LEARNER MOBILITY IN JOHANNESBURG-SOWETO, SOUTH AFRICA: DIMENSIONS AND DETERMINANTS

Julia Ruth de Kadt

A thesis submitted to the Wits School of Education, Faculty of Humanities, University of the Witwatersrand in fulfilment of the requirements for the degree of Doctor of Philosophy.

Johannesburg, 2011

Abstract and Keywords

Abstract: Many South African school children are known to travel fairly long distances to school each day, in pursuit of the best possible educational opportunities in a schooling system that is known to vary greatly in quality. This thesis documents the dimensions and determinants of the daily, education-related travel of primary school aged children in Johannesburg-Soweto, South Africa. It uses data on a sample of 1428 children drawn from the Birth to Twenty cohort study to provide the first population-based data on the extent of learner mobility in contemporary urban South Africa. Learner mobility is measured in three different ways: firstly by the straight line distance between a child's home and his or her school; secondly by whether the child's school falls into the same geographical area as his or her home; and thirdly by whether the child attends his or her nearest, grade-appropriate school.

The thesis provides clear evidence for extensive mobility using all three of these approaches to measurement. Over 25% of children were found to be travelling more than 5km each way to school and back on a daily basis. Almost 60% of children attended a school outside of the Census 2001 Sub-Place (roughly equivalent to a suburb) in which they lived, and fewer than 20% of children attended the grade-appropriate school nearest to their home. Counter to expectations, these figures were fairly stable over time, suggesting that educational mobility does not increase substantially as children age or transition to high school. Mobile children attended significantly more well-resourced and well-performing schools than their non-mobile peers, and the quality of schools attended increased with distance travelled. This substantiates the assumption that children and families make use of educational mobility to improve the quality of education that they are able to access.

The analyses presented in the thesis suggest that two distinct patterns of mobility, with different determinants, are in use in the Johannesburg-Soweto area. The first relates primarily to travel from townships to historically advantaged schools in suburban Johannesburg, and typically requires substantial economic investment and extensive parental involvement. The second form of mobility operates at a more local level, and relates to children and families making choices between a number of relatively local schools. This form of mobility is less resource intensive. Children engaging in the first form of mobility were more likely to attend a particularly advantaged school, and to have a well-educated mother. By contrast, children engaged in the second form of mobility were more likely to live in a disadvantaged area, and come from households with moderate SES levels.

The findings of this thesis provide important insights into the nature of school choice in South Africa, which have implications for educational policy, and the understanding of the nature of urban poverty as experienced by South African children. They also contribute to the international school choice literature, by providing novel information about the implications of relatively unregulated school choice for educational inequality and segregation in the South African context.

Keywords: Birth to Twenty; cohort data; Johannesburg; learner migration; learner mobility; primary school; quantitative analysis; school choice; travel to school; secondary analysis; South Africa; Soweto

Candidate's Declaration

I declare that this thesis is my own unaided work. It is being submitted for the degree of Doctor of Philosophy at the University of the Witwatersrand, Johannesburg. It has not been submitted before for any degree or examination at any other University.

Julia Ruth de Kadt

11th day of May in the year 2011

To my parents, in thanks for their support.

Publications or Presentations Emanating from this Research

De Kadt, J., Richter, L., Fleisch, B. & Norris, S. (2010) Measuring learner mobility in Johannesburg-Soweto, South Africa. 2nd Isibalo Young African Statisticians Conference, Pretoria, South Africa, December 1-3.

De Kadt, J. (2009). Learner Migration and socio-economic status in Johannesburg, South Africa. United Kingdom Forum for International Education and Training (UKFIET) Conference, Oxford, UK, 15-27 September.

Acknowledgements

My sincere thanks to Prof. Brahm Fleisch, for supervising this thesis, and for his ongoing enthusiasm on the topic.

Prof. Linda Richer played an essential role in enabling me to complete this thesis. She provided me with funding for fees, travel to conferences, and a challenging but fulfilling work environment. Despite her frantically busy schedule, she has always found time to read drafts of my work, and to provide guidance, mentorship, encouragement and support.

Prof. Shane Norris facilitated my access to the Birth to Twenty data, and has since tirelessly read drafts of my work. He has been a reliable source of valuable advice and suggestions, and his input has strengthened this thesis immeasurably.

The whole Birth to Twenty team deserves special thanks, not just for providing the data to me, and following up my endless queries, but also for the dedication that has allowed the collection of such complex data over so many years. Similarly, without the commitment of all the cohort members over many years, this rich and valuable dataset could not exist.

The National and Gauteng Departments of Basic Education, and Statistics South Africa, all provided data, without which this thesis would not have been possible.

My family has been a source of strength while working on this thesis. My father, Raphael de Kadt, provided ongoing support and encouragement, despite my stubborn independence. My mother, Elizabeth de Kadt, has somehow always known what I needed from her throughout this process, whether it was warm encouragement, a delicious dinner, or a stern talking to. She has provided support in more ways than I can count. My brothers deserve thanks too: Daniel for help with editing, motivating, and providing a steady supply of cute pictures of kittens and puppies; and Chris for generous gifts, dependable advice, and lots of sushi.

My involvement in IkamvaYouth introduced me to a group of wonderful young people, about the same age as the Birth to Twenty sample members, and made me realise how important addressing the crises in South Africa's school system is. Thank you to the organization, and all its members, for playing an important part in my life, and for understanding when I needed to become less involved. Special thanks to Joy Olivier & Khona Dlamini.

Thanks to Anita Craig, and her husband Denham, for allowing me to escape to their lovely home in Cape Town from time to time, and spoiling me with food, wine and all sorts of ideas.

Colleagues at the Developmental Pathways for Health Research Unit at Wits, from my period at the HSRC, and before that from Princeton, have been a consistent source of encouragement and ideas. Extra special thanks to Dean Janse van Rensburg, for sharing his wizardry with Stata and MS Access, and providing less technical assistance in the form of friendship.

The best group of housemates I've ever had also deserve a mention. Thanks to Tawanda Makusha, Rutendo Jaison and Thabo Letsoalo for making my home a wonderful place to be these past few years, and for always being around to listen, talk, and de-stress. Finally, a special thank you to all my friends, scattered all over the world, who have kept me company in one way or another as I worked on this thesis.

Table of Contents

List of Figures

List of Tables

Glossary

ASS – Annual Schools Survey

Bt20 – Birth to Twenty Cohort Study

DET – Department of Education and Training

DOE – Department of Education

EEA – Employment of Educators Act

EMIS – Education Management Information System

GDE – Gauteng Department of Education

GIS – Geographical Information System

MN – Municipality level

MP – Main place level

NEPA – National Education Policy Act

PCA – Principal Components Analysis

SAL – Small area level

SASA – South Africa Schools Act

SES – Socio-economic status

SGB – School Governing Body

SP – Sub-place level

SRN – School Register of Needs

Chapter 1: Introduction

1.1 Introduction

Much recent work has highlighted the generally poor state of education in South Africa. Schools, particularly those serving the less privileged, tend to be poorly performing, and the skill levels of South African children are notoriously low (Reddy 2006; Fleisch 2008; Spaull 2011; van der Berg, Burger et al. 2011). Simultaneously, the schooling system is also known for the highly variable resource levels enjoyed by different schools within the public sector, and the enormous variations in school performance that tend to accompany this. School quality is closely connected to South African history, with most well performing schools being those that historically educated white[1] children. This means that well performing schools are usually located in historically white areas, and that school quality is closely related to geography (Fiske and Ladd 2004; Woolman and Fleisch 2006; Spaull 2011; van der Berg, Burger et al. 2011).

Within this context, and in the aftermath of Apartheid's Bantu education policies, education is very highly valued by many South Africans. Completing secondary school and reaching tertiary education are core goals held by many young people, regardless of their backgrounds. That there is a strong relationship between the quality of the school attended, and future opportunities, is widely accepted. As a result, children and families are often willing to go to great lengths to ensure the best possible educational opportunities (Woolman and Fleisch 2006; Lombard 2007).

[1] In this thesis, the four race groups defined by the Apartheid-era government will be used to categorize individuals, due to South Africa's unique historical context, and the ongoing relevance of these categories to the life experiences and educational opportunities of young South Africans.

This thesis explores some of the ways in which children and families in post-Apartheid Johannesburg-Soweto pursue the highest-quality educational opportunities possible. In particular, it focuses on school choice, and the associated phenomenon of learner mobility. Learner mobility is used to denote the travel of learners to attend schools other than those closest to their homes. Anecdotal evidence suggests that learner mobility is both widespread and substantial in contemporary South Africa, but very little is actually known about its dimensions, drivers and implications. Although several studies have identified its occurrence, often at fairly substantial levels, and some have even discussed it in relative depth (Fiske & Ladd, 2004; Nelson Mandela Children's Fund, 2005; Paterson & Kruss, 1998; Soudien & Sayed, 2003; Woolman & Fleisch, 2006) there has been only one attempt to construct a broader, data-driven picture (Sekete, Shilubane, & Moila, 2001). Although this study provides valuable baseline data for a critical period of history, it was not conducted on a population based sample, meaning that the applicability of its findings to a broader urban population is not clear. Given the apparently high level of learner mobility in urban South Africa, and its potentially significant implications for the South African educational system, updating and deepening our understanding of this phenomenon is an urgent need.

While we know that learner mobility appears to be extensive, and that due to the distribution of educational opportunities in contemporary South Africa it is likely to have substantial importance to the life chances of young urban South Africans, we know very little about the actual extent of engagement in school choice and learner mobility. Although we have reason to believe that most learner mobility involves travelling to historically advantaged schools, we know very little about where children actually tend to enrol in school, relative to where they live. Our knowledge about the determinants of school choice and learner mobility is also very limited. This thesis uses secondary quantitative analysis of longitudinal data from the Birth to Twenty (Bt20) study, which tracked a sample of 3273 young people born in the highly urbanized

Johannesburg-Soweto area in a six week period in 1990, to address these questions. Establishing the dimensions, patterns and correlates of learner mobility in South Africa, and beginning to understand its implications, will fill significant gaps in the local and international literatures on school choice, and have important implications for the design and implementation of policies to ensure genuinely equitable access to high quality education in South Africa.

1.2 Defining learner mobility

Before detailing the research questions addressed by this thesis, it is critical to clearly define learner mobility. The term “learner mobility” is used to refer to the daily travel of learners to a school that is not the school nearest to their home (Karlsson 2007). It is derived from the phrase “learner migration”, which has previously been used in the South African literature on the travel patterns of learners (Sekete, Shilubane et al. 2001; Lombard 2007). This term, however, has not obtained widespread usage outside of this fairly limited literature. In addition, it does not adequately differentiate between daily travel or commuting, and genuine migration in which learners spend nights away from their permanent family residence for the purpose of attending a specific school. As both of these practices are believed to be widespread in contemporary South Africa, it is essential to distinguish between them, particularly as their implications for individual learners, the educational system and society more broadly, along with appropriate policy responses, are likely to differ substantially. In addition, the learners making use of them are expected to represent different groups, particularly with regard to residential location and family socio-economic status (SES) (Paterson and Kruss 1998). As daily travel is the focus of this dissertation, learner mobility is used as a more appropriate descriptor.

It should be stressed that learner mobility is only one expression of school choice in South Africa. Other, and likely widespread, expressions of choice occur when a family chooses their residence on the basis of proximity to specific schools, when a child leaves the public sector to attend a private school, regardless of location, or when a child lives with people other than his or her family in order to attend a specific school. These types of choice are excluded from the above definition of mobility for various reasons, including that it is not possible to measure them with the data available for this project, and that they can be understood to operate differently from the types of movement described above in terms of the resources required and used, the perceived benefits, the population groups most likely to be taking advantage of them, and their implications for public policy. While all these forms of school choice in South Africa certainly warrant further examination, they do not fall within the scope of this thesis.

This thesis takes several different approaches to the measurement of learner mobility, in order to capture a range of different aspects of the phenomenon. The first approach involves measuring the straight-line distance between a child's home, and his or her school. This is selected because it is theoretically sound, capturing to some extent the level of investment required by mobility, and because it can be accurately measured using the available data. Additionally, and for these same reasons, it has been widely used in existing work on learner mobility, although in some cases distance is supplemented or replaced by travel time (Sekete, Shilubane et al. 2001; South African Human Rights Commission 2004; South African Human Rights Commission 2006; Pendlebury and Rudolph 2008).

The second approach involves determining whether the learner is travelling to a school inside or outside of the area in which he or she lives. Unfortunately, identifying in a consistent and useful way whether a school and a home address are in fact located in the same area is very challenging in contemporary South Africa. Educational districts, as used by the Gauteng Department of Education (GDE) for administrative purposes, do not generally align with electoral wards, census districts, or neighbourhoods. Most schools do not have a clearly defined catchment area, unless it has been defined by the school itself. Electoral wards and census districts have also undergone recent and substantial changes, and in many cases bear limited resemblance to historically defined neighbourhoods. In this context, the most feasible way to measure travel between areas proved to be by locating both school and home within their relevant census districts, using GIS coordinates. Although imperfect, this data does at least provide a preliminary measure of whether a child attends school in the area in which he or she lives.

Finally, the third approach to measuring mobility is to identify whether or not a child is attending the grade-appropriate school nearest to his or her home. This is used primarily as an indicator of whether or not a child and his or her family are engaging in school choice at all. Of course, this measure is imperfect, as a child attending the school closest to his or her home may have chosen this school deliberately, while a child attending a school further from home may do so for completely involuntary reasons. Nonetheless, this measure is used as the best available indicator to provide an approximate measure of the extent of engagement in school choice amongst children in post-Apartheid Johannesburg-Soweto.

It is critical to note that many other definitions of learner mobility are of course possible, and in various contexts may indeed be more appropriate. For example, when parents and learners make school enrolment choices, they certainly take into account a wide range of factors other than distance. Given South Africa’s long and racially-defined history of learner migration and differential educational opportunities, the historical racial designations of schools is one factor that is likely to play a significant role (Paterson and Kruss 1998; Lombard 2007). Factors such as travel time, travel cost, safety, school reputation, and school fees, among others, are also likely to be of particular importance (Chisholm 2004; Maile 2004; Lemon 2005).

The measures of learner mobility used in this thesis, and discussed above, include both a continuous measure (distance from home to school), and binary measures (attending a school in the area in which a child lives, and attending the school nearest to the child's home). This is appealing, as mobility can be thought of as both binary – occurring or not occurring – and as a continuum – with the extent of mobility determined by the distance travelled to school. In addition, combining both binary and continuous measures allows for the application of a broader range of different analytical approaches.

1.3 Core research questions

The literature review, presented in Chapter 2, explores and defines the context in which school choice, including learner mobility, occurs in contemporary urban South Africa. This provides the background against which the study's research questions can be addressed. Current school choice practices in South Africa have emerged in the context of rapid, global, changes in educational systems. Key elements of these changes have included increasing demands for higher levels of education along with universal access to basic education, a focus on improving the quality of education, which is often accompanied by increased levels of testing and ICT use, and pressure to keep public sector spending to a minimum through cost recovery and privatization (Carnoy 1999). During the 1990s there were widespread debates around the extent to which market forces and an increased private role in educational production could allow these otherwise somewhat contradictory goals to be met, and establishing school choice was viewed as a critical component of getting educational markets to work properly (Hoxby 2002; Hoxby 2003). While these debates have now died down, most of the pressures that spurred them have not gone away, and some parts of the response to these pressures, such as school choice, have become entrenched (Greene, Loveless et al. 2010). Unfortunately, this entrenchment has not been combined with evidence that school choice is an effective response to these pressures, nor, indeed that these pressures should be accepted in the first place. Nonetheless, school choice, whether official or unofficial, regulated or not, is now accepted – or at least tolerated – in most countries.

In South Africa, the issue of school choice is complicated further by the country's history of massive racial inequalities. Under the Bantu Education Act, children's educational opportunities, along with the resources devoted to these, were determined entirely by the colour of their skin (Fedderke, de Kadt et al. 2000). While this policy is now a thing of the past, it has left behind a very persistent set of geographically defined inequalities in educational infrastructure and resources (Fiske and Ladd 2004; Fiske and Ladd 2005; Spaull 2011; van der Berg, Burger et al. 2011). As residential desegregation is occurring only slowly, for most black learners the only opportunity to attend a high quality school comes through a willingness and an ability to travel in order to reach one. While government policies do not legislate against school choice, they also do not facilitate choice as an option for the most disadvantaged learners, nor do they protect these learners from any harmful effects of their inability to express choice (Maile 2004; Woolman and Fleisch 2006). The ability to express choice is therefore expected to remain closely linked to SES, even as SES gradually dissociates from race.

School choice, and particularly learner mobility, has become entrenched and broadly accepted in contemporary South Africa, even though relatively little is known about its extent. In this context, it is critical to explore this phenomenon to develop an understanding of its implications, as well as how any negative effects can be mitigated. This thesis documents learner mobility between 1997 and 2003. This period of time is thought to be critical to the development and establishment of new, post-Apartheid patterns of interaction with educational opportunity. The most substantial policy changes thought to have an impact on mobility had already been made by this period.

I will now move through the two major research questions addressed in this thesis, providing critical context for each, and exploring key hypotheses.

1.3.1 What is the extent of learner mobility in South Africa?

Given good reason to believe that learner mobility, as an expression of school choice, exists at significant levels in South Africa, the next task becomes to measure it (Karlsson 2007). Determining the extent of learner mobility in post-Apartheid Johannesburg-Soweto forms the first core research question in this thesis. To date, although there is ample evidence that learner mobility is fairly widespread, there is little concrete information about just what this means. This thesis measures learner mobility in the Johannesburg-Soweto metropolitan area, using the three different definitions of learner mobility described above. The actual dimensions of mobility are both of practical and theoretical importance. Practically, the extent of learner mobility – in each of the forms described – has implications for educational policy and planning. Theoretically, the extent of the phenomenon is important as it feeds into major, ongoing debates, both domestic and international, around the implications of school choice, particularly when unregulated, for both academic and societal outcomes. Although some work has been done to measure the extent of learner mobility and school choice in particular populations (Paterson and Kruss 1998; Sekete, Shilubane et al. 2001; Msila 2005; Karlsson 2007; Lombard 2007), to my knowledge there is no existing work which provides a measurement of the extent of mobility at a population level in contemporary urban South Africa.

1.3.2 What are the patterns of learner mobility in South Africa, particularly with respect to socio-economic determinants?

There appear to be two major groups of determinants of school choice. The first set of determinants relates to the broader context in which school choice occurs, and includes factors like policy, the nature of the schooling system, the distribution of educational opportunities, and so on. This is covered in this study's literature review, and has been fairly comprehensively dealt with by the existing South African school choice literature (Woolman and Fleisch 2006).

The second set of determinants relate to the attributes of individual children, their families, and the communities in which they live, as well as the properties of the schools available to them. International work has documented the importance of these determinants, but very little theoretical work, and even less empirical work, has been done on this for South Africa. As a result, this area is where the bulk of the analytical work conducted in this thesis is focussed.

This second research question addressed by the thesis therefore revolves around identifying who, primarily, is involved in learner mobility, and exploring the patterns evident in this mobility. The evidence that is available suggests that patterns of mobility are highly mediated by race, class, gender, age and geographic location (Paterson and Kruss 1998; Sekete, Shilubane et al. 2001; Fiske and Ladd 2004; Nelson Mandela Children's Fund 2005; Karlsson 2007). The relationship between each of these variables and mobility will be explored in this thesis, as will the relationship between mobility and a range of attributes of the school a child lives nearest to, and the school which that child attends.

As the majority of documented learner mobility seems to involve children living in historically disadvantaged urban African communities (Sekete, Shilubane et al. 2001; Maile 2004), this thesis focuses particularly on the urban Johannesburg-Soweto area. In line with data provided by other work (Sekete, Shilubane et al. 2001; Chisholm 2004; Fiske and Ladd 2004), the most substantial portion of the mobility identified is hypothesised to be learners travelling from townships and informal settlements to schools in historically more advantaged areas. In addition, some evidence of choice and mobility within the more disadvantaged urban areas are also expected, with children choosing to attend schools other than those nearest to their homes, and sometimes those outside of their immediate residential area.

These forms of intra-urban learner mobility are hypothesised to be linked to family race, SES, and maternal education. Of those involved in learner mobility, the most advantaged learners and those with the most educated parents are expected to travel furthest, to attend schools in the most advantaged areas (typically, historically white schools). Somewhat less advantaged learners, and those with somewhat less educated parents, are expected to travel somewhat shorter distances to attend schools in more moderately advantaged areas (typically, historically Coloured or Indian schools). Travel to attend schools in other highly disadvantaged areas (typically, historically Black township schools) is expected to be restricted to the least advantaged among mobile learners.

The thesis also explores changes in the mobility behaviour of children over time. Children are expected to become increasingly mobile as they age, with far higher levels of mobility anticipated at the secondary school level than at the primary school level. The final analytical component of the thesis ties together all of the potential determinants of mobility analysed, and provides preliminary models for the prediction of mobility behaviours in children in the Johannesburg-Soweto area.

1.4 Methodological approach

The methodological approach taken in this thesis is quantitative secondary analysis. This is a widely accepted research method (Bryman 2004), and is well suited to answering the research questions outlined above. The thesis makes use of data on a sub-sample (n=1428) of the Bt20 cohort, which is comprised of all (n=3273) children born in a 6 week window in 1990 in the

Johannesburg-Soweto metropolitan area. In order to allow for the consideration of variables relating to schools and communities in the analysis, and the incorporation of spatial effects, the Bt20 data is combined with schools data from the Department of Education (DOE), and census district data from Census 2001. As detailed previously, a number of approaches to the measurement of mobility are taken. This is followed by a range of bivariate analyses, exploring the relationships between learner mobility and hypothesized determinants of mobility. Change in mobility over time is also examined. Finally, a regression approach is used to combine all relevant variables, and generate preliminary models of learner mobility.

1.5 Rationale

1.5.1 Rationale for a focus on learner mobility

As mentioned previously, learner mobility is only one form of school choice in South Africa. Framing this study of learner mobility as a contribution to the international school choice literature, and as a contribution to understanding school choice and its implications for policy and practice in South Africa, requires answering two questions. Firstly, why only study a part of the phenomenon, and not the entire phenomenon? Secondly, why focus on learner mobility, as opposed to any other expression of school choice?

The argument for addressing only a single form of school choice stems primarily from the complexity of school choice in contemporary South Africa, combined with the current paucity of data on the phenomenon. The review of the international, empirical literature on school choice (presented in Chapter 2) will demonstrate that the implications of school choice depend heavily on both the context of choice, as well as the exact nature of the choice that is available. If multiple forms of school choice are combined in a single study, the study will need to differentiate between these different forms of choice, and their differing implications. Because we currently know so little about school choice in South Africa, any empirical study will need to begin by identifying and measuring each particular form of choice, before it is possible to move on to examine implications. Providing such a thorough and deep examination of multiple forms of choice would be ideal, but is simply beyond what is feasible in this study.

Of the various forms of school choice in evidence in contemporary South Africa, there are several reasons for this study to focus specifically on learner mobility. Firstly, learner mobility appears to be extensive. Although the literature on learner mobility itself is extremely limited, data from studies of various other educational topics suggest that it is likely to be quite prevalent. For example, Karlsson (2007) provides evidence that the single largest group of commuters in contemporary South Africa is school children. Additionally, in a representative survey of Grade 12 learners nationwide in 2001, 24.9% were found to live more than 10km away from their school, a proportion which excluded those at boarding schools (Cosser and du Toit 2002). Even if some of this mobility can be explained by children in rural areas who have poor access to schooling, it still suggests that voluntary mobility in urban areas is likely to be prevalent.

Secondly, learner mobility is of particular theoretical interest, due to its potentially significant role in decreasing the racial segregation of schools, but simultaneously increasing their socio-economic segregation, with substantial implications for educational access and equality. This is because the costs associated with learner mobility tend to be fairly substantial, meaning that it is likely to be shaped primarily by socio-economic status, rather than race. Recent work suggests that although the role of race in determining educational outcomes has been falling in post-Apartheid South Africa, the role of SES has been becoming increasingly important (van der Berg, Burger et al. 2011). Understanding the relationship between SES and school choice is a critical precursor to understanding the nature of the shifting relationship between race, SES and educational outcomes.

Thirdly, learner mobility is in many ways the most measurable form of non-private school choice, and can in fact be measured, at least for samples of the population, using data which already exists. This means that considerable information can be obtained without requiring the time-consuming and expensive collection of specialised data. In addition, learner mobility has implications for a number of areas of educational policy, including school financing and governance. The information generated by this study will allow for a reassessment of key elements of South African educational policy, as well as existing academic work based on the assumption of limited or no learner mobility. It will also highlight the practical questions that need to be asked about South Africa's current approach to school choice, and provide some guidance regarding potential alternatives.

1.5.2 Rationale for a focus on school choice more broadly

As will be illustrated in Chapter 2, and as alluded to above, the South African literature on school choice is fairly limited, and of the work that does exist, only a very small proportion is empirical. This sparse empirical literature is in itself a strong justification for an empirical investigation of the actual extent of school choice in a fairly large urban population in South Africa. Additionally, most existing empirical work tends to be either extremely local or highly aggregated. While each of these types of work is very useful, particularly given the limited state of existing knowledge, this does leave substantial space for work at a more intermediate level, which allows for both the contextualization of data, but also the identification of more generalizable patterns.

The international literature makes very clear that school choice policy and practice are very closely connected to issues of educational quality and educational equality. Understanding the determinants of access to high quality education is of great relevance to national developmental goals of rapid and ongoing economic growth, job creation, racial equality, and socio-economic justice. Particularly in the context of the huge variations in educational quality present in South African public education, it is critical to understand who has access to good education, who does not, and how to best improve access for this second group.

Although there is a fairly broad, and growing, literature on educational quality in South Africa, there is as yet very little exploration of how school quality shapes patterns of choice, and vice versa, or of the distribution of access to high quality educational opportunities. While this project will not provide definitive information about the implications of school choice for educational quality or outcomes, it will provide preliminary information about the relationship between engagement in school choice, and the quality of the educational opportunities a child is likely to be able to access. This will facilitate the development of data-driven hypotheses to be tested by future work. Understanding patterns of learner mobility will also provide useful information around questions of access to quality education, and potential patterns of investment to address quality improvement.

Existing South African work on educational equality has tended to focus on resource and quality differentials across schools (Chisholm 2004; Fiske and Ladd 2004; Kanjee 2007). South Africa's highly unequal educational system has been strongly shaped by the country's history of racial segregation, and the race-based allocation of educational resources (Motala 1995; Fedderke, de Kadt et al. 2000; Motala, Dieltiens et al. 2009). Racial inequalities in the educational sector remain extensive, and rightfully receive a substantial amount of attention, both politically and in the academic literature (Fiske and Ladd 2004; Nkomo, McKinney et al. 2004). However, as socio-economic inequalities within race groups begin to grow (van der Berg, Wood et al. 2002), direct attention to the implications of socio-economic status for education becomes increasingly important. In focussing closely on the correlates of learner mobility, this study will bring to light a great deal of information about the extent to which SES and access to high quality schooling are related, and by extension, whether it is important to begin focussing more attention on segregation across schools on the basis of SES rather than race.

Although the South African literature on educational equality is much broader than that on school choice, once again relatively little of it is empirical, and concerns about the lack of empirical work at an intermediate level also hold. A great proportion of the existing work on equality and desegregation has also focussed on those schools generally thought of as higher quality, or more desirable – that is, generally, former Model C and independent schools. By contrast, schools which are perceived as low quality have been to a large extent ignored, even though they make up a far greater proportion of schools nationally. This project adds valuably to the literature because it will provide reliable information about who actually attends these very numerous, lower quality schools.

In addition to contributing to the local understanding of an important but poorly understood phenomenon, and the local school choice literature, this thesis has also been designed to help fill several important gaps in the international school choice literature. Firstly, it will provide insight into the extent to which school choice can become accepted in developing country context, even without the existence of intentionally pro-school-choice policies. Secondly, it will explore the implications of unplanned and largely unregulated school choice for educational equality and segregation, in a context of extremely high inequality.

1.5.3 Rationale in terms of practical and policy implications

The information generated in this project will assist in the assessment of current educational policies, most of which are based on the assumption that the large majority of children attend local, neighbourhood schools. For example, if mobility levels are substantial, and mobility is related to SES, this means that the school poverty rankings which determine the allocation of governmental funds are likely to be inaccurate, leading to underfunding of some learners, and overfunding of others. Similarly, high levels of learner mobility would imply problems with current community-centred school governance policies, in both the schools which lose learners and those which receive them. For example, if schools in disadvantaged areas are enrolling only the most disadvantaged of the children living in these areas, the managerial and fund-raising capacity of their parent bodies, and by extension their School Governing Bodies (SGBs), will probably be far more limited than would otherwise be the case. Furthermore, if more advantaged children are attending schools relatively far afield, their parents, who might otherwise be eager to play an active role in school governance, will experience greater logistical barriers to engagement in those schools. Additionally, in light of current school fee policies, which require schools to grant fee exemptions to disadvantaged learners, the extent of mobility is likely to have implications for the revenue streams of both more advantaged schools enrolling relatively disadvantaged children from further afield, and more disadvantaged schools enrolling fewer of the local children with any ability to pay fees. Information on the extent to which school choice is currently practiced, and by whom, will also allow for data-driven reflection on the appropriateness of those very policies which currently serve to shape and constrain choice.

At the level of learner outcomes, learner mobility has potential implications for academic performance, drop-out and repetition. Acting directly on learners, mobility can be hypothesised to have either positive implications, due to access to enhanced educational opportunities, or negative implications, due to travel time, sub-optimal resource allocation, and cultural differences. Mobility may also operate on learners indirectly, through its impact on the school environment. A question of particular interest in this regard is the extent to which more advantaged children do actually leave their neighbourhood schools, and what implications this has for the children remaining in those schools. While these questions are not answered explicitly in this thesis, the information generated about mobility will allow for the development of clearer, data-driven hypotheses regarding the likely implications of learner mobility for both individual children, and for South Africa's schooling system more broadly.

1.6 Outline of the thesis

Chapter 1 has provided an overview of the topic of this thesis, along with the specific research questions to be addressed, and preliminary hypotheses regarding anticipated findings. In addition, it has provided a brief rationale for the research project presented in this thesis.

In Chapter 2, the international and South African literature relating to both school choice and learner mobility is reviewed. A particular focus is placed on the literature documenting the relationship between school choice policies and practices, and educational equity outcomes. The argument is made that the study of learner mobility in South Africa can provide valuable information about the ways in which school choice may impact on educational opportunities, specifically in the context of a developing country in which choice is largely unregulated.

Chapter 3 provides an overview of the methodological approach taken in the study. Chapter 4 explores issues around sample representativity, and provides descriptive statistics for the sample. Chapter 5 uses three different approaches to measure learner mobility, and presents data on the extent of this mobility.

In Chapter 6 a range of bivariate analyses are conducted to explore the relationships between mobility behaviours, and variables at the levels of individual children, their families and households, and the communities in which they live. Chapter 7 extends these analyses to document the relationships between mobility behaviours and the attributes of the schools which children attend, as well as the schools closest to their homes. Chapter 8 documents the ways in which mobility behaviours change over time, as children age.

Chapter 9 combines variables considered in the previous three chapters to generate preliminary models for each of the measures of learner mobility that have been discussed. Distance between home and school is modelled using OLS regression, while travel between areas, and attendance at the nearest school are modelled using logistic regression. Finally, chapter 10 provides a brief overview and discussion of the findings of the study, and concludes the thesis.

1.7 Conclusion

This chapter began by providing some brief insight into the context of school choice and learner mobility in contemporary South Africa. It then defined learner mobility, and provided three different measures that will be used to document its extent in post-Apartheid Johannesburg-Soweto. Firstly, the straight-line distance between home and school will be measured. Secondly, Census 2001 district data will be used to determine whether or not a child lives and attends school in the same geographical area. Thirdly, whether or not a child attends the grade-appropriate school nearest to his or her home will be documented. The chapter then moved on to document the core research questions to be addressed in this thesis. These relate to measuring the extent of learner mobility in contemporary urban South Africa, and to identifying the correlates of engagement in learner mobility at the level of the child, the household, the community, and the school. This was followed by a brief description of the methodological approach taken in the chapter, and a rationale for the selection of the topic on both theoretical and practical, policy-relevant grounds. The chapter then concluded by providing an outline and overview of the chapters to follow.

Chapter 2: Literature review

2.1 Introduction

This chapter serves to place the key questions addressed by this thesis in the context of the relevant academic literature, and to outline the methodological and theoretical approaches used in addressing those questions. As learner mobility is a particular form of school choice, this literature comprises the international literature on school choice, and more specifically that part of this literature which explores the relationships between choice, quality, and equality. Due to learner mobility's potential implications for South African educational equality, the literature on South African inequality and segregation in the post-Apartheid era is also relevant. The first section of this chapter outlines various forms of school choice found in internationally, and argues that three key dimensions of choice policies are the extent to which they are officially legislated, the extent to which they include provisions to protect vulnerable groups from potential harm, and the extent to which they include provisions to allow members of vulnerable groups to actively engage in choice. The second section of the chapter outlines both the policy and practice relating to school choice in post-Apartheid South Africa, and makes the argument that the emergence of learner mobility in South Africa can best be understood as the outcome of an unplanned, and largely unregulated, school choice system, with very few protections for vulnerable groups.

Having described and situated South African learner mobility within the context of school choice more broadly, the review then moves on to explore two key debates in the international school choice literature: firstly, the debate about whether or not school choice improves performance in educational systems, and secondly the debate about whether school choice increases inequality and segregation in educational systems. Following this, the literature on educational equity and desegregation in South Africa is reviewed, with a particular focus on the very limited body of work which relates these issues to school choice. This illustrates that by providing data on the relationship between choice, mobility and educational equality, this thesis will make a useful contribution to the scholarly literature both locally and internationally.

The next section of the chapter places this study's methodological approach within the context of the existing scholarly literature, and illustrates the methodological contributions this thesis makes. In the final section, the conceptual framework used as the foundation of the analyses presented in this thesis is described, with reference to how it has been shaped by the literature, both South African and international.

2.2 Major forms of school choice policy

Although there is a broad and varied literature on school choice, it is a highly ideological literature, with much heated debate but fairly limited empirical data. Much of the available work has been commissioned by interested parties, or is driven by individuals with strong ideological positions (for a discussion of this phenomenon with respect to the UK school choice literature, see Gorard and Fitz (2006)). Additionally, the large majority of existing literature, particularly empirical literature, relates to school choice as implemented in the developed world, while work on school choice in the developing world is far sparser. As a result, this overview of the major forms of school choice policy draws primarily on examples from the developed world.

School choice is generally understood to be occurring when families are able to make a decision about which school a child will be enrolled in. Globally, public schooling systems range from those where there is almost complete choice (such as New Zealand and the Netherlands), to those where children are required to attend particular schools, usually on the basis of residential location (such as Cuba, France and Japan). Most systems are located somewhere between these extremes, with a global trend towards increasing levels of choice (Goldhaber and Eide 2002; Plank and Sykes 2003).

There are some forms of school choice which are virtually impossible to control. The most significant of these is when people select their residences on the basis of access to particular schools (Holmes 2002). Another significant expression occurs when people leave public schooling systems to attend private schools, although this is obviously limited to those systems in which private schooling is accepted (Rinne, Kivirauma et al. 2002). The extent to which people can exercise these types of choice depends on their location, their flexibility in terms of location, and, critically, their wealth (Hoxby 1998). Importantly, these forms of choice interact with various different legislated systems of choice.

Forms of choice provided by legislated systems of choice may include provisions for choice between a number of local schools, magnet school programmes which are accessible to any student on the basis of a lottery or academic performance, establishment of charter schools or multiple small schools in particular areas to provide alternatives to traditional schools, or simply unfettered enrolment at any school, regardless of location. Each form of choice will evidently have different implications in terms of equality of access, and consequences for particular population groups (Teske and Schneider 2001; Hoxby 2002).

While some school choice legislation simply provides a legal right to a certain amount of choice between different public schools, others are designed to counter inequalities associated with school choice, particularly when expressed through access to private schools. Most widespread here are various voucher systems, common in the US, but also emerging in other countries such as Chile, which pay a varying proportion of private school fees for less advantaged learners (Goldhaber and Eide 2002; Elacqua, Schneider et al.

2006). In addition, provisions to protect particular groups may also be included in any type of school choice policy. They might, for example, require that schools enrol a certain proportion of students from particular ethnic or economic backgrounds, or might provide free transportation for certain students to the school of their choice (Holmes 2002).

In summary, the key variations in school choice policies, shaping which people, and how many of them, can engage in school choice are the extent to which these policies:

- are planned or legislated, as opposed to unplanned or unofficial;
- include provisions to protect vulnerable groups from potential harm;
- and include provisions to ensure that vulnerable groups have equitable access to choice.

I now explore school choice policy in South Africa, with particular attention to its properties in these regards.

2.3 School choice in South Africa

2.3.1 The context of school choice in South Africa

Although the school choice literature on South Africa is quite limited, particularly with regards to empirical work, there is some strong theoretical literature documenting the ways in which existing policies shape school choice in the country. School choice in South Africa is regulated primarily through the National Education Policy Act (NEPA), the South Africa Schools Act (SASA), and the Employment of Educators Act (EEA) (Pampallis 2003; Maile 2004; Woolman and Fleisch 2006), although the more recent classification of a number of schools as no-fee schools also seems likely to be important (Ahmed and Sayed 2009).

While at first glance South African policy appears to constrain learners to attend neighbourhood schools on the basis of their residence, it combines with national history to provide both motivation and means for parents to choose their children's schools. Huge variations in both empirical and perceived school quality mean that many parents are extremely motivated to ensure that their children attend specific schools. Policies around school financing and the provision of teaching staff mean that schools are motivated to enrol as many children as possible. As public funding is limited, fee-charging schools are particularly eager to enrol large numbers of children who are able to pay fees. Decentralisation of managerial functions to school governing body (SGB) level means that schools do have some control over the design and enforcement of their admissions policies, and by extension over which learners they enrol, although this control is often de facto rather than de jure, and is subject to some legal constraints. In addition, school choice through residential selection continues to operate, as many advantaged schools prioritize the enrolment of local children. The growing independent schooling sector also provides parents with further choice. As a result, the opportunities for school choice in South Africa are substantial, but come at a fairly marked financial cost to parents. There is widespread evidence that significant numbers of parents are none-the-less exercising this choice (Sekete, Shilubane et al. 2001; Maile 2004; Lemon 2005; Nelson Mandela Children's Fund 2005; Johnson 2007; Lam, Ardington et al. 2008).

Although the policies mentioned above have played a central role in shaping the way that school choice has developed in South Africa, it should be noted that, perhaps with the exception of legislation around independent (private) schooling, this role was generally not intended (Woolman and Fleisch 2006) – instead, school choice was a largely accidental outcome of policies developed for other reasons[2]. The system of school choice in South Africa can therefore

[2] Although the fact that there have been no major efforts to reduce levels of choice in the system does suggest that the existence of choice appears to suit the political and social elite.

be understood as one which is generally, although not entirely, unplanned and unofficial. Due to the substantial freedom of choice that seems to exist, as well as the considerable inequality in both South Africa's schooling system and income distribution, the question of whether there are any protections in place for vulnerable groups in the context of large-scale school choice is clearly important. As the major determinant of the ability to exercise choice seems to be the ability to pay higher fees and pay for additional transportation, it is probable that the ability to exercise choice is strongly linked to socio-economic status.

