell Curve* (1994). They assume that the correlation between IQ and achievement test scores is very high, that intelligence is highly inherited and that IQ level is more important then socio-economic status in predicting various outcomes such as college attainment and employment status (Goldberger & Manski; 1995:762-774).

The Bell Curve has been extensively critiqued since its publication. For example, Culbertson (2001:3-9) criticises Herrnstein and Murray for attributing the race-group differences (in school and labour-market outcomes) to differences in intelligence rather than socio-economic status with which it is correlated. Goldberger and Manski (2001:765-766) criticise the authors for basing their argument on a model which assumes the correlation between genes and home environment to be zero. Parents transmit both genes and environment to their children and thus Herrnstein and Murray's concept of heritability is flawed. They also ignore the possibility that education acts as a channel that connects child home background to future adult outcomes (and they omit education from their analysis).

Hansen et al (2004: 40) find that the measures of IQ, as used by Herrnstein and Murray, are strongly affected by schooling and therefore cannot be an accurate measure of latent cognitive ability. Thus, they stress the need to deal with reverse causality between schooling and test scores when analysing the effects of cognitive ability. They postulate that latent ability affects both manifest ability and schooling but that schooling can only affect manifest ability (but not latent ability).

Neal and Johnson (1996:890) argue that the Armed Force Qualification Test (AFQT) scores used in Herrnstein and Murray's analysis reflect differences in "acquired skill" and not innate ability. In their study of identical twins, Ashenfelter and Rouse (1998:281) show that genetically equivalent individuals are able to reap positive returns to schooling and that returns to schooling may be slightly lower for higher ability individuals suggesting that

"schooling compensates for genetic differences". Their results stand in contrast to the claim put forward by Herrnstein and Murray that genetics predetermine education.

The alternative approach to conceptualising intelligence implicitly separates out IQ (latent ability) from observed/measured ability, and assumes that schooling is able to raise achievement test scores to differing degrees. This view on cognitive ability is generally the most widely accepted and is also the central assumption in this study.

While the concept of ability may be relatively consistent throughout the studies that will be reviewed here, various measures of ability have been used by different authors in their analyses. Nevertheless, the interaction between these measures and socio-economic indicators are always dealt with (explicitly or implicitly) and therefore the idea that the way in which ability is revealed is shaped by circumstance and/or life opportunities is always maintained.

Non-cognitive abilities

A number of authors argue that socialisation skills, behavioural characteristics and personality traits such as persistence, motivation, charm, docility and dependability, are important determinants of school and labour market outcomes (Heckman, Stixrud and Urzua (2006); Heckman and Rubinstein (2001); Cawley, Heckman & Vytlacil (2001); Bowles, Gintis and Osborne (2000)). These abilities have been referred to as "non-cognitive abilities" (Heckman, Stixrud and Urzua; 2006: 2).

These authors argue that non-cognitive abilities may affect labour market outcomes directly, through productivity, or through educational attainment and that these types of personality traits are rewarded separately from the usual cognitive ability, in school and in the work-place. Furthermore, cognitive and non-cognitive abilities are not necessarily highly correlated. Consequently, achievement test scores may not only reveal ability in the conventional usage of the word, but may also reflect a wide variety of other, separate individual traits. Heckman & Rubinstein (2001: 148) argue that while IQ is fairly set by age eight, these other skills can still be altered at later ages.

Home Background Impact on Cognitive Ability

There is a vast body of literature that analyses the impact of family and home background characteristics on cognitive ability as well as schooling. This literature implicitly separates out parental characteristics, home environment and context from genetic factors and thus lends support to the idea that cognitive ability is not for the most part inherited or predetermined (and thus unchangeable). Scott-Jones (1984), speaking before the publication of *The Bell Curve*, argued that it is generally agreed that a child's genetic endowment as well as their environment are responsible for influencing their cognitive capacity; however the relative importance of the two is a contentious issue because it is difficult to separate out the influences. Firstly, there are bidirectional influences between parent and child in that children may influence parental behaviour towards them through their own behaviour. Thus, children within one family may have different experiences in

the home or they may respond differently to the same experience. As a result, siblings will have different levels of intellectual development and the effects of environment (versus genetic factors) may also be underestimated (261-263). Secondly, environment and genetics will be correlated and thus difficult to separate out. It may be difficult to obtain a measure of genetic IQ; for example a mother's IQ may be influenced by home environment or broader social context (263-264).

The main factors in home background influencing cognitive and educational outcomes that are referred to in the literature, include family income (related to poverty and household resources), parental education, household structure, as well as broader neighbourhood/socio-economic context. Scott-Jones (1984:265-266) and Brooks-Gunn and Duncan (1997:60) highlight the impact of poverty on physical health. Malnutrition has adverse affects on child attention and motivation and thus may result in learning disabilities, grade repetition, limited cognitive development and lower achievement scores. However, poverty is associated with other household characteristics such as lower parental education, fewer educational resources and poorer neighbourhoods, which in turn makes it difficult to separate out the various effects. Brooks-Gunn, Klebanov and Duncan (1996:397) refer to these as "poverty co-factors" and argue that minority ethnic groups in the U.S. are likely to live in poor neighbourhoods, in single-parent households, and with parents that have lower education and literacy scores and who are more likely to be younger and unemployed.

Al-Samarrai and Peasgood (1998) and Anderson and Lam (2003) find that in Tanzania and South Africa respectively, household composition (living with mother versus father) as well as parental education have strong impacts on school-enrolment decisions and grade repetition. Similarly, Cameron and Heckman (2001:492) conclude that long-term family background (measured by parental education) is more important in explaining college attendance than short-term credit constraints. Long-term family income matters in terms of influencing learner ability (and therefore college readiness) as well as earlier grade transitions.

Cognitive ability and education as mediators of family background

Much of the literature argues that cognitive ability as well as schooling can act as channels through which home background and socio-economic status (SES) can affect labour market outcomes later on in life. For example, Kerckhoff, Raudenbush and Glennie (2001:1) argue that restricted access to education means that some people with the required skill are not exposed to the same opportunities as others. Thus, the relationship between education and earnings may reflect the degree of opportunity or power available to an individual rather than a straightforward relationship between skill and the labour market.

In their study using the US National Adult Literacy Survey (NALS), Kerckhoff et al (2001) argue that both cognitive skill and education are affected by social background and that they act as "mediators of the effects of background, ethnicity and language" (8) on labour

market outcomes.[2] They find that together, cognitive ability and educational attainment account fully for the effect of parental education on occupational status and earnings (although neither does so alone) and these effects differ between ethnic groupings (Blacks, White or Hispanic) and genders (10, 18). The authors suggest that education and cognitive skill may reflect different aspects of family background: education may represent power and economic resources in the home, while cognitive skill may be more representative of 'cultural capital' and socialisation aspects (10). Using twin data, Ashenfelter and Rouse (1998) are able to look at the total variability in educational attainment due to family background factors. Their results indicate that, "in theory, family background explains about 60 percent of the variance in schooling attainment" (275).

Neal and Johnson (1996) argue that using schooling as a measure of skill in a wage equation is highly misrepresentative as it has been shown that black children in the US have lower levels of achievement compared to other children in the same grade (870-871). Thus, the US wage-gap is primarily a skill gap (and not due to discrimination as is often argued) and this is mostly a result of family background and school-level disadvantages. Using the AFQT as an indicator of "achievement and learned skill, not of innate ability" (887), they find the score is highly significant in the wage equation and explains nearly ¾ of the racial wage gap for men and the entire gap for women (874). The results of the paper indicate that the payoff from skills between blacks and whites is equal thus differences in skill acquisition must be due to differences in costs of and therefore investment in skills. The authors substantiate this claim by showing that the AFQT scores are significantly and

[2] Cognitive skill is measured by the literacy scores, therefore can be seen as manifest cognitive ability rather than latent ability.

positively related to parental education, income and professional status as well as reading materials in the home, which reflect the availability of economic and academic resources for skills development (887-890). These results demonstrate how skill-level can act as a mediator of family background. In addition, they find that school environment measures such as student/pupil ratio, disadvantaged student/ratio, student drop-out rate and teacher turnover rate, are significant in explaining the scores. It could be argued that household SES will be highly correlated with school quality in that children from more advantaged households will attend better schools. In turn, the effects of school quality and resources on student outcomes have been shown to be highly significant.

School effects

A number of studies (for example, Simkins (2001), Ladd and Fiske (2004), Kingdon (1999), Case and Deaton (1999) and Van der Berg (2002)) have found school resources such as the pupil-teacher ratio, the presence of libraries, science laboratories, and textbook availability, as well as electricity availability, telecommunications availability, water supply and toilets, to have significant effects on educational achievement. However; more qualitative aspects of schools such as the efficiency of school management as well as the quality of teachers, which includes teacher cognitive ability and teacher qualifications, have been found to be just as, if not more important than the quantity or presence of physical resources in explaining school outcomes (see Kingdon (1996), Van der Berg (2002), Van der Berg and Burger (2003), and Crouch and Mabogoane (2001)). Liddell and Rae (2001:425) argue that in the developing world context, the in-school experience of learners

is more important than background or personal qualities in explaining achievement. They find that early academic achievement and curriculum mastery at the grade 1 level are the strongest predictors of later grade retention, and are more important than broader cognitive skills for a smooth transition through school.

In addition, the effect of school resources on educational outcomes has been found to differ between population groups in South Africa. For example, using data collected just before the end of Apartheid, Case and Deaton (1999; 1072-1080) find that pupil-teacher ratio has strong and significant negative effects on years of completed schooling, probability of enrolment, and literacy and numeracy test scores for Blacks but that the same effects are not found for Whites. They conclude that "the education of Blacks but not Whites is constrained by financial resources" (1080). These findings relating to school effectiveness will prove to be important for this study.

Linkages to labour-market theories

This study relates directly to the problems of ability bias, simultaneity bias, and measurement error discussed in much of the wage equation, or returns-to-schooling, literature. The problem of ability bias is an endogeneity problem. When labour market outcomes (such as employment status or earnings) are regressed on schooling and other determinants and one fails to control for ability, the coefficient on the schooling variable may be biased upwards. If more able students progress through school more easily and consequently obtain higher levels of school achievement, and at the same time their

relatively high productivity rewards them for example with higher wages, then the schooling coefficient will be correlated with the error term and the wage equation will suffer from endogeneity bias. In order to deal with, or measure the extent of this problem, test scores (such as literacy and numeracy or IQ test scores) are often used in such equations. However, as discussed above, these test scores may be an imperfect proxy for latent cognitive ability; for example, they may be affected by family background to a large extent. This would result in measurement error. In addition, schooling and test scores are simultaneously related thus there would continue to be an endogeneity problem in the model (Heckman, Stixrud & Urzua; 2006:26-27).

The study also relates to Signalling theory, which involves the idea that employers have limited information about employee applicants and therefore they use a worker's educational achievement as a signal or screening device for ability or productivity as well as other personal qualities such as motivation (Ehrenburg & Smith; 2003:291-295). If this is so, then it is important to determine whether educational achievement is in fact largely determined by ability, or whether it is more a reflection of an individual's socio-economic background and life opportunities.

There is a large body of literature which finds evidence of ability bias and simultaneity bias in labour market outcomes equations. Using urban wage-labour data from Kenya and Tanzania, Boissiere, Knight and Sabot (1985:1017-1018) conclude that "The direct returns to reasoning ability in the labour market are small, those to years of education are

moderate, and those to literacy and numeracy – dimensions of human capital – are large"[3] (1028). In both countries however, cognitive ability has a strong and positive impact on both educational level and reasoning ability. This reasoning ability in turn has two indirect effects on earnings: through its impact on the amount of human capital acquired in school as well as through years of education completed (secondary school attendance) (1027). Ability and education are therefore correlated and at the same time there are large direct and indirect returns to ability, thereby confirming the presence of ability bias in conventional returns to schooling equations.

The authors go on to argue that the high returns to literacy and numeracy may be because these cognitive skills are particularly scarce in the countries under consideration compared to other developing countries. This issue is elaborated on by Bowles, Gintis and Osborne (2000:24-26) who argue that it is difficult to estimate the impact of cognitive ability on production in general because the marginal effect in the earnings equation may be capturing the scarcity of the particular job skill. At the same time, cognitive skills will vary in importance for different types of occupations, thus the impact on earnings will vary.[4]

Charette and Meng (1998) find that literacy and numeracy affect employment and labour market status even when controlling for years of schooling. However, the impact differs between males and females: the inclusion of literacy and numeracy leads to a fall in the co-efficient on schooling for males and thus when excluded, results in an upward bias on the

[3] They use literacy and numeracy test scores as measures of cognitive skill.
[4] For example, general ability may be less important compared to how much the employee knows about their job (experience) for certain occupations.

male schooling coefficient. The opposite occurs for females however. The inclusion of numeracy and literacy results in a rise in the schooling coefficient, leading the authors to conclude that schooling is a poor measure of female human capital (thus resulting in attenuation bias) and that the role of literacy and numeracy skills as indicators of school quality will dominate. (515)

In a study of identical twins, Ashenfelter and Rouse (1998) hypothesise that schooling investments of genetically equivalent individuals should be equal. Their model shows that within-family differences in schooling vary only due to differences in the marginal benefit of schooling between family-members, which are in turn correlated with differences in ability. Their results show that cross-section estimates of returns-to-schooling are slightly biased upwards due to the fact that unobserved ability affects desired schooling level as well as returns to schooling. Abler individuals obtain more schooling due to lower marginal costs of schooling. They conclude that individuals with higher ability will obtain higher levels of education and thus cross-sectional estimates of returns to schooling are biased upwards due to omitted ability-variable bias.

Kerckhoff et al (2001) show that both education and cognitive skill (indicated by adult literacy scores) have significant and independent affects on labour market outcomes. However, "The relative contributions of educational attainment and cognitive skill to the explanation of occupational status and earnings vary as a function of both ethnicity and gender" (16). Together, education and cognitive skill explain the complete effect of parental education on occupational status but neither does so alone. In general however,

education explains more than cognitive skill especially in terms of occupational status (7-8; 18).

The issue of reverse causality between test scores and schooling as well as the endogeneity of schooling decisions is emphasised by Hansen et al (2004). They argue that it is important to control for simultaneity between measured ability and school choice because schooling affects measured ability while ability affects school choice. If this relationship is ignored, the role of cognitive ability in schooling decisions (for example, the decision to attend college) will be overstated. Firstly, they find that when simultaneity is controlled for, the role of cognitive ability in schooling decisions is lessened. Hence, they argue that previous literature has most likely overstated the importance of latent cognitive ability in explaining schooling decisions. Secondly, they find that latent ability affects test scores: as schooling increases, the marginal effects of ability on verbal scores decreases, while the marginal effect of ability on mathematical scores remains fairly constant or increases slightly. And thirdly, schooling has strong effects on test scores and these effects are larger for people with lower levels of ability. The effects are also larger in the earlier high school years for all ability levels (74-77).

Thus, when attempting to deal with the ability bias problem, Hansen at al (2004) stress the need to deal with the fact that test scores may be an imperfect proxy for ability and therefore may themselves be endogenous to the model. In addition, even if the test scores are a perfect proxy, because schooling affects test scores the effect of schooling on wages will remain biased (80). The authors deal with this problem by constructing a measure of

latent ability correcting for the endogenous schooling effect at the test date, family background and age. Thus, they purge out the effect of schooling and background on manifest ability. When using this measure of ability in a wage regression they find that the estimated schooling effect increases, supporting their suspicion that the simultaneous relationship between schooling and test scores may cause the bias in the schooling coefficient to remain (81).

In a later analysis, Heckman, Stixrud and Urzua (2006) solve the endogeneity problem (that schooling affects test scores) by simulating their model. They predict test scores that each individual would have received had they been in grades 9-11 at the time of the tests (given exogenous conditioning variables) by drawing the factors from the population distribution. They find that not controlling for this reverse causality in wage equations leads to an overestimation of the returns to ability. However, measurement error, due to the test score being an imperfect proxy for ability, "causes a significant downward bias, that is typically larger than the upward bias due to endogeneity and reverse causality" (Heckman, Stixrud and Urzua; 2006:27).

Concluding remarks

This study draws on the literature in a number of ways. Literacy and numeracy test scores are regarded as a reflection of what Hansen et al (2004) term 'manifest' ability, and the effects of school quality and home background factors on the scores are dealt with explicitly. The analysis uses the residual variance in these scores, after controlling for

school and household effects, as a measure of ability. Whether or not this measure includes 'non-cognitive' abilities is not dealt with here. Nor is this issue relevant; what is important is that the scores are specific to the individual. This method of using the residuals is taken from Hansen et al (2004); it is an attempt to deal with the simultaneous relationship between test scores and school quality characteristics, as well as the measurement error caused by the impact of family background factors on the scores. Using a household fixed effect, the analysis attempts to determine whether individuals in the same home may have differing levels of ability that lead to different educational outcomes. The idea that individuals who live in the same home, whether related or not, may have differing ability levels has been discussed by Scott-Jones (1984). The fixed-effect will control for school quality effects which have been argued to be highly correlated with household socio-economic status, as well as deal with other unobservable household features that may be important in determining educational decisions and outcomes.

The outcome of the analysis will not only evaluate the importance of individual ability in explaining educational attainment, but the results will also have implications for the 'ability bias' theories and 'education as a signal' hypothesis. Furthermore, the importance of these theories, or simply the importance of individual ability in general, will be determined within a developing country context, which is characterised by high levels of inequality.

METHODOLOGY

The quantitative analysis in this study involves explaining the decision of whether to pursue tertiary education after completion of matric[5]. More specifically, the analysis aims to determine the significance of individual ability in this decision, over and above socio-economic background. Although ability is likely to have a significant and positive impact on the probability of completing matric, this analysis includes as its sample only those respondents who have already successfully completed their final year of schooling. Thus, more precisely, the analysis aims to determine the significance of ability on the decision to acquire tertiary education *given* the attainment of a matric.

As reviewed in the literature, family background or 'life circumstances' largely determine educational attainment and opportunities, which in turn has a significant impact on labour market outcomes. The role of ability in determining educational attainment and labour market success has also been analysed, but the different methods of doing so have varied significantly, largely because the definition of ability varies widely. This 'definitional' issue is highly contentious; there are many socio-economic factors that work in similar patterns and which impact on the same outcomes as ability, thereby making it difficult to separate out the different effects. Thus, an important challenge in this analysis involves finding a measure of ability which measures precisely that. In other words the ability measure employed should be unrelated to socio-economic indicators as much as possible.

[5] 'Matric' refers to grade 12, or the final level of schooling in South Africa.

The data used in this analysis provides a literacy and numeracy test score. The literature surrounding the impact of cognitive ability on various schooling decisions and labour market outcomes highlights a number of shortcomings regarding the use of test scores as a measure of ability. Two important methodological issues discussed by Heckman, Stixtrud & Urzua (2006:26-27) as well as Hansen, Heckman & Mullen (2004:74-77) for example, include measurement error and simultaneity bias. Firstly, test scores may be an imperfect proxy for ability as they may be significantly influenced by family and household factors. Secondly, test scores may be a measure of what Hansen at al (2004: 40) refer to as 'manifest ability' rather than latent ability, as the scores will be highly influenced by educational attainment as well as quality of education. At the same time, latent ability will influence educational attainment thus resulting in a simultaneous relationship between test scores and schooling.

Taking these issues into account, the first step in the analysis is to create a measure of ability that is purged of the effects of both household-level and school (qualitative and quantitative) factors. This is done by regressing the standardised literacy and numeracy test scores on a range of individual, household and school variables. The individual-level variables include the age and education level at the date of the test, gender, population group, and whether or not the respondent wrote the test in their home language[6]. The household-level variables can be characterised as environmental-type variables, including parental education, whether parents helped with homework, proportion of life lived with parents, and household resource variables including per capita household income and the

[6] The test could be written in either English or Afrikaans. Thus, the majority of black respondents, who are primarily African language speakers, did not have the option of writing the test in their home language.

presence of a computer or more than five books in the home. The school-level variables include the pupil-teacher ratio and the former education department to which the school belonged under Apartheid. Note that this part of the analysis uses the full sample of respondents i.e. not just those who have completed matric.

Following the estimation, the residuals of the regression are saved and used as the measure of 'ability' for the rest of the analysis. These residuals can be thought of as representing that part of individual ability that is not explained by age, education level, household circumstance or school environment. Theoretically this implies that even individuals living in the same household and who attend (or attended) the same type of school may vary in their level of ability. In turn, they may advance to different levels of education, in this case tertiary education, and as a result have different labour market experiences in the future.

To test whether ability does in fact have an impact on the decision to pursue tertiary education, the next step in the analysis is to explain the decision to apply to higher education based on the ability measure that was created as well as other individual, household and school factors. Using the sample of respondents who have completed matric, this is done by running a probit regression with the binary dependent variable equal to 1 if the respondent has applied to tertiary education, and equal to 0 if they have not. The probit models used in the analysis allow one to identify the marginal change in the probability of y occurring due to an infinitesimal change in x if x is a continuous variable, and a discrete change in the probability of y occurring if x is a categorical variable.

The explanatory variables in the 'applied' model include similar individual, household and school variables as discussed above; however population group and home language are combined into dummy variables, and an exemption dummy as well as a control for year of matriculation are also included.

The same regression is run for each population group separately (of course, the population-group/language dummies are excluded here). The motivation behind this is to determine which variables are most important for the decision within each group, as well as to compare the degree of importance of certain variables, in particular ability, between the different groups. One may expect to find that ability is less important as an explanatory variable in poorer environments where economic constraints may dominate. For example, an individual who comes from a household with a very low per capita household income may believe that there is very little chance of them being able to afford further education, and furthermore, there may be more pressure on them to enter the labour market. As a result, they may not even consider applying to tertiary education as an option; thus, their ability level will be irrelevant in the decision. At the other extreme, very wealthy households may experience fewer financial constraints thereby making other factors such as ability more significant in the decision-making process.[7]

Alternatively, it is possible that in poorer households, the number of young adults who attend tertiary education may be smaller than in richer households due to financial constraints. In this case, the decision of who gets tertiary education, or who is allocated the

[7] The population groups are highly correlated with household wealth and thus are used as a means to divide the sample into groups of varying socio-economic status.

scarce household resources, becomes more relevant and may be based on the relative ability levels between household members. As a result, the role of ability in the decision to pursue tertiary education will in fact be more significant in poor households. Ultimately, the question of which scenario will dominate is an empirical one and will be revealed through the results of the analyses.