As a result, the major group at risk of an inability to engage in choice, or even at risk of harm by being left in the most poorly performing schoos, are those of lower socio-economic status (Pampallis 2003; Fiske and Ladd 2004). Potential protective policies might therefore include additional support for schools which primarily attract children from disadvantaged contexts, or the provision of incentives to advantaged schools to enrol less advantaged learners. However, given that there are no explicit government policies on school choice, there are also none of these types of provisions to protect or support vulnerable populations. Similarly, with regards to providing vulnerable groups with access to school choice, for example by paying their school fees or providing free transportation, there are also no policies in place[3]. This means that South Africa's school choice policy could be considered as one that is unplanned, unofficial, and unregulated, with few protective measures, while simultaneously allowing quite extensive levels of choice to certain sectors of the population.

[3] Although a school fee exemption policy exists, which exempts disadvantaged learners from the obligation to pay fees, this only applies once children have been granted admission to the school in question. Additionally, implementation is generally acknowledged to be poor. Veriava, F. (2005). Free to learn: A discussion paper on the School Fee Exemption policy. Cape Town, South Africa, Children's Institute, University of Cape Town.

2.3.2 The practice of school choice in South Africa

Due to the limited literature on South African school choice, the description of the practice of school choice in contemporary South Africa presented in this section draws primarily on the theoretical literature, although empirical studies related to school choice are cited where they exist. Currently, school choice in South Africa appears to take four major forms: **residential, private, intra-area,** and **inter-area**, where intra- and inter-area choice corresponds to learner mobility as defined in this thesis.

Residential school choice occurs when parents select homes on the basis of their proximity to particular schools. Exercising residential school choice generally requires a relatively high level of income and parental education. Due to the geographic distribution of good schools in South Africa, with most good schools located in affluent, historically white areas with high property prices, the constraints on parental ability to exercise residential choice are likely to be particularly high. Due to the private nature of residential location decisions, this type of school choice is also extremely difficult to measure.

Private school choice occurs when parents decide to exit the public schooling system altogether, instead sending their child to an independent (private) school. In South Africa, although growing, the independent schooling sector remains relatively small, accommodating just over 3 percent of learners (du Toit 2003; Hofmeyr and Lee 2004; Centre for Development and Enterprise 2010). While increasing numbers of independent schools offer relatively low fees, and access appears to have expanded greatly in recent years, these schools still serve only relatively small numbers of children (Centre for Development and Enterprise 2010). Most high quality independent schools charge high fees, and often select learners on the basis of academic capability. While the sector is increasingly diverse, racially and socio-economically (du Toit 2003; Hofmeyr and Lee 2004), choosing an independent school is still not an option for the large majority of less-advantaged parents, due to the small size of the sector. In addition, finding a space for a disadvantaged child in an independent school is likely to require fairly substantial knowledge and effort on the part of a parent, again making it less of an option for most disadvantaged families.

Intra-area choice occurs when parents are able to choose between a number of schools within their residential area, and make enrolment decisions themselves, on the basis of any particular set of factors. This type of choice is very difficult to measure, as there is no easy way to distinguish between learners who are simply attending the school most accessible to their home, and those who choosing to attend a particular school among those closest to their home for other reasons. In addition, because most residential areas in South Africa remain fairly homogenous, the extent to which this type of migration is likely to matter to socio-economic segregation may be relatively limited. Nonetheless, there is evidence from a small number of studies that parents do distinguish between local schools, and that even within disadvantaged areas, schools with better reputations tend to charge slightly higher fees and attract slightly more advantaged learners (Fiske and Ladd 2004; Msila 2009).

Finally, inter-area choice occurs when parents choose a school outside of the area of their residence. This form of choice appears to be fairly wide-spread in South Africa, with large numbers of learners in various contexts reporting that they attend school relatively far from home (Sekete, Shilubane et al. 2001; Cosser and du Toit 2002; Nelson Mandela Children's Fund 2005). In some cases, particularly in rural areas, this travel may be due to children not having schools close to their homes, rather than choice, but in urban areas this is typically not a concern.

In practice, the line between intra- and inter-area school choice is quite fuzzy, particularly in the South African context where, unlike in most developed countries, there are no consistently defined school catchment areas, and the geography of school districts rarely meshes with that of residential areas. Nonetheless, because inter-area choice generally requires parents to access some additional information, as well as fund additional travel and potentially higher fees, it seems probable that parents accessing inter-area choice are likely to be somewhat more advantaged than those only able to access intra-area choice. On the other hand, inter-area choice is also unlikely to be used by the most advantaged families, as they tend to already be living in the areas with the strongest schools.

This description of school choice practices in contemporary South Africa makes clear that the expression of school choice in the country is extremely complex and multi-layered. In some cases, multiple forms of choice may co-occur, for example with learners travelling substantial distances to attend independent schools, or parents sending a child to a local primary school, and a distant high school. As this example also illustrates, multiple forms of choice may be evident at different times during a single child's education. Each form of school choice is governed by different constraints, particularly with respect to the socio-economic attributes of those who are able to exercise them.

2.4 Debates in the international school choice literature

Two major topics of debate are evident in the international literature around school choice. The first is the relationship between choice and educational quality, and the second is the relationship between choice and educational equality. Although historically these debates have been largely theoretical, studies drawing on data are increasing in number, and these two debates will be discussed with particular reference to this type of empirical work. Nonetheless, the highly ideologically driven nature of much of the literature, and the polarized nature of these debates, does make a clear and unbiased interpretation of this literature challenging. I will explore the debates around equality in some depth, as this is the area to which this thesis contributes most directly, but will first touch briefly on the discussions around the relationship between school choice and educational quality.

2.4.1 School choice and quality

Proponents of school choice argue strongly that choice creates markets, or at the least, quasi-markets, in the educational arena, resulting in competition between schools for students and resources, and by extension, for better academic performance (Coleman 1992; Hoxby 2002; Hoxby 2003). Critics, by contrast, argue that the market in public education is too imperfect to result in the type of competition which would improve performance overall. Instead, they argue, choice will lead to a growing divide between well-resourced schools attracting good students, and poorly resourced schools attracting the weakest learners (Levin 1991; Astin 1992). These students and their schools are often termed "those left behind" in the school choice literature, and studies on how school choice policies affect them are particularly inconclusive (Teske and Schneider 2001). To date, research findings on the quality implications of school choice, both broadly and for specific populations, remain mixed. Much of the variation evident in the quality outcomes of school choice policies relates to differences in policy design, implementation, and evaluation methodologies (Henig 1994; Goldhaber 2000; Teske and Schneider 2001; Goldhaber and Eide 2002). Despite some contemporary claims of consensus around the notion that choice improves quality, an increase in the implementation of choice programmes, and a dramatic falloff in the quantity of research on the topic, there in fact remains very little agreement on this issue (Lubienski, Weitzel et al. 2009).

2.4.2 School choice and equality

The second major area of debate, closely related to the discussions around quality, is around the implications of school choice for educational equality. School choice policies have potential to impact equality of opportunity, through shaping access to particular schools, equality in school performance, equality for people from different backgrounds by way of approaches to diversity, and finally, equality with regards to whether and how parents may exercise choice in the educational realm (Godwin, Kemerer et al. 1998). While all of these aspects of equality are important, particularly in the context of rapid societal change, equality of opportunity is the most salient in contemporary South Africa, and it is here that this review is focussed.

Levin (1991) argues that while school choice may provide private benefits to individuals in the form of access to better schools, this is likely to be accompanied by social harm. Equality of educational opportunity is closely related to student sorting, which has long been recognized as an outcome of any form of school choice (Hoxby 2003). Sorting along racial and socio-economic lines, which may lead to racial or economic segregation, has been a particular focus of the literature. Segregation between schools is a well-documented phenomenon both in South Africa (Chisholm 2004) and internationally, across educational systems with a range of policies towards choice (Coleman 1992). Critics of school choice argue that choice tends to increase segregation through two key mechanisms. Firstly, the ability and willingness to take advantage of school choice varies along with demographic variables; and secondly, the bases on which choices are made also vary with demographic variables (Holmes 2002; Ladd 2003; Denessen, Driessena et al. 2005; Elacqua, Schneider et al. 2006). By contrast, proponents of school choice tend to argue that choice has a positive effect on the equality of educational opportunity, decreasing segregation along lines of race and class by allowing disadvantaged students to escape from badly-performing neighbourhood schools. They acknowledge that segregation along the lines of performance may increase, but argue that this will simply increase incentives for competition between schools, improving outcomes across the system (Hoxby 2003). While some segregation is inevitable, systems must strive to avoid segregation on the basis of race or class, and use school choice as a tool to this end (Coleman 1992). Critics counter that given the clear evidence of a close connection between class, race and academic performance, even segregation purely on the basis of performance will increase segregation on the basis of class and race (Astin 1992).

To date, the evidence as to whether, and how, school choice might influence segregation levels in schools remains mixed (Goldhaber and Eide 2002; Viteritti 2005; Godwin, Leland et al. 2006). Given that the outcomes of school choice depend very strongly on both the context of implementation, as well as the design of the policy, this is hardly surprising (Henig 1994; Hoxby 2003). Different research methods have also played a role in the variable outcomes of research into school choice and segregation. For example, in the area of parental decision-making, the bulk of research has been conducted through surveys and interviews. Studies of actual enrolment patterns, however, have demonstrated that how parents *claim* to make enrolment decisions does not always correspond with how they *actually* make enrolment decisions (Elacqua, Schneider et al. 2006). Similarly, parental answers to questionnaires may be markedly different to their answers during in-depth interviews (Bagley 1996). Data constraints are another problem. For example, many studies use data only from a single point in time, which makes it impossible for them to demonstrate a relationship between particular polices, and changes in segregation levels (Gorard and Fitz 2000). Finally, the deep-seated beliefs around social justice or free market competition held by many involved in this debate has meant that research has not always been entirely objective, making it still more difficult to unravel the genuine implications of school choice policies for inequalities and segregation (Gorard and Fitz 2006).

Nonetheless, it is worth presenting a brief overview of what has been found internationally, to provide an understanding of the current state of knowledge on the issue. In particular, these studies provide insight into how various types of school choice interact with various types of school system structure, to shape outcomes related to segregation. Particularly in the absence of much empirical work on school choice in South Africa, exploring these international variations is critical to understanding how South Africa's unplanned and unregulated choice system is likely to play out in the context of its already extremely segregated and highly varied schooling system.

Empirical investigations of the implications of choice policies for segregation

While perhaps least relevant to South Africa, the evidence from small and fairly homogenous countries, or those with a single, national policy on choice, tends to be the most straightforward. This is particularly so when, as is the case in many European countries, education policy and curricula are nationally defined, and private education is minimal. For example, in 2000 in Sweden, school admissions decisions nationwide were switched from the basis of residential location to academic performance. This resulted in a growth of differentiation between schools in terms of achievement, but also, significantly, substantial growth in segregation on the basis of race, SES and immigration status (Daun 2003; Soderstrom and Uusitalo 2005).

One of the key reasons for these types of outcomes appears to be that parents often take the racial, religious or socio-economic composition of a school's student body into account when making school choices for their children – although they are very rarely willing to admit this. School segregation in the Netherlands, for example, is clearly related to parental decision making (Karsten, Felix et al. 2006). It is not only advantaged parents who take ethnicity and class into account in decision making about schools, however. In the Netherlands, decision making by members of the typically disadvantaged Muslim minority appears to play an important role in shaping segregation (Denessen, Driessena et al. 2005). Likewise, in Germany, the provision of school choice at the primary level has been associated with ethnic segregation due to the different patterns of decision making around schooling exhibited by both German and Turkish families (Kristen 2005).

Even in these comparatively simple contexts interpretations of the data can vary substantially. Scholars concur that the introduction of school choice policies in New Zealand between 1989 and 1993 has had highly variable results for different schools. But while some argue for a clear underlying theme of increasing segregation along racial, and to a lesser extent, socio-economic, lines between schools (Waslander and Thrupp 1995; Fiske and Ladd 2000), others argue that this is not the case, and that segregation has actually decreased markedly (Gorard and Fitz 2006).

The UK's schooling system lies somewhere between most European systems and the American system, both in terms of the centralization of educational decision making and planning, and in terms of the level of choice that has traditionally been available to parents, and therefore in the complexity of the analysis of school choice outcomes. In some ways, more so in terms of policy than resource levels or history, the schooling system in the UK resembles the South African system quite closely. In particular, they both combine extremely local school management and decision making with extremely centralized curriculum planning and policy making. The debate on the implications of changes to school choice policy in the UK is particularly heated (Gorard and Fitz 2006). Some evidence of a national, short-term decrease in socio-economic segregation in response to increased school choice has been presented (Gorard and Fitz 2000; Gorard, Fitz et al. 2001; Bradley and Taylor 2002), although the implications for racial segregation as well as long-term effects are less clear (Bagley 1996; Noden 2000). Other scholars argue, however, that even the evidence for a short-term decrease in socio-economic segregation is not clear, and that segregation has actually increased with the introduction of greater parental choice. In addition, regional variations in the effects of school choice policy appear to have been very high, and aggregate changes in segregation at the national level may in fact have very little to do with choice policy (Gorard and Fitz 2000; Noden 2000). Evidence from the UK also suggests that parents take race and class into account when choosing schools. For example, white parents have been found to avoid schools with large numbers of non-white children, although the evidence for this claim is largely qualitative (Bagley 1996).

The large majority of available research evidence on the implications of school choice comes from the US. While this evidence is particularly mixed, which is to be expected given the country's wide range of choice policies as well as its substantial demographic variations, it is also particularly rich. The first form of explicit choice made available in the US was through magnet school programmes. Originally designed as a tool to combat racial segregation, there is some evidence of their working effectively in this regard. However, there is also evidence that in a number of cases magnet schools have not decreased segregation; in some cases they may have increased it, or added a new dimension, such as SES, to existing racial segregation (Henig 1994; Goldring and Hausman 1999; Saporito 2003).

Another type of choice programme designed to combat segregation, though in this case with a greater focus on socio-economic segregation, are voucher programmes. While voucher programmes were implemented starting in the 1970s, the data on their implications for segregation is extremely limited. It is clear that there are socio-economic and demographic differences between those parents who do and do not participate in voucher programmes, but what these differences mean for segregation levels is not clear. While desegregation in the private schools receiving voucher-bearing students can be expected, there are concerns about those schools 'left behind', and the students they educate (Bridge and Blackman 1978; Capell 1981; Henig 1994; Witte and Thorn 1996; Levin 1998; Goldhaber 1999; Hoxby 2003; Peterson, Howell et al. 2003). With the 2002 Supreme Court decision that voucher programmes including religious schools are constitutional, and the subsequent growth in voucher programmes, it seems likely that clearer information will gradually become available.

Currently, a particularly widespread form of choice in the US is various charter school programmes, which provide parents with a way to opt out of local public schools. Again, those who choose to participate in this choice programme differ from the general population. Participating parents tend to be better educated and somewhat better off. Race is generally found to be a strong determinant of which charter school a student will choose to attend, and there is some evidence that charter schools have had the most appeal to black parents, meaning that charter schools do in fact tend to be fairly highly segregated (Weiher and Tedin 2002; Bifulco and Ladd 2006; Garcia 2008). A final set of choice programmes has been those operating at the intra-district level, allowing parents some choice between the different schools within a particular school district. However, this type of policy can only be effectively implemented in fairly densely populated, urban, areas, and again, outcomes have been mixed, with reports of both increased and decreased segregation (Henig 1994; Godwin, Leland et al. 2006).

Across all forms of school choice, there is clear evidence that American parents do tend to take concerns about racial and socio-economic composition of schools into account, even if they do not admit this explicitly (Schneider, Marschall et al. 1998; Holmes 2002; Schneider and Buckley 2002; Saporito 2003). Furthermore, even when parents from different backgrounds vary in their school-choice preferences in terms of variables other than race or socio-economic composition, the end effect may remain one of segregation or sorting (Schneider, Marschall et al. 1998).

While the US literature is both substantial and diverse, a few clear themes stand out. The first is that in shaping outcomes, details are important, both in policy design and in the context of implementation (Henig 1994; Hoxby 2003;

Greene, Loveless et al. 2010). The second is that even when choice policies are designed to attain a particular outcome such as desegregation, their actual effects are difficult to predict. And thirdly, we do not yet have any definitive answers to questions about what choice means for educational equality.

Studies providing empirical data on school choice in developing countries are relatively few and far between. In terms of racial and socio-economic inequalities, resource levels in the educational system more broadly, and capacity to implement policy, studies from these countries are likely to be particularly relevant to South Africa. The structure of the schooling systems in most of these countries, which generally rely very heavily on private education, does however differ quite substantially from that found in South Africa.

One of the most well-documented cases of developing world school choice is the national voucher plan implemented in Chile in 1980, which spurred a rapid growth in private sector educational provision (Carnoy and McEwan 2003). There is clear evidence that those parents making use of the voucher programme to send their children to the private religious schools which produce the best educational outcomes, tend to be more advantaged, both educationally and economically, than those who do not. In addition, Chilean parents also take the socio-economic composition of a school's student body into account when making school choices, and more advantaged parents are particularly likely to enrol their children in schools with other advantaged children, prioritizing socio-economic composition over academic performance (Carnoy and McEwan 2003; Elacqua, Schneider et al. 2006). However, the extent to which the voucher programme is actually responsible for the high levels of socio-economic segregation in Chilean schools remains debated (Narodowski and Nores 2002).

In China, school choice is beginning to emerge, after a long period of entirely state-controlled schooling, but remains limited primarily to the more advantaged members of society (Tsang 2003). Evidence of increasingly choice-oriented schooling systems in post-Soviet countries is also beginning to emerge, but information on their likely implications for segregation is not yet available (Filer and Munich 2003).

While this review presents a very mixed picture, a few points are clear. Firstly, there is potential for school choice policies to influence racial and socio-economic segregation, in a range of different contexts, and in a number of different ways. Secondly, the exact nature of this influence is highly dependent on the specificities of the context of implementation, as well as on the details of policy design. This review therefore supports the contention that examining the implications of school choice in South Africa for racial and socio-economic segregation is likely to provide useful information, both for those involved in managing and improving the nation's education system, as well as for those interested in understanding more clearly the interactions between context, policy, choice and segregation.

2.5 School choice and equality in South Africa

Although the literature on school choice in South Africa is limited, there is a fairly well developed literature which speaks more broadly to the inequalities inherent in the country's educational system. This literature forms the core of the review presented below, although work with an explicit focus on choice is referred to whenever appropriate. In contemporary South Africa, numerous factors shape the access of individuals to high quality educational opportunities (Soudien, Carrim et al. 2004). For many reasons, race is the most salient of these factors, and has received a great deal of academic attention. There is now clear evidence that historically advantaged South African schools are becoming increasingly heterogeneous, although to varying degrees, with regards to race and language (Sekete, Shilubane et al. 2001; Maile 2004; Sujee 2004; Lemon 2005; Johnson 2007). While concerns remain about the inclusivity of this integration, and very few schools reflect the race distribution of the country or its regions, there has certainly been a marked change since 1994. However, while these changes in the racial composition of schools have been well-documented, much less work has looked at school socio-economic composition, either statically or over time. The information that is available tends to suggest that substantial differences exist across schools in terms of their socio-economic composition (Maile 2004; Chisholm 2005; Lemon 2005; Reschovsky 2006; Motala 2009; Hunter 2010). Given increasing levels of socio-economic inequality in South African society more broadly, it seems likely that socio-economic segregation at the school level has the potential to increase relative to racial segregation.

A major concern about the literature on school choice and equality in South Africa is closely related to this point. While the importance of race should not be understated, its high salience has tended to obscure, to some extent at least, other dimensions of inclusion/exclusion and discrimination. Particularly important in the context of school choice is SES. Given the geographically unequal distribution of good schools in South Africa, the huge variations in public school cost, and the strong relationship between school cost and quality, SES is likely to be strongly related to the quality of education a learner can access. While SES and race remain inextricably linked in South Africa, the primary pathways through which limitations on school choice operate now appear to have shifted away from race, and towards socio-economic status. This makes the paucity of the literature examining the relationship between school choice and socio-economic status particularly interesting, and alarming. Understanding whether school choice is genuinely linked to SES, and how any related negative implications of choice can be mitigated, requires urgent study.

This question is particularly critical to those historically disadvantaged communities where populations are racially fairly homogenous, but increasing socio-economic diversity is becoming evident. In part, this is because of the risk that certain schools and learners will be ‘left behind’ by school choice. While the current system of largely unregulated choice may enable well-performing schools to attract more resources and further improve their performance, it may occur at the cost of schools which are struggling due to historically low resource levels, and learners who do not have the resources to exercise their genuine preferences for schools (Motala 2009). This division of schools and learners is of even more concern in light of the extremely high levels of socio-economic inequality in South Africa.

This thesis makes an important contribution to the South African literature on school choice firstly by documenting how wide spread school choice is in the Johannesburg-Soweto metropolitan area. Additionally, it fills theoretical and empirical gaps by examining a range of potential contributors to school choice, including both race and SES. Finally, by examining SES explicitly, and including both advantaged and disadvantaged individuals, it will shed light on how broadly school choice is available as an option for members of all racial and socioeconomic groups. In the following section, the methodological contributions of thesis are discussed with reference to the existing body of work.

2.6 Methodological approach

As already mentioned, empirical work on school choice in South Africa is extremely limited. With a few exceptions, those empirical studies which do touch on either school choice or on educational inequality tend to belong to one of two extremes. Either, they focus on a small sample, providing deep, rich data, usually but not always qualitative (Soudien 2003; Msila 2005; Msila 2009; Bray, Gooskens et al. 2010; Hunter 2010), or they are quantitative and extremely broad, providing only highly aggregated statistics (van der Berg, Wood et al. 2002; Sujee 2004). Each of these types of study clearly has great value, particularly when knowledge about a phenomenon is limited. However, as knowledge increases, particularly in a country as diverse as South Africa, it becomes important that a more rounded literature develops and provides information at an intermediate level of aggregation, allowing data to become increasingly contextualized. Very little is currently available at this level, and even when it is, it tends to focus on either school choice or on educational inequalities (Sekete, Shilubane et al. 2001). To my knowledge, no studies exist which combine the examination of choice and educational inequality at an intermediate level. This is important gap, as it prevents us from understanding, for example, potentially substantial variations between rural, urban and peri-urban areas, or between different parts of the country at substantially different levels of development.

In light of this gap, there is considerable scope for an empirical analysis of the relationship between school choice and socio-economic status, and particularly one which makes use of a relatively large sample from a clearly specified context. The two core questions addressed by this thesis, relating to the scope and the correlates of learner mobility in Johannesburg-Soweto, South Africa, seek to fill this gap. Although these questions are fundamentally empirical in nature, addressing them requires that this thesis also tackles the methodological and theoretical gaps evident in the literature, and proposes some novel solutions. In addition, as will have become clear from the overview of international literature, this thesis will also enrich the international debate around school choice, and provide particularly valuable information for the discussions about how school choice, inequality and segregation interact.

This study offers three further methodological contributions. Firstly, most existing empirical work around learner mobility, whether quantitative or qualitative in nature, comes from school-based studies (Fiske et al., 2004;

Sekete et al., 2001). While the school-based approach has many advantages, particularly with regards to understanding how levels of mobility impact school performance and functioning, it does also have drawbacks, most notably with respect to understanding population-based levels and patterns of mobility. This is partly because mobility appears to be highly clustered around specific schools. Once a school first begins to enrol learners from outside the local community, particularly when those learners are from a different race group, members of the local community tend to begin to avoid that particular school (Fiske et al., 2004). Focusing on particular schools, rather than on particular populations or communities, may therefore provide either inflated or deflated data, depending on whether the school is one that caters primarily to mobile learners or not.

In order to overcome some of these difficulties, this thesis makes use of population level data. As far as I am able to determine, it is the first study of learner mobility in South Africa to do so. Unfortunately, the type of national data necessary to generate this understanding is not available for South Africa. The best available alternative is to draw on large-scale datasets that sample populations in particular parts of the country, and use this information to draw conclusions at an intermediate level. This can in turn guide future data collection and analysis, toward a more complete understanding of the phenomenon. In addition, the use of data at an intermediate level has advantages of its own, such as allowing for control for geographic and other associated variation. The reasons for the selection of the Birth to Twenty dataset are documented in Chapter 3. Using this type of data, however, provides an additional advantage, which is that changes in mobility can be explored over time. Again, this thesis is the first study of which I am aware which looks at mobility behaviour at more than one time point.

A second methodological innovation is that the project makes use of data from a number of different sources, in order to overcome the limitations imposed by each individual data source. This allows for the simultaneous examination of school-level, child and family-level, and community-level determinants of mobility. Again this is the first empirical South African study of which I am aware that is able to examine this full range of potential determinates of mobility.

Thirdly, this thesis takes a completely novel approach to the measurement of learner mobility. Whereas previous studies have used only one approach to measuring mobility, typically either travel distance or travel time, this study uses three different measures of mobility, each capturing a different aspect of the phenomenon. These are straight-line distance between home and school, whether a child attends a school in the same area in which he or she lives, and whether a child attends the grade-appropriate school nearest to his or her home. This is critical in that it allows for the unpacking of different forms of choice and mobility, as well as their varying determinants, and raises questions as to whether learner mobility in South Africa is a unitary concept. These measures are documented more fully in Chapter 3.

2.7 Conceptual framework

Although the international literature on school choice is fairly large, there is a very limited body of work which actually explores, empirically, the determinants of the choices made by individual learners and their families (Bosetti 2004). Where empirical data does exist, it is often focused on choice of a particular type of school, for example independent or religious, as opposed to the choice of a particular school (Le and Miller 2003; Elacqua 2006). While a range of variables have fairly consistently been found to be important – particularly those relating to social class, as discussed above – there is no clear model which is systematically used to predict engagement in choice. As a result, I draw on work in other areas to develop the conceptual framework used in this thesis. In particular, the framework draws on work on decision-making around mobility more broadly (De Jong 2000), and the literature on the process by which students choose the higher education institutions they apply to (Hanson and Litten 1982). The framework that I have derived, presented in Figure 2.1 below, therefore also serves as a contribution to the theoretical literature around school choice in South Africa.

The framework conceives of school choice as always occurring within a particular historical, geographical and policy context. This is guided by the evidence presented in the literature review above that context is critical in determining the ways in which school choice play out. Examples of relevant aspects of historical context in contemporary South Africa include the enormous variability of school quality, and the ways in which access to resources, including education, are distributed across the population, on the basis of both race and class (Fiske and Ladd 2004; Msila 2005; Msila 2009; Spaull 2011; van der Berg, Burger et al. 2011). Geographical context includes the variable nature of different residential areas, and the ways in which these are physically located, as well as the geographical distribution of schools of differing levels of quality (Hunter 2010; van der Berg, Burger et al. 2011). Finally, examples of the ways in which policy shape school choice include the high levels of variation in the cost of attending different public schools, the ways in which policy allows and constrains school choice, and the extent to which families have access to reliable information about particular schools (Fiske and Ladd 2004; Woolman and Fleisch 2006).

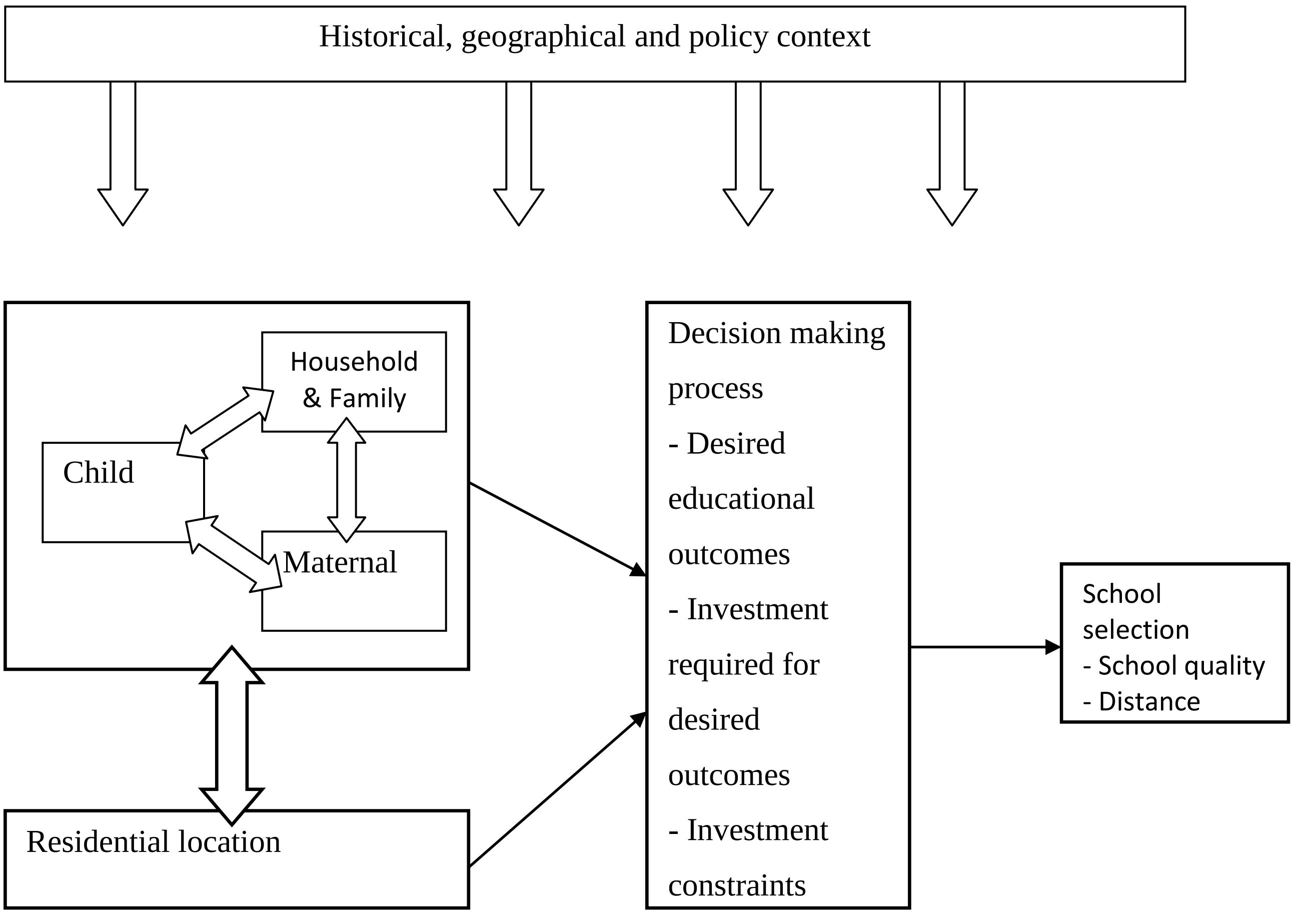

Figure 2.1: Conceptual framework, based on De Jong (2000) and Hanson and Litten (1982)

In this framework, variables at the level of the individual child, the family, and the community in which the family lives, all interact in shaping each other, and simultaneously all feed into the decision making process (De Jong 2000). In this thesis, a range of variables at each of these levels are tested for their relationship with mobility. The selection of variables for testing is guided by the international and local literatures on school choice where available.

Potentially important child-level variables include gender, race, school performance and academic aptitude, psychological adjustment, and child location within the family (for example, number and relative ages and genders of siblings). In this thesis, due to data limitations, race, gender, and a few indicators of academic aptitude are examined. Race is suggested for inclusion by the wide range of international literature suggesting that it is a strong predictor of school choice behaviours (Glazerman 1998; Fiske and Ladd 2004; Fiske and Ladd 2005; Karsten, Felix et al. 2006; West and Hind 2007). Gender is included due to the literature documenting differential parental investment in children's education on the basis of their gender (Alderman and King 1998; Klasen 2002; Unterhalter 2005), although concerns about this in the contemporary South African context are fairly limited. Academic aptitude is also included as there is some reason to believe that it may also shape parental willingness to invest in education, or contribute to school choice in other ways (Glazerman 1998; Zietz and Joshi 2005; West and Hind 2007).

At the household level, variables such as education levels of members of the household, the structural stability of the household, the wealth of the household, and the household's residential stability are considered. Additionally, maternal attributes are likely to be of particular importance to schooling decisions. Relevant variables, very similar to household level variables include the mother's marital status, and the stability of her relationship status, her education level, her income, and her age. In this thesis, the variables tested are maternal education, maternal marital status (as a proxy for household stability), and household SES. Maternal education is included on the basis of evidence for a relationship between maternal education and educational choices made for children (Magnuson 2007; Andrabi, Das et al. 2009; Greenberg). Marital status is included as a proxy for household stability, as there is evidence that both indicators are related to child outcomes and wellbeing (Osborne and McLanahan 2007). Finally, household SES is included due to evidence for a relationship with school choice (Glazerman 1998; Msila 2005; Andrabi, Das et al. 2009; Msila 2009; Hunter 2010).

At the community level, relevant variables might include the quality and attributes of local schools, the affluence and education levels of the community, as well as the coherence of the community itself, and the extent to which it suffers from problems such as crime. This thesis includes a range of measures of school quality, such as poverty quintile rating, school fees charged, matric examination performance, and historical level of advantage, along with the racial composition of the student body and the size of the school, all of which have been demonstrated or hypothesized to relate to school choice in the South African context (Paterson and Kruss 1998; Fiske and Ladd 2004; Fiske and Ladd 2005; Msila 2005; Woolman and Fleisch 2006; Lombard 2007; Msila 2009). The thesis also includes a measure of community poverty, as this may influence the willingness of parents to send children to local schools (Msila 2005; Lombard 2007).

All of these variables at the child, family, community levels are expected to feed into the decision making process, in which children and families weigh their desired outcomes in terms of school level attributes, with the investment required, and the constraints that they face. Required investments are likely to depend on desired outcomes, the geographical location of both the household and the schools considered, and the socio-economic status of the household. Constraints on investment are likely to depend on household access to human capital, social and economic resources, the structure of the household, and the extent to which the household prioritizes educational or other forms of investment. This decision making process is likely to result in the identification of a small group of schools which the child and family considers appropriate and feasible. The school at which a child finally enrols is likely to be shaped by some extent by school-level constraints, such as whether the school still has space available, and also probably by some degree of chance.

2.8 Conclusion

This chapter has provided an overview of school choice practices and policy, both internationally and in South Africa, and has also provided a review of the school choice literature relevant to educational equality and segregation. It has identified a number of gaps in the scholarly literature, which this thesis aims to fill. At the international level, these gaps include the general lack of empirical data relating firstly to the determinants of school choice, and secondly to the implications of school choice for educational segregation and equality, and the absence of an appropriate conceptual framework for the empirical investigation of these issues. At the South African level, gaps include the absence of information about the dimensions of learner mobility, as well as a general shortage of methodological tools with which to explore the issue.

The chapter has also highlighted the ways in which this thesis will make original contributions towards filling these gaps, at methodological, theoretical and empirical levels. Methodologically, contributions will include the use of a longitudinal, population-based dataset, at a level that provides both some generalizability, but also fairly detailed information at the level of the individual, the use of data from a range of different sources to explore potential determinants of mobility at a range of different levels, and finally the use of three different operational definitions of learner mobility to allow the exploration of a range of different dimensions of the phenomenon. Theoretically, the contribution will include a preliminary theoretical model of the determinants of school choice at the child, household, community and school levels. In addition, it will provide data and insight regarding long-standing debates around the relationship between school choice and educational equality and segregation. Finally, empirically, it will provide the first population-level data documenting the scope and dimensions of learner mobility and school choice in post-Apartheid urban South Africa, along with preliminary data on the determinants of these phenomena.

Chapter 3: Methods

3.1 Methodological approach: quantitative secondary data analysis

This project makes use of quantitative, secondary analysis of pre-existing data to explore questions related to the extent and nature of learner mobility in contemporary urban South Africa. Secondary analysis of pre-existing data is a well-accepted research method with a long history of use in educational research as well as the study of mobility (McMillan and Schumacher 2005; Smith 2006; Fleisch and Schindler 2008). As will be described later, taking a quantitative approach to secondary analysis is particularly well suited to answering the questions posed in this thesis.

Secondary data analysis is an approach to research that is based on the analysis, or in some cases the reanalysis, of pre-existing data (Bryman 2004; McMillan and Schumacher 2005). Typically, this data, which may be quantitative or qualitative, and may consist of primary or secondary sources, was originally collected for a particular purpose other than the research project under consideration. Secondary data analysis allows this data to be reused, to answer a different set of research questions. A major strength of this methodology is therefore the ability to make use of pre-existing data, eliminating the need for time-consuming and expensive data collection, and allowing for more time and effort to be dedicated to analysis. Eliminating the need to collect data also allows time and resources for the analysis of a greater volume of data, possibly covering a longer period of time, and greatly improving the breadth and reliability of work. Additionally, it enables research projects to make use of data from multiple sources, increasing the depth of findings, or to explore a particular historical era, generating period-specific conclusions. The volume of data available for analysis is likely to be far greater, and potentially of higher quality, than the data that could be collected during the limited time, and with the limited resources, available for most doctoral research. Quality of data is also less of a concern, as most large, pre-existing datasets have already gone through multiple levels of quality checking.

Quantitative analysis typically refers to the use of statistical approaches to deriving meaning from numerical data. Strengths of quantitative research include its potential for extracting meaningful and non-obvious information from large pools of data, and the efficacy with which it can be used on large sample sizes. Weaknesses include an inherent assumption that the principles of the scientific method apply to human phenomena, the inability to incorporate qualitative contextual information, and a deceptive sense of accuracy generated by the availability of numerical results (Bryman 2004). Its strength in aggregation, which makes it so valuable in providing an overall measure of a phenomenon, does, however, often also result in a substantial loss of individual detail.

Taking a quantitative approach to secondary data analysis provides a research method well-suited to the major empirical questions the project answers, particularly in the context of an extremely limited and almost entirely qualitative pre-existing empirical literature. Measuring the scope of learner mobility in contemporary urban South Africa – the first major empirical task undertaken in this thesis – is essentially a quantitative question, and requires a quantitative approach. While we already have some information about children travelling to particular schools, and about learner mobility within particular, fairly constrained, communities (Sekete, Shilubane et al. 2001; Fiske and Ladd 2004; Msila 2005; Msila 2009), we don't currently have a broader understanding of the scale of this mobility. Qualitative approaches have proved informative in exploring some reasons for learner mobility, as well as documenting the behaviours of individuals, but they cannot give us an overview of overall levels of learner mobility in a major urban area. For this, a quantitative approach to the analysis of data drawn from a fairly large sample is required.

Answering questions about the scale of mobility also requires the use of data collected at a population level. To date, the large majority of research on learner mobility has explored the question either by focusing on particular schools, or by making use of a non-representative sample, typically drawn from a fairly geographically constrained area (Sekete, Shilubane et al. 2001; Fiske and Ladd 2004; Msila 2005; Msila 2009; Hunter 2010). While these approaches provide valuable data, particularly with regards to the causes and implications of the phenomenon, learner mobility appears to be highly clustered around particular schools, and amongst particular groups of people. This means that any sample that is not drawn to be relatively representative of a fairly sizeable and varied population is unlikely to provide an accurate measure of the overall scope of learner mobility. Unfortunately, collecting data on a relatively representative sample of a substantial population, such as that found in major urban hubs, is an extremely complex and time-consuming process, particularly if data is wanted for more than one point in time. Drawing on a dataset that has already been collected offers a way to gain access to a volume of reasonably representative, high-quality data that could not be otherwise be obtained in the context of a PhD project.

A second major empirical question posed by this thesis relates to the patterns and correlates of learner mobility in contemporary urban South Africa, particularly with respect to socio-economic status. In answering this question, using pre-existing data is particularly valuable, as it allows for access to a wider range of variables, often over a wider interval of time, than would be feasible to collect for a single thesis. In particular, using secondary data provides access to data from a range of different time points. It also allows the researcher to tap into data from a range of different sources, and combine these to enable the exploration of dimensions of the phenomenon that might otherwise not be possible. Identifying patterns in learner mobility also requires access to data for a large and reasonably representative sample of individuals. This provides further support for the use of quantitative secondary data analysis for this project.

3.2 Dataset selection

Making use of quantitative secondary data analysis requires access to an appropriate data set. Although it would be ideal to draw data from a nationally representative sample, the type of data needed for this study (specifically data on the school at which children are enrolled) has unfortunately not been collected in the national census, or other nationally representative household surveys such as the October Household Survey, the Labour Force Survey or the Community Survey. Fortunately, South Africa has a number of other studies tracking fairly large numbers of people of varying ages, in different parts of the country, for different lengths of time. Making use of one of these studies was therefore the most feasible way of obtaining the necessary data.

Deciding which of these various datasets would be most appropriate to use in answering the questions this thesis poses was an important step in developing the project. Key considerations were the extent of data focused on school-aged children, and in particular the availability of residential addresses and school enrolment information, the extent to which this data was available longitudinally, and the extent to which the children included in the study were representative of a fairly well-defined and large population. After substantial consideration, the Birth to Twenty (Bt20) study, based in the Soweto-Johannesburg area of South Africa was selected. Appendix A contains details of each of the other datasets considered for use, and documents the reasons that these sources were decided against in the context of this particular project. In summary, however, there were three main reasons for the selection of Bt20. These are explored in more detail below, but relate firstly to the availability of particular variables, secondly to the availability of data for particularly important time points both in South African history and in the educational lives of the sample members, and thirdly to the fact that the data was made accessible to me.

3.2.1 Birth to Twenty

The Birth to Twenty (Bt20) cohort study started in 1989 with pilot studies to test the feasibility of a long-term follow-up study of children's health and wellbeing (Yach, Cameron et al. 1991). Women were enrolled in their second and third trimester of pregnancy through public health facilities and interviewed regarding their health and social history and current circumstances. Singleton children (n=3 273) born between April and June 1990 and resident for at least 6 months in the municipal area of Soweto-Johannesburg were enrolled into the birth cohort and have been followed up 16 times between birth and 20 years of age (Richter, Norris et al. 2004; Richter, Norris et al. 2007). During the last 7 years, young people have been seen twice a year, at the Bt20 offices and at home. Attrition over two decades has been comparatively low (30%), mostly occurring during children's infancy and early childhood, and approximately 2 300 children and their families currently remain in contact with the study (Norris, Richter et al. 2007). The sample is roughly representative of the demographic parameters of South Africa with equal numbers of male and female participants. Assessments across multiple domains have been made of children, families, households, schools and communities during the course of the study, including growth, development, psychological adjustment, physiological functioning, genetics, school performance, and sexual and reproductive health. The third generation, children of Bt20 children, began to be born in 2004. The Bt20 research programme, including all data collection, has received clearance by the Ethics Committee on Human Subjects at the University of the Witwatersrand (M010556). The Federal-Wide Assurance registration number of the Committee is FWA00000715.

Of particular importance to the selection of this dataset for use in this thesis, data on both residential address and on school enrolment was available for each Bt20 participant at multiple time points, providing a uniquely longitudinal record of schooling and residence, and allowing for an exploration of variation in mobility over time. The residential data was particularly promising, as substantial work in cleaning the data had already been conducted for another PhD project (Ginsburg, Norris et al. 2009; Ginsburg, Richter et al. 2010), and GIS coordinates had been collected for most residential addresses at three recent time points. As the Bt20 study was designed as a cohort of children born in the Johannesburg-Soweto metropolitan area, and has achieved fairly limited attrition, this also meant that its use would permit the development of a roughly representative understanding of the extent and nature of learner mobility for a substantial urban area. In addition, in contrast to the Cape Town metropolitan area, the Johannesburg-Soweto area is substantially more similar to the rest of South Africa, and particularly urban South Africa, in terms of the performance of the schooling system, as well as the demographic makeup of the population. Bt20 also had the advantage of providing data on the earlier years of schooling, rather than only focusing on adolescence and later, where much South African research on schooling behaviour tends to be focused. A final important consideration was that Bt20 was able to make all residential and schooling enrolment data available to me for the purposes of this study, on the understanding that no information that could allow for the identification of participants would be made public. Finally, the Principal Investigator of the project, Prof. Linda Richter, and the Project Director, Prof. Shane Norris, both expressed interest in this project, and indicated a willingness to provide intellectual support for the work.

Concerns about making use of the data from Bt20 included the limited amount of previous work making use of the more detailed schooling data collected. While data regarding the school at which each participant was enrolled in each year had been collected, this had not previously been used, and was therefore likely to require substantial cleaning. While substantial work on the residential addresses and residential mobility of the cohort had already been conducted (Ginsburg, Norris et al. 2009), the GIS data had also not previously been worked with. A second concern related to the cohort nature of the database, particularly given the historically unique period of time, marking South Africa's transition to democracy, during which the participants were born and grew up. In a society as rapidly changing as South Africa during this period, it is possible that conclusions derived on the basis of this particular cohort might not apply to children born and attending school slightly later. A final concern relates to the extent to which the study sample is indeed representative of the youth population of the Johannesburg-Soweto metropolitan area. In particular, as will be detailed in the next section, decisions around data collection, combined with different response rates, has led to variable levels of enrolment, and subsequently attrition, for individuals from different racial and socio-economic backgrounds. Minority racial groups are therefore underrepresented in the study, as are both the most advantaged and the most disadvantaged individuals (Norris, Richter et al. 2007; Ginsburg, Norris et al. 2009; Richter, Panday et al. 2009). However, the study does appear to remain representative for the middle-income African population of the area, which is the group of greatest interest for the questions asked in this thesis. These concerns notwithstanding, Bt20 was the most feasible and suitable dataset for use in this thesis, and was therefore selected.

3.3 Ethical considerations

Ethical clearance for this thesis was received from the University of the Witwatersrand. The letter of approval is attached to this thesis as Appendix B. As the project relies only on secondary analysis of existing data, there was no data collection conducted for this project, and there were by extension no ethical issues related to data collection or study instruments for this study. As detailed previously, all data collection for the Bt20 study received clearance by the Ethics Committee on Human Subjects at the University of the Witwatersrand (original Birth to Ten protocol ethics clearance: 24/1/90; extended Birth to Twenty protocol ethics clearance: M01-05-56). Bt20 data is owned by Birth to Twenty, which is located in the Department of Paediatrics in the Faculty of Health at the University of Witwatersrand. Access to the data required for the purposes of this project was provided by Birth to Twenty.

All data was made available on the basis of unique identifying numbers attached to each individual, with names and other identifying information removed. The sole exclusion to the removal of identifying information was with regards to residential addresses, which were essential to the project. All residential address data was stored securely at all times, on a password protected computer. Care was taken to ensure that exact residential addresses of individuals were not disclosed in writing this thesis and related work, and that addresses were only ever presented at the level of suburb. Similarly, graphics are presented at a level of detail which ensures that an address cannot be identified. In all other instances, only aggregated data is presented.

The Bt20 data was supplemented by data from two additional sources. Data from the South Africa National Census 2001 was used to provide information regarding the demographics and socio-economic conditions of the communities in which Bt20 participants lived and schooled. Census 2001 data was used aggregated at the small-area level[4] and higher, and did not contain any information which might identify individuals or households. Data provided by the South African National Department of Education (DOE, renamed the Department of Basic Education in 2009) was also used to supplement Bt20 data. This data consisted both of publicly available data, and data made available specifically for this project. All DOE data was provided at school

[4] Detailed explanations of the various geographic levels used, including the small area level, are provided in the section detailing the Census 2001 data used, below.

level, and was linked to particular schools. Non-public data was made available on the understanding that data would not be reported in such a way as to link it to an identifiable school.

3.4 Overview of data and variables used

In this section, I provide information on the construction of each of the variables used during my analysis, including details of the data sources used in their generation, and any cleaning or modifications to the data that were required. Descriptive statistics for each variable are presented in Chapter 4. As discussed in Chapter 2, variables relating to potential child-level, household and maternal-level and community-level determinants of mobility are included in this thesis. Figure 3.1, below, shows the location of each of the variables considered within the conceptual framework used in this study.

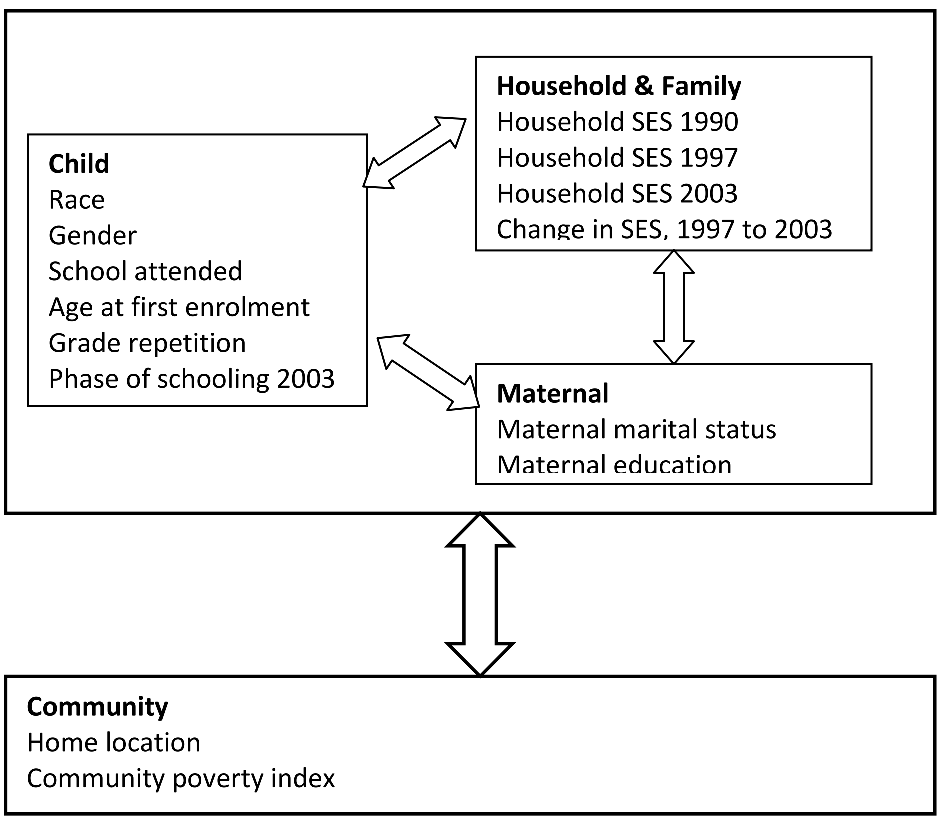

Figure 3.1: Location of study variables within the conceptual framework presented in Chapter 2

3.4.1 Child level variables

Race

All children were recorded as either white, black African, coloured, or Indian, on the basis of information provided by caregivers shortly after birth. These groupings are not meant to denote biological categories, but represent socially meaningful categories in the South African context, as they are the groupings on which Apartheid-era policy was based, and which continue to shape to a significant extent the life experiences and opportunities open to South Africans. This variable has been extensively used in existing analyses, and did not require any cleaning or manipulation.

Gender

All children were recorded as either male or female on the basis of information provided by caregivers after birth. As with race, this variable has been extensively used, and required no cleaning.

School attended

The Bt20 education data required substantial manipulation prior to use. Firstly, as data collection waves did not mirror the academic year, it was necessary to restructure much of this data so that school name and grade could be attached to a particular calendar year, rather than a data collection wave. Secondly, data was available from two different types of source: prospective data, collected during each wave, and retrospective data collected for all calendar years to date during study Year 14. Thirdly, to maximise the value of the school name data, it was necessary to match each school's name to the correct Education Management Information System (EMIS) number, as it is through the EMIS number that information about a school, such as it GIS coordinates, enrolment and resource levels, can be obtained. The school level variables derived on the basis of EMIS numbers are discussed later, in the section on school variables.

This section focuses on the education-related variables obtained from the Bt20 data.

The cleaning and reorganization of Bt20 education data followed a three-step process. Firstly, all prospective schooling data was reorganized by calendar year. Secondly, whenever the prospective record was incomplete, retrospective data, where available, was used to fill in these gaps. Thirdly, using information on the school name, location, and grades, schools were matched to the appropriate EMIS number. Due to the variable spellings used for school names, and the existence of several pairs of schools with the same or similar names, this process had to be done manually. All cases in which there were inconsistencies or a lack of clarity around which school the child attended were checked against the original data collected. When it was not possible to identify definitively the particular school attended, the school was coded as missing.

School attended variables were generated for two points in time, 1997 and 2003. These two points were selected respectively as the earliest point at which all children could be expected to be enrolled in primary school, and the end of primary schooling. Data for 1997 was drawn from the Year 7 round of interviews, combined with retrospective data from Year 14. Over all, for 1997, schools were identified and matched to EMIS numbers for 1241 of the 1428 study sample[5] members. Data for 2003 was drawn from the Year 13 and Year 14 interview rounds, and schools were identified and linked to EMIS numbers for 1311 of the study sample members.

Age at first enrolment in school

Age at first enrolment was calculated using the child's grade in 1997, controlling for repetition. If a child started formal schooling at the earliest

[5] Details for the selection of the 1428 children included in the study sample are provided in a subsequent section of this chapter.

possible point, they would have been in grade 1 in 1996, and therefore, barring repetition or failure, would be in grade 2 in 1997. These children, as well as a few who appeared to start particularly early and were already in grade 3 by 1997, were classed as early-starters. Children in grade 1 in 1997 (with the exception of those who had repeated a grade), as well as a handful not yet enrolled in school by 1997, were classed as late-starters. Note that most of the children classified as late-starters by this variable do actually start school on time with regards to official policy, which only requires children to start school by the year in which they turn seven. However, in order to obtain any variation on this variable, it was necessary to look at the timing of first enrollment within the bounds specified by policy.

Phase of schooling in 2003

Phase of schooling is a binary variable indicating whether a child was still attending a primary school in 2003, or whether he or she had progressed to high school already.

Grade repetition

Grade repetition is a binary indicator, coded 1 if the child repeated any grades between 1997 and 2003, and 0 if the child did not repeat any grades.

3.4.2 Household and maternal level variables

Maternal marital status

The maternal marital status variable is based on self-reported maternal data, and refers to 1990, the year of the child's birth. Mothers selected from the options married; unmarried but living together, single, and divorced/separated/widowed. For the purpose of this analysis, in which marital status was used as a proxy for household stability, all options other than married were combined into a single unmarried category.

Maternal education

The maternal education variable is derived from each cohort member's mother's self-reported highest completed level of education in 1990, at the birth of the child. Mothers selected from the following options: no formal education; up to and including grade 5; grade 6 or 7; grade 8, 9 or 10; grade 11 or 12; and post-school education. When mothers reported post-school education, more detailed information as to the nature of this education was collected. As very few mothers reported no formal education (n=13 in the study sample), this category was merged with the group of mothers completing up to grade 5. As education beyond Grade 5 is used as the cut-off point for determining functional literacy, this group of mothers can be classified as functionally illiterate. Due to the relatively low numbers of mothers who had completed any specific type of post-school education, all forms of post-school education were combined into a single category.

1990, 1997 and 2003 household SES (raw PCA score and quintile)

During each wave of data collection, a varying number of different indicators related to socio-economic status were collected. Using each of these indicators independently is not really feasible. Firstly, they tend to be very highly correlated, and secondly, the value of ownership of a particular asset, or access to a particular service, tends to change substantially over time. For this reason, it is more appropriate to combine those indicators appropriate to a particular point in time into a single SES index score, and use this in analyses. Grouping sample members into quintiles on the basis of their scores at each time point provides a straightforward means for comparison over time.

Due to a combination of changing societal context over time, and the use of different indicators during each data collection wave, it is challenging to construct an SES variable comparable across different study time-points. The best option available within the Bt20 data was to construct an SES score drawing on asset ownership and housing quality data for three different points in time (Filmer and Pritchett 2001; Rutstein and Johnson 2004; Howe, Hargreaves et al. 2008). Data from study years 0, 7, 12, and 13 were used to construct scores for 1990 (around the time of the child's birth), 1997 (around the time of enrolment in primary school), and 2003 (around the time of completion of primary school). See Table 3.1 below for details of the variables used in composing each SES index score.

Time point	1990	1997	2003
Variables used in score creation	All variables collected during pregnancy or the first two years of the child's life (1989-1992): **Home type** (1=house or flat; 0=anything else) **Home ownership** (1=owned; 0=anything else) **Water type** (1=indoor running water; 0=anything else) **Water use** (1=sole use; 0=shared) **Toilet type** (1=indoor flush; 0=anything else) **Toilet use** (1=sole use; 0=shared) **Electricity in the home** (1=available; 0=not available) **TV ownership** (1=owned; 0=not owned) **Car ownership** (1=owned; 0=not owned) **Fridge ownership** (1=owned; 0=not owned) **Washing machine ownership** (1=owned; 0=not owned)	All variables collectd in study year 7 (1997-1998): **Home type** (1=house or flat; 0=anything else) **Home ownership** (1=owned & fully paid; 0=anything else) **Water type** (1=running indoor; 0=other) **Toilet type** (1=indoor flush; 0=other) **Radio ownership** (1=owned; 0=not owned) **Car ownership** (1=owned; 0=not owned) **Washing machine ownership** (1=owned; 0=not owned) **VCR ownership** (1=owned; 0=not owned) **Microwave ownership** (1=owned; 0=not owned)	Variables collected in study years 12 (2002-2003) or 13 (2003-2004): **Home type** (Year 13; 1=house or flat; 0=anything else) **Water type** (Year 13; 1=hot and cold indoor running water; 0=anything else) **Toilet type** (Year 13; 1=indoor flush toilet; 0=anything else) **Electricity in the home** (Year 12; 1=yes; 0=no) **TV ownership** (Year 12; 1=yes; 0=no) **Radio ownership** (Year 12; 1=yes; 0=no) **Motor vehicle ownership** (Year 12; 1=yes; 0=no) **Fridge ownership** (Year 12; 1=yes; 0=no) **Washing machine ownership** (Year 12; 1=yes; 0=no) **Telephone ownership** (Year

			12; 1=yes; 0=no) **VCR ownership** (Year 12; 1=yes; 0=no) **Microwave ownership** (Year 12; 1=yes; 0=no) **MNet ownership** (Year 12; 1=yes; 0=no) **Satellite TV ownership** (Year 12; 1=yes; 0=no) **Cellphone ownership** (Year 12; 1=yes; 0=no)

Table 3.1: Variables used in the creation of SES scores

After the relevant variables for each time point had been identified, all non-binary variables were recoded to become binary variables (see Table 3.1 for the final coding scheme used). A process of manual imputation was then conducted for the 1990 and 2003 data, to reduce the number of missing values. Imputation was done on the principle that if one variable predicted the value of another correctly for 75% or more of the cases without missing data, it could be used to impute the other variable where it was missing. As data for 1997 was either uniformly present, or uniformly missing for all variables, it was not possible to impute any values in this year.

Once the amount of missing data had been minimized as far as possible, principal components analysis (PCA) was run for each year, to determine the appropriate weighting for each component variable (Vyas and Kumaranayake 2006). The literature is somewhat divided as to whether it is appropriate to use PCA on binary data, as has been done here. While some studies have shown it to perform as well as, or better than, alternatives (Filmer and Pritchett 2001; Howe, Hargreaves et al. 2008), others have found that its performance is suboptimal (Kolenikov and Angeles 2008). Nonetheless, it remains the accepted standard approach used for working with binary SES data (Filmer and

Pritchett 2001; Rutstein and Johnson 2004; Vyas and Kumaranayake 2006), and as such, has been used here. Depending on the nature of analysis for which the SES data is used in this thesis, either the raw scores generated by the PCA process are used, or the appropriate sample is divided into quintiles, which are then used.

Change in household SES from 1997 to 2003

An additional variables, change in SES over time, was constructed for analyses exploring changes in mobility behaviour over time. This was constructed by taking the sample member's SES quintile in 2003, and subtracting the sample member's SES quintile in 1997.

3.4.3 Community level variables

Home location

Address information was collected for all cohort members during each wave of data collection, as this information was critical to maintaining contact. GIS coordinates for home addresses, however, were not captured until Year 13 of the study, when home visits were initiated. During Year 13, 15 and 16 home visits, data collectors stood outside of the participant's home, and used a mobile GPS device to record the coordinates at that location. These coordinates were then either manually captured on the home visit instrument, or downloaded into a study database at a later point. Unfortunately, while working with this data, it became evident that between a third and a half of the coordinates captured during each data collection wave were incorrect. These problems were traced to data collectors not resetting the GPS device correctly between uses, and therefore not always collecting the correct coordinates for each location. Unfortunately this meant that the GIS coordinates available could not be used without being checked manually for accuracy. Additionally, due to limited available street-name and number data for the Soweto area, and the vague nature of many addresses provided, it was not possible to generate accurate coordinates for the majority of addresses using software or maps.

Due to time constraints, a decision was made to pursue the correct GIS coordinates only for those children who had not moved home between 1996 and 2004, as this meant that only one set of coordinates would be required per child. Details of sample generation, and implications of the decision to limit the study sample on the basis of residential mobility are discussed in the section on sample selection and bias below. Of the initial BTT cohort of 3273, 66% (n=2158) completed a residential history questionnaire in 2005 or 2006. These individuals comprise the cumulative non-attrition cases. Data from the questionnaire was used to generate a longitudinal dataset containing all address and residential movement information for these sample members, from birth through to 2004 (Ginsburg, Norris et al. 2009). This dataset was used to distinguish between those individuals who had and had not changed residence during the period 1996-2004. 1470 individuals reported stable residential addresses during this period, and these formed the basis for the study sample. Very few of these children had moved between 2004 and the end of Year 16 data collection, meaning that 3 different sets of GIS data, corresponding to three different study waves, were available in most cases.

The residential GIS coordinates for these remaining children were manually checked for accuracy using Google Earth, and fortunately, for the majority of participants, at least one of the sets of coordinates collected did appear to be correct, and could be used. For the remaining children with traceable addresses, GIS data was generated using Google Earth where possible, or otherwise re-collected through an additional visit to the address. Reliable GIS coordinates for residential addresses could not be obtained for 27. These instances of missing information were largely due to street name changes, house numbering systems, and redevelopment of areas, and were distributed throughout the greater Soweto area.

Once residential GIS coordinates had been obtained for sample members, a number of additional community-level variables were generated for each set of coordinates using data from Census 2001 – the most recent national census, and that most relevant to the time period under consideration. This data was used for two purposes. Firstly, Census 2001 geographical area boundary data was used to delineate the boundaries of residential areas, suburbs and municipalities. These boundaries were then used to identify the area in which each child lived. This allowed for the later examination of whether children were living in and schooling in the same areas, or not. Secondly, I also made use of Census 2001 data to provide contextual information, in the form of a poverty index, for the areas in which learners live, and the areas in which they attend school.

Census geographical area delineation

Census geography was used to define area boundaries as census geography generally – although not always – corresponds fairly closely to local perceptions of area boundaries. They are sensitive to the ways in which socio-economic factors and history, along with geographical features, have shaped perceptions of areas. In addition, the Census 2001 provides 4 different levels of geography, which allow for the exploration of mobility and context at a range of different levels.

The smallest level used here is the Small Area Level (SAL). Each SAL typically corresponds to between one and three enumerator areas, and contains approximately 200 households. This level of geography is the lowest level at which census data is released. Given the relatively small size of most SALs, most ‘areas’, as typically perceived, tend to contain several of them. The next level of geography is Sub-Place name level (SP). This level typically corresponds with residential suburbs, or small but distinct areas of a city. Examples might be Pimville or Diepsloot. The third level, Main-Place name

(MP) corresponds roughly to small cities or towns, or large but distinct areas within a large city, for example Soweto. Finally, the largest level of geography used in these analyses is the Municipal level (MN), which represents entire municipalities or districts. Each SAL is entirely contained within a particular SP, each SP within a particular MP, and each MP within a particular MN.

The GIS coordinates home address for each sample member was linked to the SAL, SP, MP and MN in which the address falls, using gvSIG software. Similarly, the GIS coordinates for the school attended by each child, in both 1997 and 2003, were also linked to the SAL, SP, MP and MN within which they fall. Once these linkages have been made, it is possible to test whether children live and attend school in the same area or not. It is also possible to use these linkages to obtain information about the area in which the home or school is situated, as is discussed in the next section.

Community poverty rating

To explore the nature of the areas in which the BT20 families live, small area level data from Census 2001 was used. For each area of geography (SAL, SP, MP and MN), a PCA was conducted using a range of variables related to affluence, and obtained from the Census 2001 SAL dataset. Variables used were the percent of the working age population employed; average household income; the percent of households living in informal dwellings; percent of adults who had no secondary schooling; the percent of the area's population who were black Africans; and the percent of households who did not have access to services such as running water, electricity, hygienic toilets, refuse removals and landline telephones. The results of the PCA were used to generate a poverty level score for each area. This score has been used in its raw form, and has also been used to divide sample members into quintiles on the basis of the poverty of the area in which they live or attend school, where appropriate.

3.4.4 School variables

In addition to the child, family and community-level variables above, which are hypothesised to shape and constrain decision making around school choices, attributes of potential and selected schools are also expected to play a role in shaping decisions. In addition, data on the properties of the school a child attends also provides valuable information on the quality of education that child is likely to receive. As a result, the variables documented above are supplemented with additional data on Gauteng province schools, obtained from the South African National Department of Basic Education (formerly the Department of Education). The bulk of the data comes from the Educational Management Information System (EMIS) 2008, 2009 and master schools lists for Gauteng, and the 2002 and 2003 Annual Schools Survey (ASS). This data includes the GIS coordinates of schools, historical classification of schools, school poverty quintile rating, enrolment data, and data on school fees. Although there are some important concerns about the quality of data coming from the Department of Education, particularly given that most data is self-reported by schools, it is by far the best available source of school information. This self-reported school-level data is supplemented with matric pass rates for secondary schools in 2002, which were obtained from the Department of Basic Education.

Although this thesis examines schooling patterns from 1997, the schools data used is for 2002 or later. This is for a few different reasons. Firstly, data available for the post-Apartheid period prior to 2002 is primarily through the School Register of Needs (SRN), conducted in 1996 and 2000. Despite several requests, I was unable to obtain a copy of the school-level data collected in the SRN. Secondly, even had access been possible, issues of compatibility between SRN and ASS data would have limited the utility of earlier data (Yamauchi 2004). Finally, there is reasonable evidence to believe that the variables of primary interest in this thesis are unlikely to have changed substantially for individual schools over the period under consideration. These reasons, along with the data source used, are discussed in the following paragraphs. Data sources are also summarized in Table 3.2 below.

Variable	Source	Year
School location (street address & GIS coordinates)	Gauteng Schools Master List	2008 (The list includes information on schools operating in 2002, but which subsequently closed)
School sector (Public or private)	Gauteng Schools Master List	2008
Section 21 Status	Gauteng Schools Master List	2008
Phase (Primary, Secondary or Combined)	Gauteng Schools Master List	2008
School resource levels: School fees charged	Annual Schools Survey	2002
School resource levels: Poverty Quintile rating	Gauteng Schools List	2008
Historical Department of Education	Gauteng Schools List	2008
School enrolment	Annual Schools Survey	2002
Racial composition of the student body	Annual Schools Survey	2002
Matric pass rates by school	Department of Basic Education	2002

Table 3.2: Sources for school-level variables used

School location, school sector, school phase, and Section 21 status

School location, school sector (whether the school is public or private), school phase (primary, secondary or combined) and Section 21 status (whether or not the school is allowed to manage its own finances) are all typically fairly stable, and the data for these variables was obtained from the Gauteng schools master list for 2008. Using this data source meant that the variables were available not just for schools operating in 2008, but also for those schools which had closed previously. Although it would have been ideal to obtain these variables, with the exception of school location, from an earlier source, limited access to data meant that this was not possible. Given the stability of school locations, the 2008 data was given preference over earlier sources due to the generally increasing accuracy of GIS data over time.

School quintile rating

Data on school resource levels proved somewhat complicated to obtain. The resource variable which is clearest and most widely used for South African schools is probably the school quintile rating. The quintile system, which ranks schools in categories from 1-5, depending on their resource levels, was first instituted in 1998 by the Department of Education's Norms and Standards for School Funding. At this point, schools were assigned to quintiles on a provincial basis. This was amended in 2006, after which schools were assigned to quintiles at the national level. This meant that many of the poorer school in the wealthier provinces were moved to higher quintiles at this point (Pampallis 2008). In using school quintiles as a measure of resource levels in 2002, it was necessary to decide between the old, provincial quintiles, and the newer, national quintiles. Unfortunately it did not prove possible to obtain the quintile ratings under the older provincial system, and so the ratings used here are those for 2008, and therefore developed under the national system. Although this is not ideal, it should be noted that despite changes in the quintile system, the relative resource levels and of schools with respect to each other, and related to this, their performance, has remained relatively constant (Fiske and Ladd 2004; Fiske and Ladd 2005; Fleisch 2008). That is, those schools which were most advantaged in 2002 are, more or less, the same schools that were most advantaged in 2008. Therefore, although the actual quintiles assigned to schools may have changed somewhat between 2002 and 2008, the rating of schools relative to each other is unlikely to have changed substantially.

Using the 2008 poverty quintile ratings for the 1997 time point, when cohort members were entering primary school is even more problematic than its use for the 2003 time point. Between 1996 and 2000, substantial infrastructural investments were made in historically under-resourced schools, particularly in upgrading the basic infrastructure and facilities at schools, for example, sanitation, telecommunications, water and power (Department of Education 2000). However, most of this investment was focussed on rural areas, and therefore was of less relevance with regards to schools in the Gauteng province. More importantly, the redistribution of resources occurred largely between provinces. By contrast, within provinces, even if the overall levels of resources changed, the relative proportions going to schools which historically served different race group remained fairly similar (Fiske and Ladd 2005). Historically advantaged schools continued to receive higher levels of state funding than historically disadvantaged schools, and this discrepancy was further exacerbated by the ability of parents at historically advantaged schools to supplement state funding through higher school fees (Fiske and Ladd 2005). Overall, then, those schools which were most advantaged in 2002 were also those that were most advantaged 1997, while those with the fewest resources in 2002 were those which had always received the fewest resources.

School fees

Due to the age of the data, as well as various concerns that have been raised about the quintile system more broadly (Kanjee and Chudgar 2009; Kanjee and Chudgar 2009), this thesis also makes use of two additional measures of school resources. Firstly, the school fees charged by schools in 2002, as reported by schools in the ASS 2002, is used as an indicator of the school's access to resources. It is widely accepted that schools with higher resource levels charge higher fees (Fiske and Ladd 2004; Pampallis 2008). This data is also valuable in that it dates to the time at which the children in this study were actually attending the schools in question.

Historical department of education

Secondly, as much of the performance of South African schools today continues to be explained by the educational department under which they fell during the Apartheid era, a variable indicating whether the school was operated by the Department of Education and Training (DET) during the Apartheid era was used. The DET was responsible for the running of urban schools serving black children, and DET schools received far fewer resources than all other urban schools, and typically continue to be under-resourced and poorly performing to the present time (Fiske and Ladd 2004; Fiske and Ladd 2005; Fleisch 2008). This variable was obtained from the 2008 schools master list. Not all currently disadvantaged schools operated under the DET during Apartheid, and any schools which have been opened since the end of Apartheid will obviously not fall under historical DET status. Nonetheless, given the large number of education departments in place during the apartheid era, the binary variable used here to indicate the historical status of the school is the most feasible available option, and contains valuable information about the history of a school. As it is a historical variable, the use of the 2008 data in its composition is not problematic.

School enrollment

The next group of variables used related to school size and the composition of the student body. For school size, the enrolment reported by the school in the ASS 2002 was used. As this data is self-reported by schools, concerns have been raised that the figures may be inflated, particularly for less well-managed schools. However, this is the best available data, and as such, is used here. With regards to figures around school size, Fiske and Ladd (2005) and The Department of Education (2000) note some fluctuations in enrolment levels, but these do not appear to be very substantial. Importantly, school choice had already been possible for a number of years by 1997, suggesting that any initial surge in changing school enrolments post-Apartheid had probably already largely stabilised. Nonetheless, the data for 2002 was the earliest that could be obtained, and is therefore used as a proxy for 1997.

Racial composition of student body

The variable used to describe the racial composition of the student body of the school was also derived from ASS 2002 data. For this variable, the reported number of black African students enrolled at a school was divided by the schools total enrolment, to obtain the proportion of the student body who were black African. Again, the use of 2002 data for the 1997 time point raises some concerns, as discussed above.

Matric pass rate

The final variable examined here is the Matric pass rate – the proportion of a school's students who write the national school-leaving examinations (the Matric examinations) and pass – which is used here as a proxy for school performance. This data was obtained for 2002 from the Gauteng Department of Education, and is available for the majority of public secondary schools in Gauteng. Private schools can choose between writing independent matric exams, or writing those that public sector schools write. Results are only available here for those schools which chose to write the public sector exams. It should be noted that well performing private schools typically choose the independent option, and so the results presented here for private schools are likely to be extremely biased. Unfortunately, no measure of the academic performance of primary schools is currently available for South Africa. In order to generate a proxy of the likely performance of primary schools in Gauteng, each primary school was matched to its nearest secondary school using gvSIG software. The primary school was then provided with the matric pass rate of its nearest secondary school[6]. Obviously this method is highly imperfect[7], but given the geographical clustering of school performance in

[6] Thank you to my examiners for suggesting that a three-year average of matric pass rate be used in future work, as this will help to counter concerns about the high annual variability in performance of many historically disadvantaged secondary schools.

[7] Particular concerns about this measure relate to the fact that within any given context, there will always be a few schools which perform particularly well or particularly poorly. This approach has no way of differentiating these schools from those performing at more expected levels, which means that proxy pass rates generated for some schools will be inaccurate.

South Africa, it is at least plausible, and given the available data it is the only real option.

School name and EMIS number

Once all variables to be used in the analyses had been sourced, the next challenge was to create a list of the schools operating in the Gauteng province during the 1997-2003 period, and combine the variables with these schools appropriately. The best available base list of schools obtained was the 2008 Gauteng master schools list. This includes all registered schools in the province, including those which had previously been registered, and had subsequently closed. This list was therefore used as the base for the list of schools used in analyses. All special schools, as well as non-school institutions (FET colleges, pre-primary schools, exam centres, administrative offices, and so on) were dropped from this list, leaving 2604 institutions, both open and closed. All other variables were then merged with this list.

Aggregated school quality variables

Due to the strong correlations identified between many of the school related variables, and particularly those relating to school quality, it was necessary to combine these variables into an index for use in model generation. This was done by the use of PCA on all school attribute variables that were consistently found to be significantly related to mobility (school quintile, school fees, school enrolment, percent black learners, school sector, historical DET status, and pass rate). This process was repeated for the school attended by each child in 1997 and 2003, as well as for the nearest grade-appropriate school to the child's home in 1997 and 2003. In all cases, the eigenvalues of the first two components of the PCA were both greater than 1, and were therefore both retained.

3.5 Sample selection and creation of the analytical database

As indicated previously, when conducting secondary data analysis, it is critical to develop a clear understanding of the datasets being used, and the way that the methodological decisions involved in creating the data may impact on data analysis and findings. In this section, I begin by providing an overview of how the study sample was composed, and detail the construction of the study's analytical database.

A major concern with regards to the study sample is the extent to which it is representative firstly of the Bt20 cohort in general, and secondly of the youth population of urban Johannesburg-Soweto more broadly. These issues are explored extensively in Chapter 4, which includes a review of the way in which the Bt20 study was designed and conducted, and how this has influenced sample composition. It also explores how the study population, participant attrition, and composition of the sub-sample used in this thesis are likely to have impacted on the outcomes of the analysis, and on the extent to which findings are likely to be representative of the population of contemporary Johannesburg-Soweto youth. In the current chapter, however, the focus remains on the way in which the study sample and the analytical databases were constructed.

3.5.1 Sample selection

As discussed previously, due to unanticipated challenges in preparing the GIS data for use, the study sample was limited to the 1470 cohort members who did not change address between 1996 and 2004. Subsequently, 28 additional children who were either not attending school in 2003, were enrolled in a special school at any time from 1996 to 2003, were enrolled in a school outside of the Gauteng province, or were known to be boarding at a school within the Gauteng province, were also excluded from the study sample. Children with special educational needs were excluded as they and their families are unlikely to experience the same degree, if any, of choice around where they will be educated, due to their particular educational needs. Children attending boarding schools or extremely distant schools were excluded, as they are not travelling on a daily basis. An additional 14 cases of children who had changed address between 1996 and 2004, or who were spending substantial time a different address from their home, we also identified and removed from the sample. This left a sample of 1428 individuals, which formed the sample on which all subsequent analysis for this thesis was conducted. Members of all race groups were retained in the study sample. However, due the very small numbers of Indian and white participants, no results are presented for these groups. See Figure 3.2, below, for a flow chart illustrating this process of sample composition.