An important explanatory variable in the decision-to-apply model is the 'exemption' dummy which has a value of 1 if the respondent passed matric with an exemption and a value of 0 if they passed without an exemption. This variable is expected to have a large impact on the decision to apply. Obtaining an exemption requires an explicit decision as certain marks are required and certain subjects need to be written in the matric exams; thus, the decision to apply to university education may have already been made by the individual. More importantly for this study however, is the indirect effect that ability may have on the decision to apply, via the attainment of a matric exemption. In other words, ability may be a significant explanatory factor in the decision to acquire (and the attainment of) a matric exemption, which in turn will have a strong impact on the decision to pursue tertiary education.

In South Africa, a matric exemption (or endorsement) is obtained if a certain number of subjects are chosen from particular subject-categories, and are done on the higher grade as opposed to standard or lower grade. Details of these requirements can be found in Appendix A. Passing matric with an exemption is necessary for university entrance; however matriculants who do not obtain an exemption are still able to study at other tertiary

institutions such as technikons or other colleges. Out of the sample of respondents used in the analyses which follow, a very small percentage of respondents who obtain an exemption do not apply for tertiary education. Furthermore, out of those who do apply for tertiary education with an exemption, just over 60 percent enrol in a university while approximately 37 percent enrol in a technikon or study for a diploma or certificate from some other post-matric institution. Thus, obtaining a matric exemption can be regarded as an important gate-way to tertiary education in general, and is also a strong indicator of university study.

The possibility of an indirect effect is analysed through probit regressions with the exemption variable on the left-hand-side. The explanatory variables include the same individual, household and school factors that are used in the 'applied' model. Once again regressions are run for population groups separately. The purpose of doing so involves a similar logic as discussed above. In this case, individuals who come from a poor background (most of the Black sample) will attend low-quality schools, which in South Africa are characterised by poor management and teaching, lack of resources and, as a result, high rates of grade repetition and poor school outcomes. In such an environment individual ability may be 'crowded out' by these unfavourable conditions, and as a result its impact on school outcomes will by limited. The opposite may occur for those who come from wealthy backgrounds and attend good-quality schools.[8]

[8] The discussion here focuses on school environment however home environment, such as limited household resources (including both parental academic support and physical resources), will have similar effects.

Fixed-effects analysis

In order to refine the analysis and obtain a more conclusive result, the 'applied' and 'exemption' models are rerun using a household fixed effect. Essentially, this allows one to identify whether individuals living within the same household may have differing educational outcomes due to differing levels of ability. Most importantly, using the fixed effect model allows one to control for all unobservable household characteristics that were previously not dealt with in the regression analysis. Thus, it enables one to identify whether individual ability is significant in explaining the decision to apply for tertiary education while being confident that the effect is in fact purely individual, i.e. that it is not linked to any socio-economic factors. The assumption here is that individuals who live in the same household have grown up under similar circumstances, have been exposed to similar opportunities, and have attended schools of similar quality.

The fixed effects model can be represented as follows:

1. $y_{i1} = x_{i1}\beta + f_i + u_{i1}$ Sibling 1

2. $y_{i2} = x_{i2}\beta + f_i + u_{i2}$ Sibling 2

Equations 1 and 2 are for each household member (for example, two siblings) within each household i, where y is the outcome variable (in this case, one or zero), x is a vector of explanatory variables, f is an unobserved family effect, and u is the error term, which is

assumed to be uncorrelated with the explanatory variables in *both* equations (known as the strict exogeneity assumption). In this particular analysis, the unobserved family effect is assumed to include household, school quality as well as neighbourhood observable and unobservable characteristics (Wooldridge; 2002:265-329).

In order to remove *f*, the model differences across the individuals (1 – 2):

3. $y_{i1} - y_{i2} = (x_{i1} - x_{i2})\beta + f_i - f_i + u_{i1} - u_{i2}$, to get

4. $\Delta y_{is} = \Delta x_{is}\beta + \Delta u_{is}$,

where s indexes siblings or members of a household. This first-differencing transformation eliminates the unobserved household effect, *f*.

It is important to note that "**x** cannot contain common observable family background variables, as these are indistinguishable from f_i" (Wooldridge; 2002:329). Thus, the fixed effects regression for the 'applied' model includes as its explanatory variables only those which vary between members within the same household: ability, age, gender, attainment of an exemption, and year of matriculation. The 'exemption' model includes the same ability variable and age, gender and year of matriculation dummies.[9]

[9] Only variables that do not vary between individuals in the same household for *any* households are not allowed. In other words, as long as the vector **x** does not contain a column of zeros for all *i* households.

One draw-back of using the fixed-effect model is that it requires one to include only households which include more than one respondent. The CAPS data includes information for up to three young adults within each household; the majority of households however include only one young adult. Thus, when the single young-adult households have been dropped from the sample, the sample size decreases considerably. In this part of the analysis, a linear probability model is used instead of a probit model. Katz (2001: 379-384) argues that unconditional fixed-effects logits (or probits) will produce biased maximum-likelihood estimators when T is small. He explains that, "The inconsistency arises because the number of incidental parameters increases without bound, while the amount of information about each incidental parameter remains fixed...The unconditional estimator of the structural parameters is also, in general, inconsistent..." (Katz; 2001: 380) In this study, T is the number of respondents within a household (rather than 'time'), which is small (T=2 or 3). Furthermore, the number of incidental parameters (which are the household specific parameters) is relatively large; thus, if a logit or probit were to be used in the fixed-effect analysis, the estimates would be biased and inconsistent. Katz (2001: 384) concludes that "when T is close to 2, the bias is substantial" and the estimator reaches approximately 2β. In order to avoid this problem, a linear probability model is used in the fixed-effect analysis; however the beta coefficients are interpreted in exactly the same way as before (i.e. they measure the marginal or discrete change in the probability of y occurring due to a change in x). [10]

[10] Katz (2001: 380) argues that using a *conditional* likelihood function would produce consistent estimates; however he mentions four reasons why the conditional estimator would not be practical for the applied researcher. For example, it is more computationally complex and does not deal well with missing data. Therefore, the linear probability model is chosen as the preferred option.

DESCRIPTION OF DATA

The data used in the analysis comes from the Cape Area Panel Study (CAPS) (2002-2005) survey and the School Register of Needs Survey (2000). The CAPS survey is a longitudinal study of youth in the Cape Town metropolitan area and, over a period of seven years, aims to track the respondents' lives as they progress from school or other educational institutions and into the labour market. The data from the survey include demographic information as well as information regarding the respondent's household (such as living arrangements, information about other household members, and household resources), education, work, sexual lives, and general attitudes and behaviour. The first interview required the respondents to complete a literacy and numeracy evaluation; these test scores have been used in this analysis as a measure of manifest ability.

The School Register of Needs Survey is a country-wide survey of schools in South Africa and aims to collect data on school management and resources. The analysis which follows uses two variables from this survey – pupil-teacher ratio and former education department of the school under Apartheid. The departments include: Department of Education and Training (DET – former Black schools); House of Assembly (HOA – former White schools); House of Delegates and House of Representatives (HOD-HOR – former Indian and Coloured schools respectively); former-Homeland schools; and 'New Education Department' schools (established between 1994 and 2000). School name information in the CAPS data allows one to link these school-level variables to each respondent. Thus,

together, the CAPS and School Register of Needs surveys provide data at the individual, household and school level.

This analysis uses data from the first wave (2002) and third wave (2005) of the CAPS survey. The respondents were aged 14-22 years in 2002. The creation of the latent ability measure uses the full sample in the first wave along with the wave 1 individual and household data. The rest of the analysis however uses a combination of the wave 1, 2 and 3 data. Wave 1 household and demographic data is used for home-background and individual-level variables respectively, while wave 1, 2 and 3 data are used to determine the respondent's year of matriculation and whether or not they passed with an exemption. Only respondents who had completed matric by wave 3 (2005) are included in the analysis.

Attrition

The wave 1 sample included 4697 respondents; however in wave 3 only 3492 of these respondents were re-interviewed:

Table 1: Attrition in CAPS data set

	Freq.	Percent	%Attrition
wave 1 (2002)	4,697	100	0
wave 2 (2003/2004)	4,106	87.42	12.58
wave 3 (2005)	3,492	74.35	25.65

Reasons for incomplete interviews can be found in Appendix B. Almost a quarter of the attrition is due to respondents moving to other parts of the country. Out of all the provinces however, including the Western Cape, the Eastern Cape is the most common destination.

Those who move to the Eastern Cape are primarily Black, African-language speaking, have less than grade 12 education (in 2002), and have attended either ex-DET or ex-homeland schools. Two thirds of this group are female.

Table 2 below compares the mean characteristics of the wave 1 sample versus the wave 3 sample. It is clear that by wave 3 the proportion of Black and White respondents has decreased while the proportion of Coloured respondents in the sample has increased relative to the wave 1 sample. The movers are more likely to be African language speaking and attend ex-homeland schools, compared to those who remained in the sample. This is not surprising considering the high proportion of respondents that move to the Eastern Cape, suggesting that they are originally from that area and are likely to still have family living there. [11]

Table 2: Attrition in CAPS data set - variation in characteristics (means)

	Wave 1	**Wave 3**
Black	0.45	0.43
Coloured	0.42	0.47
White	0.12	0.10
Male	0.45	0.46
Education level in wave 1 (2002)	9.26	9.16
English	0.21	0.21
Afrikaans	0.35	0.37
African Language	0.44	0.42
Age in 2005	21.08	20.87
DET	0.26	0.26
HOA	0.11	0.10
HOD_HOR	0.38	0.43
Homeland	0.17	0.14
New education department	0.04	0.04
Education department missing	0.16	0.14
Per capita household income	1096.68	1024.62
Sample size	4,697	3,492

[11] All means are significantly different from one another except for the English, DET and New Education Department dummies.

Because the changes in the sample make-up are fairly marginal, the results of the analyses are not likely to be distorted. Furthermore, because population groups are analysed separately, the change in population group make-up of the sample will pose even less of a problem. One possibility however is that by including fewer respondents who have attended ex-homeland schools this may lead to an understatement of the school-type/quality effect.