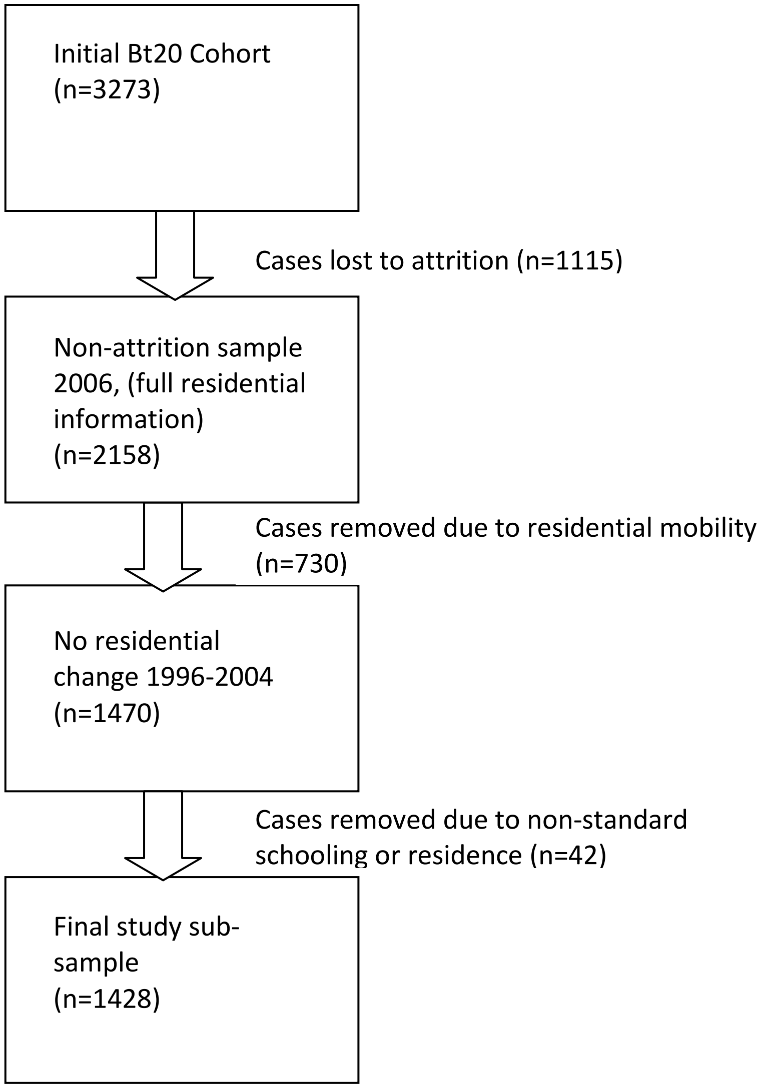

Figure 3.2: Flow chart illustrating selection of sub-sample for use in thesis.

3.5.2 Creation of the analytical dataset

Selection of time points

1997 was selected as the initial year within the analysis as it was the earliest point for which reliable schooling information was available for the majority of cohort members. It also reflected the first point in time by which all cohort members, then aged 6-7, could safely be expected to be enrolled in formal education. 2003 was selected as the point closest to the end of primary schooling for the majority of the sample. In addition, it was selected as the year for which the most reliable address data was available and during which GIS coordinates were first collected, but this became irrelevant once the problems with the GIS coordinates had been discovered.

Selection of these two time points, 1997 and 2003, allows this thesis to explore both ends of the primary school experience, as well as to address questions about both changes and consistency in schooling experience over time. One drawback of the use of these time points, however, is that by 2003, roughly a third of the sample had already completed primary school, and are enrolled in secondary schooling. This progression to secondary schooling is strongly related to socio-economic status, age at initial school enrolment, and academic performance. As these variables are expected to relate to learner mobility, and as mobility behaviour is also expected to differ between primary and secondary school children, this makes it particularly challenging to draw definitive conclusions about the cause of changing mobility at the end of primary schooling. This concern is discussed where appropriate in the results chapters.

Details of the 1997 dataset

Of all the schooling data encountered during this project, that for 1997 was, expectedly, the most problematic, for a few different reasons. The prospective data was not as reliable as in subsequent years, as in 1997 the children were still too young to report their own school names. School names were therefore generally provided by caregivers, who were not always aware of which school the child attended, or knew of the school only by an informal name. Secondly, data capturing for 1997 schooling data was only done in 2009, and was done using a drop down menu filled with school names as of 2009. This induced a number of data capture errors, particularly when several schools had similar names, as the data capturers would simply choose the first name that appeared to match on the drop-down menu. It also introduced problems when schools had closed or changed their names between 1997 and 2009, as the school name would not appear on the drop down list. In these cases, data capturers were also inclined to simply choose the most similar name on the drop down menu – which would be incorrect. For this reason, all cases in which the school name provided for the child was similar to other contemporary or historical primary schools in the Gauteng area were double checked, and corrections made whenever possible. However, it is quite probable that there do remain a small number of incorrect school attributions in the data. The major reasons for concern about the quality of the retrospective data relates to the substantial length of time between 1997, and 2005-2006 when the retrospective schooling data was collected, and the fact that if children had changed schools frequently, they may have struggled to remember accurately which school they attended in 1997, when they were only 6 to 7 years old.

As a result, two different home and schooling datasets were created for 1997. The first dataset was created by combining the 1997 address and GIS data with the prospective schooling data for that year. While this schooling data was highly accurate, it was only available for 760 of the sample individuals, just over 50%. Given this extremely high level of missing information, a second dataset was constructed using the less reliable, but more comprehensive retrospective data to fill in as many gaps as possible. This second dataset contained schooling information for 1244 sample members, with EMIS data missing for only 184 individuals. Given the complementary nature of these two datasets, with one being far more complete, but the data in the other being far more reliable, initial analyses were conducted using both datasets. As both datasets provided similar results, the more comprehensive one was used for the majority of analysis presented in this thesis.

Details of the 2003 dataset

For the same reasons as described above, two home and schooling datasets were also constructed for 2003. The dataset using only the more reliable prospective data contained information for only 760 cases, again just over 50% of the sample. When the retrospective data was integrated, schooling information became available for 1310 sample members, with data missing for only 118 individuals. Given that the retrospective schooling data was collected in 2005-2006, the relatively lower number of cases with missing data is as expected. The accuracy of the retrospective data is also expected to be greater than for 1997, more closely resembling the prospective data collected in 2003. Again, analyses using both sets of data provided similar results, and the more comprehensive data including retrospective information is therefore used.

3.6 Operationalization of learner mobility

Given the poorly developed state of theory around the concept of learner mobility, particularly within the South African context, as discussed in Chapters 1 and 2, I make use of data from a range of different sources to explore a number of alternative ways in which learner mobility might be operationalized. I begin by using GIS coordinates to calculate the straight line distance between each participant's home and school. I then explore definitions shaped by movement between different areas as defined by census geography. Finally I explore whether a child attends his or her nearest grade-appropriate public school.

3.6.1 A distance-based definition of mobility

As discussed previously, a number of distance based definitions of what constitutes learner migration or mobility have been advanced. Typically, these work on the principle of assessing what constitutes excessive travel, on the basis of assumptions around the age of the child, the safety of the area, the availability of safe and affordable public transportation, and possibly other context-dependent concerns. Internationally, the literature suggests a range of maximum distances, ranging up to 10km. In the South African context, the maximum distance that a child should need to travel has tended to be fixed from between 2.5km up to this maximum of 10km. In current official policy, a school's catchment area is defined as the area within a 3km radius of the school, suggesting that this is felt to be, at the policy level, the maximum distance a child should travel (Martin 2010).

In working with a distance as an indicator of mobility, it is possible to either create a binary variable, using a particular distance as an indicator of whether or not mobility is occurring, or to work with distance as a continuous measure of the extent to which mobility is occurring. The binary, all-or-nothing indicator approach is the one typically implied by existing distance-based definitions of mobility, where mobility either occurs, or does not, on the basis of a specified cut-off point, with no middle ground. While a binary approach to measurement is therefore most useful from a policy assessment point of view, mobility can also be understood as something which always occurs, but at variable levels. As distance travelled has a close relationship to the resources required to be dedicated to that travel, a continuous definition of mobility can be seen as more closely related to reality as experienced by individuals engaged in mobility. As these two different approaches to measurement are complementary, this analysis makes use of both of them.

A second concern when working with distance-based definitions of mobility relates to the way in which the distance from home to school should be determined. Straight-line measurements are both methodologically most simple, and are what tends to be used in mobility-related policy and assessment. However, it does have shortcomings, the most obvious being that children do not travel to school in a straight line, but make use of roads, footpaths, and transportation networks. Furthermore, the relationship between straight-line distance, and actual distance, is likely to be quite variable. Calculating actual distance travelled for this sample would have required the collection of substantial historical route data, and was not practicable. A more practical alternative is using GIS software to calculate the shortest feasible rout between two points, drawing on road network information. Unfortunately, road network data for the Johannesburg-Soweto area was prohibitively expensive, which meant that this alternative could not be pursued for this project. As road network and similar data becomes increasingly available, however, this is an avenue that could be valuably explored.

For the purposes of this project, however, all distances are straight line distances. These are calculated using the Haversine formula applied to the GIS coordinates of the child's home and the child's school (Sinnott 1984). The Haversine formula is calculated as follows, where R is the radius of the earth in kms:

dlon = lon2 - lon1

dlat = lat2 - lat1

a = sin^2(dlat/2) + cos(lat1) * cos(lat2) * sin^2(dlon/2)

c = 2 * arcsin(min(1,sqrt(a)))

d = R * c

3.6.2 An area based definition of mobility

The second approach to the operationalization of learner mobility draws on tests of whether the learner attends school in the same area, using a range of definitions, in which he or she lives. This is motivated by the concern that in some cases, a child's 'local' school may not be the closest school, but the school that is located in the same community in which a child lives. Additionally, in some cases, barriers such as rivers, hills, busy roads or train tracks may mean that a child is cut off from the school that is closest to him or her on the basis of straight-line distance. In these cases, it would also be more natural for a child to attend a school that is slightly further away, but is located in the same geographic community.

The various definitions of area which are used in this thesis are those developed for Census 2001, and discussed previously. For each of these levels, SAL, SP, MP and MN, a binary indicator was created for each of 1997 and 2003, coded one if the child lived and attended school in the same area, and zero if the child did not live and attend school in the same area. If the child is mobile at the smallest level of geography, the SAL level, all other tests will also categorize a child as mobile. Similarly, if a child is mobile at the SP level, the MP and MN tests will also categorize him or her as mobile.

3.6.3 Mobility defined by attendance at the nearest school

The final operationalization used for mobility is based on whether or not the child is attending his or her nearest grade-appropriate school. The two approaches to measuring mobility discussed so far tend to focus on those learners who are travelling particularly substantial distances. While travelling a substantial distance is an important indicator of the amount of effort and money invested in school choice, and certainly identifies the forms of school choice that are most salient in a gradually integrating post-apartheid South Africa, it does not reveal much about individuals who may only be able to participate in school choice at a relatively local level. One way to measure mobility without losing these individuals is to determine whether a child is attending the age-appropriate school nearest to their home, or whether they are choosing to travel slightly further to attend a different school. This is a particularly important aspect of mobility and school choice to explore in an area such as Soweto, where the density of schools is extremely high, with most children living within easy walking distance of more than one school.

Obviously, the figures obtained using this operationalization will be an imprecise reflection of school choice. One particular concern is around those children who are explicitly choosing to attend the school closest to their home, and not simply attending it because it is closest. The extent of this phenomenon is unfortunately not measurable with the existing data, and if it is substantial, would mean that levels of school choice are actually higher than is reflected in the available data. A second concern is around the possibility that there may be children who would like to attend their nearest school, but are unable to do so, perhaps because the closest school is over-enrolled, the child is discouraged by higher school fees at this school, or the child is (illegally) refused admission due to poor academic performance. In these cases, this definition would suggest that the child is engaging in school choice and mobility, when in fact the child is actively being prevented from exercising their choice. Thirdly, it is possible that for reasons of geography, the closest school on the basis of a distance calculation is not actually the closest school for a child on foot. In this instance, it is possible that children who are simply attending the nearest school on the basis of available roads are being misclassified as engaging in school choice and mobility. Nonetheless, this nearest-school analysis should provide a more accurate reflection of the extent of school choice when local school choice is included, than any other definitions explored this far.

In order to measure the proportion of children attending their nearest age-appropriate school, a spatial join was conducted in gvSIG to identify, for each child, the relevant school nearest to their home. This was done twice, firstly for all schools, both public and private, and secondly using only public schools. For 1997, only primary, intermediate and combined schools were included in the analysis. For 2003, both the nearest primary phase school and the nearest secondary phase school was calculated. Then, on the basis of the child's grade, the most appropriate of these two schools was selected. Once the nearest grade-appropriate school for each child, at each time point, was obtained, this school was then compared to the school actually attended by the child at that time point. A binary indicator was created and coded one when the child did attend his or her nearest grade-appropriate school, and zero when the child did not attend that school.

3.7 Analysis

Data management was conducted using Microsoft Access. GIS analysis was done with gvSIG, and statistical analysis with Stata Standard Edition 11. The analytical procedures followed to generate each set of findings are described in the relevant analytical chapters. In general, however, as most variables were non-normally distributed, non-parametric analyses were used. In particular, analyses involving only categorical variables were conducted with chi-square tests, unless the numbers in any category dropped below 5, in which case a Fisher exact test was used. Analyses involving both categorical and continuous variables were conducted using Mann-Whitney Wilcoxon rank-sum or Kruskal-Wallis tests, and analysis involving only continuous variables was conducted using Spearman rank correlation. For multivariate analysis, a multiple regression approach was used. Distributions are displayed as kernel density plots, which show a smoothed curve based on the estimated non-parametric probability densities of a variable's distribution (Scott 1992; van der Berg, Wood et al. 2002).

Chapter 4, the first results chapter, presents a range of descriptive data, covering sample representativeness and bias, and properties of schools attended by sample members in relation to the universe of schools available to them. Chapter 5 documents the overall extent of mobility using each of the three different operationalizations detailed earlier in this chapter. In Chapter 6, the relationship between mobility behaviour and characteristics of learners, their families and households, and their communities is explored. Chapter 7 details the relationship between learner mobility and properties of the schools that children attend, and the schools that are closest to their homes. Chapter 8 explores the extent to which mobility behaviours are subject to change over time. Finally, Chapter 9 uses the findings of all previous chapters to generate preliminary models of learner mobility. Chapter 10 summarizes and discusses the findings, and serves as a conclusion to the thesis.

3.8 Conclusion

This chapter has outlined the methodological approach taken in this thesis, that of quantitative secondary data analysis, as well as the rationale that lay behind this choice. It has documented the range of datasets that were considered for use in the study, and justified the selection of the Bt20 longitudinal cohort study as the most appropriate. The composition of this cohort, particularly with respect to demographic variables, SES, and residential mobility were discussed. Existing data was used to demonstrate that while the overall Bt20 cohort does under-represent the most advantaged and disadvantaged children in Soweto-Johannesburg, and particularly those in minority racial groups, it remains highly representative of black low and middle-income township residents – the group of primary interest to this thesis. The creation of the sub-sample of the Bt20 cohort used in this thesis was explained, and the tests for sample representativity and bias, which are will be presented in Chapter 4, were outlined. Finally, the three different approaches to the operationalization of learner mobility that are used in this thesis were presented, and approaches to data analysis were described.

Chapter 4: Sample descriptive statistics and representativity

4.1 Introduction

This chapter begins by providing descriptive statistics for the sample with respect to the variables hypothesised to have a relationship to engagement in learner mobility. It then moves on to ask whether the study sample is representative firstly of the Bt20 cohort as a whole, and secondly of the youth population of Johannesburg-Soweto metropolitan area more broadly. These questions around representativity are answered both through a range of statistical tests, and through a discussion of the process by which the Bt20 cohort and the study sample were created. Finally, to provide some context for the discussions of school choice to follow in subsequent chapters, some basic descriptive statistics for schools in the Gauteng province (within which the study area falls) are presented.

4.2 Sample descriptive statistics

This section presents descriptive statistics for key child, household and community variables for which there are theoretical grounds to anticipate a relationship with educational mobility. At the child level, race, gender, age at first school enrolment, school phase in 2003, and grade repetition between 1997 and 2003 are explored. At the household level, maternal education, maternal marital status, and household SES in both 1997 and 2003 are considered. Finally, at the community level, the poverty of the area in which the child lives is documented, for three different levels of geography. The relationship between these variables and mobility behaviour will be tested in Chapter 6.

4.2.1 Child level variables

Race

As noted previously, the study sample is mostly black. While coloured children are reasonably well represented, white and Indian children are under-represented, and their numbers are also extremely small. The exact breakdown of study sample members across race groups is presented in Table 4.1 below.

Race	Black African	White	Coloured	Indian
Number of children (n=1428)	1,145 (80.18%)	41 (2.87%)	192 (13.45%)	50 (3.50%)

Table 4.1: Breakdown of study sample members by race

Gender

The study sample is approximately evenly split between males and females (see Table 4.2 below).

Gender	Male	Female
Number of children (n=1428)	711 (49.79%)	717 (50.21%)

Table 4.2: Breakdown of study sample members by gender

Age at first school enrolment

Overall, a very slight majority of sample members enrolled in school for the first time either early or on time for their age, while a slight minority enrolled late (see Table 4.3 below). The extent of late enrolment, at over 47%, is striking. However, as noted in Chapter 3, although the late-starters being their schooling a year later than their peers, the majority of them do not start their schooling outside of the two-year window for enrolment specified by policy.

Age at first enrolment	Earlier	Later
Number of children (n=1275)	673 (52.78%)	602 (47.22%)

Table 4.3: Breakdown of study sample members by age at first school enrolment

Schooling phase in 2003

Table 4.4, below, illustrates that by 2003, just under one third of children had progressed to high school, while just over two thirds remained in primary school. A sample member who started their primary schooling on time, and who had not repeated a grade, would be expected to have reached high school by 2003, whereas those who started late, or who had repeated a grade, would typically not be expected to have reached high school.

Schooling phase (03)	**Primary**	**High**
Number of children (n=1330)	897 (67.44%)	433 (32.56%)

Table 4.4: Breakdown of study sample members by phase of schooling in 2003

Grade repetition

As shown in Table 4.5 below, slightly more than one third of children repeated a grade between 1997 and 2003, while the remainder did not. This figure is similar, though slightly higher than that reported for other work on the Bt20 cohort (Fleisch and Schindler 2009).

Grade repetition between 1997 and 2003	**No repetitions**	**One or more repetitions**
Number of children (n=1240)	778 (62.74%)	462 (37.26%)

Table 4.5: Breakdown of study sample members by whether or not they have repeated at least one grade between 1997 and 2003

4.2.2 Household level variables

Maternal education

The distribution of maternal educational levels is shown in Table 4.6 below. The largest proportion of mothers have completed some secondary school, while relatively few are have grade 5 education or less, which is equivalent to functional illiteracy. The proportion with post-school education is also low.

Maternal Highest Completed Education Level	Grade 5 or below	Grade 6 or 7	Grade 8, 9 or 10	Grade 11 or 12	Post-school education
Number of children (n=1305)	86 (6.59%)	94 (7.20%)	610 (46.74%)	399 (30.57%)	116 (8.89%)

Table 4.6: Breakdown of study sample by highest level of maternal education attained at the time at which the study sample member was born

Maternal marital status

Slightly over one third of mothers were married at the time of the birth of study sample member, while just less than two thirds were unmarried (see Table 4.7 below).

Maternal marital status in 1990	Married	Unmarried
Number of children (n=1418)	506 (35.68%)	912 (64.32%)

Table 4.7: Breakdown of study sample members by maternal marital status

Household SES: 1997

The grouping of households into quintiles on the basis of SES in 1997 is shown in Table 4.8 below. Due to several clusters of households with similar scores, it was not possible to create completely even quintiles.

Household SES quintile 1997	1 (poorest)	2	3	4	5 (wealthiest)
Number of children (n=1205)	254 (21.08%)	240 (19.92%)	233 (19.34%)	246 (20.41%)	232 (19.25%)

Table 4.8: Breakdown of study sample by household SES in 1997

Household SES: 2003

Similarly, the household SES quintiles for 2003 are also not completely even, as evident in Table 4.9 below. Additionally, the small proportion of sample member for whom SES data is available for 2003 should be noted.

Household SES quintile 2003	**1 (poorest)**	**2**	**3**	**4**	**5 (wealthiest)**
Number of children (n=887)	181 (20.41%)	174 (19.62%)	179 (20.18%)	177 (19.95%)	176 (19.84%)

Table 4.9: Breakdown of study sample members by household SES in 2003

4.2.3 Community level variables

Small Area Level poverty

An index of community poverty was calculated for each level of census geography, as described in Chapter 3. Households were broken down into quintiles on the basis of poverty level of the SAL in which they lived. The distribution of households across the quintiles of SAL poverty level is illustrated in Table 4.10 below.

SAL poverty quintile	**1 (lowest poverty)**	**2**	**3**	**4**	**5 (highest poverty)**
Number of children (n=1399)	280 (20.01%)	280 (20.01%)	280 (20.01%)	281 (20.09%)	278 (19.87%)

Table 4.10: Breakdown of study sample members by the poverty level of the SAL in which they live

Sub Place poverty

Table 4.11, below, shows the distribution of households across quintiles based on the poverty level of the SP in which they are located.

SP poverty quintile	1 (lowest poverty)	2	3	4	5 (highest poverty)
Number of children (n=1399)	284 (20.30%)	287 (20.51%)	288 (20.59%)	273 (19.51%)	267 (19.09%)

Table 4.11: Breakdown of study sample members by the poverty level of the SP in which they live

Main Place poverty

Due to the small number of MPs represented in the data, with most sample members concentrated in just a few MPs, attempts to create poverty quintiles based on this level of geography were unsuccessful. The creation of tertiles was slightly more successful, and is illustrated in Table 4.12 below, although substantial clustering is still evident.

MP poverty tertile	1 (lowest poverty)	2	3 (highest poverty)
Number of children (n=1399)	491 (35.07%)	776 (55.43%)	133 (9.50%)

Table 4.12: Breakdown of study sample members by the poverty level of the MP in which they live

4.3 Relationships between variables

Tests were conducted to explore the relationships between each of the variables discussed above within the study sample. All relationships operated in the expected directions, and are documented in Appendix C.

4.4 Study sample representativity

Understanding the representativity of the study sample consists of two different elements. Firstly the representativity of the full Bt20 cohort with regards to the broader population of similarly aged children in the

Johannesburg-Soweto area needs to be understood. This requires a discussion of the initial sample composition, as well as of subsequent sample attrition. Secondly, the representativity of the study sample, with regards to the full, non-attrition Bt20 cohort must be explored. This section begins by describing Bt20 cohort composition and attrition, and implications for cohort representativity. This is followed by series of tests to determine whether the decision to limit the study sample to only those children who did not change residential address between 1996 and 2004 introduced any additional representativity concerns.

4.4.1 How representative is Birth to Twenty?

Cohort composition

The Bt20 study enrolled and collected longitudinal data on a birth cohort of 3273 singleton children, born to mothers resident in the Johannesburg-Soweto area between April 28 and June 8, 1990. These children have been followed up 16 times to date, meaning that data is available for them at a range of points throughout their schooling. Along with data regarding home situation, caregivers, and a range of health and psychometric measures, Bt20 data relevant to schooling includes each child's school's name, grade, repetition, drop out, and academic performance as captured on standardized tests and by the children's school reports. Although the quality and depth of the data collected by Bt20 is high, and attrition is generally low, concerns about how representative the Bt20 cohort is of Johannesburg-Soweto children and youth, particularly over time, remain important. For this study, this is particularly relevant when the cohort may differ from the broader population with respect to variables such as SES that are expected to have a substantial influence on schooling choices and learner mobility.

The Bt20 cohort was intended to consist of all the singleton children born in the Johannesburg-Soweto area between 23 April and 8 June, 1990. For various reasons, however, not all eligible children were actually enrolled. A subsequent study undertaken to identify these ‘missing’ children – children eligible to be in the cohort but who had either never been identified by the study, or had been identified but had not enrolled– revealed that non-enrolment was largely for two reasons. Firstly, for practical and resource-related reasons, study recruitment was concentrated in public sector health facilities. This meant that members of the more advantaged groups, who typically used private facilities, were less likely to come into contact with study recruiters, and were therefore less likely to be enrolled (Richter, Norris et al. 2004). Secondly, a number of individuals who had initially been identified by study recruiters either declined to enrol, or agreed to study participation but could not subsequently be traced for full enrolment. Reasons for non-enrolment included mobility, often combined with incorrect or incomplete address data, incorrect or incomplete recording of caregiver’s names, particularly when multiple names were in use, and participant concerns about invasion of privacy or about participation being overly time-consuming. The majority of eligible children who did not enrol again came from relatively affluent backgrounds compared to the rest of the cohort, although some were also relatively disadvantaged (Richter, Norris et al. 2004; Richter, Panday et al. 2009).

As a result of these two sets of factors, the initial Bt20 cohort underrepresented white and Indian children, along with more affluent children more generally, but was largely representative of the predominantly black African population of similarly aged children living in the area in 1990. This under-representation of more advantaged children, particularly white and Indian, may result in an under-representation of those children who live very close to high quality schools, and therefore do not experience any pressure to travel to attend schools further afield. It may also, however, result in an under-representation of those children who are able to travel particularly great distances in order to access the most desirable schools. It is therefore unclear how, in aggregate, the under-representation of more affluent children is likely to influence the outcomes of this study. Given, however, that the population of primary interest to this study is the largely black African lower-middle-class population of Johannesburg-Soweto, and that the cohort was representative in this regard, these concerns are not of substantial significance. The relatively small size of the affluent population of Johannesburg-Soweto, as well as of the white and Indian populations, also means that any impact that their under-representation is likely to have on study outcomes will be fairly minor.

A second concern relating to overall Bt20 cohort composition relates to the fact that, with the passage of time, the population of the greater Johannesburg-Soweto area has changed. This means that regardless of how representative the cohort was when it was launched in 1990, over time it is likely to have become less representative of the population of same-aged children actually living in the area. This is, of course, a concern with any cohort study, but while it is not specific to Bt20, Bt20 might be particularly seriously affected due the historical era which the study covers. The post-Apartheid era, during which the Bt20 participants grew up and attended school, has been characterized by substantial changes in the residential patterns of people, including children and youth. In part, this has been a response to the demise of Apartheid's strictly enforced segregationist residential rules, in which black South Africans were only permitted to live in urban areas if they were employed there. The perception that people living in urban areas, and particularly in Gauteng, are better off than those in rural areas, particularly with regards to access to services and economic resources, combined with the repeal of Apartheid-era segregationist policies, has triggered a substantial influx of new residents to the Soweto-Johannesburg area during the past 20 years (Richter, Norris et al. 2006).

A sub-study of Bt20, the 2002 Children's School Survey, collected data on all children born between April 23 and June 8 1990 enrolled at 81% of the primary schools in the greater Johannesburg-Soweto area (Richter, Norris et al.

2006; Richter, Panday et al. 2009). With a 92% response rate, detailed demographic data was collected on 5367 children. Almost half of these children had not been born in the greater Johannesburg-Soweto area, indicating a substantial level of in-migration amongst children of this age. Analysis of this data revealed significant differences on a number of key indicators between those children who had been born in the greater Johannesburg-Soweto area and those who had moved there at a later point. Generally, in-migrating children appeared to be living in more adverse circumstances and were at greater risk of poverty than their peers who had been born in the area. They were less likely to live in formal housing, and their parents were less likely to own their dwellings. They had lower levels of ownership of most household assets, and poorer access to basic household services such as running water, electricity, sanitation, and refuse removal. Parents of in-migrating children were more likely to be unemployed, and a higher proportion of those who were employed were in unskilled employment. The in-migration of children appeared to be closely connected to school attendance, with the majority of children migrating into the area doing so before commencing their schooling. In-migrating children were also more likely to have started their schooling late, although no evidence of any impact of this delayed start on academic performance in 2002 was found (Richter, Norris et al. 2006). While we don’t have access to information about out-migration, it is likely that the effects of this on the population have been captured fairly well by attrition from the cohort, which is discussed in the next section.

From the information collected from the Children’s School Survey, however, it is clear that the children migrating into the Johannesburg-Soweto area do tend to differ in terms of their home environments from those children born in the area, and represented in the Bt20 cohort. Specifically, cohort members are likely to be more advantaged, and therefore may be more likely to have access to the necessary resources to participate in school choice and learner mobility. As such, we may again expect to see higher levels of learner mobility in the cohort than would be found in the overall population of same-aged children in Johannesburg-Soweto. Given, however, that we know very little about how income shapes learner mobility, it is not possible to estimate the scale of this effect. It appears, nonetheless, that our cohort remains largely representative of the lower-middle-income black people who make up the bulk of the population of Johannesburg-Soweto.

Finally, as is the case with any cohort study, it cannot be assumed that the findings for children born in 1990 can necessarily be extrapolated to children born at other points in time. For the same reasons as those outlined above, relating to the political and social transformations that South Africa has undergone during the lives of the cohort members, it is also likely that Bt20 is particularly sensitive to these sorts of changes. Nonetheless, given the strengths of the cohort, and the absence of alternative sources of data, it remains the best source from which to derive findings relevant to primary school children in contemporary South Africa. Additionally, it seems likely that the period during which Bt20 cohort members were attending primary school was the period during which new, post-Apartheid patterns of school enrolment were defined and stabilized. As a result, the data presented documents a critical period in the evolution of South Africa's post-Apartheid schooling system, and can probably, with some caution, be extrapolated to more recent points in time.

Bt20 cohort attrition

Changes in the composition of the cohort over time, predominantly due to the non-random attrition of participants, will also affect the extent to which it is representative of the population that it was designed to represent. Attrition in Bt20, while low for a study of this length and magnitude, is known to be related to certain variables such as race and socio-economic status (Richter, Norris et al. 2004; Richter, Norris et al. 2007; Ginsburg, Norris et al. 2009). As these variables are expected to be related to learner mobility and school choice more broadly, it is important to explore the nature of this attrition, along with the ways in which it might influence study findings.

As mentioned previously, the Bt20 study covers a particularly eventful era in South African history, during which the Apartheid system of controls and regulations governing where people could live, work and be educated was dismantled, with huge implications for population distribution in South Africa. Along with in-migration, as discussed above, the Gauteng province was also affected by out-migration, particularly amongst women and young children, driven by political tension, and overstretched public services. High levels of circular mobility between urban and rural areas, along with mobility within urban areas, became features of the Gauteng province. Migration out of the study area, and as well as mobility within it, caused attrition of cohort members from the study. Attrition has also been caused by maternal or child death, child abandonment or adoption, and study fatigue (Richter, Norris et al. 2004).

Even in times and areas of relative stability, the maintenance of a longitudinal sample is difficult, and the internationally accepted norm for sample attrition is between 10 and 20% per annum. In the early post-Apartheid South African context, the Bt20 study was greatly challenged to find ways to maintain the birth cohort, and minimize attrition (Richter, Norris et al. 2004). Using a combination of approaches at both the community and individual level, Bt20 succeeded in keeping sample attrition at an extremely low level, averaging below 3% per year. At the community level, efforts included cooperation with a Community Advisory Board, the use of local fieldworkers, and strict adherence to ethical guidelines including confidentiality, to build a strong relationship with participant communities, and by extension, trust. The provision of some limited social and health services to cohort members also encouraged them to maintain contact. At the individual level, participants were contacted regularly by post, telephone and in person, and were followed up extensively if contact was lost.

While attrition has been greatly limited by the efforts described above, it has nonetheless been non-random in nature. Most notably, attrition has been substantially higher among white participants, as well as participants with higher socio-economic status, exacerbating the existing under-enrolment of these groups (Richter, Norris et al. 2004). By contrast to more advantaged groups, retention amongst more vulnerable members of the sample has been extremely high, with black African mothers and their children being particularly likely to remain in the study.

Ginsburg et al. (2009) present a number of mobility-related analyses comparing all cases lost to the study prior to 2005, and those cases remaining in the cohort. The attrition cases had experienced significantly higher frequency of residential movement, with 81.3% of them having moved at least once by 2005, and 13.3% moving within any of the documented intervals. By contrast, amongst children remaining in the cohort, only 55.5% had experienced any residential movement. The attrition group also contained significantly more white participants, those born in private hospitals, and those residing in the inner city or suburbs, as opposed to the townships. Mothers in the attrition group were more likely to have been married at the time of child's birth, and to have either no formal education, or to have attained some level of post-school education. As suggested by the maternal education information, attrition children were also more likely to live in particularly highly or poorly resourced households. These finding echo expectations that the most advantaged children, particularly white children and those living in the most affluent areas are likely to be underrepresented, as are the most disadvantaged children. As neither of these two groups are likely to be participating substantially in learner mobility (the disadvantaged due to inability, the advantaged due to living close to good schools), we may again expect that figures on mobility are likely to be somewhat inflated for urban populations as a whole. However, as attrition amongst the predominantly African, township dwelling majority of the cohort has remained low, we can expect results fairly representative of this particular population, which is fortunately the group of most interest for this particular project.

4.4.2 How representative is the study sub-sample?

As documented in Chapter 3, unanticipated problems with the residential GIS coordinate data, requiring a substantial and time-consuming cleaning process, meant that it was infeasible to use the full, non-attrition Bt20 cohort for this thesis, as had originally been planned. Of the initial Bt20 cohort of 3273, 66% (n=2158) completed a residential history questionnaire in 2005 or 2006 (Ginsburg, Norris et al. 2009), and forms the cumulative non-attrition cohort as of 2006. Of this group, 1470 individuals reported no changes in residential address during this period. Once children attending special schools or boarding schools, those attending schools outside of the Gauteng province, and those who were not attending school at all were removed from this group, along with a small number of children who were resident at multiple addresses during the study period, this left 1428 individuals. The data on these 1428 individuals forms the basis for all analysis presented in this thesis. While not ideal, the decision to focus on this sub-sample of cohort members was made to maximize the available sample size in light of problems with the residential GIS data which required a lengthy cleaning procedure. However, as residential stability between 1996 and 2004 appears likely to be closely related to SES and other variables which may influence school choice, it is important to measure and document the differences between this non-random sub-sample and the full, non-attrition cohort, and to think about how this is likely to influence study findings.

Given that residential mobility levels are greatest amongst the most advantaged and most disadvantaged sectors of the population (Ginsburg, Norris et al.

2009), a sub-sample constructed primarily on the basis of mobility behaviours is clearly unlikely to be representative of the full sample from which it is drawn. Ginsburg et al. (2009) provide some valuable data on the differences between Bt20 participants who had experienced some residential mobility prior to 2005, and those who had not. While these figures are valuable as indicators, it should be noted that the sub-sample used in this thesis excludes only those children who experienced mobility between 1996 and 2004, and not those who moved earlier or later. The highest levels of mobility, however, were seen in the earliest years, with the commencement of primary schooling typically having a stabilizing effect on children's residence. For this reason, the differences between residentially mobile and non-mobile children presented in Ginsburg et al (2009) are likely to be somewhat more substantial than the differences between those children included and excluded from the sample used in this thesis. Additionally, sample construction seems less likely to have influenced representativeness for black African children with mid-range SES levels, the population of primary interest in this study.

In order to better understand the nature of the sub-sample used in this thesis, and in particular the ways in which included cases may differ from the excluded, I compare both groups of cases. Firstly, I compare the members of the study sub-sample (n=1428) to all cohort members not included in the sub-sample (n=1845). This provides an indication of how different this sub-sample is from the full Bt20 cohort, including all those individuals lost to contact, as well as those excluded from this sub-sample for any other reasons. Secondly, I focus on the group of cohort members not lost to attrition (n=2158), and contrast those included in my sub-sample (n=1428) with those excluded from my sub-sample (n=730). This provides an indication of how different my sample is from those members of the full Bt20 non-attrition cohort who were not included in this study.

Thesis sub-sample compared to all excluded cohort members

For SES, and each available demographic variable, the group of individuals included in the study sub-sample was compared to the group of all cohort members excluded from the sub-sample for any reason. Chi-squared tests were conducted to determine whether the distribution of individuals was significantly different across groups for each of the variables. As evident in Table 4.13, below, the included and excluded cohort members differed to a statistically significant degree on all tested variables, with the exception of gender. Importantly, and as hypothesized, the different distributions support the contention that the study sample appears to under-represent those at either extreme of the socioeconomic scale. Under-representation of the historically more advantaged race groups, those born in private hospitals, and outside of the Soweto-DiepMeadow area all suggest that the most advantaged are likely to be underrepresented, but can provide little information about the most disadvantaged section of the cohort. However, an examination of the SES data, as well as the data on maternal educational level suggests that the most disadvantaged are also underrepresented. This is clearest with regards to the SES variable, which shows that the study sample is biased towards the 3 middle quintiles, while the group of excluded individuals contains a higher proportion of individuals falling into the first (poorest) and fifth (wealthiest) quintiles. As discussed previously, the exclusion of the most advantaged and disadvantaged, who for various reasons are hypothesized to be least likely to engage in learner mobility, may lead to somewhat inflated findings regarding levels of learner mobility. However, given that the study sample does appear to be reasonably representative of the black, township-based middle-class, the group which is of greatest interest in this examination of school choice behaviours, this is not anticipated to be likely to be a major problem.

Variable	Value	Included # (% of included)	Excluded # (% of excluded)	Chi-squared results
Child gender	Male	711 (49.79%)	880 (47.70%)	$\chi^2_{(1)}$ =1.412 not significant n=3273
	Female	717 (50.21%)	965 (52.30%)	
Race	Black	1145 (80.18%)	1423 (77.13%)	$\chi^2_{(3)}$ =55.307 p<0.001 n=3273
	Coloured	192 (13.45%)	191 (10.35%)	
	Indian	50 (3.50%)	65 (3.52%)	
	White	41 (2.87%)	166 (9.00%)	
Maternal age	18 or younger	162 (11.35%)	180 (9.76%)	$\chi^2_{(2)}$ =20.733 p<0.001 n=3271
	19-34	1084 (75.96%)	1511 (81.94%)	
	34 or older	181 (12.68%)	153 (8.30%)	
Place of birth	Soweto/DiepMeadow	1134 (79.41%)	1295 (70.19%)	$\chi^2_{(3)}$ = 135.345 p<0.001 n=3273
	Historically Indian/Coloured area	221 (15.48%)	211 (11.44%)	
	Inner city JHB	5 (0.35%)	64 (3.47%)	
	Suburban JHB	68 (19.83%)	275 (14.91%)	
Hospital of birth	Public	1255 (87.89%)	1576 (85.47%)	$\chi^2_{(1)}$ = 4.038 p<0.05 n=3272
	Private	173 (12.11%)	268 (14.53%)	
Maternal marital status	Married	506 (35.68%)	696 (37.97%)	$\chi^2_{(3)}$ = 42.236 p<0.001 n=3251
	Cohabiting	53 (3.74%)	160 (8.73%)	
	Separated/Divorced/Widowed	28 (1.97%)	19 (1.04%)	
	Single	831 (58.60%)	958 (52.26%)	
Maternal highest educational level	No formal education	13 (1.00%)	34 (2.09%)	$\chi^2_{(3)}$ = 23.812 p<0.001 n=2932
	Primary schooling	167 (12.80%)	241 (14.81%)	
	Secondary schooling	1009 (77.32%)	1140 (70.07%)	
	Post-school education	116 (8.89%)	212 (13.03%)	
SES at birth	Quintile 1 (most poor)	216 (18.56%)	393 (27.14%)	$\chi^2_{(4)}$ = 36.803 p<0.001 n=2612
	Quintile 2	215 (18.47%)	223 (15.40%)	
	Quintile 3	286 (24.57%)	284 (19.61%)	
	Quintile 4	238 (20.45%)	251 (17.33%)	
	Quintile 5 (least poor)	209 (17.96%)	297 (20.51%)	

Table 4.13: Differences between cohort members included in the study sample, and those excluded from the study sample with regards to all available demographic variables collected at birth

Thesis sub-sample compared to other non-attrition cases: historical data

This second analysis explores the extent to which the study sub-sample differs from the group of cohort members from which it is drawn; that is, the full non-attrition sample. This captures the way in which those cases excluded because the children moved home between 1996 and 2004 differ from those who did not move during this period. The same 1990 data was used for this set of analyses as in the section above. In particular, the SES estimates and quintile allocations for each case were not re-calculated, but were used as generated in the previous set of analyses, on the basis of the data from the full cohort. The results are presented in Table 4.14, below.