Descriptive statistics of sample used in the analysis

Table 3: Descriptive statistics: means with standard deviations using the CAPS and School Register of Needs data

Variable	Total	Black	Coloured	White
White English	0.11 (0.31)			
White Afrikaans	0.06 (0.24)			
Coloured English	0.21 (0.41)			
Coloured Afrikaans	0.27 (0.45)			
Black	0.34 (0.47)			
Applied	0.60 (0.49)	0.59 (0.49)	0.53 (0.50)	0.83 (0.38)
Ability	0.11 (0.88)	0.08 (1.03)	0.16 (0.82)	0.05 (0.68)
Age 2005	21.84 (1.99)	22.21 (2.00)	21.77 (1.96)	21.38 (1.86)
Male	0.42 (0.49)	0.41 (0.49)	0.41 (0.49)	0.43 (0.50)
Matric pre-2002	0.46 (0.50)	0.42 (0.49)	0.49 (0.50)	0.48 (0.50)
Matric 2002	0.17 (0.38)	0.16 (0.37)	0.18 (0.38)	0.18 (0.39)
Matric 2003	0.19 (0.39)	0.19 (0.39)	0.18 (0.39)	0.21 (0.41)
Matric 2004	0.15 (0.36)	0.17 (0.38)	0.12 (0.33)	0.17 (0.38)
Exemption	0.29 (0.46)	0.18 (0.39)	0.21 (0.41)	0.68 (0.47)
DET	0.23 (0.42)	0.67 (0.47)	0.02 (0.14)	0.00 (0.00)
HOA	0.21 (0.41)	0.03 (0.18)	0.13 (0.33)	0.95 (0.23)
HOD/HOR	0.51 (0.50)	0.15 (0.36)	0.85 (0.36)	0.01 (0.09)
New Edu. department	0.03 (0.17)	0.07 (0.25)	0.00 (0.04)	0.04 (0.19)
Homeland	0.14 (0.35)	0.38 (0.49)	0.02 (0.14)	0.04 (0.19)
Pupil-teacher ratio	29.62 (5.57)	32.62 (5.97)	29.56 (3.60)	23.11 (4.79)
Education of mother	9.70 (3.12)	8.84 (3.23)	9.14 (2.69)	12.66 (2.04)
Education of father	9.98 (3.62)	8.09 (4.06)	9.54 (2.91)	13.07 (2.15)
Prop. life lived mother	0.88 (0.24)	0.80 (0.27)	0.91 (0.22)	0.96 (0.16)
Prop. life lived father	0.64 (0.42)	0.43 (0.42)	0.72 (0.40)	0.85 (0.29)
Per capita HH income	1561.38 (2025.75)	504.89 (600.61)	1326.84 (1235.12)	4147.80 (3029.01)
>5 books in home	0.89 (0.31)	0.76 (0.43)	0.94 (0.24)	1.00 (0.00)
Computer in home	0.30 (0.46)	0.04 (0.19)	0.29 (0.45)	0.82 (0.38)
Exemption missing	0.06 (0.23)	0.09 (0.29)	0.05 (0.21)	0.01 (0.12)
Edu department missing	0.24 (0.43)	0.27 (0.45)	0.16 (0.37)	0.40 (0.49)
P-T ratio missing	0.24 (0.43)	0.27 (0.45)	0.16 (0.37)	0.39 (0.49)
Edu. mother missing	0.07 (0.26)	0.12 (0.32)	0.05 (0.22)	0.03 (0.18)
Edu. father missing	0.27 (0.44)	0.42 (0.49)	0.24 (0.43)	0.06 (0.24)
# Observations	1279	430	619	218

Source: School data is taken from the School Register of Needs Survey (2000); all other data is from the Cape Area Panel Study (2002 – 2005).

Table 3 above gives the mean values of the variables used in the analyses, for the sample as a whole as well as for each population group separately. Out of the wave 3 sample, almost 37% of the respondents are recorded as having completed matric. This includes 1281 respondents, two of whom have missing data regarding whether they have applied to tertiary education. Thus, the sample reduces to 1279 respondents.

In wave 3, all respondents were asked whether or not they had ever applied to a post-matric institution. This includes universities, technikons, or any other college that requires a matric. Out of the sample, just over 60% of the respondents have applied to some kind of tertiary education. However, not all of those who applied were necessarily admitted, and not all of those who were admitted decided to enrol. Nevertheless, the decision to apply to education rather than 'who attends tertiary education' is used as the dependent variable in the analysis; if the latter were to be used the 'decision' would be endogenous to the model. Furthermore, out of those who do apply, less than 9% are not admitted and out of those who are admitted, 8% do not enrol.[12] Thus, the decision to apply would be a highly significant explanatory factor in who enrols in tertiary education.

It is clear that out of the three population groups, the White sample has the highest rate of tertiary application as well as the highest percentage of respondents passing matric with an exemption. While 68% of the White sample passes matric with an exemption, only 18% of the Black sample does so. Note that out of the total sample the exemption variable is missing for 72 respondents; these respondents have a mean ability measure that is 0.12

[12] The most common reason for not enrolling when admitted is not being able to afford the tuition fees. At the same time, over 90% of those who are not admitted and 83% of those who were admitted but did not enrol, say they will return to attend classes in the future.

standard deviations below the mean for those who do not have a missing exemption dummy. This is also below the mean ability level of those who do not obtain an exemption, which suggests that many of those who are missing this variable are likely to not have obtained an exemption. Furthermore, the Black sample, which has the highest percentage of observations with an 'exemption missing' dummy, also has the lowest mean exemption attainment. These missing observations may cause measurement error to the 'exemption' dummy and as a result the coefficient may understate the effect of obtaining an exemption on the decision to apply to tertiary education.[13]

The ability measure is a latent, individual measure and has been created using the unexplained variation in the CAPS literacy and numeracy test scores. In order to determine the net impact of ability on the decision to attend tertiary education, it is necessary to separate out the effects of ability versus the effects of other individual, school and home background effects. Therefore, the first step in the analysis is to create a measure of ability that is unrelated to the individual's age or current education level as well as their SES. SES is assumed to include the individual's household or living circumstance as well as the quality of their education. The literacy and numeracy score in the CAPS data set provides a measure of manifest cognitive ability that is influenced by the above mentioned factors to a large extent. Thus, in order to create a more accurate measure of latent ability, these factors need to be purged out of the test scores. This is done by regressing the standardised total score on individual, household and school level characteristics and then saving the residuals, or the unexplained variation in the scores, as the measure of latent ability. Note

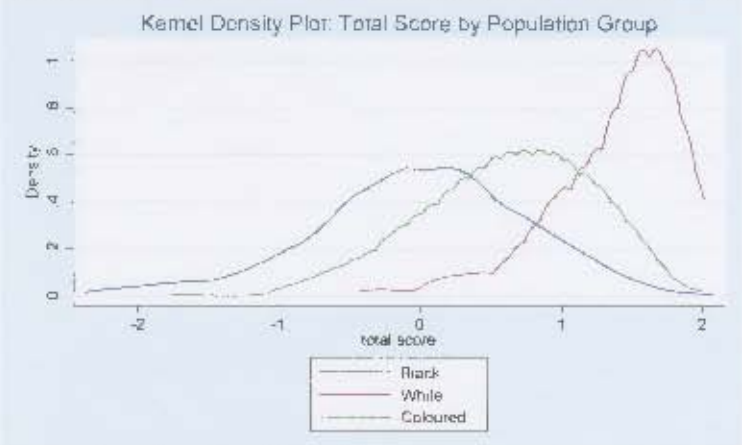
[13] However, one would not expect this measurement error to have a drastic effect considering that the 'exemption' variable is missing for less than 6% of the sample. The effect may also only be relevant for the Black sample as it has the highest percentage of observations with a missing exemption dummy.

that this part of the analysis includes the full sample of wave 1 respondents and uses the wave 1 data (2002), which is when the test was written. The regression results can be found in Appendix C.

The regression output indicates that individual, household and school characteristics are all relevant in explaining the test scores. The explanatory variables explain just over 50 percent of the variation in the test scores, thus confirming that the scores are largely a reflection of acquired skill rather than innate ability. The other half of the variation remains unexplained. These residuals are saved and standardised and for the rest of the analysis they are used as a measure of latent ability.

Figure 1 displays the kernel densities of the standardised literacy and numeracy test scores for each population group. It is clear that the White sample has the highest mean score and the distribution is highly skewed to left, while the Black sample has the lowest mean and widest distribution.

Figure 1: Standardised literacy and numeracy test score by population group

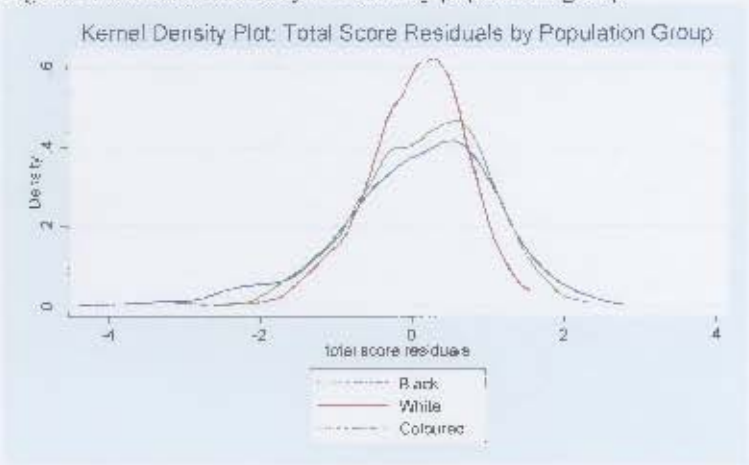


Source: Cape Area Panel Study (2002)

However, the residuals, or measure of latent ability, show a somewhat different variation between the three groups. Figure 2 shows that once age, education level and

socio-economic background variables have been controlled for, the variation in individual ability between the population groups is much less dispersed. Note that the White kernel density function of the residuals has a relatively narrow distribution thus indicating that the White literacy and numeracy scores are best explained by the model. On the other hand, the model has the least explanatory power in terms of the Black test scores.

Figure 2: Standardise ability measure by population group



Source: Cape Area Panel Study (2002)

The other explanatory variables used in the analysis are at the individual, school and household level. Individual-level variables include age in 2005, gender, year of matriculation and population-group/home-language. These last two characteristics have been combined into dummies as almost all of the Black respondents report an African language (primarily Xhosa) as their home language. Thus, it makes more sense to categorise these respondents into one group. Similarly, Whites and Coloureds are categorised as either English or Afrikaans speaking. The household variables can be classified into 'household-composition' including proportion of life lived with biological

mother and father, and into 'household-resources' including parental education (knowledge resources) and per capita household income (financial resources). Finally, the school-level variables include the pupil-teacher ratio and the former education department of the school. For the respondents who had completed matric by 2002, these variables refer to the school where they matriculated. For those who were still enrolled in secondary school in 2002, the variables refer to the school which they were attending at that point in time. It is assumed however that individuals are likely to attend schools of similar quality throughout their life. Thus, the school-level variables can be understood to represent the individual's overall in-school experience.

The variation in mean household and school characteristics between the population groups reveals that the Black group is the most disadvantaged in terms of economic resources as well as family support. The black respondents are more likely to have attended ex-DET or former homeland schools, which are characterised by a lack of resources and poor teaching relative to ex-HOA schools for example (which is where the majority of the White respondents matriculated). The Black sample also has a higher pupil-teacher ratio, lower mean parental education, lower mean per capita household income, fewer household resources (indicated by the 'more than five books in the home' dummy and presence of a computer dummy), and are also likely to have spent a smaller proportion of their lives living with their biological parents. In terms of these variables, the White group is most well off while the Coloured group falls in the middle. These apparent differences provide strong motivation for analysing the population groups separately. Because the books and computer dummies do not show much variation within the population groups, these

variables are not included in the regression analysis. However, it is important to bear in mind that population group is a strong indicator of household resources when analysing the regression results.