Variable	Value	Included # (% of included)	Excluded # (% of excluded)	Chi-squared results
Child gender	Male	711 (49.79%)	342 (46.85%)	$\chi^2_{(1)} = 1.672$ not significant n=2158
	Female	717 (50.21%)	388 (53.15%)	
Race	Black	1145 (80.18%)	601 (82.33%)	$\chi^2_{(3)} = 3.442$ not significant n=2158
	Coloured	192 (13.45%)	86 (11.78%)	
	Indian	50 (3.50%)	18 (2.47%)	
	White	41 (2.87%)	25 (3.42%)	
Maternal age	18 or younger	162 (11.35%)	92 (12.62%)	$\chi^2_{(2)} = 19.940$ $p<0.001$ n=2156
	19-34	1084 (75.96%)	590 (80.93%)	
	34 or older	181 (12.68%)	47 (6.45%)	
Place of birth	Soweto/DiepMeadow	1134 (79.41%)	584 (80.00%)	$\chi^2_{(3)} = 9.950$ $p<0.05$ n=2158
	Historically Indian/Coloured area	221 (15.48%)	90 (12.33%)	
	Inner city JHB	5 (0.35%)	7 (0.96%)	
	Suburban JHB	68 (4.76%)	49 (6.71%)	
Hospital of birth	Public	1255 (87.89%)	642 (88.07%)	$\chi^2_{(1)} = 0.015$ not significant n=2157
	Private	173 (12.11%)	87 (11.93%)	
Maternal marital status	Married	506 (35.68%)	247 (34.12%)	$\chi^2_{(3)} = 3.144$ not significant n=2142
	Cohabiting	53 (3.74%)	31 (4.28%)	
	Separated/Divorced/ Widowed	28 (1.97%)	8 (1.10%)	
	Single	831 (58.60%)	438 (60.50%)	
Maternal highest educational level	No formal education	13 (1.00%)	4 (0.60%)	$\chi^2_{(3)} = 7.886$ $p<0.05$ n=1971
	Primary schooling	167 (12.80%)	64 (9.61%)	
	Secondary schooling	1009 (77.32%)	521 (78.23%)	
	Post-school education	116 (8.89%)	77 (11.56%)	
SES	Quintile 1 (most poor)	223 (19.16%)	135 (22.06%)	$\chi^2_{(4)} = 6.8454$ not significant n=1776
	Quintile 2	253 (21.75%)	104 (16.99%)	
	Quintile 3	241 (20.70%)	137 (22.39%)	
	Quintile 4	238 (20.45%)	122 (19.93%)	
	Quintile 5 (least poor)	209 (17.96%)	114 (18.63%)	

Table 4.14: Differences between members of the non-attrition sample included in and excluded from the study sub-sample, with respect to variables collected at birth

The results of this second set of analyses suggest that the study sub-sample, selected on the basis of not having changed residence between 1996 and 2004, while differing from the full non-attrition sample in some regards (maternal age, place of birth, and maternal education), is not significantly different in others. Additionally, for those variables which are significantly different, the levels of significance are lower, with only maternal age remaining significant at $p<0.001$. These results suggest that children with mothers under the age of 35 were significantly more likely to be excluded from the study sub-sample than those with older mothers. Children with mothers with particularly high levels of education were also significantly more likely to be excluded from the sub-sample, as were children born in the typically more affluent suburban areas of Johannesburg. By contrast, children born in the historically Indian and coloured areas were particularly likely to be included in the sub-sample.

While these figures do suggest that more advantaged children may be somewhat under-represented in the study sub-sample, compared to in the non-attrition sample, the lack of any statistical significance on this variable suggests that any genuine differences are likely to be fairly minor. The absence of any significant difference on race, hospital of birth, and maternal marital status is also encouraging. It therefore seems reasonable to conclude that while the composition of the thesis sub-sample selection may under-represent both the most advantaged and disadvantaged children, this effect is not as substantial as that anticipated on the basis of Ginsburg et al. (2009), and is certainly less severe than that caused by sample attrition.

Thesis sub-sample compared to other non-attrition cases: 1997 and 2003 data

A second question with regards to study sample representativity is whether, despite initial similarities, those included in and excluded from the study sample have changed over time in systematically different ways. Ideally, one would also want to ask this question with regards to emergent differences between the study sample and the full cohort, but this is not feasible as, due to attrition, data for later time points is not available for all cohort members. For this reason, this second exploration of potential sample bias is restricted to testing for differences between the study sample and those non-attrition sample members excluded from it. SES data is available for both 1997 and 2003, and this is used to test whether the study sub-sample differs from the non-attrition sample in this regard at each of these time points.

Socio-economic status in 1997

Full SES data for 1997 was available for 1758 of the 2158 cases in the non-attrition sample, and this is the data that was used for the following analyses. It is worth noting that cases removed from the study sub-sample due to residential mobility were substantially more likely to be missing SES data ($\chi^2_{(1)}$= 23.828; pr<0.001) than those that were retained in the study sub-sample (see Table 4.15 below). While this makes sense, in that children who were mobile were probably harder to locate during any particular round of data collection, the implications of this difference in levels of missing data for the validity of the following analysis are not clear.

	Included: n (% of included)	**Excluded: n (% of excluded)**
Yr7 SES data available	1205 (84.38%)	553 (75.75%)
Yr7 SES data missing	223 (15.62%)	177 (24.25%)

Table 4.15: Availability of 1997 SES data for members of the non-attrition sample included in and excluded from the study sample

As described in Chapter 3, PCA was used to estimate SES scores for each individual using asset ownership data (see Table 3.1). These scores were then used to generate poverty quintiles. Distribution across the quintiles differed significantly between those included in the study sample, and those excluded from the study sample on the basis of residential mobility (see Table 4.16

below). Substantially more cases in the very lowest quintile were removed in generating the study sample, while a relatively lower proportion of cases in the other quintiles were removed. This suggests that the study sample includes a somewhat higher proportion of children living in middle-class and affluent families in 1997 than the full non-attrition sample. Reasons why this might be the case are not obvious, but may relate to differences in the timing of mobility for families with different levels of SES, as only children who moved during their primary school years were excluded.

Variable	Value	Included # (%)	Excluded # (%)	Chi-squared results
SES quintile	1 (poorest)	254 (21.08%)	160 (28.93%)	$\chi^2_{(4)} = 14.587$ p<0.01 n=1758
	2	204 (16.93%)	87 (15.73%)	
	3	261 (21.66%)	97 (17.54%)	
	4	236 (19.59%)	108 (19.53%)	
	5 (least poor)	250 (20.75%)	101 (18.26%)	

Table 4.16: Differences between members of the non-attrition sample included in and excluded from the study sub-sample, with respect to SES in 1997

Socio-economic status in 2003

As described in Chapter 3, SES for 2003 was estimated using an assets index collected during study year 12, and housing quality data collected during study year 13 (see Table 3.1). Following manual imputation of missing values, complete data was available for 1296 individuals, or approximately 60% of non-attrition cases. PCA was used to estimate an SES variable for each sample member. These were additionally used to categorize individuals into 5 poverty quintiles. The extent to which data was missing for cases included in the study sample, and for the non-attrition cases excluded, were compared (see Table 4.17 below). Once again, a substantially higher proportion of those excluded from the study sample on the basis of residential mobility were missing SES data ($\chi^2_{(1)} = 7.462$; pr<0.01).

	Included in study sample: n (% of included)	Excluded from study sample: n (% of excluded)
2003 SES data available	887 (62.11%)	409 (56.03%)
2003 SES data missing	541 (37.89%)	321 (43.97%)

Table 4.17: Availability of 2003 SES data for members of the non-attrition sample included in and excluded from the study sample

A chi-squared analysis was conducted to test whether the distribution of cases across the SES quintiles was different for those participants included in the study sample, and those excluded (see Table 4.18 below). The test revealed no significant differences between these distributions. On the basis of these 2003 SES scores then, sample selection does not appear to have created any additional sample bias in favour of children with mid-range or high SES scores. This combines with the analyses of SES scores at other points in time to suggest that while sample selection on the basis of residential mobility may reduce the representation of the most disadvantaged and the most advantaged participants, this effect is relatively minor, and is weakest for the most contemporary data.

Variable	Value	Included # (%)	Excluded # (%)	Chi-squared results
SES quintile	1 (poorest)	181 (20.41%)	79 (19.32%)	$\chi^2_{(4)} = 1.1916$ not significant n=1296
	2	173 (19.50%)	86 (21.03%)	
	3	180 (20.29%)	80 (19.56%)	
	4	181 (20.41%)	78 (19.07%)	
	5 (least poor)	172 (19.39%)	86 (21.03%)	

Table 4.18: Differences between members of the non-attrition sample included in and excluded from the study sub-sample, with respect to SES in 2002/2003

4.4.3 Sample selection & bias: Conclusion

In summary, it seems likely that cohort enrolment, attrition over time, and the non-random selection of the sub-sample used for this thesis are likely to have each played a small role in contributing to a somewhat biased sample. In particular, there seems to be some indication of an over-representation of black children, and t hose from middle-income families. By contrast, there is some under-representation of members of minority racial groups, as well as children whose families are amongst the richest or poorest 20% of the population. Overall, however, the extent of this under-representation does not appear to be extreme, particularly when the length of time for which data is available is considered. However, caution should of course be used in generalizing the findings of this study to the broader population, and particularly to children at extreme ends of the socio-economic continuum. In particular, it is possible that the level of mobility found in this cohort may be slightly higher than that found in the population more broadly, as it seems possible that children from medium-income families may be the most likely to engage in learner mobility.

4.5 Descriptive schools data: all Gauteng schools

The final section in this chapter presents a range of descriptive statistics for the schools found in the Gauteng province of South Africa. As detailed in Chapter 3, the data presented here comes from a variety of time points between 2002 and 2008, but covers all registered schools, public and independent, known to have operated in the Gauteng province in the post-Apartheid era. This information is presented to provide an overview of the nature of the educational opportunities available to children growing up in the Johannesburg-Soweto area, which is essential to understanding the findings presented in subsequent chapters. Properties of schools with regards to each of the school level variables considered are described, and bivariate relationships with other school properties are described. Additional data relating solely to that subsample of schools attended by study sample members is presented in Chapter 7, as is information relating different school attributes to mobility behaviours.

4.5.1 School types and sectors

The final schools dataset used for the analyses in this thesis contains data on 2604 schools, both public and independent, known to have operated in the Gauteng province in the post-Apartheid era. Of these, 1570 (60.29%) are primary schools, covering grades 1-7, and 289 (11.10%) are combined schools, running all the way from grade 1 through to grade 12. There are 656 (40.90%) secondary schools, covering grades 8-12, along with small numbers of intermediate schools (n=73; 4.55%), and finishing schools (n=16; 0.61%).

Just below 20% of the schools in the dataset are independent schools. The majority of the combined (n=229, 79.24%) and finishing schools (n=15; 93.75%) are independent. Amongst public schools, 1406 (67.27%) are primary schools, 552 (26.41%) are secondary schools, and only 132 (6.32%) are other school types. This fact, that Gauteng contains a substantially larger number of public primary schools than public high schools will be revisited in subsequent chapters, as it relates to changes in mobility behaviour as children move from primary to high school. Of particular note is the fact that, on average, a child's nearest secondary school will be somewhat further away from his or her home than his or her nearest primary school. This means that, all else held constant, a child should be expected to travel somewhat further to school on enrolling at a high school. The smaller number of high schools also means that the range of schools which children are choosing between is more limited, reducing the extent of choice available to children.

4.5.2 School Quintile

The quintile rating system[8] applies only to public schools. As discussed in Chapter 3, it rates schools from 1 (being the poorest) to 5 (the most affluent), primarily on the basis of the community within which the school is located

[8] It should be noted that despite its name, the quintile rating system does not divide either schools or learners evenly into five different groups. Very little information is publicly available as to how exactly the poverty quintile ratings for schools were arrived at, or why the quintiles are so variable in size. Available information is summarized in Chapters 2 and 3.

(Kanjee and Chudgar 2009). Within Gauteng province, schools are not very evenly distributed between the different quintiles, and the majority of schools are in quintiles 3 or above, which is in line with Gauteng being a primarily urban, and comparatively affluent province. Secondary schools appear to be somewhat more likely to be in higher quintiles than primary schools. The distribution of Gauteng public schools across the poverty quintiles is shown in Table 4.19 below.

Quintile	**Primary schools** (% of public primary schools)	**High school** (% of public high schools)	**Total** (% of all public schools)
1	n=176 (13.02%)	n=42 (8.11%)	n=218 (11.66%)
2	n=118 (8.73%)	n=43 (8.30%)	n=161 (8.61%)
3	n=433 (32.03%)	n=145 (27.99%)	n=578 (30.91%)
4	n=355 (26.26%)	n=155 (29.92%)	n=510 (27.27)
5	n=270 (19.97%)	n=133 (25.68%)	n=403 (21.55%)

Table 4.19: Numbers of schools in each quintile in Gauteng province

4.5.3 Section 21 Status

Section 21 status is also only relevant to public schools, and is based on 2008 data. Any public school can apply to operate as a Section 21 school, which places responsibility for the management of school finances at the school level, substantially increasing autonomy. If schools are not Section 21, their finances are operated by their provincial Department of Education, which is typically much less efficient, and can be substantially more expensive. As a result, schools typically pursue Section 21 status whenever they have any managerial capacity, although in some cases applications for Section 21 status are turned down. The large majority of the Gauteng schools on which data is available, 87%, have Section 21 status. High schools appear less likely to have Section 21 status (84%) than primary schools (90%), and a chi square test confirms that this is a significant difference ($\chi^2_{(1)}$=13.6702, Pr=0.000). Predictably, there is a strong relationship between quintile rating and Section 21 status, with more affluent schools being significantly more likely to manage their own finances ($\chi^2_{(4)}$=35.1538, Pr=0.000).

4.5.4 School enrolment

Due to the substantially larger number of public primary schools in Gauteng, it seems likely that they would tend to be smaller than high schools, even though primary schools cover 7 years of schooling, as opposed to 5 years covered by high schools. The available schools data bears this out, with the mean enrolment for primary schools in 2002 being 687, compared with 979 for secondary schools (see Table 4.20 below). Combined schools had the smallest mean size of any school type, probably because only a small proportion of them were public. Independent schools were typically smaller than public schools, with the mean size for a independent primary school at just 301 learners, while the mean for public primary schools was 711 learners. Similarly, independent secondary schools were a mean size of 344, whereas public secondary schools had a mean enrolment of 1067. The mean size overall was 525 learners, rising to 776 for public combined schools, and falling to 448 for independent combined schools.

School Type	Number in Gauteng	Average number of learners
Combined	297	620
Intermediate	76	782
Primary	1807	744
Secondary	702	1003

Table 4.20: Average number of learners for different types of schools in Gauteng

Although school enrolment does vary significantly by quintile rating, according to a Kruskal-Wallis test (Pr=0.00), the nature of this relationship is not entirely clear. Quintile 2 schools are on average the largest, quintile 1 schools the smallest, and the average enrolments of schools in quintile 3, 4 and 5 are between these two extremes. This pattern holds when all schools are examined and when primary schools only are examined, but becomes less extreme when only secondary schools are examined. Overall, there is no evidence that enrolment rises or falls strictly in line with quintile rating. An explanation for the particularly small size of quintile 1 schools may be their predominantly rural locations.

The relationship between school size and Section 21 is clearer, with schools with Section 21 status tending to be substantially larger than those without. This may again relate to many of those schools without Section 21 status being located in rural areas, and therefore having lower enrolments. On the other hand, it may also be the case that those few schools without Section 21 status are a particularly poorly performing subset of schools, and therefore particularly unattractive to learners, leading to lower enrolment.

4.5.5 Percentage of black learners

Although the mean proportion of black African learners in Gauteng schools was just over 73%, much in line with the population of the province as a whole, this figure obscures the actual distribution of black learners across the province's schools. When the data is broken down, it becomes clear that over 50% of schools have an enrolment that is over 99% black, while a full 10% of schools have fewer than 5% black learners. Schools that are meaningfully integrated, and representative of the racial composition of the province's population as a whole, are extremely rare. These figures are in line with those presented by Sujee (2004). Figure 4.1 below illustrates the distribution of the proportion of black learners across all Gauteng province schools.

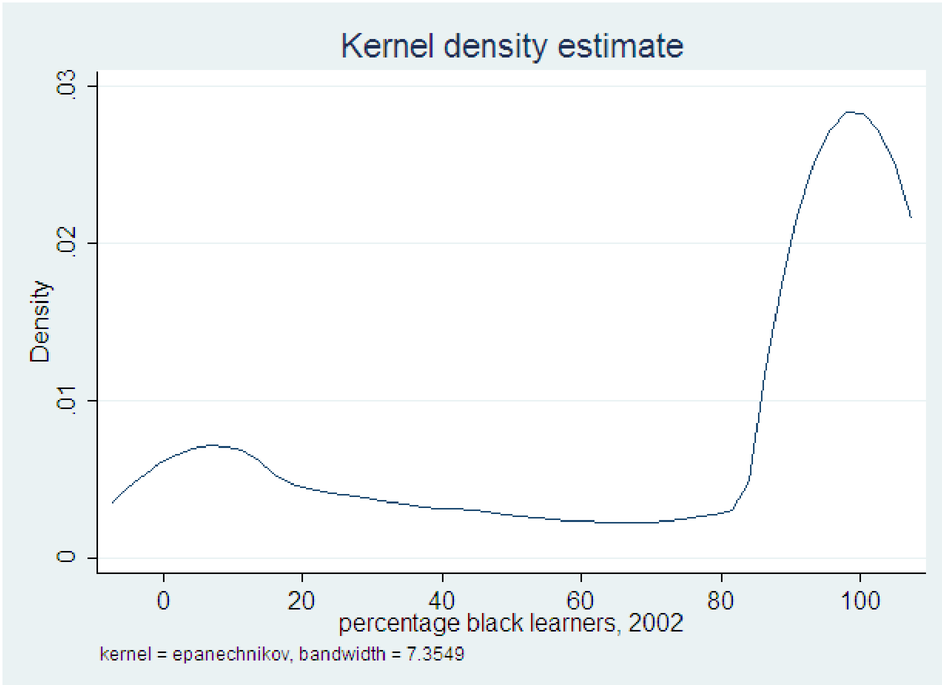

Figure 4.1: Distribution of Gauteng schools by the proportion of their learners who are black

Exploring the racial distribution of learners by schooling sector, public or independent, reveals that overall, independent schools have a lower proportion of black students. For public schools, the mean proportion of black students is 77%, while for independent schools it falls to 55%. In addition, a Wilcoxon rank-sum test indicates that the distributions of the proportion of black learners across schools is significantly different (P=0.000) across independent and public schools. Public schools are substantially more likely than independent schools to be 100% black, while independent schools are much more likely than public schools to have very low proportions of black children. Figure 4.2 below illustrates the different ways in which black learners are distributed across public and independent schools.

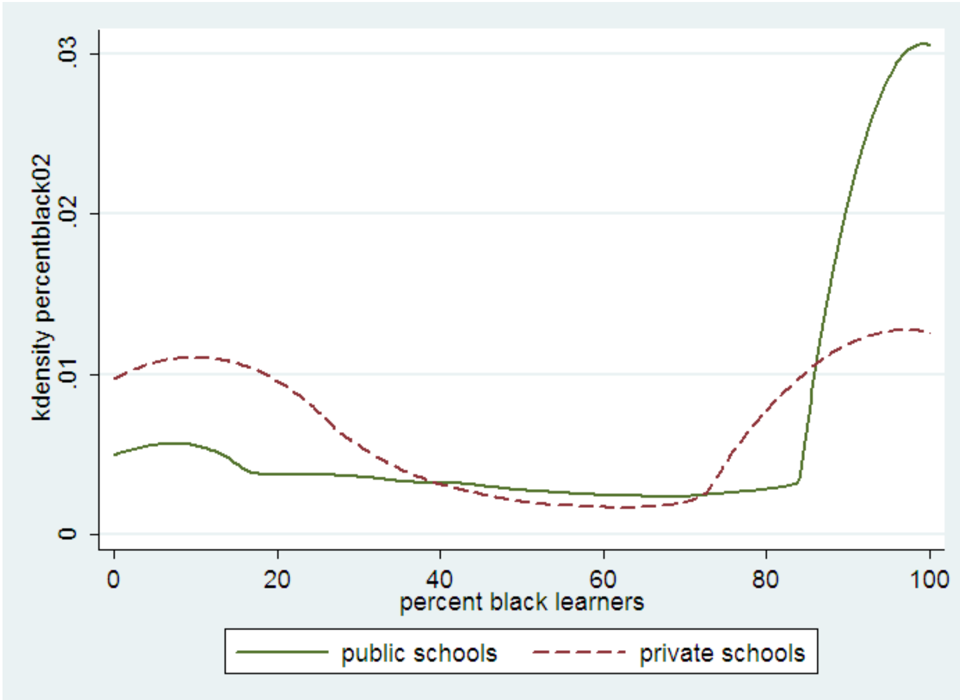

Figure 4.2: Distribution of Gauteng schools by the proportion of their learners who are black

These figures highlight that there remains some accuracy to the perception of many independent schools as 'white' institutions, even as there is evidence of the development of a substantial sub-group of independent schools which serve an entirely black student body (Centre for Development and Enterprise 2010). It seems likely that the disaggregation of independent schools into two groups, one entirely black, and one almost entirely white, is likely to fall largely along the same lines as the division of independent schools into two groups on the basis of performance – one excellent, and the other extremely poor (although the Centre for Development and Enterprise report referred to above does contest this hypothesis).

Within the public sector, the distribution of black children across schools differs between primary and secondary schools, with primary schools being more likely to be entirely black than secondary schools, while secondary schools are more likely to be almost entirely white than primary schools (see Figure 4.3 below). A Wilcoxon rank-sum test finds that the distributions of the proportion of black learners across primary and secondary schools are statistically significantly different (P=0.0112), and an examination of the differences in distributions reveals that while for primary schools the mean proportion of black learners is 78%, this falls to 71% for secondary schools. The differences in the racial composition of student bodies at primary and secondary schools seem likely to be an artifact of the high number of relatively small primary schools and the smaller number of relatively large secondary schools found in historically black areas. Unfortunately it is not possible to test this here. It is also important to point out that the data reflected here is only for one year (2002), during a period of extremely rapid change in South African society and schools, and that the distributions may well have since shifted.

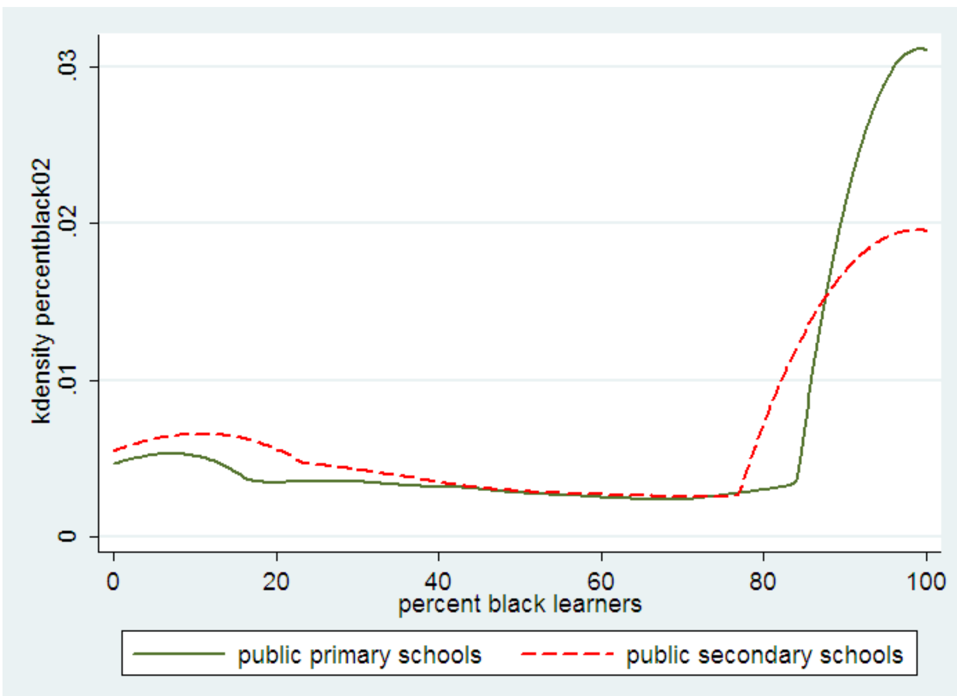

Figure 4.3: Distribution of the proportion of black learners for public primary and secondary schools in Gauteng

The relationships between the proportion of black children in a school, and that school's quintile rating and Section 21 status are more straightforward than those relationships discussed previously, and are highly statistically significant (Kruskal-Wallis test, Pr=0.000). Almost all schools in quintiles 1-3 are entirely black African. In quintile 4, the mean proportion of black children falls to 78%, and in quintile 5, the most advantaged schools, the mean proportion of black children is only 31%. Equally predictably, amongst schools without Section 21 status, almost all are entirely black. By contrast, amongst those with Section 21 status, the distribution is very similar to the distribution across all public schools as described previously.

The evidence around the relationship between proportion of black learners and total school size is mixed. When all schools are considered, there is a very weak positive correlation, which does not reach statistical significance. However, when independent schools are removed from the sample and only public schools are considered, a statistically significant – but not extremely strong – negative correlation emerges. A scatter plot suggests that schools which are predominantly black have a much wider range of sizes than schools which are predominantly white, which tend to be middle-sized. Interestingly, the negative correlation between school size and percent black learners is particularly strong amongst public-sector primary schools – that is, public primary schools with a higher proportion of black learners tend to be smaller than those with a lower proportion of black learners. By contrast, among public secondary schools, there is evidence of a (more weakly) statistically significant positive correlation between school size and percent of learners that are black, suggesting that secondary schools with a higher proportion of black learners are larger than those with a smaller proportion. This tends to suggest that historically disadvantaged primary schools are likely to be particularly small, while historically disadvantaged secondary schools are likely to be particularly large. This will have implications for the average distance from a child's home, in a historically disadvantaged area, to his or her nearest schools – the distance to the nearest primary school is likely to be substantially shorter than the distance to the nearest secondary school. In addition, the child is likely to have access to a larger number of local primary schools than secondary schools.

4.5.6 School fees

Of all the school data considered, the data on school fees is one of the most problematic, and is probably the least reliable. That said, school fees serve as an extremely useful measure of how accessible a school is, and are worth exploring even in the context of imperfect data. Overall, the recorded school fees for 2002 range from R0 to R9900. It seems likely that some schools, especially in the independent sector, were charging higher fees, but for some reason the captured figure was capped at R9900. In the full sample, the mean school fee charged was R1117, however, this obscured an extremely skewed distribution, as revealed by the median school fee of R120. When the schools are separated on the basis of public and independent, it is evident that while a small number of independent schools charged very low or no fees, the majority charged substantial fees, with the median figure being R2500. By contrast, when looking only at public schools, the median falls to R100, and the maximum to R8600.

Within the public sector, school fees also vary substantially by phase, with secondary schools tending to be substantially more expensive than primary schools. For public primary schools, fees range from R0 to R6500, with a mean of R683, and a median of R70. For public secondary schools, the range is from R0 to R8600, with a mean of R1302, and a median of R200. Somewhat predictably, school fees also vary significantly by school quintile rating, with more affluent schools charging substantially higher fees, although there is some anomalous data for quintile 1 schools, which probably reflects poor reporting by those schools (Kruskal-Wallis test, Pr=0.0001). Along similar lines, fees are significantly higher (Wilcoxon rank sum test) at those schools with Section 21 status, as opposed to those without.

There is a statistically significant positive relationship between school fees, and school size (Spearman correlation, Pr=0.0009). Much of this is probably explained by the higher fees typically charged by secondary schools, which are also larger than primary schools. When only public primary schools are examined, however, the relationship between fees charged and enrolment becomes even stronger. By contrast, within the group of public secondary schools, fees tend to fall as enrolment increases. These divergent patterns are consistent with previous data showing that smaller primary schools typically have more black learners, and are therefore likely to be historically disadvantaged and less affluent, while in the secondary phase it is the larger schools that typically have more black learners.

Also in line with this data, direct analysis of the relationship between school fees and the proportion of a school's learners that are black reveals an extremely strong relationship (Spearman correlation, Pr=0.0000). As schools become increasingly black, school fees fall substantially. This pattern is consistent across all schools, public and independent, and within the groups of primary and secondary schools as well. The relationship is stronger within the public sector than the independent sector, however, suggesting that a proportion of black learners attending independent schools may well be buying out of the public sector. Of all the relationships described so far, this negative relationship between school fees and black enrolment is by far the strongest. In the South African context where economic disadvantage and race are so strongly conflated, this is not in the least surprising.

4.5.7 Historical racial status of the school

Of the Gauteng schools for which Apartheid-era department data is available, just under 60% fell under the DET. Although the majority of these schools remain in the public sphere, there are a small proportion of them – about 10% – that have subsequently become independent schools. When only public schools are considered, a significantly higher ($\chi^2_{(1)}=4.1224$, $Pr<0.042$) proportion of primary schools are historically DET schools than the proportion of secondary schools. However, when independent schools are included, this distinction disappears. It is also noticeable that the majority of public schools which are intermediate or combined are historically DET schools. Together, this suggests that fewer DET schools have become public secondary schools than would be expected, and that this may be because a number have moved into the independent sector, while others remain categorized as combined or intermediate schools.

Predictably, historical DET status is strongly linked to quintile ($\chi^2_{(4)}$=961.5693, Pr=0.000), with very few Quintile 4 or 5 schools having historically operated under the DET. Surprising, at first glance, is the fact that by far the largest proportion of historically DET schools are found in Quintile 3. However, given that the DET was an urban department, and that schools in the more rural areas of Gauteng were typically run by other departments, this does in fact make sense. Clearly, historical DET status cannot be used purely as a proxy for low resource levels at a school, as few of the quintile 1 and 2 schools were operated by the DET. However, it can probably operate as a useful proxy for poor schools located in urban contexts. Schools historically operated under the DET were substantially less likely to have obtained Section 21 status by 2002, compared to schools operated under other departments ($\chi^2_{(1)}$=42.5644, Pr=0.000).

Public schools historically operated by the DET typically had lower enrolments in 2002 than other schools (Wilcoxon rank-sum test, Pr=0.0000). When these schools are broken down into primary and secondary schools, however, it become clear that historically DET operated high schools are larger than other high schools (Wilcoxon rank-sum test, Pr=0.0000), while it is only the primary schools that actually tend to be smaller (Wilcoxon rank-sum test, Pr=0.0000).

Historical DET status also provided a strong predictor of the 2002 proportion of a school's learners who were black (Wilcoxon rank-sum, Pr=0.0000). For ex-DET schools, the mean proportion of black students was 99%, while for others it was 39%. This was not substantially different when further disaggregated by phase. In line with all data presented previously, historical DET status was also a strong predictor of school fees, with ex-DET schools charging substantially lower fees than others (Wilcoxon rank-sum, Pr=0.0000).

4.5.8 Matric pass rate

As detailed in Chapter 3, due to the absence of school performance data for primary schools, the matric pass rate of the nearest secondary school was used to approximate performance for primary schools. While highly suboptimal, particularly given recognition that the matric pass rate is often a dubious indicator of school performance in secondary schools, this was the best available option, and as a result is presented here. Figure 4.4 below provides the kernel density plots for the matric pass rates at secondary schools, and the rates extrapolated onto primary schools as discussed in Chapter 3. Overall, both distributions are fairly similar. The one aberration is the much higher proportion of primary schools with pass rates around the 70% level. This is probably related to the trend of a larger number of smaller primary schools in historically black urban areas, where the typical pass rates of high schools are around the 70% level. The similarity of the distributions of matric pass rates applied to primary and secondary schools is also evident when details of the two distributions are examined. Both have means in the low seventies, although the median score in the secondary school distribution is substantially higher (78%) than that for the primary school distribution (72%). Standard deviations and scores at more extreme percentiles are also very similar.

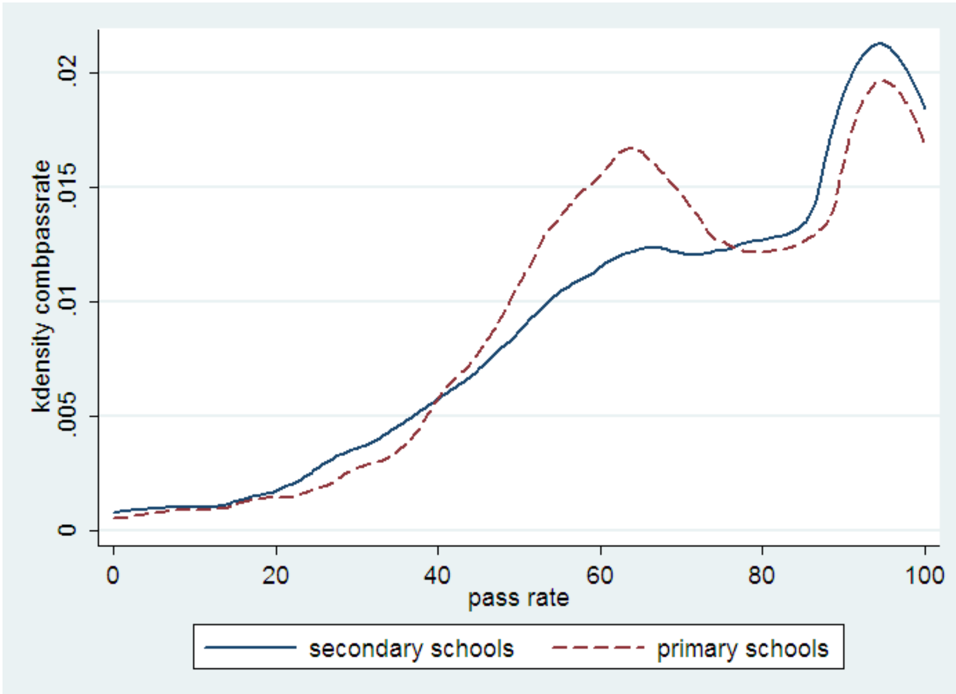

Figure 4.4: Kernel density plots of pass rates at primary and secondary schools

For various reasons, when matric pass rate is explored by school status, public or independent, the data is not very consistent. At the primary school level, the distinction does not make sense, as the imputed performance bears no relationship to whether the school is independent or public, but only to the performance of the secondary school closest to it. At the secondary level, the data is more meaningful, but, as mentioned in Chapter 3, performance data for the more highly performing secondary schools is largely missing. As a result, although performance varies significantly on the basis of whether a school is public or independent (Wilcoxon rank sum test, Pr=0.0951), with independent schools performing more poorly, this is due to a bias in the data and cannot be taken at face value. Due to concerns about the validity of the data for independent schools, the remaining analyses will be conducted on public schools only, unless otherwise specified.

Examining the pass rates of public schools on the basis of their poverty quintile rating reveals, unsurprisingly, a strongly significant difference (Kruskal-Wallis test, Pr=0.0001). Schools in the lowest quintiles (that is, the poorest schools) perform substantially more poorly than those in the higher quintiles, particularly at the extremes. The mean pass rate for quintile 1 schools is 64%, for quintile 2 schools it is 69%, for quintile 3 schools a slightly inconsistent 65%, for quintile 4 schools 72%, and for quintile 5 schools 92%. When primary and secondary schools are examined in isolation, this pattern does not change substantially.

A strongly significant relationship was also identified between section 21 status in public schools and matric pass rates, with section 21 schools performing substantially better than those without section 21 status (Wilcoxon rank-sum test, Pr=0.0001). The mean pass rate was 66% for schools without section 21 status, and 74% for those with this status. The difference between schools with and without section 21 status is greater at the secondary school level.

There is also a statistically significant relationship between pass rate and the size of a school, with larger schools tending to perform somewhat better (Spearman correlation, Pr=0.0005). However, when schools are broken down into primary and secondary, this relationship changes. At the primary school level, the positive relationship between school size and performance persists, whereas at the secondary level this relationship reverses, with smaller schools out-performing larger schools. This almost certainly relates to the high numbers of comparatively smaller primary schools found in less affluent urban areas, and the tendency for the secondary schools in these same areas to be larger than their counterparts in more advantaged areas.

The strongest and most statistically significant relationship identified with pass rate is that with the proportion of a school's learners who are black (Spearman correlation, Pr=0.0000). As the proportion of black learners in a school rises, performance falls. The relationship is strongest for public sector secondary schools, and somewhat weaker at the primary school level. The relationship between school fees and pass rates is almost as strong, with performance rising substantially in the schools charging the highest fees (Spearman correlation, Pr=0.0000). Again this relationship is weaker – though still strong – at the primary school level.

The final analysis undertaken explored pass rate by whether the school was historically a DET school or not. Predictably, performance between DET and non-DET schools was significantly different, with DET schools substantially underperforming all other schools (Wilcoxon rank-sum test, Pr=0.0000). This was consistent across public and independent schools, as well as primary and secondary schools.

4.5.9 Descriptive schools data: discussion

In summary, this descriptive schools data suggests that schools in Gauteng are strongly clustered, with a relatively small group of highly performing, well-resourced and expensive schools, and a much larger group of less well-performing, and more poorly resourced schools. Resource levels, school fees, racial composition of the student body, and school-level academic performance all remain closely related, even in the context of post-Apartheid South Africa. Additionally, these analyses provide evidence that the distribution of primary and high schools with respect to each other is somewhat different in different areas of the Gauteng province. In particular, there are substantially higher numbers of public primary schools than public high schools. In less advantaged urban areas (typically township areas), primary schools tend to be fairly small, and rather densely distributed. In these same areas, high schools tend to be particularly large, and far more sparsely distributed. In more advantaged areas, and amongst independent schools, primary and high schools are far more similar in terms of size, as well as in the numbers of schools available at each phase.

4.6 Conclusion

This chapter began by providing an overview of the study sample with regards to the key explanatory variables examined in this thesis. It then moved on to address issues related to sample representativity and bias. By comparing the study sub-sample to both the full Bt20 cohort, as well as the 2005/6 non-attrition cohort, it was possible to establish that while the study sub-sample does under-represent children at the most advantaged and disadvantaged extremes of the population, this was largely the result of sample attrition, rather than the a result of the method of sub-sample construction. While caution should therefore be taken in applying the findings of the study to the full urban population of Johannesburg-Soweto, the sub-sample appeared to remain largely representative of the black African majority with mid-range levels of SES, who are the group of primary interest with regards to the questions asked in this thesis.

Finally, the third component of the chapter presented descriptive statistics for the schools in the Gauteng province. It provided evidence that schools across the province are far from comparable, varying widely in terms of their student bodies, their fees, their access to resources, and their performance. Historically advantaged schools, typically located in historically advantaged areas, continue to outperform historically disadvantaged schools. All findings were very much in line with other analyses of the South African educational system (Fiske and Ladd 2004; Sujee 2004). In addition, this section provided evidence that in historically disadvantaged urban areas, a large number of relatively small primary schools are found, along with a fairly small number of much larger secondary schools. By comparison, in more advantaged areas, as well as in the independent sector, primary and secondary schools are much closer in size, and more evenly distributed. This has clear implications for the range of schools between which children in different areas are choosing, as well as the distances they are likely to need to travel from home to school.

Chapter 5: Measuring the extent of learner mobility in contemporary urban Johannesburg-Soweto

5.1 Introduction

This chapter answers the first major empirical question asked in this thesis, by presenting data on the extent of leaner mobility in contemporary Johannesburg-Soweto. As discussed in previous chapters, due to the limited levels of knowledge and theory about learner mobility, particularly in the South African context but in the international literature as well, the best approach to measuring learner mobility is not immediately clear. While the majority of existing studies have looked at travel distance or time (Sekete, Shilubane et al. 2001), there is reason to believe that other approaches to measuring mobility, such as whether home and school are in the same area (Msila 2005; Karlsson 2007; Hunter 2010), or whether children attend their nearest school (Msila 2009), may also be important. Therefore, this chapter uses the three different operationalizations of learner mobility discussed previously to measure learner mobility amongst members of the study sample in both 1997 and 2003. Firstly, straight line distance between home and school is used as an indicator of distance. Secondly, whether or not the child lives and attends school in the same 'area' is used as an indicator of whether a child attends a local school. Thirdly, whether or not a child attends his or her nearest grade-appropriate school is used as an indicator of choice.

5.2 Distance-based operationalization of learner mobility

In this section, two different approaches to the use of the straight line distance between a child's home and school in measuring mobility are presented. The first approach simply looks at the distance between home and school, while the second is to use this distance to create binary indicators coded one if a child is travelling further than a particular distance, and zero if he or she is not. This allows us to answer two different question – firstly, how far children are travelling, and secondly, how many children are actually mobile.

5.2.1 Actual straight-line distance from home to school

Comparing datasets

As discussed in Chapter 3, two different variables for the school attended were created for each time point. The first variable for each time point was created using purely prospective schooling data, collected at that particular point in time. The second variable was based on the first, but used additional retrospective data collected at a later point to fill in gaps. For both timepoints, the prospective dataset therefore has a much higher number of missing cases, but is likely to have greater accuracy than the retrospective data, which makes use of recollection at a later point. For this reason, initial explorations of distance travelled to school focused on comparing these two different datasets for each point in time, to establish whether or not they provided satisfactorily similar results. The figures obtained using the different datasets are presented in Table 5.1 below.

Sample	No. of Observations	Mean (km)	Standard Deviation (km)	Minimum (km)	Maximum (km)
1997, small sample	746	5.623	10.798	.007	79.867
1997, small sample	742	5.249	9.544	.007	57.766

constrained at 60km					
1997, large sample	1221	5.901	10.955	.007	105.038
1997, large sample constrained at 60km	1214	5.493	9.524	.007	57.766
2003, small sample	745	5.479	8.725	.045	57.551
2003, large sample	1285	5.625	9.768	.045	105.948
2003, large sample constrained	1281	5.355	8.462	.045	57.551

Table 5.1: Comparison of findings on distances travelled from home to school for different datasets, for 1997 and 2003.

It is immediately apparent that although the minimum and mean distances from home to school using the prospective and retrospective variables are extremely similar, there is a substantial difference in the maximum distances, with much higher maximums presented in the retrospective data. For both time points, however, this is due to a small number of cases in the retrospective variable that travelled further than the maximum distance reported in the prospective dataset. Re-examination of the raw data in these cases did not provide any definitive information as to whether these distances were correct, but in most cases it seemed implausible that a child would travel that distance to attend the school identified, suggesting that perhaps an incorrect school name had been provided, or that the child was in fact not actually resident at the reported home address.

Given that a few extreme outliers would bias the results of any further analysis, it was decided to constrain the data by recoding as missing any reported distances of over 60km between home and school. This removed 7 cases from the 1997 data, and 4 cases from the 2003 data. Once these outliers had been removed, the figures and distributions for the prospective and retrospective data became substantially more similar, as is evident in Table 5.1 above, and Figure 5.1 below.

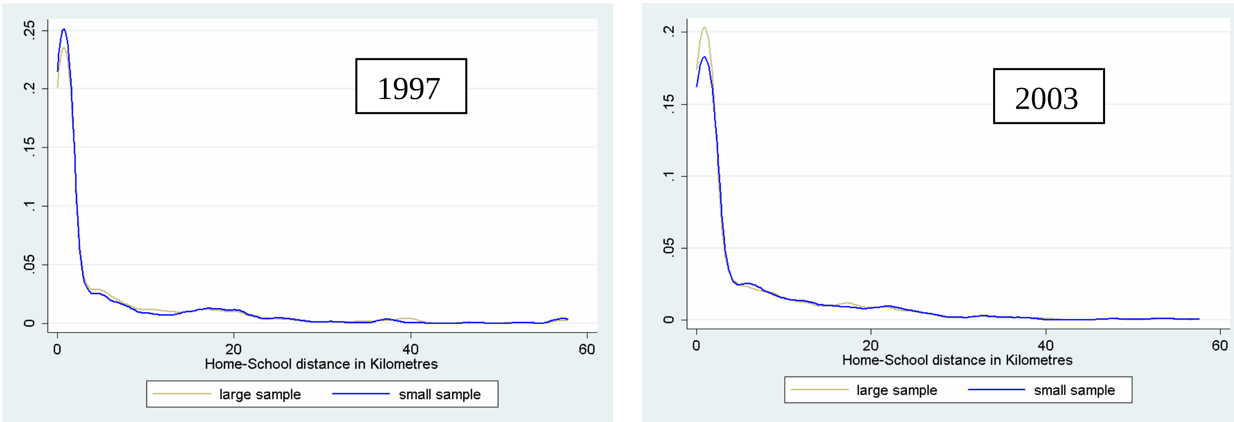

Figure 5.1: Kernel density plots of the distribution of distances to school for the small and large samples, curtailed at a maximum distance of 60km, for both 1997 and 2003.

Examination of the properties of both distributions for 1997, constrained at 60km, reveals that they are extremely similar. The same holds true for both distributions for 2003. Conducting an unmatched t-test on the two different samples for each time point also fails to reject the hypothesis that the means are the same in both cases, further supporting the argument that the distributions are indeed similar. Given the similarity in each year between the small sample and the large sample when the most extreme cases have been removed, all subsequent analysis makes use of these constrained larger samples, unless otherwise specified.

1997 distance from home to school

The distance data for 1997 show a very high concentration of learners at the lowest levels of mobility (see Figure 5.2 and Table 5.2 below). 25% of learners are travelling less than half a kilometre to school, and almost half travel less than a full kilometre. As the distance of travel increases above 1km, however, the distribution begins to spread out substantially. The 75th percentile is reached at just below 6km, and the remaining 25% of the sample forms a long, thin tail reaching out to 60km. In practical terms, this means that although almost half of learners attend a school that is extremely close to their home, there are also almost 25% of learners who travel over 6km – a fairly substantial distance for a 7 year old, and one that almost certainly indicates that these children or their families are making use of school choice, and are investing some financial resources into this choice, at the very least in terms of paying for transportation. These children are also likely to be travelling to schools in communities that differ substantially from those in which the live, particularly with regards to community affluence, resource levels, and historical racial designation.

Given the shape of the distribution of distances travelled, taking a log transformation provides a useful way to compress the tail, and makes the distribution somewhat more normal. Key data regarding the distributions of the distance and transformed distance are provided in Table 5.2 below, and the kernel density plots are provided in Figure 5.2. The log transformation is particularly interesting in that it pulls together the cases spread out over the tail of the untransformed distribution. The second peak, around 3, is the effect of concentrating all these cases, and demonstrates that despite their low density, they do actually form a significant proportion of the distribution when considered together. Although the log transformation still fails standard tests for normality (Shapiro-Wilk in Stata 11), it is far closer to a normal distribution than the original data.

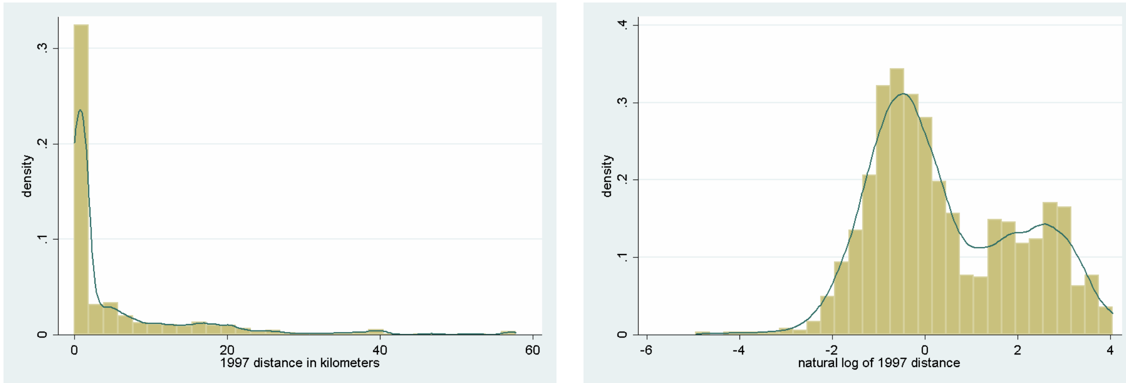

Figure 5.2: Kernel density plot overlaid on histogram illustrating the distribution of distances travelled by sample members in 1997. The log transformation of the distribution is also provided.

1997	25%	50%	75%	95%	Mean
Distance (km)	0.471km	1.032km	5.825km	24.686km	5.493km
Natural log of distance	-0.753	0.031	1.762	3.206	0.454

Table 5.2: Distribution of distances and log distances travelled by sample members

2003 distance from home to school

The distribution of the 2003 data resembles the distribution of the 1997 data very closely, even though the children are substantially older at this time point (13 years, as opposed to 7 years), with a number already enrolled in high school (see Figure 5.3 and Table 5.3 below). The distribution is however slightly more compressed, indicated by the lower mean even as the values at the percentile levels are generally slightly higher. This may relate to the better quality of the data, which has cleared out spurious cases from the tail of the distribution, or it may relate to actual differences in the behaviour of children or the distribution of the relevant schools.

The 2003 data is analysed more closely, controlling for schooling phase (primary or secondary) in Chapter 6, and a detailed comparison of the data for 1997 and 2003 is presented in Chapter 8. For the moment, however, the key points to note are that around 50% of children are attending schools less than

1.25km away from their home, but also that over a quarter of 13 year old children are travelling more than 6.5km to get to school on a daily basis. This is very similar to the data for 1997, which is unanticipated as it was hypothesized that mobility would increase substantially as children aged. Instead, it suggests that a fairly similar (although still high) proportion of children and their families are participating and investing resources in school choice in 2003 as was the case in 1997.

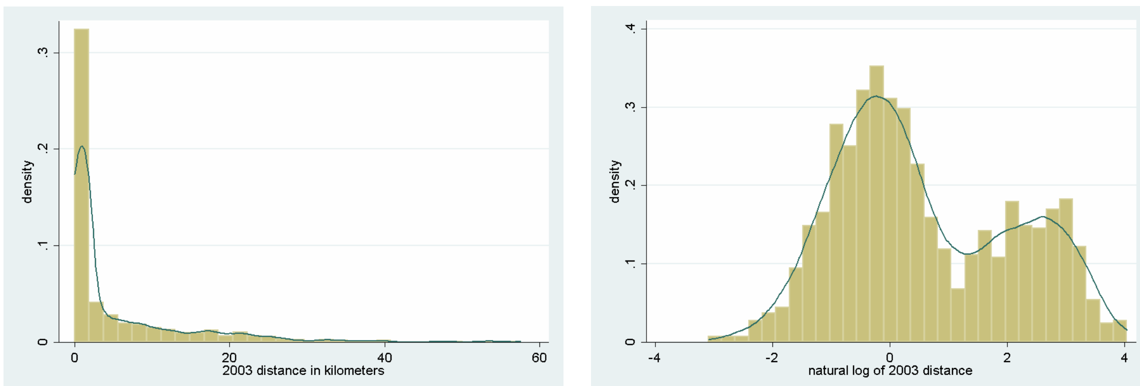

Figure 5.3: Kernel density plot overlaid on histogram illustrating the distribution of distances travelled by sample members in 2003. The log transformation of the distribution is also provided

2003	25%	50%	75%	95%	Mean
Distance (km)	0.578km	1.243km	6.879km	23.362km	5.355km
Natural log of distance	-0.549	0.218	1.928	3.151	0.584

Table 5.3: Distribution of distances and log distances travelled by sample members

Again, taking the log transformation of the distances travelled is helpful in compressing the distribution, and revealing the extent to which cases are concentrated in the tail end of the distribution. Once again, although it still fails tests for normal distribution, the transformed distribution is closer to a normal curve.

Actual straight line distance to school: Conclusion

Overall, calculation of the distances between children's homes and schools reveals two key points. Firstly, the average (one-way) distance from home to school, regardless of which year or sample size is examined, is somewhere between 5 and 6km. Secondly, the distances are distributed as an approximately normal curve centred somewhere between 2 and 3km, with an extremely long and narrow tail to the right representing the roughly 25% of children who appear to live particularly far away from their schools.

5.2.2 Binary definitions of mobility

While examining the actual distances between children's homes and their schools provides more detailed information about how far children are travelling, the use of binary definitions of learner mobility can facilitate the development of policy around learner mobility and school catchment areas, as well as the assessment of the implementation of existing policies. While some information is lost in moving from a continuous measure to a binary definition, analysis and interpretation are also simplified. As discussed in Chapter 3, various cut-off points for the binary definition of learner mobility suggest themselves on the basis of the existing literature and information on the topic, and the analysis presented here makes use of a number of them. 3km is used as this is the maximum distance a learner can travel and still be considered to attend a local school in South African policy (Martin 2010). It is also probably the maximum distance that a young child can be expected to walk to school. 5km and 10km cutoffs are also used, as they are frequently encountered in the local and international literature (Sekete, Shilubane et al. 2001; South African Human Rights Commission 2004). Working with these definitions, and exploring cumulative density plots suggested that various other definitions, particularly around 1, 1.5, 2 and 2.5km would also provide useful information. In all instances, the variable is defined by coding all children travelling up to and including the cut off distance as 0 (not mobile), and all those travelling more than the cut off distance as 1 (mobile).

As the data for 1997 and 2003 were again extremely similar, both time points are discussed together here, and data for both are presented in Table 5.4, below, which provides the numbers and percentages of children who are classified as mobile and non-mobile for each of the various binary definitions of mobility considered. The first important outcome of exploring the various binary definitions of mobility as proposed above is that there is remarkably little difference between the proportions of children defined as mobile across the different definitions, once a distance of 2.5km has been exceeded. After this point, the most striking shift occurs in the interval from 3 to 5 km, where the proportion of children classified as mobile falls from 33.53% to 27.59% (1997) and 33.96% to 28.96% (2003) with an increase in travel distance of a full 2km. While the absolute decreases in mobility from the 5km to the 10km definition are greater, this is spread over an interval of 5km. Given the steepness of the distribution curve prior to 2.5km, the relative flatness of the curve at the 2.5km to 3km interval is somewhat surprising, and indicates that this interval may have some significance.

Mobility definition	**1997: Number (%) mobile**	**2003: Number mobile (%)**
Travel more than 1 km	613 (50.49%)	727 (56.75%)
Travel more than 1.5 km	505 (41.60%)	574 (44.81%)
Travel more than 2km	451 (37.15%)	503 (39.27%)
Travel more than 2.5 km	418 (34.43%)	458 (35.75%)
Travel more than 3km	407 (33.53%)	435 (33.96%)
Travel more than 5km	335 (27.59%)	371 (28.96%)
Travel more than 10km	226 (18.62%)	239 (18.66%)

Table 5.4: Numbers and percentages of children classified as mobile in 1997 and 2003, for each binary definition of mobility considered.

Looking at a plot of the distribution of distances (see Figure 5.4 below) substantiates this indication that something important is happening around the 2-3km interval. The initial, parabolic distribution ends here, and the long flat tail of the distribution seems to begin. Similarly, the slope of the cumulative density function shifts from steep to flat during this interval. That this shift in distributions occurs in the interval between 2 and 3km is fairly compelling for both empirical and theoretical reasons. Empirically, it corresponds roughly to the distance that that a young, school-age child could reasonably be expected to walk to school on a regular basis. Theoretically, the 3km endpoint of this interval corresponds to the South African definition of a local school as being within 3km of a child's home.

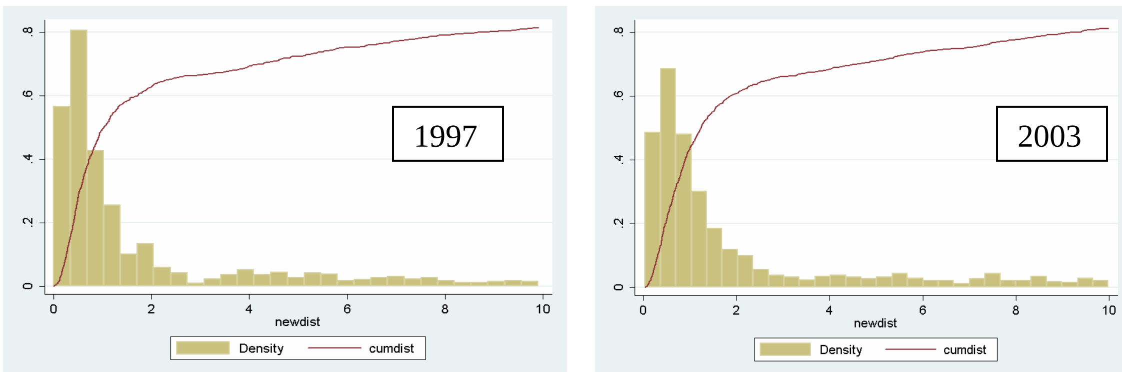

Figure 5.4: Cumulative density plot of distance between home and school, up to 10km, laid over a histogram illustrating the density distribution of distance

The second important outcome of these analyses is that they do indicate that for any of the proposed definitions of mobility, particularly those that are guided by the learner mobility literature, a substantial proportion of children are actually travelling substantial distances on a daily basis. In particular, roughly one third of children are travelling more than 3km. This is pretty clear indication that they are not attending local schools, particularly in an area such as Soweto where there is an extremely high density of public schools. It also suggests – at least in an urban area – that at least a third of children are making use of transportation, whether public or independent, to access schooling. This entails a substantial additional level of family investment in the schooling of these children.

5.3 Area-based operationalization of learner mobility

As described in Chapter 3, the second approach to the operationalization of learner mobility draws on various levels of geographic areas as defined by Census 2001. The smallest area is the Small Area Level (SAL), followed by the Sub-Place (SP), the Main Place (MP), and finally the largest area, the Municipality (MN). Preliminary investigations of the mobility at the SAL level revealed that, due to the small size of SALs, very few children (less than 7%) attended school in the same SAL in which they lived. This makes sense, as very few urban schools are small enough to serve a community of only 200 households, suggesting that SALs are likely to be sharing schools, and as a result, data for mobility at the SAL level is not presented here. Once again, due to the strong similarity between the data for 1997 and 2003, findings for both time points are presented together.

The numbers and proportions of children who are mobile for each of SP, MP and MN are presented in Table 5.5 below, for both 1997 and 2003. The SP level of analysis shows that just over 40% of children attended school in the same SP as they lived in in 1997, and just below 37% in 2003. Given that SP geography is roughly equivalent to residential suburbs, this suggests that around 40% of children are attending a local school within their suburb, while the other 60% are travelling to schools outside of their suburb. At the MP level, which corresponds to major areas of the city (for example Soweto, Meadowlands, Johannesburg, and so on), the proportion of children attending school within the MP where they live rises to over 70% for both 1997 and 2003. Interestingly, the proportion of children travelling across MP boundaries is very similar to the proportion travelling over 5km. Finally, at the MN level, very few children are travelling to a different MN area for school. Given the size of MN areas, and the fact that each MN includes a wide range of schools in terms of performance and cost, this is unsurprising.

	1997: number (%) not mobile (i.e. school and home in same area)	2003: number (%) not mobile (i.e. school and home in same area)
SP	494 (40.63%)	473 (36.90%)
MP	882 (72.53%)	901 (70.28%)
MN	1157 (95.15%)	1,236 (96.41%)

Table 5.5: Number and percent of children who live and attend school in the same SP, MP and MN areas, in 1997 and 2003

The figure for mobility at the MP level is particularly significant, because the boundaries at the MP level of geography correspond most closely to the historical boundaries between areas designated for different race groups. This is critical, because the historical racial designation of a school remains one of the strongest predictors of school performance in contemporary South Africa, and is likely to be one of the major determinants of school choice. Additionally, historical racial group is also a strong predictor of the cost of attending a school. For this reason, those children crossing MP boundaries can be roughly equated to the group that are choosing to attend schools that were historically restricted to white, Indian or coloured children. Those children crossing SP, but not MP, boundaries, can by contrast be roughly equated to those children exercising some degree of school choice but without travelling to areas that were historically designated for other racial groups. This group of just under 30% of children should be roughly equivalent to those who are exercising school choice within historically disadvantaged areas, without leaving those areas.

Determining the correlations between the different possible measures of mobility reveals that there is a very strong overlap between the MP definition, and the distance based definition using travel greater than 5km (see Tables 5.6 and 5.7 below). This substantiates the notion that both of these measures are identifying roughly the same group of children, those travelling fairly substantial distances to attend historically more advantaged schools, and that these measures are therefore likely to be of particular significance.

1997	Travel >2.5km	Travel >3km	Travel >5km	Travel>10km
SAL	0.1963	0.1924	0.1672	0.1296
SP	0.5897	0.5847	0.5114	0.3962
MP	0.7650	0.7700	0.8036	0.7131
MN	0.3063	0.3043	0.3334	0.4241

Table 5.6: Correlation coefficients between distance-based and area-based measures of mobility for 1997

2003	Travel >2.5km	Travel >3km	Travel >5km	Travel>10km
SAL	0.1503	0.1445	0.1287	0.0965
SP	0.5674	0.5486	0.4885	0.3664
MP	0.7850	0.8010	0.8212	0.6980
MN	0.2558	0.2482	0.2614	0.3440

Table 5.7: Correlation coefficients between distance-based and area-based measures of mobility for 2003

5.4 Nearest school based operationalization of learner mobility

The final approach to measuring learner mobility involves determining whether or not children are enrolled at their nearest grade-appropriate school. As discussed in Chapter 3, while this is not a perfect indicator of engagement in school choice, the proportion of children not attending their nearest school is expected to provide a fair approximation of the proportion of children engaging in choice. Again, due to substantial similarities over time, the data for 1997 and 2003 is presented together.

The first key finding using this approach to measuring mobility is that less than 20% of children are actually attending the grade-appropriate school nearest to their homes in both 1997 and 2003 (see Table 5.8 below). This figure is surprisingly low, and suggests that over 80% of children are travelling further than strictly necessary in order to attend school. One possible reason that children might not be attending their nearest grade-appropriate school could be that the school in question is an independent (private) school. Due to this, two sets of figures are presented, one including only public schools, and one including independent schools as well. As is clear from the data in Table 5.8, this makes very little difference to the results.

However, when the 2003 data is disaggregated by schooling phase – that is, when children enrolled in primary school are separated from those enrolled in secondary school – an interesting pattern is revealed. Despite hypotheses that mobility should be higher amongst high school children, a substantially higher proportion of these children are attending their nearest school (just under 22%). The overall proportion of children attending the nearest school remains the same because the proportion of primary school children attending the nearest primary school actually falls fairly markedly to just over 15%. While the higher proportion of high school children attending the closest school may just be due to a smaller number of available high schools (see Chapter 4), the lower proportion of primary school children attending their closest school at age 13 is more intriguing. One potential explanation is that children who are attending schools further afield perform more poorly, making them more likely to still be in primary school at age 13. An alternative, and somewhat more plausible explanation may be that when children fail a grade, their parents are more likely to try sending them to different schools, which may be further from their homes. These hypotheses, and others, will be explored in the subsequent chapters.

	Number (%) of learners attending the school closest to their home	**Mean distance to nearest school**	**Maximum distance to nearest school**
1997 public schools only	219 (17.92%)	0.417km	3.142km
1997 public and independent schools	217 (17.76%)	0.398 km	2.767km
2003, public schools only	235 (18.58%)	0.489km	4.230km
2003, public schools only;	141 (16.49%)	0.428km	3.142km

primary school learners only			
2003, public schools only; high school learners only	94 (22.33%)	0.616km	4.230km
2003, public and independent schools	222 (17.40%)	0.466km	4.230km
2003, public and independent schools; primary school learners only	131 (15.32%)	0.410km	2.767km
2003, public and independent schools; high school learners only	91 (21.62%)	0.583km	4.230km

Table 5.8: Number and percentage of learners attending the school closest to their home in 1997 and 2003, and the mean and maximum distances to the schools nearest to sample members' homes

The data on the distance from children's homes to their nearest schools provides an additional interesting finding: the mean distance a child needs to travel to attend their nearest primary phase school is just approximately 400m, and less than 5% of children need to travel over 1km. When contrasted to the actual distances children are travelling – previous calculations indicated over 50% of children travelling over 1km – this highlights the extent to which travel, even of moderate levels, appears to be due to children attending schools further from home than is strictly necessary.

In 2003, however, not all children are still within easy walking distance of a grade-appropriate school. Although 95% of children in 2003 have to travel less than 1.15km to reach their nearest school, there are a small number of children who have to travel over 3km. This is probably primarily due to the fact, discussed in Chapter 4, that there are substantially fewer high schools in the Johannesburg-Soweto area, due to their typically having a somewhat larger size. Disaggregating the children by schooling phase supports this hypothesis, as the data generated for the primary school children remains similar to that generated in 1997. Nonetheless, even when children are in high school, the average distance travelled remains substantially greater than the distance a child would need to travel to access his or her nearest school.

5.5 Conclusion:

This chapter has explored three different approaches to defining and measuring learner mobility, and provided data about the extent of learner mobility in Johannesburg-Soweto on the basis of each of these definitions. Each definition is likely to prove particularly valuable for certain purposes, and in certain contexts. Using a distance-based measure provides both a binary and a continuous measure of mobility. The distance-based binary measure is of the form that is typically used in school choice related policy, and is therefore particularly valuable in assessments of the appropriateness or applicability of policy. The continuous measure of distance is particularly useful in exploring the actual extent of mobility and what it entails for particular learners in terms of the investments they are required to make, both financially and in terms of time. In addition, it allows for the examination of the distribution of the distances travelled by the entire sample, and, as it is the measure that has been most commonly used in the existing literature, it also allows for comparison with previous findings.

The definition of mobility based on census geography is particularly useful in that it makes use of generally accepted geographical areas to explore the extent to which mobility is occurring within and between these areas. This is helpful in identifying whether learners are travelling between areas historically designated for different race groups, and thereby significantly enhancing the quality of education they are likely to receive. Additionally, it is, and thus identifying those learners who are likely to be making the most substantial economic investments in their education. Finally, the definition based on whether or not the learner is attending the age-appropriate school closest to his or her home is useful in highlighting the extent to which even learners with relatively low levels of mobility may be engaging in more travel than strictly necessary or anticipated, and may also be engaging in school choice, particularly within the historically disadvantaged areas.

This chapter has made two key contributions to the literature. Firstly, in providing three different approaches to the conceptualization and measurement of learner mobility, it has significantly enhanced the methodological tools available to the study of this practice. Secondly, it has, for the first time, provided population-based data on the extent of learner mobility in contemporary urban South Africa. In so doing, it has identified preliminary evidence to suggest that there may in fact be two patterns of school choice and mobility in operation in Johannesburg-Soweto. Firstly, there is a group of approximately 25% of the sample who are engaged in substantial travel from home to school on a daily basis, and who seem likely to be making significant investments in this mobility. Secondly, and somewhat less expectedly, there is also evidence that a large proportion of children who are not travelling substantial distances to school are still engaging in mobility and school choice. Even though they are attending schools relatively close to home, they are not attending the nearest grade-appropriate school to their home, and are often travelling to schools that are not located in the same residential areas as their homes. These patterns, and their importance to understanding the implications of learner mobility to educational access and equality, are explored in greater detail in the subsequent chapters.

Chapter 6: Individual, family and community characteristics of mobile learners

6.1 Introduction

The previous chapter illustrated the extent of learner mobility in post-Apartheid Johannesburg-Soweto. Even using the most stringent definitions of mobility, the numbers of children engaged in mobility are substantial. Having developed an understanding of the extent to which learner mobility is taking place amongst school-age children in Johannesburg-Soweto, along with a clearer idea of exactly what this mobility entails, two further questions arise. The first relates to which children in particular are most likely to be engaging in learner mobility. The second relates to the characteristics of schools these children are choosing to attend, and the types of schools they are travelling to avoid. This chapter addresses the first of these questions, while the second will be addressed in Chapter 7.

A child's educational mobility is expected to be closely related to the level of investment the child's family makes in his or her education. For this reason, it is useful to explore variation in both a family's access to resources to invest in education, and in variables which might be associated with a family's propensity to invest in education. This chapter explores the relationship to mobility of a range of variables at the levels of the individual child, the child's household, and his or her community. The child characteristics considered are race, gender, age at first enrolment in school, grade repetition and schooling phase. Family characteristics examined are maternal education, maternal marital status, and household SES, in both 1997 and 2003. At the community level, SAL, SP and MP poverty levels are examined.

Child level characteristics

6.2 Race

Given the strong relationship between race, access to resources, and area of residence in South Africa, it is anticipated that race is related to educational mobility behaviours (Fiske and Ladd 2004; Fiske and Ladd 2005). While white children are most likely to have access to the resources required to engage in educational mobility, they are also least likely to need to engage in it to access good schools, given that most highly performing schools are located in historically white areas. Black children, by contrast, are likely to have the greatest incentives to engage in mobility, typically living in the areas with poorest schools, but are simultaneously least likely to have access to the necessary resources. Indian and coloured children are likely to fall somewhere in between the black and white children in terms of both incentives and ability to engage in mobility. As the numbers of white and Indian children present in the study sub-sample (28 and 25, respectively) are extremely small, findings for these groups are unlikely to be broadly representative, and are therefore not presented here. Discussion will be limited to the behaviour of black and coloured children.

6.2.1 1997

Straight-line distance

Examining the distance between home and school on the basis of race reveals a strong relationship with race. As is evident in Table 6.1, black children tend to travel substantially further to school than coloured children (Wilcoxon rank-sum, Pr=0.0000). A kernel density plot (see Figure 6.1 below), illustrates just how different the distances from home to school are for black and coloured children. The kernel density plot for the coloured children is far more concentrated at very low levels of travel for coloured children than for their black peers.

Race	Number of children	Mean distance to school	Standard deviation	25th percentile distance	50th percentile distance	75th percentile distance
Black African	1002	5.773km	9.589km	0.492km	1.138km	6.904km
Coloured	159	3.935km	9.533km	0.360km	0.561km	1.834km

Table 6.1: 1997 Distance between home and school on the basis of race

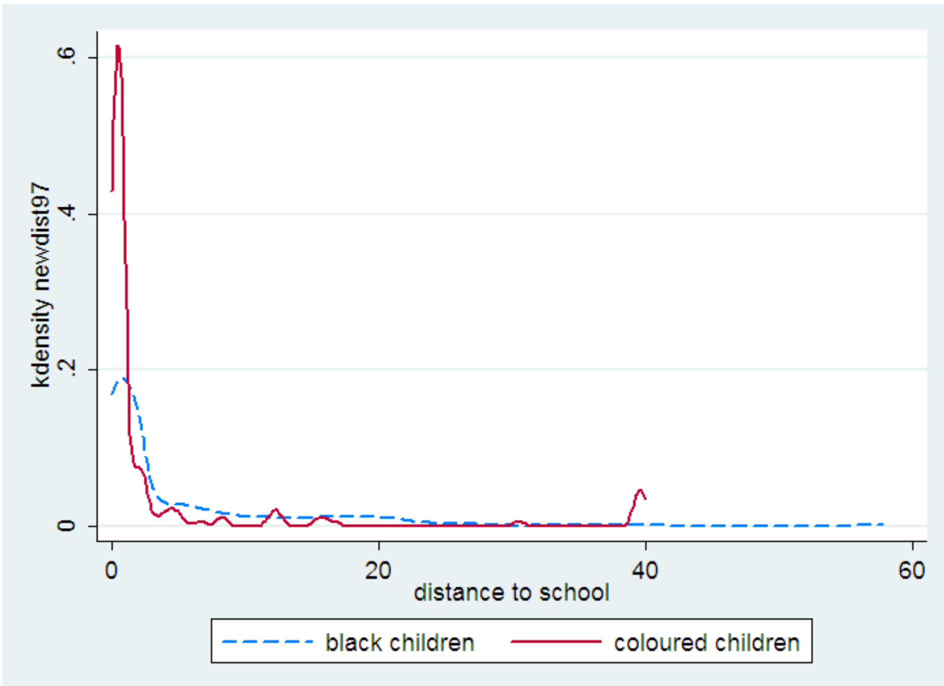

Figure 6.1: Kernel density plot of distance to school in 1997, on the basis of race

Census geography

An area-based approach to measuring mobility finds similar patterns (see Table 6.2 below). At all levels of geography other than MN, coloured children are substantially more likely to live and attend school in the same area than black children.

	Black (n=1003)	Coloured(n=160)	χ^2
School and home in same SAL	48 (4.79%	29 (18.31%)	$\chi^2_{(1)}$ = 39.7146, Pr=0.000
School and home in	373	105	$\chi^2_{(1)}$ = 46.0932,

same SP	(37.19%)	(65.63%)	Pr=0.000
School and home in same MP	696 (69.39%)	141 (88.13%)	$\chi^2_{(1)}$ = 24.0038, Pr=0.000
School and home in same MN	957 (95.41%)	150 (93.75%)	$\chi^2_{(1)}$ = 0.8334, N.S.

Table 6.2: 1997 mobility at different levels of census geography, by race

Nearest school

Mobility analyses exploring whether or not children attended their nearest grade-appropriate schools again provided similar results, with coloured children being substantially more likely to attend their nearest school than black children (see Table 6.3). The analysis was conducted firstly using only public schools, and secondly using both public and independent schools, and in both cases similar figures were obtained. The finding that black children are the least likely to attend their nearest school makes sense given that they are likely to live in the poorest areas (as described in Appendix 3), and their nearest school is therefore more likely to be particularly poorly performing.

	Black (n=1009)	**Coloured (n=155)**	**χ^2**
Child attends nearest school (public or independent)	146 (14.47%)	60 (38.71%)	$\chi^2_{(1)}$ = 54.2007, Pr=0.000
Child attends nearest school (public only)	143 (14.17%)	62 (38.75%)	$\chi^2_{(1)}$ = 57.6862, Pr=0.000

Table 6.3: Children attending their nearest grade-appropriate school, by race, for public schools only, and for all schools

6.2.2 2003

Straight-line distance

As evident in Table 6.4, the distances from home to school in 2003 are slightly different from those in 1997, with an increase in the difference between the mean distances travelled by black and coloured children (Wilcoxon rank-sum, Pr=0.0000). In Figure 6.2, the extent to which coloured children are more likely to live very close to their school than black children is highly evident.

Race	Number of children	Mean distance to school	Standard deviation	25th percentile distance	50th percentile distance	75th percentile distance
Black African	1065	5.834km	8.617km	0.635km	1.380km	8.090km
Coloured	163	2.625km	7.121km	0.402km	0.668km	1.315km

Table 6.4: 2003 Distance between home and school on the basis of race

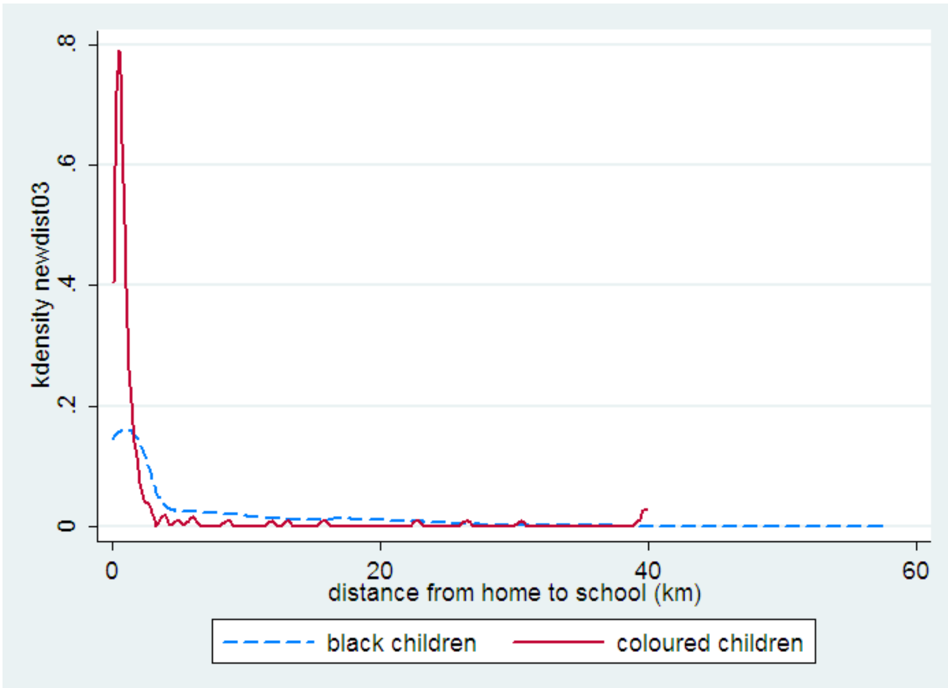

Figure 6.2: Kernel density plot of distance to school in 1997, on the basis of race

Census Geography

The 2003 area-based analysis (see Table 6.5 below) provides very similar results to the 1997 analysis. Once again, coloured children are much more like to attend school in the same SAL, SP and MP areas in which they live than black children.

	Black (n=1066)	Coloured (n=163)	χ^2
School and home in same SAL	28 (2.63%)	16 (9.82%)	$\chi^2_{(1)} = 72.4870$ Pr=0.000
School and home in same SP	350 (32.83%)	110 (67.48)	$\chi^2_{(1)} = 72.4870$ Pr=0.000

School and home in same MP	702 (65.95%)	150 (92.02%)	$\chi^2_{(1)} = 45.5357$ Pr=0.000
School and home in same MN	1026 (96.25%)	159 (97.55%)	$\chi^2_{(1)} = 0.6904$ N.S.

Table 6.5: 2003 Mobility across different levels of census geography, by race

Nearest school

Coloured children remain significantly more likely to attend their nearest grade-appropriate school than black children in 2003 (see Table 6.6 below).