Table 4: Summary of variables by 'applied' (means)

applied	0	1	Total	Significantly different
White English	0.03	0.16	0.11	*
White Afrikaans	0.04	0.08	0.06	*
Coloured English	0.20	0.22	0.21	
Coloured Afrikaans	0.38	0.20	0.27	*
Black	0.34	0.33	0.34	
Ability	0.03	0.17	0.11	*
Age 2005	21.98	21.75	21.84	*
Male	0.42	0.42	0.42	
Matric pre-2002	0.44	0.48	0.46	
Matric 2002	0.14	0.20	0.17	*
Matric 2003	0.18	0.20	0.19	
Matric 2004	0.17	0.14	0.15	*
Exemption	0.12	0.40	0.29	*
DET	0.22	0.23	0.23	
HOA	0.12	0.29	0.21	*
HOD/HOR	0.59	0.44	0.51	*
New education department	0.03	0.03	0.03	
Homeland	0.17	0.11	0.14	*
Pupil-teacher ratio	30.37	28.96	29.62	*
Education of mother	8.59	10.42	9.70	*
Education of father	8.56	10.82	9.98	*
% life lived with mother	0.86	0.89	0.88	*
% life lived with father	0.62	0.66	0.64	*
Per capita HH income	943.44	1968.78	1561.38	*
Edu. department missing	0.11	0.33	0.24	*
Pupil-teacher ratio missing	0.11	0.32	0.24	*
Exemption missing	0.09	0.03	0.06	*
Education of mother missing	0.08	0.07	0.07	
Education of father missing	0.31	0.24	0.27	*

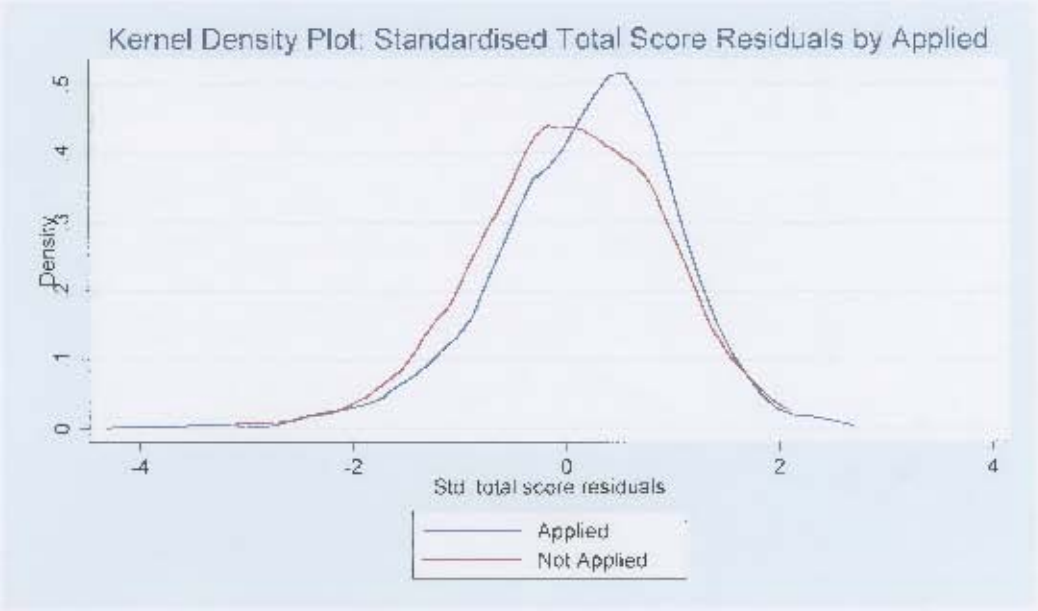
Source: School data is taken from the School Register of Needs Survey (2000); all other data is from the Cape Area Panel Study (2002 – 2005).

Table 4 above shows how the explanatory variables vary in mean values according to the 'applied' dependent variable dummy. Most noteworthy is the significant variation in mean ability as well as the exemption dummy between the zero and one category. One may expect the regression analysis to show a positive impact of both these variables on the decision to apply for tertiary education. Furthermore, out of the respondents who do obtain an exemption, less than 16% do not apply for tertiary education. Not surprisingly, the White-English dummy as well as the HOA dummy also show marked deviations in their mean values across the 'applied' and 'not applied' categories. In general, the 'applied' category has significantly higher mean values of both household and school resources compared to the zero category.

Kernel density graphs

Figures 3 and 4 show the kernel densities of the ability measure for the 'applied' and 'exemption' outcomes. The graphs indicate that the mean ability for those who have applied to tertiary education and for those who attained a matric exemption is greater than the mean ability of those who have not applied and who did not obtain exemptions respectively.

Figure 3: Ability by 'applied'



Source: Cape Area Panel Study (2002 – 2005)

Figure 4: Ability by 'exemption'

Source: Cape Area Panel Study (2002 – 2005)

Household fixed-effect sample

The fixed-effect regressions include only households with more than one respondent. This means that only 353 respondents, making up 171 households, are included in this part of the analysis (27.6% of the original sample). Approximately 9% of these households have three respondents, while for the most part the sample includes two respondents in each household.[14] Table 5 shows the number of households in the total sample with either 1, 2 or 3 respondents.

Table 5: summary of # of Young Adults in households (total sample)

Young Adults	Freq.	Percent
1	926	72.4
2	320	25.02
3	33	2.58
Total	1,279	100

It is also important to investigate whether or not this reduced sample differs in its 'explanatory variable' characteristics compared to the single-respondent households. The table below shows mean values for each variable across the two samples and indicates whether or not these means are significantly different from one another:

[14] Note that the majority of respondents who live in the same household are either siblings or cousins, while a very small percentage of respondents who live together are unrelated. See Appendix D for details.

Table 6: Comparison of samples: means

	single YA households	2-3 YA households	Significantly different
Applied	0.60	0.62	
Ability	0.08	0.19	*
Age 2005	21.84	21.84	
White English	0.11	0.10	
White Afrikaans	0.06	0.07	
Coloured English	0.21	0.20	
Coloured Afrikaans	0.29	0.24	*
Black	0.32	0.37	
Prop. life lived with mother	0.88	0.88	
Prop. life lived with father	0.65	0.64	
Per capita HH income	1539.32	1619.52	
Pupil-teacher ratio	29.65	29.52	
Education of mother	9.55	10.09	*
Education of father	9.77	10.55	*
DET	0.20	0.29	*
HOA	0.21	0.21	
HOD/HOR	0.53	0.47	*
Homeland	0.14	0.15	
New education department	0.03	0.02	
Exemption	0.27	0.34	*
Exemption missing	0.06	0.06	
Edu. mother missing	0.07	0.07	
Edu. father missing	0.26	0.28	
Pupil-teacher ratio missing	0.23	0.25	
Edu. department missing	0.24	0.25	
No. Observations	926	353	

Source: School data is taken from the School Register of Needs Survey (2000); all other data is from the Cape Area Panel Study (2002 – 2005).

It is evident that while the sample size has dropped significantly, the proportion of the sample in the 'applied' category has remained relatively unchanged, increasing only slightly by approximately 2% (the full sample has a mean tertiary application rate of 60%).

However, note that the reduced sample has a mean ability level that is above the average. These respondents also have parental education levels slightly above the average, and are more likely to have obtained a matric exemption. At the same time however, they are more

Table 6: Comparison of samples: means

	single YA households	2-3 YA households	Significantly different
Applied	0.60	0.62	
Ability	0.08	0.19	*
Age 2005	21.84	21.84	
White English	0.11	0.10	
White Afrikaans	0.06	0.07	
Coloured English	0.21	0.20	
Coloured Afrikaans	0.29	0.24	*
Black	0.32	0.37	
Prop. life lived with mother	0.88	0.88	
Prop. life lived with father	0.65	0.64	
Per capita HH income	1539.32	1619.52	
Pupil-teacher ratio	29.65	29.52	
Education of mother	9.55	10.09	*
Education of father	9.77	10.55	*
DET	0.20	0.29	*
HOA	0.21	0.21	
HOD/HOR	0.53	0.47	*
Homeland	0.14	0.15	
New education department	0.03	0.02	
Exemption	0.27	0.34	*
Exemption missing	0.06	0.06	
Edu. mother missing	0.07	0.07	
Edu. father missing	0.26	0.28	
Pupil-teacher ratio missing	0.23	0.25	
Edu. department missing	0.24	0.25	
No. Observations	926	353	

Source: School data is taken from the School Register of Needs Survey (2000); all other data is from the Cape Area Panel Study (2002 – 2005).

It is evident that while the sample size has dropped significantly, the proportion of the sample in the 'applied' category has remained relatively unchanged, increasing only slightly by approximately 2% (the full sample has a mean tertiary application rate of 60%).

However, note that the reduced sample has a mean ability level that is above the average. These respondents also have parental education levels slightly above the average, and are more likely to have obtained a matric exemption. At the same time however, they are more

likely to have attended ex-DET schools compared to the single-respondent households. If one assumes that these characteristics are proxies for socio-economic status, the differences between the single versus 2-3 respondent households seem to flow in opposite directions; the fixed-effect sample is more endowed with household resources (in terms of parental education) and are more likely to have passed matric with an exemption, however they are also more likely to have attended poor-quality (ex-DET) schools. As a result, sample selection bias does not appear to be a problem for the household fixed-effect analysis.

An important assumption of this part of the analysis is that young adults who reside in the same household are likely to have attended (or attend) schools of similar quality. In other words, they have been exposed to similar life opportunities. Out of the 100 households in which the former school department is known for more than one respondent, in only 10 have respondents attended schools of a different type. Thus, one can assume that the household fixed-effect is able to control for school-quality unobservables to a large degree.[15]

[15] See Appendix E for details of variation in school-type within households.

ANALYSIS OF DATA

Regressions 1-4: 'Applied' probit model

	1. Total	2. Black	3. Coloured	4. White
	applied	applied	applied	applied
Ability	0.064 (0.018)***	0.041 (0.025)	0.078 (0.029)***	0.058 (0.022)***
Age 18	-0.014 (0.256)	-0.664 (0.025)***	0.509 (0.026)***	
Age 19	-0.084 (0.252)	-0.742 (0.026)***	0.653 (0.047)***	0.086 (0.033)**
Age 20	-0.032 (0.238)	-0.835 (0.027)***	0.753 (0.057)***	0.068 (0.081)
Age 21	-0.064 (0.243)	-0.883 (0.025)***	0.777 (0.060)***	0.092 (0.047)*
Age 22	-0.067 (0.244)	-0.912 (0.023)***	0.759 (0.055)***	0.084 (0.066)
Age 23	-0.039 (0.242)	-0.892 (0.024)***	0.742 (0.053)***	0.066 (0.040)*
Age 24	-0.039 (0.243)	-0.891 (0.023)***	0.694 (0.047)***	0.054 (0.068)
Age 25	-0.132 (0.258)	-0.869 (0.024)***	0.654 (0.044)***	0.06 (0.039)
Age 26	-0.358 (0.265)	-0.666 (0.025)***		
Male	-0.031 (0.031)	-0.023 (0.052)	0 (0.045)	-0.041 (0.037)
White Afrikaans	-0.171 (0.102)*			
Coloured English	-0.018 (0.081)			
Coloured Afrikaans	0.003 (0.087)			
Black	0.157 (0.080)**			
DET	0.186 (0.077)**			
HOA	0.193 (0.095)**			
HOD/HOR	0.168 (0.109)			
New ed. department	0.148 (0.098)			
Ed. Dpt. Missing	0.365 (0.079)***			
Exemption	0.221 (0.037)***	0.092 (0.064)	0.186 (0.057)***	0.223 (0.081)***
Exemption missing	-0.159 (0.090)*	-0.298 (0.101)***	-0.126 (0.129)	
Pupil-teacher ratio	0.003 (0.004)	-0.002 (0.005)	0.005 (0.007)	-0.017 (0.006)***
P-T ratio missing	0.108 (0.134)	0.16 (0.164)	0.421 (0.122)***	-0.531 (0.298)*
Education mother	0.001 (0.020)	0.011 (0.026)	0.016 (0.036)	0.039 (0.039)
Ed. Mother squared	0.001 (0.001)	0 (0.002)	0 (0.002)	-0.001 (0.002)
Ed. Mother missing	0.077 (0.091)	0.127 (0.115)	0.112 (0.166)	
Education father	0.031 (0.020)	-0.003 (0.026)	0.07 (0.044)	0.003 (0.070)
Ed. Father squared	-0.001 (0.001)	0.001 (0.002)	-0.002 (0.002)	-0.001 (0.003)
Ed. Father missing	0.206 (0.073)***	-0.019 (0.109)	0.443 (0.136)***	-0.259 (1.256)
Matric 2002	0.111 (0.048)**	0.038 (0.077)	0.263 (0.066)***	-0.051 (0.093)
Matric 2003	0.032 (0.056)	0.063 (0.078)	0.129 (0.085)	0.057 (0.041)
Matric 2004	-0.062 (0.082)	-0.226 (0.090)**	0.129 (0.117)	-0.147 (0.314)
% life lived mother	0.057 (0.069)	0.122 (0.100)	0.093 (0.112)	0.04 (0.131)
% life lived father	0.027 (0.053)	-0.064 (0.071)	0.004 (0.078)	0.07 (0.069)
Log p.c. HH income	0.107 (0.023)***	0.062 (0.034)*	0.111 (0.037)***	0.047 (0.020)**
Observations	1259	428	600	204