	Black (n=1009)	**Coloured (n=155)**	**χ2**
Child attends nearest school (public or independent)	154 (14.50%)	61 (39.35%)	$\chi^2_{(1)}= 57.4420$ Pr=0.000
Child attends nearest school (public only)	161 (15.16%)	66 (40.99%)	$\chi^2_{(1)}= 61.7256$ Pr=0.000

Table 6.6: 2003 Children attending their nearest grade-appropriate school, by race, both for public schools only, and for all schools

6.2.3 Race and mobility discussion

There is strong evidence that both in 1997 and 2003, race is closely related to mobility behaviour, regardless of the way in which mobility is measured. Overall, black children appear to be substantially more engaged in all forms of learner mobility than coloured children. There are a range of possible explanations for this, including that coloured children live in areas with better schools, that the coloured community is more cohesive and prefer to keep their children at local schools, or that coloured families are less likely to want to make substantial investments in their children's education for various reasons.

6.3 Gender

It is possible that families approach the education of girls and boys differently. Certainly, in contemporary South Africa, girls are known to remain in formal education longer, and also tend to outperform boys (Unterhalter 2005; Fleisch and Schindler 2009). During the Apartheid era, both policy and practice favoured different approaches to education on the basis of gender (Fiske and Ladd 2004), and some legacy of this might be expected to persist, particularly around the levels of investment in the education of children of different genders. To determine whether mobility, and by extension, educational investment, differs on the basis of gender, male and female populations were compared, using different definitions of mobility.

6.3.1 1997

Straight-line distance

Examining the distribution of distance by gender does suggest girls, on average, travel slightly further than boys (see Table 6.7 and Figure 6.3 below). Closer examination of the data, however, seems to suggest that this difference, particularly evident in the means, may be caused primarily by a cluster of girls travelling fairly substantial distances, particularly between 20 and 60km, pulling the overall mean for girls (along with the standard deviation and percentile breaks) upwards. A Wilcoxon rank-sum (Mann-Whitney) test fails to find any significant difference in the distribution of distance from home to school on the basis of gender.

Gender	Number of children	Mean distance to school	Standard deviation	25th percentile distance	50th percentile distance	75th percentile distance
Boys	592	4.976 km	8.782 km	0.466 km	0.941 km	4.997 km
Girls	622	5.985 km	10.163 km	0.475 km	1.086 km	6.928 km

Table 6.7: 1997 distance from home to school, by gender

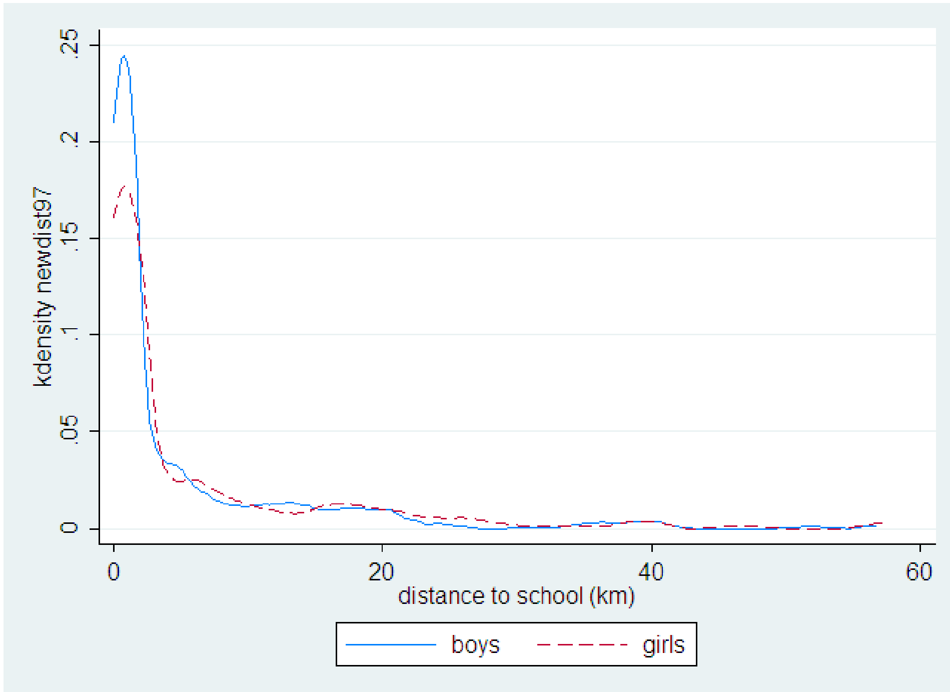

Figure 6.3: Kernel density plot of 1997 distance from home to school, by gender

If distance from home to school is grouped into categories, a chi-square test does, however, reveal a significant difference between boys and girls (see Table 6.8 below). Overall, while distributions are fairly similar, boys are somewhat more likely to be travelling extremely short distances, and girls are somewhat more likely to be travelling distances over 20km.

	Up to 1km	**1km-2.5km**	**2.5km-5km**	**5km-10km**	**10km-20km**	**Over 20km**	
Boys (n=592)	306 (51.69%)	88 (14.86%)	50 (8.45%)	47 (7.94%)	66 (11.15%)	35 (5.91%)	$\chi^2_{(5)}$ =12.156, Pr=0.033
Girls (n=622)	295 (47.43%)	107 (17.20%)	33 (5.31%)	62 (9.97%)	68 (10.93%)	57 (9.16%)	

Table 6.8: Gender breakdown of 1997 categories of distance from home to school

Census geography

Girls were significantly more likely to attend a school in the same SAL in which they lived than boys (see Table 6.9 below). At all other levels of geography, however, there was no evidence that girls and boys behaved differently.

	Boys (n=592)	Girls (n=622)	χ^2
School and home in same SAL	33 (5.56%)	50 (8.03%)	$\chi^2_{(1)}=2.893$ Pr = 0.089
School and home in same SP	242 (40.81%)	252 (40.45%)	$\chi^2_{(1)}=0.0163$ Not significant
School and home in same MP	439 (74.03%)	443 (71.11%)	$\chi^2_{(1)}=1.303$ Not significant
School and home in same MN	570 (96.12%)	587 (94.22%)	$\chi^2_{(1)}=2.376$ Not significant

Table 6.9: 1997 mobility across different levels of census geography, by race

Nearest School

Chi-squared tests provided no indication for any gender differences in the likelihood of children attending their nearest school, regardless of whether independent schools were included in the analysis or not.

6.3.2 2003

Straight-line distance

Overall, patterns of mobility by gender in 2003 remained largely consistent with those identified in 1997, although the mean distance travelled to school by girls did decrease slightly (see Table 6.10). Nonetheless, girls still continue to travel, on average, almost a kilometre further than boys. Interestingly, the standard deviation on the distances travelled by girls has fallen quite substantially, approaching fairly closely the standard deviation on the distances travelled by boys. For both genders, the percentile distances have increased slightly, with the effect more noticeable for girls, particularly from the 75th percentile up. In sum, this suggests that distribution of travel distances for girls may have spread out slightly towards the tail end (representing greater distances), with proportionally fewer girls continuing to travel particularly short distances, although this effect is not large enough to be evident on the kernel density plot of the distributions of distance travelled by gender (see Figure 6.4 below). A Wilcoxon rank-sum test indicates that girls travel further than boys in 2003 (Pr= 0.0489).

Gender	Number of observations	Mean distance to school	Standard deviation	25th percentile distance	50th percentile distance	75th percentile distance
Boys	631	4.953469	8.224416	0.522km	1.175km	5.524km
Girls	650	5.745164	8.67542	0.619km	1.303km	7.446km

Table 6.10: 1997 distance from home to school, by gender

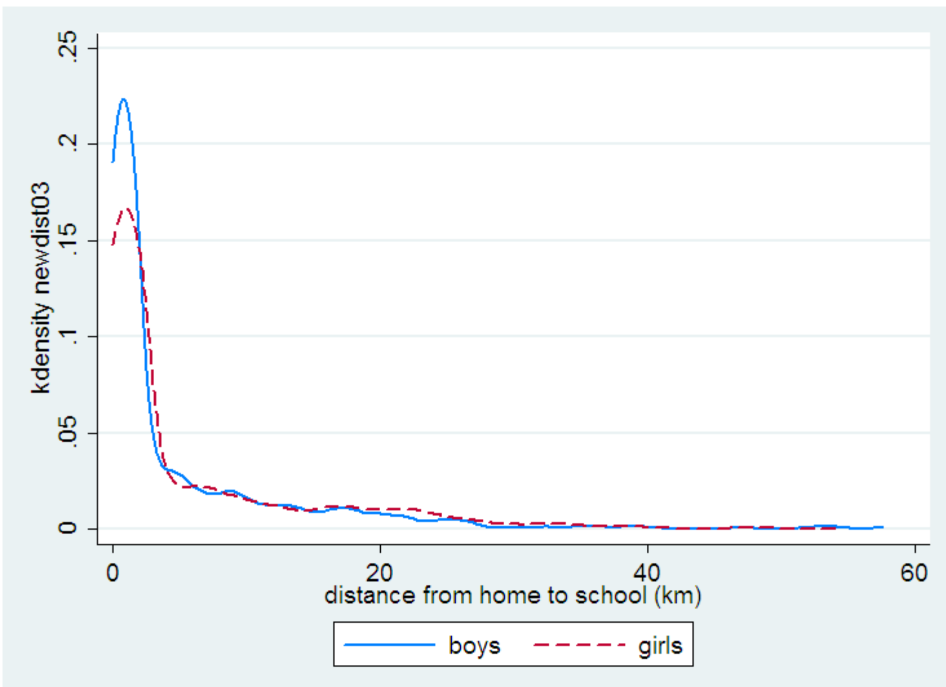

Figure 6.4: Kernel density plot of 2003 distance from home to school, by gender

In 2003, the chi-square analysis of distance grouped into categories no longer reveals any significant difference between boys and girls, although a higher percentage of girls continue to travel particularly great distances (see Table 6.11 below).

	Up to 1km	1km-2.5km	2.5km-5km	5km-10km	10km-20km	Over 20km	
Boys (n=631)	290 (45.96%)	123 (19.49%)	47 (7.45%)	65 (10.30%)	66 (10.46%)	40 (6.34%)	$\chi^2_{(5)}$ =8.428; Not significant
Girls (n=650)	264 (40.62%)	146 (22.46%)	40 (6.15%)	67 (10.31%)	71 (10.92%)	62 (9.54%)	

Table 6.11: Gender breakdown of 2003 categories of distance from home to school

Census geography

Examining mobility by census area in 2003 reveals that the gender difference previously evident at the SAL has disappeared (see Table 6.12). A gender difference at the level of SP level has emerged, although in the opposite direction, with boys being more likely to attend school in the same SP where they live. There is no evidence of any gender difference at the MP or MN levels.

	Boys (n=631)	Girls (n=651)	χ^2
School and home in same SAL	24 (3.80%)	26 (3.99%)	$\chi^2_{(1)}$ =0.031 Not significant
School and home in same SP	247 (39.14%)	226 (34.72%)	$\chi^2_{(1)}$ =2.699 Pr =.100
School and home in same MP	452 (71.63%)	449 (68.97%)	$\chi^2_{(1)}$ =1.087 Not significant
School and home in same MN	609 (96.51%)	627 (96.31%)	$\chi^2_{(1)}$ =0.037 Not significant

Table 6.12: 2003 mobility across different levels of census geography, by gender

Nearest school

As with 1997, there is no evidence in that children of either gender were more likely to attend their nearest school in 2003, regardless of whether independent schools are included in the analysis or not.

6.3.3 Gender and mobility discussion

There was some evidence that girls travelled, on average, further than boys in both 1997 and 2003, but this was extremely sensitive to the way in which mobility was measured. If girls are travelling further, one possible explanation is that parents value the education of girl children more highly, and therefore tend to invest more in their education. An alternative explanation may be that girl children are more likely to travel with parents, and attend school close to a parents' work place, perhaps due to safety concerns. Overall, however, the data presented here does not provide conclusive evidence for any substantial difference in mobility on the basis of gender.

6.4 Age at first school enrolment

In South Africa, children have a legal window of two years during which to start their schooling. It is possible that the point during this window at which children begin their formal schooling relates to the level of interest or commitment that parents feel towards their child's schooling, with more committed parents enrolling children earlier. By contrast, it may also relate to a parent's ability to fulfil care-giving responsibilities, in which case parents with fewer resources may be more likely to pursue the earliest possible enrolment of their children to reduce their care-giving burden. It may also relate to different enrolment and application policies applied in different schools, with more selective schools preferring to enrol older and more independent children. It is therefore conceivable that the distance a child travels to school is connected to their age at first school enrolment, although the expected direction of this relationship is not evident.

6.4.1 1997

Straight-line distance

Table 6.13 and Figure 6.5 show that children who start school at a later age travel significantly further than those who start at an earlier age (Wilcoxon rank-sum test; Pr= 0.0004).

	Number of children	Mean distance to school	Standard deviation	25th percentile distance	50th percentile distance	75th percentile distance
Earlier starters	646	4.704km	9.216km	0.449km	0.880km	4.569km
Later starters	557	6.482km	9.866km	0.498km	1.221km	9.044km

Table 6.13: 1997 distance from home to school, by age at first enrollment

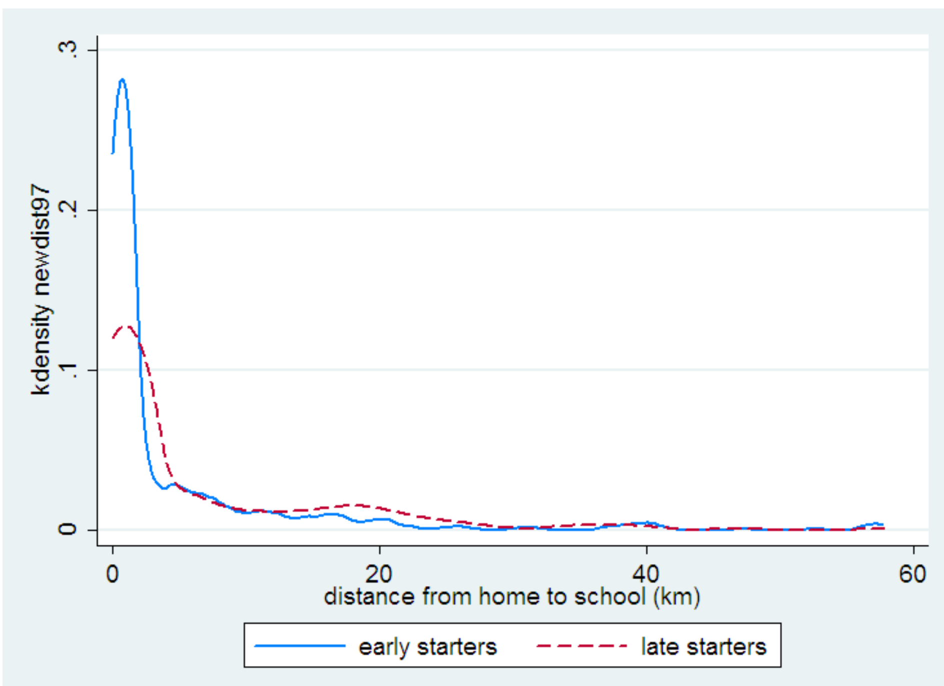

Figure 6.5: Kernel density plot of 1997 distance from home to school, by age at first school enrolment

Census geography

Although there was no relationship between starting school late and travelling between SALs or MNs for schooling, children who started school early were significantly more likely to school within their residential SP ($\chi^2_{(1)}$=10.8425; Pr=0.001), as well as in their residential MP ($\chi^2_{(1)}$=9.8651; Pr=0.002).

Nearest school

There was no significant relationship between age at first enrolment and whether or not a child attended their closest grade-appropriate school, regardless of whether independent schools are included or excluded.

6.4.2 2003

Straight-line distance

In 2003, although late starters still travel further on average than early starters, this difference is no longer statistically significant (Wilcoxon rank-sum test). Additionally, although the distance for the 75th percentile of late starters is still substantially higher than for early starters, at the 25th and 50th percentile, the early starters are actually travelling further. The distribution for late starters is therefore wider, but slightly flatter, than that for early starters (see Table 6.14 and Figure 6.6 below). One potential explanation for this change over time is that by 2003, more of the children who started school early are enrolled in high school, and that this requires them to travel somewhat further. This hypothesis is explored in the next section of this chapter, on the relationship between schooling phase in 2003 and mobility behaviours.

	Number of children	Mean distance to school	Standard deviation	25th percentile distance	50th percentile distance	75th percentile distance
Earlier starters	644	5.122km	8.143km	0.615km	1.288km	6.187km
Later starters	563	5.750km	8.867km	0.567km	1.181km	8.077km

Table 6.14: 2003 distance from home to school, by age at first enrolment

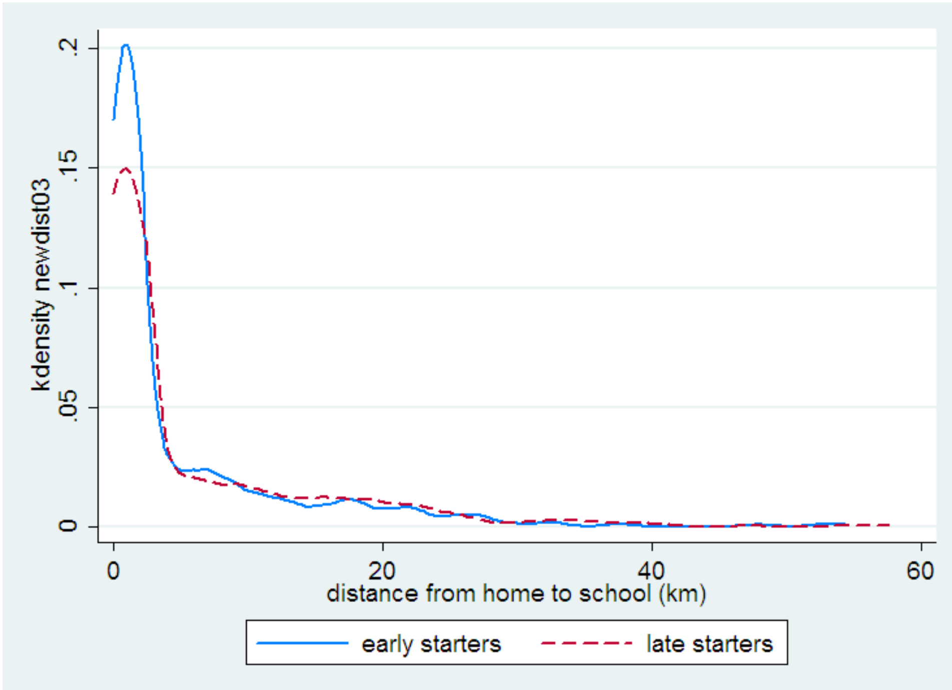

Figure 6.6: Kernel density plot of 2003 distance from home to school by age at first school enrolment

Census geography analysis

By 2003, there is no longer any evidence for differential mobility at any level of census geography on the basis of whether a child started school early or late.

Nearest school

There is a weakly significant relationship between whether a child starts school late, and whether he or she attends his or her nearest grade-appropriate school in 2003, but only when both public and independent schools are considered ($\chi^2_{(1)}$= 2.9904, Pr = 0.084), with children starting late being less likely to be attending their nearest school in 2003. When only public schools are considered, there is no significant relationship.

6.4.3 Age at first enrolment and mobility discussion

In 1997, there is a relationship between mobility and whether a child starts school earlier or later, with children starting later being more likely to travel further. By 2003, this relationship has however largely disappeared. This may relate to the fact that a proportion of those who started school early have begun attending secondary schools, while those who started late are almost all still in primary schools. The relationship between mobility and schooling phase is explored further in the next section of this chapter.

There is no clear and obvious explanation for why, in 1997, children who started school late are likely to have a greater distance between their homes and their schools. It may be the case that parents who plan to send their children to schools further afield, necessitating independent travel, as well as in many cases the ability to adapt to a different cultural environment, are waiting until children are slightly older before enrolling them in school. It may also relate to enrolment practices at more advantaged schools, were children are in some instances required to take entrance tests or undergo interviews. A third potential explanation is that less affluent parents, who are not able to send children to schools far from home, may also not be able to afford pre-school or child care for their children, and therefore prefer to send them to school as early as possible. More affluent parents, by contrast, may be less pressed to enrol children in primary school, preferring to ensure that children are genuinely school-ready.

6.5 School phase in 2003

As discussed in Chapter 4, there is evidence that those children who have reached secondary school in 2003 differ systematically from those that have not on the basis of race, gender, age at first enrolment, grade repetition, maternal education, and 1997 SES. All of these variables have been hypothesized to have some relationship to mobility, and as a result, children who have reached high school in 2003 may be exhibiting different mobility behaviours to other children simply because of this. On the other hand, high school status may in itself have implications for mobility behaviour for a range of reasons.

As described in Chapter 4, the Gauteng province contains a fairly small number of large high schools, and a much larger number of much smaller primary schools. Due to this, all else being constant, we would expect to see high school children travelling slightly further to school on average, and we would also expect a higher proportion of high school children to attend their nearest school. The descriptive mobility data presented in Chapter 5 suggests that this is indeed the case.

An additional reason to anticipate changes in mobility behaviour between primary and secondary schooling is that the costs and benefits of mobility may change, in turn changing preferences around the selection of schools. For example, children of high school age can travel greater distances, on their own, more safely than younger children, and may also be able to walk further, lowering the cost of attending a more distant school. A factor in favour of stability in mobility behaviours between the primary and secondary schooling years, however, is any element of path dependence. For example, secondary schools may give preference to children from local primary schools, and those primary schools might likewise encourage children to enrol in local high schools. The difficulty of moving between schools in different areas may be much higher for older children.

Preferences may also be shaped through different criteria at the secondary school level. For example, the academic performance of a secondary school is more immediately salient than the performance of a primary school, largely due to the availability of some information about the matric exam pass rate. The greater availability of evidence with regards to school academic performance may affect the importance attributed to school academic performance in school selection. At the high school level, children may also be far more actively involved in the selection of their school, and may well be driven by different priorities than those used by their parents in selecting a primary school. By contrast, however, secondary schooling is typically more expensive than primary schooling, and this may influence parents to maintain or even increase their role in school selection. Finally, secondary schooling may simply be attributed greater value than primary schooling, changing the level of investment which families and children are willing to make in education. It is not immediately evident, however, in which direction these potential changes in school preferences should influence mobility when aggregated.

Differences between the mobility behaviours in 2003 of children who have reached high school, and those who haven't, could be attributed either to their being different with regards to variables associated with mobility, or alternatively simply to the fact that they have reached high school. It is also possible that any difference is due to a combination of these factors. Due to the nature of the sample used in this study, in which only a relatively small, non-random group of children has reached high school by 2003, the relative contributions of individual and household variables on mobility cannot be separated out from any independent effect of high school status. Really untangling the extent to which schooling phase shapes mobility independent of socio-economic and other individual and family-level variables could be approached either through the use of data for years beyond 2003 for the current sample, or through the use of a broader, or differently structured sample.

However, the available data does provide one way of obtaining some insight into this issue, by looking at whether the two groups of children had similar mobility behaviours in 1997 or not. If they behaved similarly in 1997, it seems likely that more of the difference in 2003 can be attributed to schooling phase. By contrast, if behaviour was already different in 1997, this suggests a more important role for individual, family and community variables associated with mobility. This data is presented below.

6.5.1 1997

Straight-line distance

Children who were still in primary school in 2003, and those who had reached secondary school in 2003, had very similar distances from home to school in 1997, as evident in Table 6.15 and Figure 6.7. A Wilcoxon rank-sum test confirmed that distance from home to school in 1997 is not statistically significantly different for the two groups.

	Number of children	**Mean distance to school**	**Standard deviation**	**25^{th} percentile distance**	**50^{th} percentile distance**	**75^{th} percentile distance**
Primary school in 2003	780 (65.60%)	5.729km	9.350km	0.469km	1.132km	6.739km
High school in 2003	409 (34.40%)	5.285km	10.073km	0.466km	0.910km	5.414km

Table 6.15: 1997 distance from home to school, by schooling phase in 2003

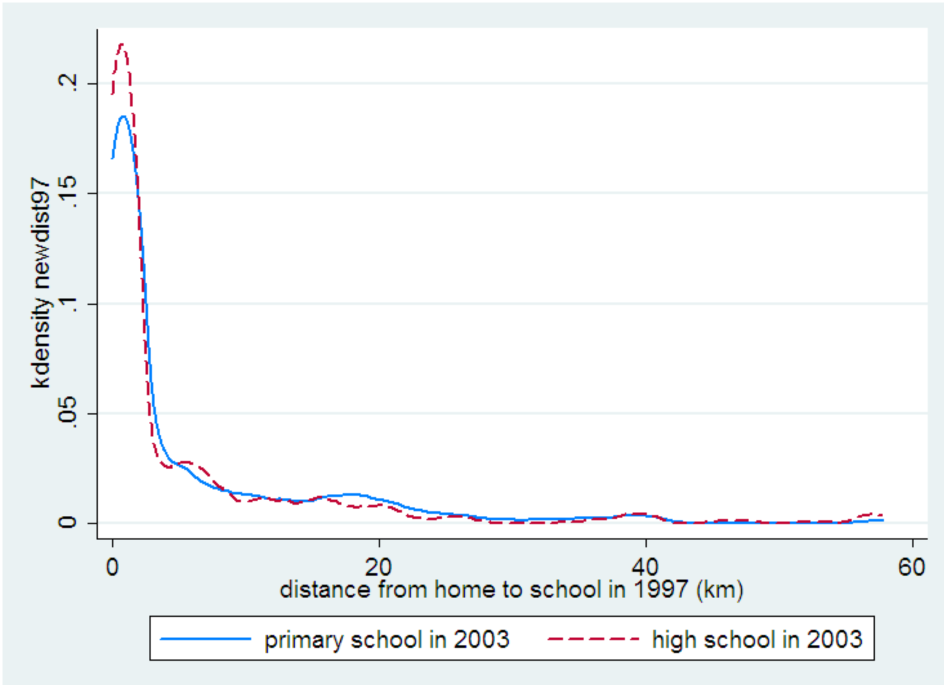

Figure 6.7: Kernel density plot of 1997 distance from home to school by 2003 schooling phase

Census Geography

Children who have reached high school by 2003 are slightly more likely to attend a school in 1997 that is in the same SAL as their home ($\chi^2_{(1)}$=4.3214, Pr=0.038), and the same SP as their home ($\chi^2_{(1)}$=7.6029, Pr=0.006). There is, however, no evidence of any differences in mobility on the basis of 2003 schooling phase at either the MP or MN levels.

Nearest School

There is no evidence that schooling phase in 2003 is associated with a child's likelihood of attending his or her nearest grade-appropriate school in 1997, regardless of whether independent schools are included in the analysis or not.

6.5.2 2003

Straight-line distance

As evident in Table 6.16 and Figure 6.8 below, children attending high school travel significantly further to school than those still in primary school in 2003 (Wilcoxon rank-sum test; Pr=0.0263).

	Number of children	Mean distance to school	Standard deviation	25th percentile distance	50th percentile distance	75th percentile distance
High school in 2003	418	5.791km	9.216km	0.712km	1.421km	7.367km
Primary school in 2003	853	5.164km	8.085km	0.533km	1.175km	6.475km

Table 6.16: 2003 distance from home to school, by progression to high school by 2003

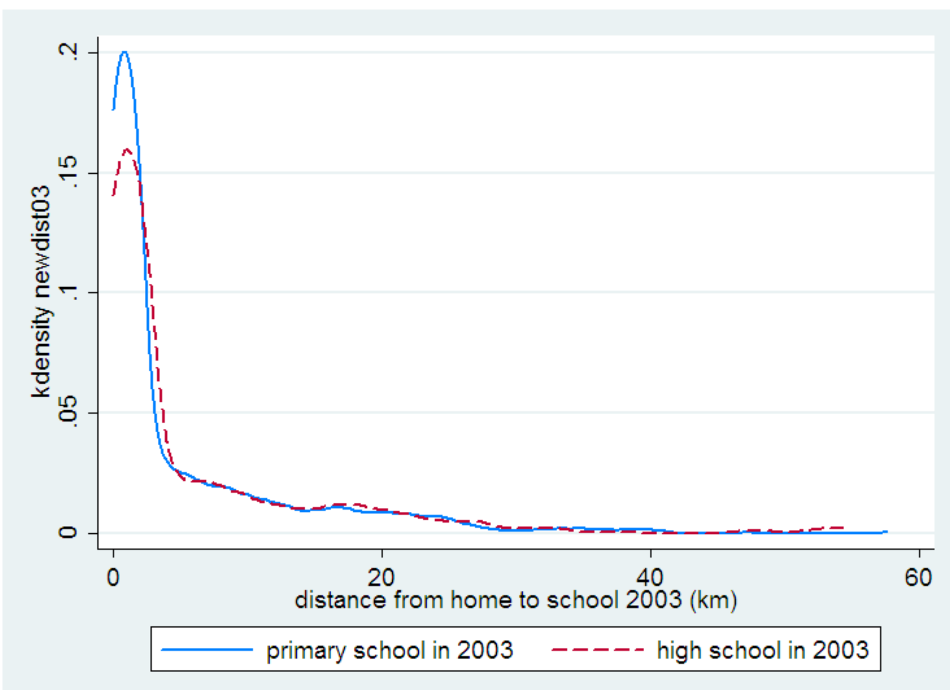

Figure 6.8: Kernel density plot of 2003 distance from home to school by phase of education in 2003

Given the fairly strong relationships between gender, age at first enrollment, and schooling phase in 2003 (see Appendix 3 for details), it is worth exploring the combined interactions of these variables with distance from home to school. Table 6.17 and Figure 6.9, below, illustrate that distance from home to school is different for each group when broken down by both gender and schooling phase. In all cases, boys travel less far than girls, with primary school boys travelling the shortest distances of all. High school girls are travelling further than any other group, including primary school girls. Both girls and boys at high school level are less likely to be travelling short distances to school, and it is only at the particularly high distances that the distributions actually differ on the basis of gender. At shorter distances (as is evident in Table 6.18), the distributions for high school girls and boys are fairly similar, as are the distributions for primary school girls and boys. A Wilcoxon rank-sum test indicates a weakly significant positive relationship between high school status and distance travelled for girls (Pr= 0.0689), but not for boys. There is, however, no evidence that girls travel significantly further than boys at either the primary or high school level.

	Number of children	**Mean distance to school**	**Standard deviation**	**25th percentile distance**	**50th percentile distance**	**75th percentile distance**
Primary school boys	459	4.778km	7.620km	0.509km	1.144km	5.816km
High school boys	166	5.564km	9.805km	0.655km	1.260km	5.524km
Primary school girls	394	5.614km	8.583km	0.567km	1.183km	7.425km
High school girls	252	5.941km	8.823km	0.758km	1.481km	8.211km

Table 6.17: 2003 distance from home to school, by gender and phase of education in 2003

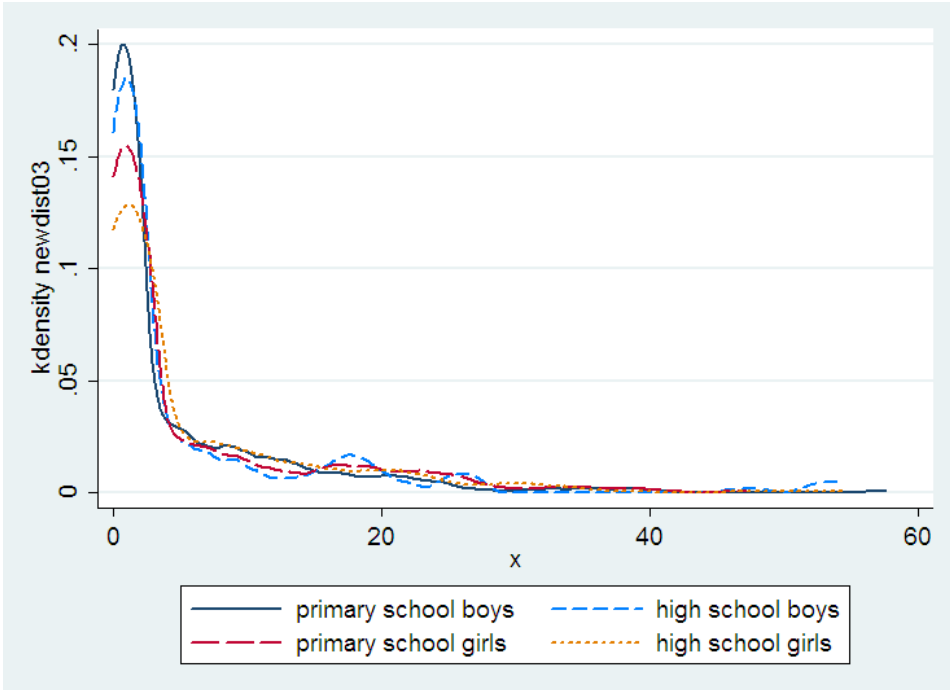

Figure 6.9: Kernel density plot of 2003 distance from home to school by gender and phase of education in 2003

Census geography

There is no evidence for any significant relationship between mobility at any level of census geography in 2003, and whether a child is in primary or high school.

Nearest school

Children who are in high school in 2003 are significantly more likely to be attending their nearest grade-appropriate school, whether independent schools are included ($\chi^2_{(1)}$= 7.9497, Pr = 0.005) or excluded ($\chi^2_{(1)}$= 6.3958, Pr = 0.011).

6.5.3 School phase in 2003 and mobility discussion

The results presented in this section have suggested that those children who had reached secondary school by 2003, and those who had not, behaved fairly similarly with respect to educational mobility in 1997. In 2003, by contrast, those children who have reached secondary school do travel further than those still in primary school. This makes sense as high schools tend to be larger and therefore less densely distributed, meaning that the nearest high school will on average be slightly further from a child's home than the nearest primary school. School distribution is also likely to explain why children in high school are more likely to be attending their nearest school – there are simply fewer options, particularly when a child faces financial constraints. The fact that these differences only emerge in 2003, once these children have reached high school, also supports the hypothesis that the changes in behaviour are linked more to high school status itself, than to the individual, family and community attributes of the learners in question.

While it would be ideal to explore these patterns further, the relatively small number of children in high school by 2003, and the non-random constitution of this group, raises problems. These questions could, however, be usefully explored in future work making use of data from subsequent years, or a broader sample. The data presented in this section also suggest that gender-based differences in mobility may increase at the secondary school level. Again, however, additional data will be required before these relationships can be more conclusively tested.

6.6 Grade repetition

Once enrolled in formal schooling, children have different experiences with grade progression. While many children do pass smoothly through the grades, a fairly substantial number are forced to repeat a particular grade, or even a number of grades (Fleisch and Schindler 2009). This section explores whether grade repetition between 1997 and 2003 has any relationship to distance travelled to attend school. This relationship is explored both in 1997, prior to repetition, and in 2003, after repetition.

Grade repetition may be thought of as an indicator of a child's inherent academic capabilities. In this case, it is possible that it might influence parental decisions on investment in schooling in both 1997 and 2003. Parents might choose to invest less in an academically less gifted child, or they might choose to invest more in the hopes of ensuring that child's success. To the extent that grade repetition reflects inherent academic capacity, the direction in which it should be expected to impact mobility is not clear. However, grade repetition may also simply reflect the quality of the school which a child attends (Lam, Ardington et al. 2008). In this case, grade repetition would be expected to be negatively associated with mobility.

6.6.1 1997

Straight-line distance

Children who repeated a grade between 1997 and 2003 are significantly more likely to have shorter distances between home and school than those who did not repeat any grades (see Table 6.18 and Figure 6.10 below), as confirmed by a Wilcoxon rank-sum test (Pr=0.0000).

	Number of children	**Mean distance to school**	**Standard deviation**	**25th percentile distance**	**50th percentile distance**	**75th percentile distance**
Repeaters	440	4.082km	7.997km	0.398km	0.823km	3.751km
Non-repeaters	730	6.516km	10.364km	0.506km	1.195km	8.152km

Table 6.18: 1997 distance from home to school, by grade repetition status

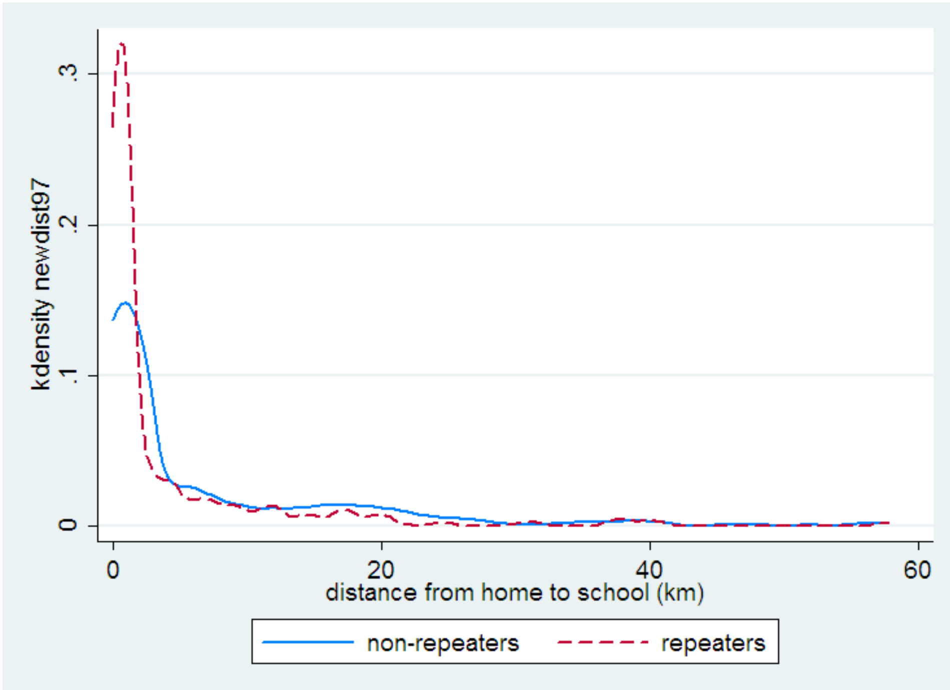

Figure 6.10: Kernel density plot of 1997 distance from home to school by grade repetition

Census geography

Children who repeated a grade at least once between 1997 and 2003 were less likely to attend schools outside of the SP ($\chi^2_{(1)}$= 5.1148; Pr = 0.024) and MP ($\chi^2_{(1)}$= 12.6441; Pr = 0.000) in which they lived. There was no relationship between repetition and mobility at the SAL and MN levels.

Nearest school

There is no significant relationship between grade repetition and whether or not a child attends their closest grade-appropriate school in 1997, regardless of whether independent schools are included or excluded.

6.6.2 2003

Straight-line distance

The relationship between grade repetition and distance travelled in 2003 is similar to that with distance in 1997. Again, as illustrated in Table 6.19 and

Figure 6.11 below, children who have repeated grades have a shorter distance from home to school (Wilcoxon rank-sum test; Pr=0.0000).