Probit model. * significant at 10%; ** significant at 5%; *** significant at 1%. Regressions are estimated with robust standard errors (in brackets) to allow for correlation between observations from the same household. Coefficients indicate marginal effects. The Ability variable is a standardised measure. Base categories are: Age 17, White-English for population-group-language, 'homeland' school for education department, and matric pre-2002 for matric year. School-level variables are taken from the School Register of Needs Survey (2000); other variables come from the Cape Area Panel Study Survey (2002–2005).

Regression 1

Regression 1 models the decision to apply for tertiary education as explained by the individual's ability, as well as other individual, household and school characteristics. These other explanatory variables include age dummies, gender, race-language dummies, pupil-teacher ratio, former education department of school, whether or not the individual passed matric with an exemption, year of matriculation, parental education, proportion of life spent with parents and per capita household income. The school-level variables refer to the individual's last school attended as of 2002, and the household variables refer to the individual's 2002 household. Thus, these explanatory variables act as proxies for the individual's general school-quality and home-background experience.

The results indicate that the measure of ability is significant at the 1% level in explaining the decision to apply to tertiary education. The coefficient also has the expected positive sign – on average, an individual with a level of ability one standard deviation above the mean has a 6.5% greater probability of applying to tertiary education. Other significant explanatory variables include the White-Afrikaans dummy, which indicates a negative impact compared to the base category, White-English. The Black-African-language dummy however is positive suggesting that, holding all else constant, the Black African-language speakers would be more likely to apply to tertiary education compared to any other race-language group. In other words, if the Black African-language speakers had the mean characteristics of the other groups, they would be more likely to decide to pursue further education compared to the other groups. The DET, HOA and HOD/HOR dummies are also

significant and positive indicating that individuals who attended these schools are more likely to apply for tertiary education compared to those who attended former homeland schools. The HOA dummy has the largest co-efficient, which suggests that the former White schools continue to produce the most tertiary-level candidates. [16]

The exemption dummy is also highly significant and positive. This is an important explanatory variable for two reasons. Firstly, it is a proxy for the individual's matric results[17] (which would be highly important in the decision to pursue further education), and secondly it is the only right-hand-side variable that one may expect to be influenced by the ability measure. The latter suggests a possible *indirect* effect of ability on the decision to apply for tertiary education, via the attainment of a matric exemption (or simply, passing matric well). This possibility is explored later. If it is found that ability is in fact significant in explaining exemption, the coefficient on ability in explaining the decision to apply for tertiary education is likely to be biased downwards to some extent.

Not surprisingly, per capita household income is also significant and positively related to the decision to apply, reflecting the high direct financial costs of tertiary education in South Africa. Members of poorer households may also be under more pressure to enter the labour market soon after completing high school in order to help support the household; therefore

[16] Recall that the attrition in the CAPS survey is mostly due to respondents moving to the Eastern Cape and who are more likely to have attended ex-homeland schools. Thus, the coefficients on the school-department variables may be understated. If more ex-homeland school respondents were included, the effect of the poor quality of these schools (relative to the other school-types) may be more prominent.

[17] The data on matric results for each subject is missing for a large number of respondents, especially for those who matriculated pre-2002. Therefore, this data is not included in the analysis. The 'exemption' dummy is assumed to capture this information to a fairly large extent.

they would be less able to substitute labour for studying as the opportunity cost of studying is higher in their case.

The exemption-missing dummy is significant at the 10% level and negative in sign. A closer examination of this category of respondents reveals that they are more likely to have attended ex-DET schools or former homeland schools compared to those who do not have this variable missing and who are more likely to have attended former White schools (HOA). Therefore, this may explain the negative association. Similarly, the highly significant and positive coefficient on the school-department-missing dummy may be explained by the higher mean per capita income of this group of people. The education-of-father-missing dummy is also highly significant and positive; this may be because this group is made up primarily of Black African-language speakers and thus it may be picking up the positive impact of this category on the decision to apply to tertiary education, holding all else constant.

Regressions 2-4

In order to test whether the factors that explain the decision to apply to tertiary education operate differently across different population groups, regressions for each group have been run separately. Regression 2 includes the Black sample and reveals that neither ability nor exemption is significant in explaining the decision to apply to tertiary education. Age at matriculation however is highly significant - the dummies are all large and negative, the base category being 17 years of age. Although the age dummies are for age at 2005, the

year of matriculation is controlled for, thus one is comparing different age-groups within a single matric year. The results suggest that grade repetition (a common trend in black schools) and the resulting later age of matriculation are responsible in determining whether the individual decides to apply for tertiary education. Considering the poor quality of most African schools, it is not surprising that those who still manage to matriculate at a relatively young age, transcending these 'poor-school' disadvantages in a sense, are those who then decide to pursue higher education. In turn, this may explain why the coefficients on both ability and exemption are not significant (as their effects may be picked up in the age dummies). Furthermore, not only are those who complete matric at a younger age more likely to cope with tertiary education, but they are also under less pressure to enter the labour market compared to their older counterparts. [18]

The insignificant coefficient on the ability variable may also be interpreted as a reflection of the disadvantaged school and home environments that are characteristic of the Black sample. It could be argued that ability is made redundant when the conditions of school and home environment fall below a certain point of inadequacy; in other words ability may be crowded-out by these poor conditions.

Other significant explanatory variables in the Black regression include matriculants of 2004 (the negative sign indicating that those who matriculated more recently are less likely to have applied), and per capita household income (which has the expected positive sign) and

[18] A summary of mean age at matriculation by population group can be found in Appendix F.

the exemption-missing dummy[19]. Thus, for the Black population, having financial resources is more important in the decision to apply for tertiary education compared to individual ability.

Regressions 3 and 4, which are for the Coloured and White samples respectively, show somewhat different results to the Black regression. Most noteworthy is that both the ability and exemption coefficients are significant at the 1% level in both the Coloured and White regressions. This suggests that for these two population groups, individual ability has an important impact on the decision to apply for tertiary education, independent from school and home background variables. It could also reflect the relatively advantaged schooling and home background conditions compared to the Black sample, allowing individual ability to play a role in educational decision-making for the White and Coloured samples. For the White group, pupil-teacher ratio is also significant and has a negative impact, and for both population groups per capita household income is significant and positive.

In order to test the equality of coefficients of the same variable across the three groups, the regressions are rerun simultaneously; the new estimates are then comparable using chi-squared tests. The results from these tests can be found in Appendix H. Regarding the Coloured results, the coefficient on ability is not significantly different from both the Black and White coefficients. However, the Black ability coefficient is significantly different from the White coefficient at the 5% level. Thus, the impact of ability on the decision to apply to tertiary education is significantly larger within the White sample compared to the

[19] This may be due to the lower mean per capita household income, which is characteristic of this group of people (within the Black population group). See Appendix G for details.

Black sample. In terms of the exemption dummy, while the Black and Coloured coefficients are not significantly different from one another, they are both significantly different from the White coefficient at the 1% level. In other words, obtaining a matric exemption has a greater impact on the decision to apply to tertiary education for Whites compared to Coloureds and Blacks. The per capita household income coefficients are not significantly different between any of the groups. Thus, pure financial constraints act in a similar manner across the population groups. [20]

The effect of ability on 'exemption'

The next step is to determine the effect of ability on the attainment of a matric exemption. Table 7 displays the impact of ability on passing matric with an exemption for the sample as a whole and for population groups separately. The same individual, household and school factors are controlled for (as in the 'ability' model); however these results are not displayed here. The total sample regression shows that the ability coefficient is positive and significant at the 1% level – on average, a one standard deviation above the mean ability level is associated with a 14.7% higher probability of obtaining an exemption on completion of matric. This result confirms the suspicion above that ability may have an indirect effect on the decision to pursue tertiary education via its impact on matric results. As a result, the coefficient on ability in the overall regression as well as the separate regressions may be slightly biased downwards.

[20] A heteroskedastic probit model was also run on the three regressions and the chi-squared test reveals that the variance of the error structures are not significantly different from one another i.e. there is evidence of homoskedasticity. (Prob > chi2= 0.57)

Other noteworthy results include the negative coefficients on the race-language dummies indicating that White English-speakers are the most likely group to obtain a matric exemption; the significant and negative coefficient on the male dummy; and the positive impact of per capita household income once again.

Table 7: Ability coefficient results in the 'Exemption' probit model

Sample	Total	Black	Coloured	White
	exemption	exemption	exemption	exemption
Ability	0.147 (0.024)***	0.039 (0.023)*	0.113 (0.024)***	0.244 (0.057)***

Probit model. * significant at 10%; ** significant at 5%; *** significant at 1%. Regressions are estimated with robust standard errors (in brackets) to allow for correlation between observations from the same household. Coefficients indicate marginal effects. The Ability variable is a standardised measure. Source: Cape Area Panel Study Survey (2002–2005).
See Appendix I for full results.

Running the regressions simultaneously and using a chi-squared test of equality of coefficients, it is found that overall the coefficients on ability are significantly different from each other across the population groups. However, testing the hypothesis in pairs reveals that only the Black coefficient on ability is significantly different from the other two while the coefficients in the Coloured and White regressions are not significantly different from each other. Thus, once again we see that for the Black group, ability is relatively less important in determining their educational outcomes. This may be attributable to the poor quality of Black schools as well as less resourceful home environments (as shown above), which may limit the effect of individual ability and as a result make it more difficult for Black learners to achieve high educational outcomes.

Fixed-Effects Model

Regression 5 is a linear probability model of the outcome 'applied', using a household fixed effect. As mentioned above, the explanatory variables should only include characteristics which vary between members of the same household. Thus, included on the right-hand side is: the ability measure, age dummies, gender, exemption (and missing dummy), and year of matriculation.

The results show that ability is significant at the 10% level and has a positive impact on the decision to apply to tertiary education. This suggests that individuals living in the same household differ from one another in terms of ability and those with higher levels of ability will be more likely to pursue higher education. Thus, individual ability is significant in explaining educational attainment even after socio-economic background and other unobservable household characteristics are taken into account.

Regression 5: 'Applied' linear probability model using household fixed-effect

	Applied
Ability	0.109 (0.059)*
Age 19	-0.08 (0.223)
Age 20	-0.007 (0.250)
Age 21	-0.034 (0.252)
Age 22	-0.183 (0.283)
Age 23	-0.177 (0.282)
Age 24	-0.122 (0.286)
Age 25	-0.25 (0.293)
Age 26	-0.989 (0.418)**
Male	-0.058 (0.065)
Exemption	0.094 (0.100)
Exemption missing	-0.284 (0.153)*
Matric 2002	-0.11 (0.118)
Matric 2003	-0.15 (0.143)
Matric 2004	-0.475 (0.172)***
Constant	0.867 (0.289)***
Observations	344
R-squared	0.68

Linear probability fixed effect model. * significant at 10%; ** significant at 5%; *** significant at 1%. Regressions are estimated with robust standard errors (in brackets). The Ability variable is a standardised measure. The base categories are: 18 years for age, and matric pre-2002 for matric year. Source: Cape Area Panel Study Survey (2002 – 2005).

A similar model is run but with exemption on the left-hand side. Table 8 below displays the coefficient for the impact of ability on passing matric with an exemption. Once again, ability is significant at the 10% level and has a positive impact on exemption attainment. Thus, not only does ability have a direct effect on whether the individual progresses to tertiary education, but it is also important in explaining matric outcomes and again this relationship is separate from family background or household circumstance. However, one cannot conclude that ability has an indirect effect on the decision to apply to tertiary education as the exemption dummy is not significant in regression 5.[21]

[21] This may be because in only 28% of the households in this sample do the members differ in their 'exemption' outcome i.e. the variation in exemption attainment is fairly low within households.

Table 8: Ability coefficient in 'Exemption' linear probability model using household fixed-effect

	Exemption
Ability	0.113 (0.062)*

*See Appendix I for full results

Comparing ability coefficients between models

In order to compare the relative sizes of the ability coefficients across the different sections of the analysis, the initial probit regressions are rerun as linear regressions for the full sample as well as for the fixed effect sample. The results for the ability and exemption coefficients are displayed below for these two regressions as well as for the fixed-effect model. The reason for such a comparison is to explore why the coefficients in the fixed-effect model may be larger than in the non-fixed-effect models.

The non-fixed-effect regressions suggest that the effect of ability on the decision to apply is largely indirect and operates primarily through the attainment of a matric exemption. In the fixed-effect model however, the impact of exemption on 'applied' is no longer significant; this may be explained by the fact that there is very little variation in exemption attainment within households. Thus, the exemption dummy is indicative of socio-economic status to a fairly large extent. In turn, this may explain why exemption attainment has such a large impact on 'applied' in the non-fixed-effect models. Furthermore, the effect of ability on 'applied' may be picked up largely by the exemption effect, resulting in only small direct effects of ability on the decision to apply to tertiary education. Once household socio-economic status is controlled for in the fixed-effect model, the direct effect of ability becomes more evident.

Therefore, one can conclude that the relatively large ability coefficient in the fixed-effect model may be because in the non-fixed effect models, the effect of ability is largely indirect (through the exemption attainment) while in the fixed-effect model, the effect is largely direct due to the small variation in exemption attainment within households.

Table 9: Linear regressions

	Full Sample		Fixed-effect sample			
	Non-fixed effect linear model		Non-fixed effect linear model		Fixed-effect linear model	
	applied	exemption	applied	exemption	applied	exemption
Ability	0.05 (0.016)***	0.109 (0.016)***	0.018 (0.037)	0.139 (0.038)***	0.109 (0.059)*	0.113 (0.062)*
Exemption	0.182 (0.035)***		0.151 (0.074)**		0.094 (0.100)	
# obs	1259	1187	344	324	344	324

* significant at 10%; ** significant at 5%; *** significant at 1%. Regressions are estimated with robust standard errors (in brackets). Source: Cape Area Panel Study (2002-2005). Full results are not shown here.

CONCLUSION

The results of the analysis have shown that overall, individual ability is significant in explaining educational attainment over and above socio-economic context. One can be certain that this measure of ability is highly independent of school and home background firstly, because the observable characteristics of these factors have been purged from the test scores and secondly, because the fixed-effects analysis controls for any unobservable home-background or other contextual factors. This part of the analysis confirms that even individuals living within the same household are able to advance to different levels of educational attainment due to different levels of individual ability.

In terms of labour market or returns-to-schooling literature, the results suggest that wage or other labour market outcomes equations may suffer from endogeneity bias if ability is not controlled for. Even though the bias is likely to be small, this analysis has shown that latent ability is indeed significant in explaining educational attainment; this type of ability may influence labour market decisions and outcomes too.

Despite these results however, one cannot conclude that ability plays a role in educational achievement and decisions for all segments of the population. It has been shown that ability is not significant for the Black population in determining the decision to pursue tertiary education. Furthermore, it is significantly less important compared to the Coloured and White samples in influencing the attainment of a matric exemption. The regression analysis has shown that for the Black sample, age at matriculation has a large and negative impact

on the attainment of an exemption and the decision to apply for tertiary education. Thus, grade repetition (commonly regarded as a result of poor quality schools) and the resulting later age at matriculation make these individuals less likely to advance to tertiary education, perhaps because there is also more pressure on them to enter the labour market. This effect is reinforced by the fact that these individuals come from poor backgrounds and have few financial resources, making the prospect of attending tertiary education less of a reality.

Therefore, the more accurate conclusion for the Black population is that individual ability is not able to play a significant role in educational attainment over and above socio-economic context; rather, schools and home environments which are characterised by lack of adequate resources and unfavourable learning environments lead to a 'crowding out' of individual ability. For the Coloured population and even more so for the White, the results suggest a more meritocratic process whereby ability is able to play a small part; however educational attainment and outcomes are still largely dependent on financial resources. In comparison however, the unequal circumstances of the Black population place them at a distinct disadvantage long before they enter the labour market.

Appendix A: Requirements for a matric exemption

There are four groups of subjects as outlined by the Rosebank House Damelin College website:

1. Languages (first and second)
2. Mathematics subjects (if not chosen, matriculant must do a science (3) to obtain an exemption)
3. Sciences (Science and Biology)
4. Third languages
5. Social Sciences (eg. History, Geography, Economics)
6. Other (eg. Accounting, Art, Business Economics, Computers)[22]

In order to obtain a matric exemption/endorsement, six subjects must be chosen and passed from at least 4 different groups as follows:

- English 1st Language Higher Grade and Afrikaans 2nd language Higher Grade.
- The language policy is changing, but for South African students a 1st and 2nd language chosen from the official languages is required.
- At least 2, but preferably 3 subjects, on the Higher Grade chosen from at least 2 different groups in addition to the two languages.
- There are special regulations for foreign students with immigrant status who only want to do one language

[22] Note that subjects offered will vary by school or college.

For a non-university entrance certificate, six subjects must still be taken, though they may all be taken on the Standard Grade.

(Source: http://www.rosebankhousedamelin.co.za/html/prospectus.html (21/05/07))

Appendix B: Attrition

Reasons for incomplete interviews

Final result of Wave 3 interview	Freq.	Percent
Not Available	209	17.61
Refused	187	15.75
Deceased	28	2.36
Moved within Cape Town	134	11.29
Moved within South Africa	295	24.85
Moved Abroad	66	5.56
Moved No details	230	19.38
Temporary Living Situation	28	2.36
Mentally Unfit/Disabled	2	0.17
No contact	8	0.67
Total	1,187	100

Source: Cape Area Panel Study (2002-2005)

Province moved to (using wave 2 & 3 data)

Province moved to within SA:	Freq.	Percent
Eastern Cape	189	41.63
Free State	6	1.32
Gauteng	56	12.33
Kwa-Zulu Natal	9	1.98
Limpopo	2	0.44
Mpumalanga	2	0.44
Northern Cape	9	1.98
North-West Province	1	0.22
Western Cape	180	39.65
Total	454	100

Source: Cape Area Panel Study (2002-2005)

Appendix C: Creation of latent cognitive ability measure

	Standardised test score
Age 15_2002	-0.055 (0.042)
Age 16_2002	-0.19 (0.045)***
Age 17_2002	-0.257 (0.046)***
Age 18_2002	-0.3 (0.048)***
Age 19_2002	-0.361 (0.052)***
Age 20_2002	-0.432 (0.056)***
Age 21_2002	-0.391 (0.057)***
Age 22_2002	-0.45 (0.058)***
Male	0.091 (0.021)***
Black	-0.625 (0.062)***
Coloured	-0.369 (0.045)***
Wrote test in home language	0.065 (0.037)*
Education level in wave 1	0.201 (0.008)***
Education of mother	0.008 (0.015)
Education of mother squared	0.001 (0.001)
Education of mother missing	0.068 (0.076)
Education of father	-0.026 (0.013)**
Education of father squared	0.003 (0.001)***
Education of father missing	0.007 (0.061)
Proportion of life lived with mother	0.096 (0.046)**
Proportion of life lived with father	0.022 (0.033)
Mother helped with homework	-0.135 (0.032)***
Father helped with homework	0.094 (0.038)**
Log per capita household income	0.053 (0.015)***
More than five books in the home	0.062 (0.032)*
Computer in the home	0.112 (0.037)***
Pupil-teacher ratio	-0.004 (0.003)
Pupil-teacher ratio missing	0.058 (0.114)
DET	0.022 (0.061)
HOA	0.384 (0.072)***
HOD/HOR	0.136 (0.068)**
New education department	0.101 (0.079)
Education department missing	-0.025 (0.092)
Constant	-1.872 (0.175)***

Observations	4622
R-squared	0.51

Probit model. * significant at 10%; ** significant at 5%; *** significant at 1%. Regressions are estimated with robust standard errors (in brackets) to allow for correlation between observations from the same household. Coefficients indicate marginal effects. The dependent variable is the standardised literacy and numeracy score. Base categories are: Age 14 (2002), White for population group, and 'homeland' school for education department. School-level variables are taken from the School Register of Needs Survey (2000); other variables come from the Cape Area Panel Study Survey (2002/wave 1).

Appendix D: Relationships between respondents in same household

The majority of respondents (young adults) living in the same household in 2002 are siblings. Very few co-residents are unrelated, indicated by 'friend' or 'other unrelated':

#Young adult 1	**158**	
Relationship to YA#1	**Freq.**	**Percent**
Brother/ sister	138	71.5
Brother-in-law/ sister-in-law	3	1.55
Step brother/sister	4	2.07
Half brother/sister	5	2.59
Nephew/ niece	4	2.07
Cousin	30	15.54
Other family	1	0.52
Friend	5	2.59
Other unrelated	3	1.55
Total	193	100

#Young adult 2	**158**	
Relationship to YA#2	**Freq.**	**Percent**
Brother/ sister	146	75.65
Brother-in-law/ sister-in-law	1	0.52
Step brother/sister	2	1.04
Half brother/sister	4	2.07
Uncle/ aunt	4	2.07
Cousin	28	14.51
Friend	6	3.11
Other unrelated	2	1.04
Total	193	100

#Young adult 3	**35**	
Relationship to YA#3	**Freq.**	**Percent**
Brother/ sister	60	53.57
Brother-in-law/ sister-in-law	1	0.89
Step brother/sister	1	0.89
Half brother/sister	1	0.89
Uncle/ aunt	7	6.25
Cousin	33	29.46
Friend	4	3.57
Other unrelated	5	4.46
Total	112	100

Source: Cape Area Panel Study (2002)

Appendix E: Respondents in same household tend to go to the same school types (fixed effect sample)

Number of households with respondents who have former education department of school variable missing

Total # in HH with School Educ. Department Missing	Freq.	Percent
0	204	57.79
1	109	30.88
2	40	11.33
Total	353	100

Source: Cape Area Panel Study (2002)

Number of respondents in same household with 'same' or 'different' former school department

school type	# young adults	Percent
different	20	9.8
same	184	90.2
Total	204	100

Source: Cape Area Panel Study (2002)

Number of households with respondents with 'same' of 'different' former school department

school type	# households	Percent
different	10	10
same	90	90
Total	100	100

Source: Cape Area Panel Study (2002)

Appendix F: Mean age at matriculation by population group

Black individuals tend to finish school at a later age compared to Coloured and White individuals. The table below shows mean age at matriculation by population group for the years 2002-2004.

	black	coloured	white
year of matriculation	**mean age**	**mean age**	**mean age**
2002	19.30	18.14	18.13
2003	19.15	18.22	18.11
2004	19.08	18.25	18.00
Total	19.18	18.20	18.08

Source: Cape Area Panel Study (2002-2005)

Appendix G: Mean per capita household income of those who have exemption data missing, within Black sample

Exemption missing	mean pc HH income
0	517.8846
1	378.175
Total	504.8884

Source: Cape Area Panel Study (2002-2005)

Appendix H: Chi-squared tests

Results for chi-squared tests of equality of coefficients between population group regressions:

Applied regressions						
Test (null)	Ability		Exemption		Log pc HH income	
	chi2	Prob>chi2	chi2	Prob>chi2	chi2	Prob>chi2
Black=Coloured	0.85	0.3575	1	0.3164	0.86	0.3536
Black=White	4.15	0.0416	6.89	0.0087	1.28	0.2578
Coloured=White	2.37	0.124	4.23	0.0397	0.31	0.5757

Exemption regressions				
Test (null)	Ability		Log pc HH income	
	chi2	Prob>chi2	chi2	Prob>chi2
Black=Coloured	4.01	0.0453	3.4	0.0653
Black=White	8.16	0.0043	5.55	0.0185
Coloured=White	2.18	0.1399	0.86	0.3542

Appendix I: Exemption regressions

	Total	Black	Coloured	White
	exemption	exemption	exemption	exemption
Ability	0.147 (0.024)***	0.039 (0.023)*	0.113 (0.024)***	0.244 (0.057)***
Age 18	-0.207 (0.192)	-0.106 (0.101)	0.302 (0.230)	
Age 19	-0.248 (0.182)	-0.15 (0.056)***	0.32 (0.160)**	-0.111 (0.334)
Age 20	-0.368 (0.117)***	-0.194 (0.051)***	0.056 (0.104)	-0.994 (0.008)***
Age 21	-0.426 (0.097)***	-0.19 (0.062)***	-0.091 (0.057)	-0.998 (0.003)***
Age 22	-0.419 (0.089)***	-0.221 (0.063)***	-0.069 (0.057)	-1 (0.001)***
Age 23	-0.405 (0.069)***	-0.256 (0.053)***	-0.091 (0.052)*	-0.967 (0.023)***
Age 24	-0.383 (0.085)***	-0.228 (0.051)***	0.03 (0.075)	-0.992 (0.010)***
Age 25	-0.362 (0.074)***	-0.211 (0.050)***		-0.927 (0.032)***
Age 26	-0.31 (0.018)***	-0.162 (0.023)***		
Male	-0.109 (0.035)***	0.045 (0.042)	-0.104 (0.032)***	-0.061 (0.075)
White Afrikaans	-0.143 (0.056)**			
Coloured English	-0.21 (0.050)***			
Coloured Afrikaans	-0.312 (0.055)***			
Black	-0.247 (0.055)***			
DET	0.17 (0.167)			
HOA	0.092 (0.159)			
HOD/HOR	0.038 (0.149)			
New ed. department	0.095 (0.203)			
Ed. Dpt missing	0.071 (0.212)			
Pupil-teacher ratio	-0.003 (0.004)	0 (0.004)	-0.011 (0.005)**	-0.024 (0.012)**
PT ratio missing	-0.021 (0.196)	0.005 (0.127)	-0.135 (0.080)*	-0.516 (0.283)*
Education mother	0.007 (0.024)	0.001 (0.021)	0.005 (0.027)	-0.029 (0.097)
Ed. Mother squared	0 (0.001)	0.001 (0.001)	0 (0.001)	0.004 (0.004)
Ed. Mother missing	0.198 (0.151)	0.173 (0.163)	0.06 (0.171)	0.199 (0.262)
Education father	-0.061 (0.021)***	-0.027 (0.019)	-0.001 (0.026)	-0.159 (0.198)
Ed. Father squared	0.005 (0.001)***	0.002 (0.001)	0.002 (0.001)	0.008 (0.008)
Ed. Father missing	-0.096 (0.094)	-0.103 (0.078)	0.148 (0.157)	-0.742 (0.435)*
Matric 2002	-0.002 (0.062)	0.046 (0.056)	0.119 (0.076)	-0.282 (0.190)
Matric 2003	-0.218 (0.052)***	-0.088 (0.044)**	-0.135 (0.051)***	-0.983 (0.019)***
Matric 2004	-0.278 (0.050)***	-0.15 (0.042)***	-0.132 (0.055)**	-0.999 (0.001)***
% life lived mother	0.078 (0.086)	0.156 (0.075)**	-0.014 (0.077)	0.281 (0.327)
% life lived father	0.01 (0.053)	0.009 (0.052)	0.002 (0.050)	0.066 (0.152)
Log pc HH income	0.045 (0.025)*	-0.004 (0.022)	0.059 (0.025)**	0.132 (0.047)***
Observations	1187	388	572	210

Probit model. * significant at 10%; ** significant at 5%; *** significant at 1%. Regressions are estimated with robust standard errors (in brackets) to allow for correlation between observations from the same household. Coefficients indicate marginal effects. The Ability variable is a standardised measure. Base categories are: Age 17 (2005), White-English for population-group-language, 'homeland' school for education department, and matric pre-2002 for matric year. School-level variables are taken from the School Register of Needs Survey (2000); other variables come from the Cape Area Panel Study Survey (2002–2005).

	exemption
Ability	0.113 (0.062)*
Age 19	-0.271 (0.238)
Age 20	-0.483 (0.282)*
Age 21	-0.604 (0.253)**
Age 22	-0.678 (0.278)**
Age 23	-0.63 (0.285)**
Age 24	-0.477 (0.286)*
Age 25	-0.57 (0.295)*
Age 26	-0.947 (0.377)**
Male	-0.151 (0.069)**
Matric 2002	0.078 (0.099)
Matric 2003	-0.156 (0.115)
Matric 2004	-0.298 (0.165)*
Constant	1.032 (0.271)***
Observations	324
R-squared	0.75

Linear probability fixed effect model. * significant at 10%; ** significant at 5%; *** significant at 1%. Regressions are estimated with robust standard errors (in brackets). The Ability variable is a standardised measure. The base categories are: 18 years for age (2005), and matric pre-2002 for matric year. All variables come from the Cape Area Panel Study Survey (2002 – 2005).

REFERENCES

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.

Boissiere, M., Knight, J. B. and Sabot, R. H. (1985), "Earnings, Schooling, Ability, and Cognitive Skills", *The American Economic Review*, 75:5, pp. 1016-1030.

Bowles, S., Gintis, H. and Osborne, M. (2000), "The Determinants of Earnings: Skills, Preferences, and Schooling", http://e:\papers\JEL paper\Determinants of Earnings.tex May 18, 2000.

Brooks-Gunn, J. and Duncan, G. (1997), "The Effects of Poverty on Children", *The Future of Children*, 7:2, pp. 55-71.

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.

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

Cawley, J., Heckman, J. and Vytlacil, E. (2001), "Three Observations on wages and measured cognitive ability", *Labour Economics*, 8, pp. 419-442.

Charette, M. and Meng, R. (1998), "The Determinants of Literacy and Numeracy, and the Effect of Literacy and Numeracy on Labour Market Outcomes", *The Canadian Journal of Economics / Revue Canadienne d'Economique*, 31:3, pp. 495-517.

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.

Culbertson, J. (2001), "Herrnstein, Richard and Murray, Charles. (1994). *The Bell Curve: Class Structure and the Future of America.* New York: The Free Press. Reviewed by John C. Culbertson", University of Kansas, December 25, 2001. *Education Review.* Available from: http://edrev.asu.edu/reviews/rev153.htm [Accessed 18 February 2007]

Ehrenberg, R. and Smith, R. (2003), *Modern Labor Economics: Theory and Public Policy, Eighth Edition.* New York: Addison Wesley.

Fiske, E.B and Ladd, H.F. (2004) *Elusive Equity – Education Reform in Post-Apartheid South Africa.* Washington D.C: Brookings Institution Press.

Goldberger, A. and Manski, C. (1995), "Review Article: The Bell Curve by Herrnstein and Murray. Reviewed Work(s): *The Bell Curve: Intelligence and Class Structure in American Life.* by Richard J. Herrnstein; Charles Murray" *Journal of Economic Literature*, 33:2, pp. 762-776.

Hansen, K. Heckman, J. and Mullen, K. (2004), "The effect of schooling and ability on achievement test scores", *Journal of Econometrics*, 121, pp. 39 – 98.

Heckman, J. and Rubinstein, Y. (2001), "The Importance of Noncognitive Skills: Lessons from the GED Testing Program", *The American Economic Review*, 91:2, Papers and Proceedings of the Hundred Thirteenth Annual Meeting of the American Economic Association (May, 2001), pp.145-149.

Heckman, J., Stixrud, J., and Urzua, S. (2006), "The effects of cognitive and noncognitive abilities on labour market outcomes and social behaviour". Working Paper No. 12006, NBER, Cambridge, MA 02138, London.

Kerckhoff, A., Raudenbush, S. and Glennie, E. (2001), "Education, Cognitive Skill, and Labor Force Outcomes", *Sociology of Education*, 74:1, pp. 1-24.

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.G. (1996), "The Quality and Efficiency of Private and Public Education: A Case-Study of Urban India", *Institute of Economics and Statistics*, University of Oxford.

Kingdon, G.G. (1999) "How Much Do Schools Matter to Pupil Achievement in India?" *Institute of Economics and Statistics*, University of Oxford.

Liddell, C. and Rae, G. (2001), "Predicting early grade retention: A longitudinal investigation of primary school progress in a sample of rural South African children", *British Journal of Educational Psychology*, 71, pp. 413-428.

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.

Rosebank House Damelin College website. Available from: http://www.rosebankhousedamelin.co.za/html/prospectus.html [Accessed 21 May 2007]

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

Simkins, C. (2001), "The Determinants of Mathematics and Physical Science Results in the Senior Certificate Examinations of 1998 and 2000", Draft report, unpublished.

Van der Berg, S. (2002), "Education, Poverty and inequality in South Africa", *Paper to the conference of the Centre for the Study of African Economies on Economic growth and poverty in Africa,* March 2002, Oxford.

Van der Berg, S. and Burger, R. (2003) "Education and Socio-Economic Differentials: A Study of School Performance in the Western Cape". DPRU Working Paper No 03/73. University of Cape Town: Development Policy Research Unit.

Wooldridge, J. (2002), *Econometric Analysis of Cross Section and Panel Data*. London: The MIT Press.