	Number of children	Mean distance to school	Standard deviation	25th percentile distance	50th percentile distance	75th percentile distance
Repeaters	439	4.249km	7.236km	0.464km	1.013km	4.640km
Non-repeaters	750	6.145km	9.121km	0.664km	1.447km	8.583km

Table 6.19: 2003 distance from home to school, by grade repetition status

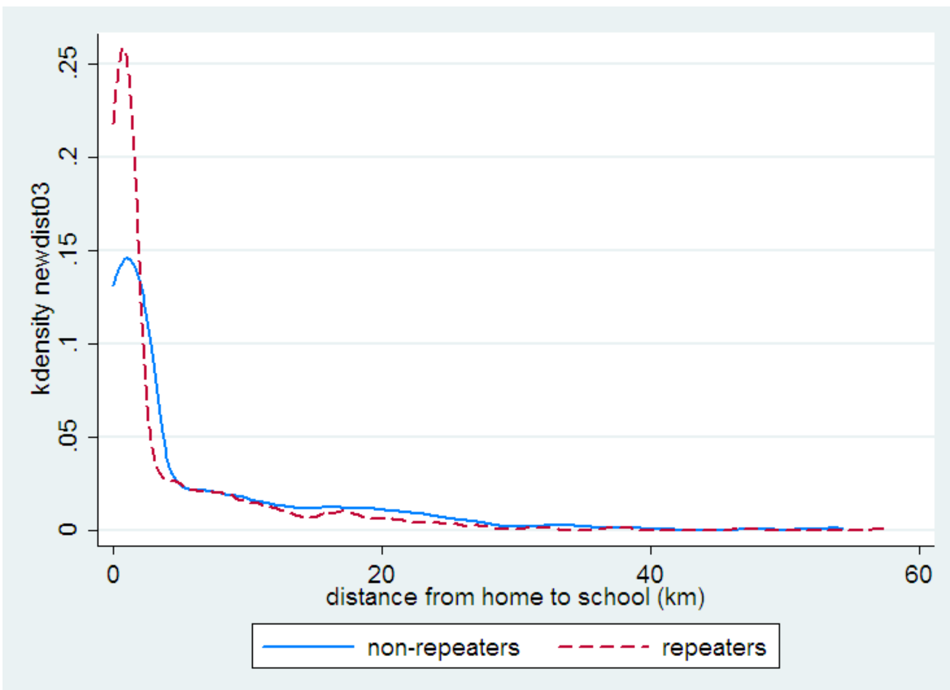

Figure 6.11: Kernel density plot of 2003 distance from home to school by grade repetition

Census geography

In 2003, children who have repeated grades are again significantly more likely to attend a school in their residential SP ($\chi^2_{(1)}$= 8.3371, Pr = 0.004) or MP ($\chi^2_{(1)}$= 8.3578, Pr = 0.004). There remains no evidence for any relationship at the SAL or MN levels.

Nearest school

There is no evidence for any significant relationship between grade repetition, and whether or not a child attends his or her nearest school, public or independent, in 2003.

6.6.3 Grade repetition and mobility discussion

Although there is a strong and consistent relationship between repetition and mobility, the available data says nothing about the causal direction of this relationship – does mobility shape repetition, or the reverse? Because most sample members live in relatively disadvantaged areas, mobility is usually associated with attendance at more advantaged schools, which tend to have lower levels of repetition. By extension, attending a local, less-advantaged school is likely to be associated with higher levels of repetition. For this reason, it seems more likely that mobility predicts repetition, than the reverse. A similar explanation seems plausible for the relationship between repetition and mobility at SP and MP levels of census geography. Of course, repetition is not determined purely by the school a child attends, and there is likely to also be an interaction effect operating, with children whose family circumstances favour better academic performance also being more likely to travel. The next set of analyses presented will explore the relationship between family attributes often associated with academic performance, and mobility.

Household level characteristics

6.7 Maternal education

Maternal education is anticipated to have a positive relationship with mobility, as more educated mothers are expected to place a higher premium on educational investment, and to have access to more resources to invest in their children’s education.

6.7.1 1997

Straight-line distance

As expected, mean distance travelled to school increases with maternal education level (see Table 6.20 and Figure 6.12 below). The mean figures for children with mothers with lower levels of education, up to grade 7, are skewed upwards by a few children who are travelling extremely substantial distances. This effect is evident in the high standard deviations and the very low distances at the 25th, 50th and 75th percentile.

Maternal Education level	Number of children	Mean distance to school	Standard deviation	25th percentile distance	50th percentile distance	75th percentile distance
Up to Grade 5	64	4.591km	13.075km	0.353km	0.708km	1.456km
Grade 6 or 7	78	5.572km	11.411km	0.468km	1.121km	3.304km
Grade 8, 9 or 10	530	4.207km	8.353km	0.434km	0.81km	3.646km
Grade 11 or 12	349	6.828km	9.400km	0.608km	1.491km	9.769km
Post-school education	92	7.224km	9.392km	0.642km	2.148km	11.879km

Table 6.20: 1997 distance from home to school, by maternal education level

Figure 6.12, below, provides a kernel density plot of the distances travelled by children to school, grouped by maternal education. The graph illustrates very clearly the variable nature of mobility by educational level, with children with functionally illiterate mothers being most likely to travel very short distances. As maternal education increases, however, a smaller proportion of children can be seen to travel very short distances, and the distributions gradually spread out, becoming slightly more normal. There is, however, substantial overlap between those children whose mothers have either grade 6 or 7, and those whose mothers have either grade 8, 9 or 10. A Kruskal-Wallis test indicates that distance from home to school varies significantly on the basis of maternal education (Pr=0.0001).

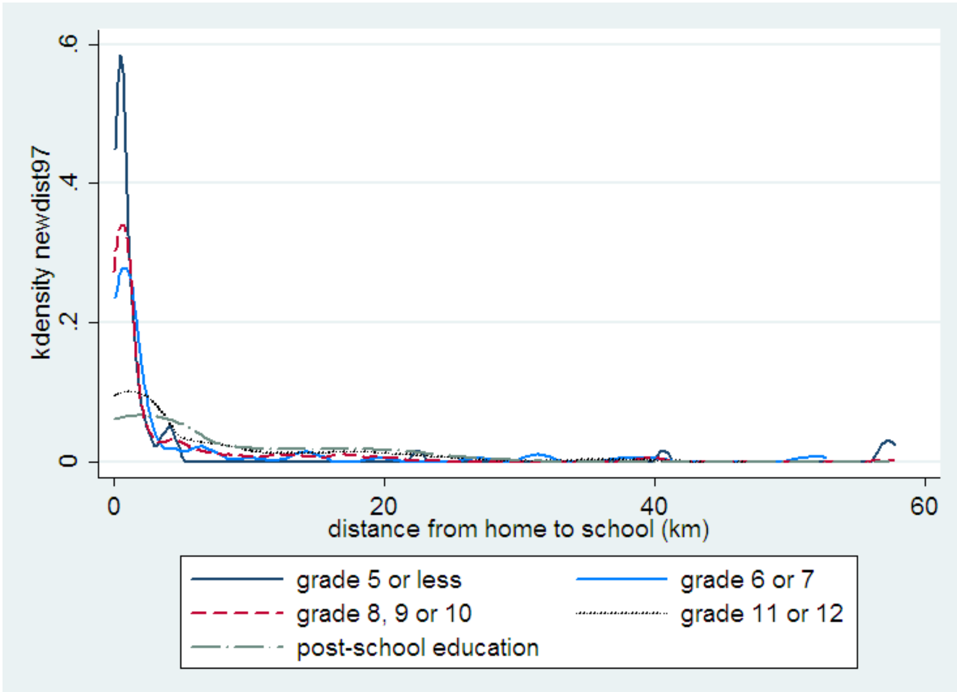

Figure 6.12: Kernel density plot of 1997 distance from home to school by maternal education level.

Census geography

Using chi-squared tests, whether or not a child attended school in the same MP as their home was the only geographical measure significantly linked to the distribution of maternal education for all levels of maternal education (see Table 6.21 below). When the children were divided on the basis of whether their mothers had completed primary education or not, significantly different levels of mobility at the SAL, SP and MP levels were identified.

	Grade 5 and lower vs. Grade 6 or higher (functionally illiterate vs. functionally literate)	Grade 7 and lower vs. Grade 8 or higher (primary school vs. higher than primary school education)	Grade 10 and lower vs. Grade 11 or higher	Grade 12 and lower vs. any post-school education (no post-school education vs. any post-school education)
School in same SAL as home	$\chi^2_{(1)}$= 1.902 Not significant	$\chi^2_{(1)}$= 4.202 Pr = 0.040	$\chi^2_{(1)}$= 0.815 Not significant	$\chi^2_{(1)}$= 0.153 Not significant
School in same SP as home	$\chi^2_{(1)}$= 2.578 Not significant	$\chi^2_{(1)}$= 15.201 Pr = 0.000	$\chi^2_{(1)}$= 2.1208 Not significant	$\chi^2_{(1)}$= 1.734 Not significant
School in same MP as home	$\chi^2_{(1)}$= 6.409 Pr = 0.011	$\chi^2_{(1)}$= 41.05 Pr = 0.000	$\chi^2_{(1)}$= 9.363 Pr = 0.002	$\chi^2_{(1)}$= 8.216 Pr = 0.004
School in same MN as home	$\chi^2_{(1)}$= 0.299 Not significant	$\chi^2_{(1)}$= 0.081 Not significant	$\chi^2_{(1)}$= 3.13 Pr = 0.077	$\chi^2_{(1)}$= 0.693 Not significant

Table 6.21: 1997 mobility across different levels of census geography, by maternal education level

Nearest school

A chi-square test indicated that children whose mothers were functionally illiterate were significantly more likely to attend their nearest public school than children whose mothers had higher levels of education (see Table 6.22 below). However, this relationship did become less linear at the highest levels of maternal education. When independent schools were included in the analysis, figures were largely similar, although somewhat more statistically significant ($\chi^2_{(4)}$ = 12.707, Pr=0.013).

Maternal education	Up to grade 5	Grade 6 or 7	Grade 8, 9 or 10	Grade 11 or 12	Post-school education	χ2 test results
Child attends nearest public school (n=198)	19 (28.79%)	13 (16.67%)	103 (19.36%)	48 (13.64%)	15 (16.13%)	$\chi^2_{(4)}$= 10.7994 Pr = 0.029

Table 6.22: Children attending their 1997 nearest grade-appropriate public school, by maternal education

6.7.2 2003

Straight-line distance

As was the case in 1997, distance from home to school in 2003 is also closely connected to maternal education (Kruskal-Wallis test, Pr=0.0001). As evident in Table 6.23 below, the gap between children whose mothers have completed up to Grade 10, and those who have completed Grade 11 or higher appears to have grown, while the distributions for children whose mothers have completed Grade 11 or 12, and those whose mothers have some post-school education, have become more similar. This is even more evident in Figure 6.13 below, in which the distributions of distance travelled are shown on the basis of maternal education.

Maternal Education level	Number of children	Mean distance to school	Standard deviation	25th percentile distance	50th percentile distance	75th percentile distance
Up to Grade 5	73	3.417km	8.443km	0.397km	0.778km	1.532km
Grade 6 or 7	83	3.021km	5.826km	0.398km	0.788km	2.332km
Grade 8, 9 or 10	560	3.855km	7.372km	0.506km	1.002km	2.692km
Grade 11 or 12	364	7.792km	9.403km	0.835km	2.870km	12.175km
Post-school education	97	7.811km	9.189km	0.938km	4.410km	11.858km

Table 6.23: 2003 distance from home to school, by maternal education level

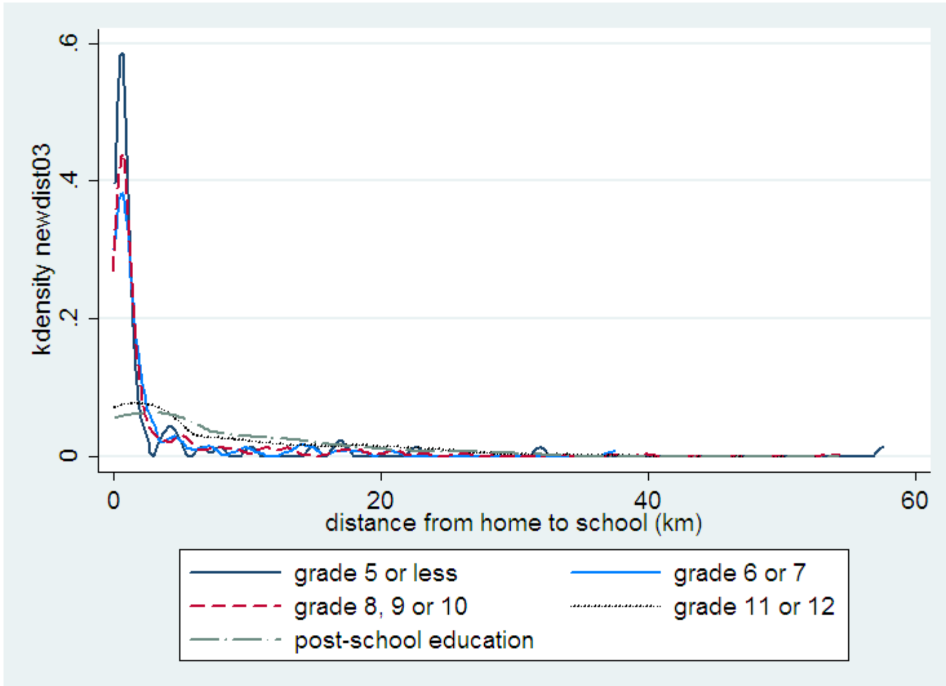

Figure 6.13: Kernel density plot of 2003 distance from home to school by maternal education level

Census geography

Again, as in 1997, chi-squared tests indicated that children with more educated mothers were more likely to attend a school outside of the area in which they lived for all levels of census geography other than MN (see Table 6.24 below). The 2003 data generally seems to indicate a stronger relationship between maternal education level, and whether or not children live and school within the same geographic area than was evident in 1997.

	Grade 5 and lower vs. Grade 6 or higher (functionally illiterate vs. functionally literate)	Grade 7 and lower vs. Grade 8 or higher (primary school vs. higher than primary school education)	Grade 10 and lower vs. Grade 11 or higher	Grade 12 and lower vs. any post-school education (no post-school education vs. any post-school education)
School in	$\chi^2_{(1)}$= 7.417	$\chi^2_{(1)}$= 5.215	$\chi^2_{(1)}$= 3.579	$\chi^2_{(1)}$= 1.765

same SAL as home	Pr = 0.006	Pr = 0.022	Pr = 0.059	Not significant
School in same SP as home	$\chi^2_{(1)}$= 3.925 Pr = 0.048	$\chi^2_{(1)}$= 52.025 Pr = 0.000	$\chi^2_{(1)}$= 12.829 Pr = 0.000	$\chi^2_{(1)}$= 9.414 Pr = 0.002
School in same MP as home	$\chi^2_{(1)}$= 4.135 Pr = 0.042	$\chi^2_{(1)}$= 78.232 Pr = 0.000	$\chi^2_{(1)}$= 17.673 Pr = 0.000	$\chi^2_{(1)}$= 10.817 Pr = 0.001
School in same MN as home	$\chi^2_{(1)}$= 0.215 Not significant	$\chi^2_{(1)}$= 2.229 Not significant	$\chi^2_{(1)}$= 0.686 Not significant	$\chi^2_{(1)}$= 0.592 Not significant

Table 6.24: 2003 mobility across different levels of census geography, by maternal education level

Nearest school

The distribution of levels of maternal education is significantly different between children attending their closest grade-appropriate school, public or independent, and those travelling further afield (Table 6.25). Children whose mothers have completed schooling up to grade 7 appear to be substantially more likely to attend their nearest public school than children whose mothers have higher levels of education. As with all previous analyses of maternal education, this relationship again appears to be stronger in 2003 than 1997.

Mother's education level	Up to grade 5	Grade 6 or 7	Grade 8, 9 or 10	Grade 11 or 12	Post-school education	χ^2 test results
Child attends nearest public school (n=213)	20 (27.03%)	26 (31.71%)	111 (19.89%)	43 (11.85%)	13 (13.54%)	$\chi^2_{(4)}$= 26.285 Pr=0.000

Table 6.25: Children attending their 2003 nearest grade-appropriate school, by maternal education

6.7.3 Maternal education and mobility discussion

In 1997, at the lowest levels of maternal education, there is a very high level of variability around the distance travelled. This effect seems largely to have disappeared by 2003, and may relate to poor quality data around these children's education and residence, resulting in inaccurate mobility data, or alternatively, lack of residential stability resulting in substantial travel.

Overall, however, there is evidence that as maternal education increases, so too does educational mobility. The effect of maternal education on mobility also appears to increase over time, strengthening as children reach the end of primary school and start to enrol in secondary school. This relationship between mobility and maternal education is not strictly linear, and children whose mothers have attained varying levels of education between grade 6 and 10 all appear to behave quite similarly rather than as distinct groups. Additionally, the relationship between distance and maternal education appears to stop holding at the very highest levels of maternal education. One possible explanation for this is that children with the most highly educated mothers are more likely to live in affluent areas, and by extension, close to good schools.

An interesting feature of the maternal education analyses is the fairly strong inverse relationship between maternal education, particularly at intermediate levels, and the likelihood that children will attend their nearest grade-appropriate school. This raises questions around the relationship between maternal education and the decision to engage in school choice, even if the resources available allow only for choice between fairly local schools. Of course, with both this relationship between maternal education and local school choice, and with the broader relationship between maternal education and distance travelled, it is critical to explore the role of SES. As SES is highly correlated with maternal education, it is important to attempt to separate the roles of resource availability and maternal education in shaping schooling decisions. This is explored both in the section on household SES later in this chapter, and in Chapter 9.

6.8 Maternal Marital status

The next potential determinant of learner mobility to be examined is maternal marital status at the time of the child's birth. This provides an indicator of the home environment into which a child is born, with married mothers typically being associated with a more stable, socio-economically advantaged home environment than non-married mothers. It is therefore plausible to expect that the children of married mothers may be more likely to travel further to attend school.

6.8.1 1997

Straight-line distance

As shown in Table 6.26 and Figure 6.14 below, children of married mothers have, on average, a slightly greater distance from home to school than the children of unmarried mothers (Wilcoxon rank-sum test; Pr= 0.0090).

	Number of children	Mean distance to school	Standard deviation	25th percentile distance	50th percentile distance	75th percentile distance
Married mothers	410	6.284km	10.102km	0.510km	1.280km	8.102km
Unmarried mothers	797	5.125km	9.228km	0.448km	0.934km	5.298km

Table 6.26: 1997 distance from home to school, by maternal marital status

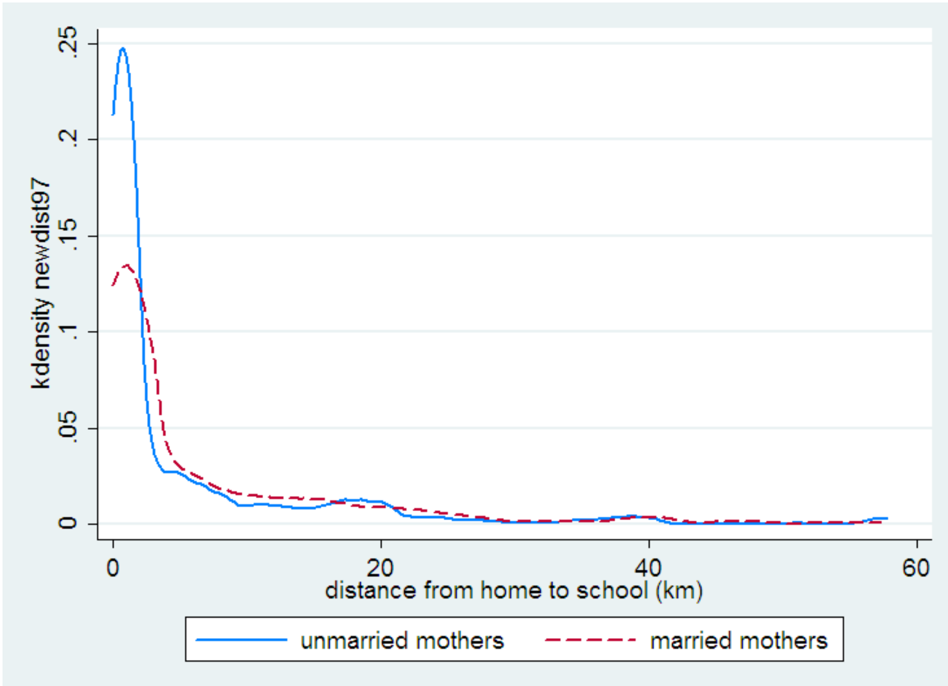

Figure 6.14: Kernel density plot of 1997 distance from home to school by maternal marital status

Census geography

There was no evidence of a significant relationship between maternal marital status at birth, and a child's mobility at any level of census geography in 1997.

Nearest school

There was no evidence of a significant relationship between maternal marital status at birth and whether a child attended his or her nearest grade-appropriate school in 1997.

6.8.2 2003

Straight-line distance

Although it is evident from Table 6.27 and Figure 6.15, below, that in 2003 children with married mothers continued to travel slightly further than children of unmarried mothers, this difference is no longer statistically significant (Wilcoxon rank-sum test).

	Number of children	Mean distance to school	Standard deviation	25th percentile distance	50th percentile distance	75th percentile distance
Married mothers	426	5.511km	8.235km	0.639km	1.264km	7.446km
Unmarried mothers	848	5.292km	8.603km	0.554km	1.238km	6.327km

Table 6.27: 2003 distance from home to school, by maternal marital status

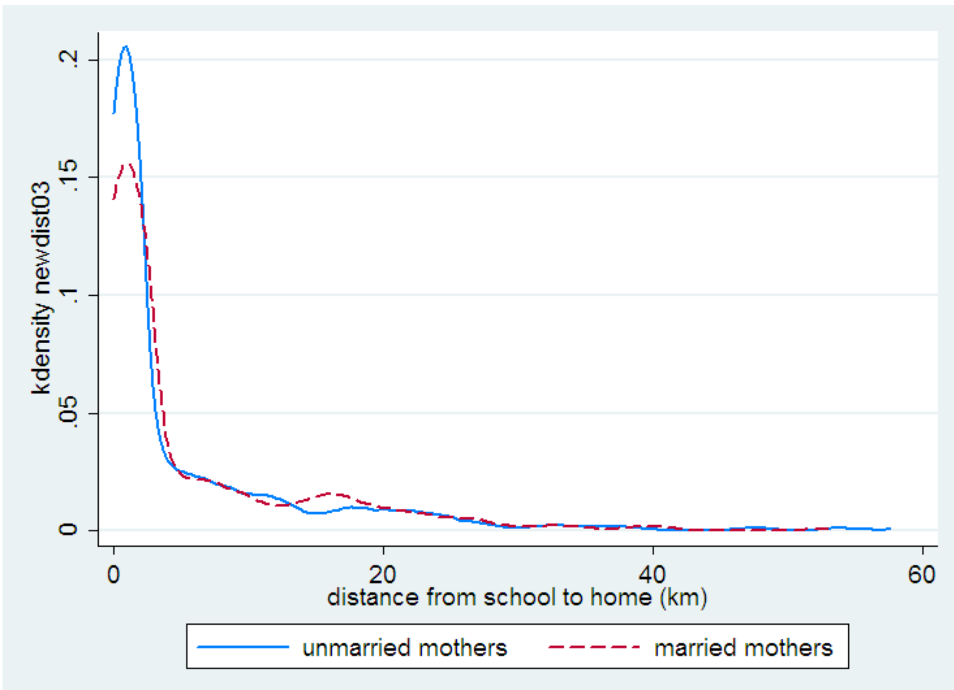

Figure 6.15: Kernel density plot of 2003 distance from home to school, by maternal marital status

Census geography

There was no evidence for a significant relationship between maternal marital status at birth and a child's mobility at any level of census geography in 2003.

Nearest school

There was no evidence for a significant relationship between maternal marital status at birth and whether a child attends his or her nearest grade-appropriate school in 2003.

6.8.3 Maternal marital status and mobility discussion

Although there does seem to be a weak relationship between distance from home to school and maternal marital status in 1997, this effect seems to have disappeared by 2003. There is no evidence that maternal marital status is related to any other measures of mobility. This may suggest that to the extent that maternal marital status does influence schooling choices, this effect is strongest when the children are young, and have fairly limited independence. It may also simply reflect the stronger relationship between marital status at birth and 1997 SES, as opposed to 2003 SES. Indeed, the extent to which any effect of maternal marital status on schooling choices is operating through the relationship between marital status and SES also merits further investigation. This is covered in Chapter 9.

6.9 Household SES

Household SES is likely to play a core role in shaping school choice. Access to resources determines how much a family can afford to spend on travel to school, as well as how much they can afford to contribute to school fees and related expenses. It may also be highly correlated with determinants of the value the family places on education, such as parental education levels. Generally, it is anticipated that as SES increases, so too will, on average, the distance travelled to school.

6.9.1 1997

Straight-line distance

Examining the means and distributions of distance from home to school on the basis of 1997 SES (see Table 6.28 and Figure 6.16 below) confirms the hypothesis that children from wealthier families do tend to travel further to school. Interestingly, the pattern for quintile 1 is quite distinct, while those for quintiles 2 and 3 are quite similar, as are those for quintiles 4 and 5. There is also some non-linearity in the relationship between SES and distance, particularly in quintiles 2 and 3.

SES Quintile	Number of children	Mean distance to school	Standard deviation	25th percentile distance	50th percentile distance	75th percentile distance
1 (most poor)	232	3.896km	10.559km	0.370km	0.694km	1.380km
2	217	4.789km	8.988km	0.449km	0.842km	4.230km
3	217	4.460km	8.226km	0.465km	0.946km	3.995km
4	225	6.256km	9.430km	0.483km	1.445km	8.654km
5 (least poor)	187	6.993km	8.865km	0.611km	2.416km	11.549km

Table 6.28: 1997 distance from home to school, by 1997 household SES quintile

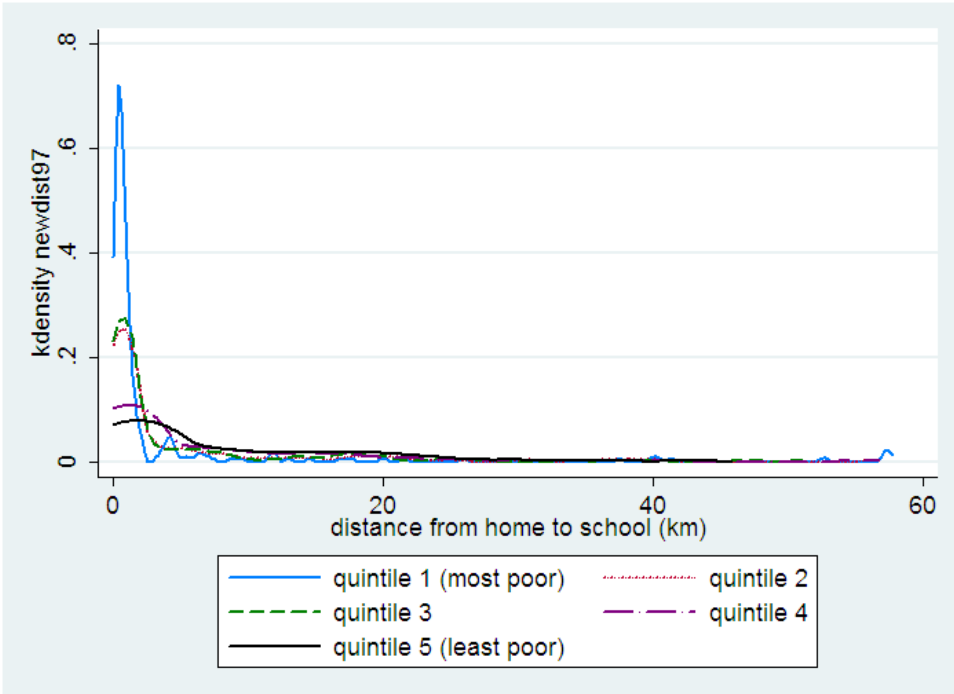

Figure 6.16: Kernel density plot of 1997 distance from home to school, by 1997 household SES quintile

Census Geography

As is evident in Table 6.29 below, more children become mobile at both the SP and MP levels of census geography, as SES increases. At the SAL level, the proportion of children who are mobile is extremely high for all SES groups, and the chi-squared test does not provide evidence that mobility varies on the basis of SES. At the MN level so few children are mobile that there is no difference in behaviour on the basis of SES. Due to the very small numbers of children involved the MN level is not shown in the table below.

Quintile	1 (most poor)	2	3	4	5 (least poor)	χ^2 test results
School in same SAL as home	23 (9.87%)	11 (5.07%)	16 (7.37%)	17 (7.56%)	13 (6.95%)	$\chi^2_{(4)}$= 3.8529 N.S.
School in same SP as home	126 (54.08%)	89 (41.01%)	94 (43.32%)	79 (35.11%)	64 (34.22%)	$\chi^2_{(4)}$= 23.2285 Pr=0.000
School in same MP as home	200 (85.84%)	164 (75.58%)	167 (76.96%)	143 (63.56%)	122 (65.24%)	$\chi^2_{(4)}$= 38.2041 Pr=0.000

Table 6.29: 1997 mobility across different levels of census geography, by 1997 household SES

Nearest school

Table 6.30, below, illustrates the proportions of children from each quintile who attend their nearest school. Although a higher proportion of children from quintiles 1 and 5 do appear to be attending their nearest school in both analyses, this is not statistically significant. The U-shaped nature of the relationship may explain the absence of significant result.

Quintile	1 (most poor)	2	3	4	5 (least poor)	χ^2 test results
Attends nearest public school	51 (21.89%)	31 (14.22%)	37 (16.97%)	40 (17.70%)	38 (20.32%)	$\chi^2_{(4)}$=5.2706 N.S.
Attends nearest public or independent school	51 (21.89%)	33 (15.21%)	37 (16.97%)	38 (16.81%)	38 (20.65%)	$\chi^2_{(4)}$=4.6710 N.S.

Table 6.30: Children attending their 1997 nearest grade-appropriate school, by 1997 household SES

6.9.2 2003

Straight-line distance

The relationship between distance from home to school and household SES in 2003 remains very similar to that for 1997 (see Table 6.31 and Figure 6.17

below). However, the distances travelled have increased, particularly for children in the higher quintiles. This may in part relate to the higher likelihood that more advantaged children have reached high school by 2003. A Kruskal-Wallis test confirms that distance from home to school varies significantly on the basis of 2003 SES (Pr=0.0001).

SES Quintile	**Number of children**	**Mean distance to school**	**Standard deviation**	**25th percentile distance**	**50th percentile distance**	**75th percentile distance**
1 (most poor)	170	1.422km	2.480km	0.420km	0.749km	1.395km
2	169	4.717km	9.098km	0.560km	1.019km	4.267km
3	172	5.106km	8.476km	0.457km	1.185km	5.145km
4	170	7.664km	9.548km	0.637km	2.443km	12.466km
5 (least poor)	162	9.501km	9.433km	1.315km	6.392km	14.994km

Table 6.31: 2003 distance from home to school, by 2003 household SES quintile

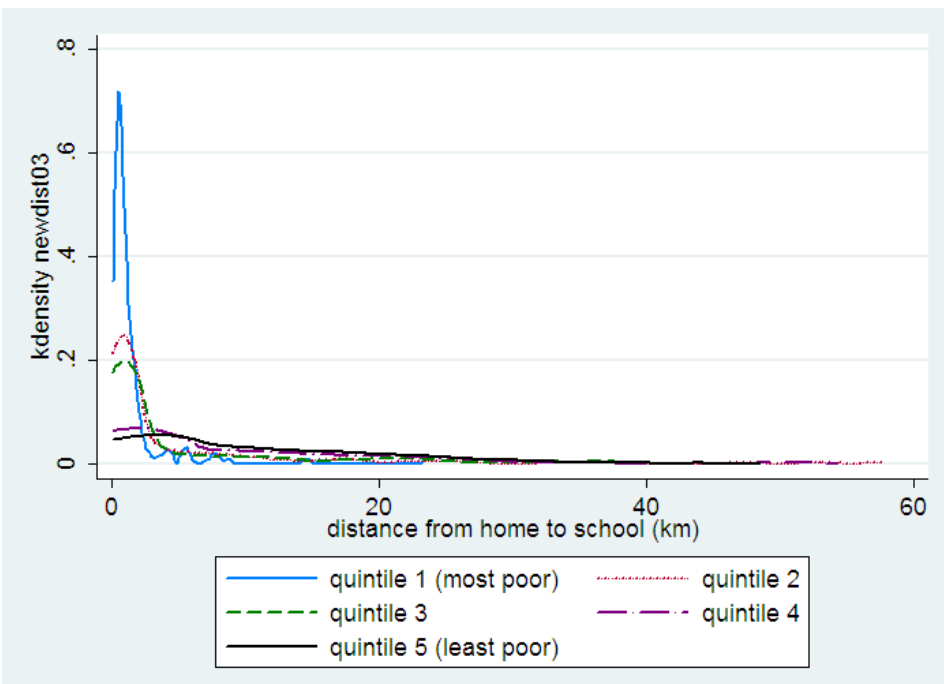

Figure 6.17: Kernel density plot of 2003 distance from home to school, by 2003 household SES quintile

Census Geography

In 2003 there is a fairly straightforward relationship between SES and mobility at the SP and MP levels (see Table 6.32 below), with children from higher income families more likely to be mobile. At the SAL level, there is again no relationship between mobility and SES. At the MN level, by contrast, in 2003 there is a significant relationship ($\chi^2_{(4)}$=7.8098, Pr=0.099), with more wealthy children being more likely to travel between MNs for schooling. However, the actual numbers of children travelling between MNs remains very small (n=31), and this level is therefore not included in the table below.

	Quintile 1	Quintile 2	Quintile 3	Quintile 4	Quintile 5	χ2 test results
School in same SAL as home	5 (2.94%)	4 (2.37%)	5 (2.91%)	9 (5.29%)	4 (2.47%)	$\chi^2_{(4)}$=3.1465 N.S
School in same SP as home	95 (55.88%)	63 (37.28%)	61 (35.47%)	46 (27.06%)	33 (20.37%)	$\chi^2_{(4)}$=52.6549 Pr=0.000
School in same MP as home	154 (90.59%)	128 (75.74%)	123 (71.51%)	95 (55.88%)	75 (46.30%)	$\chi^2_{(4)}$=92.3332 Pr=0.000

Table 6.32: 2003 mobility across different levels of census geography, by 2003 household SES

Nearest school

In 2003 there is evidence that the poorest children are significantly more likely to attend the school nearest to their home (see Table 6.33 below).

	Quintile 1	Quintile 2	Quintile 3	Quintile 4	Quintile 5	χ^2 test results
Attends nearest public school	41 (24.26%)	26 (15.38%)	28 (16.28%)	27 (15.88%)	18 (11.04%)	$\chi^2_{(4)}$=11.0515 Pr=0.026
Attends nearest public or independent school	40 (23.67%)	25 (14.79%)	26 (15.20%)	25 (14.71%)	17 (10.49%)	$\chi^2_{(4)}$=11.6124 Pr=0.020

Table 6.33: Children attending their 2003 nearest grade-appropriate school, by 2003 household SES

6.9.3 Household SES and mobility discussion

In both 1997 and 2003, SES is a strong predictor of the distance between a child's home and school. The relationship becomes slightly more direct in 2003, when children are older, and at least for wealthier children appears to become stronger. There is also a fairly strong relationship between SES and mobility at both the SP and MP levels, although not at the SAL or the MN levels. In 2003, poorer children are more likely to attend their nearest grade-appropriate schools than their wealthier peers.

As noted in Chapter 4, SES is strongly correlated with maternal education, and the similarities in the relationships between maternal education and distance and SES and distance are therefore unsurprising. One of the most interesting features of this set of analyses is the absence of a significant relationship between SES and whether or not a child attends their closest school in 1997. Given the significant nature of the relationship between maternal education and enrolment at the closest school, this is unexpected. It suggests one way in which the relationship between mobility and maternal education may differ from the relationship between mobility and SES. One potential explanation for these results is that more educated mothers are more likely to engage in school choice. To the extent that they have access to additional resources (which is reflected in the SES data), they may choose to access more advantaged schools outside of their immediate area. This is evident in the relationship between SES and distance travelled. However, to the extent that these mothers do not have access to additional resources, and have lower SES, they may be constrained to choose from schools closer to home. However, they do continue to exercise choice, which is demonstrated by the lower levels at which their children attend the nearest schools, even when they aren't able to travel far. Whether this is indeed the explanation, and if so, why this effect disappears by 2003, is not clear, and is explored further in Chapter 9.

Community level characteristics

6.10 Residential area poverty

The area in which a child and his or her family live is also likely to influence school choice, and particularly whether travel over a substantial distance is a seen as a potentially beneficial option. An area's affluence is likely to be related to, but not identical with, the affluence of any particular household within this area. A wealthier family is more likely to be able to afford to live in an affluent area, but within well-off areas, for a range of reasons such as historical accident, an employer providing somewhere to live, or the rapid emergence of an informal settlement, a number of comparatively and absolutely disadvantaged families can typically be found. Similarly, some wealthy families can typically be found in even the most disadvantaged communities, choosing to live there for various reasons which might be historical, social or even economic (lower expenditure on rent, purchasing a home, or rates and property taxes may allow for greater expenditure on consumer goods or investments such as education).

An area's relative affluence is expected to play a role in the educational choices available to families, and the costs of those choices. This is particularly true in South Africa, where the affluence of areas is closely connected to their historical racial designation, which simultaneously also shaped the quality of the schooling available in those areas. Simply put, historically white areas are typically more affluent than other areas, and also offer higher quality educational opportunities locally. A child living in a particularly affluent community is therefore not likely to need to travel substantial distances to access a good school. By contrast, a child living in a poor and historically disadvantaged community is likely to need to travel a great distance. We can therefore expect that a family's SES, which shapes ability to travel, and the nature of the area in which they live, which shapes the extent to which travel is beneficial, will operate together in determining levels of learner mobility. As a result, the more affluent children living in less affluent areas are those expected to travel the greatest distances.

The relationship between the poverty level of the area in which a child lives, and the distance that they are likely to travel to school seems likely to be considerably more complex than any of the other relationships explored thus far. The nature of the relationship seems likely to depend strongly on the level of geography which is being considered – SAL, SP or MP. For this reason, the results will be explored level by level.

6.10.1 1997

Small Area Level (SAL)

Straight-line distance

In the two lowest poverty quintiles of SAL (that is, the wealthiest areas), children typically travel fairly substantial distances to school (see Table 6.34 and Figure 6.18 below). Children in quintile 3 travel substantially shorter distances than children in any other quintile. The distances travelled by children in quintiles 4 and 5 (the poorest areas), are between these two extremes. Given this non-linear relationship between SAL poverty and distance travelled, it is understandable that the correlation between SAL area poverty and distance from home to school is fairly low, although negative overall (-0.0591 for raw scores). Overall, children in wealthier SALs with lower poverty levels travel somewhat further than others, as confirmed by both a Kruskal-Wallis test (Pr=0.0001). However, this relationship is clearly not linear, and particularly worth notice is that the children living in the very poorest SALs are actually typically travelling further than those in only moderately poor SALs.

Area Poverty Quintile	**Number of children**	**Mean distance to school**	**Standard deviation**	**25th percentile distance**	**50th percentile distance**	**75th percentile distance**
1 (wealthiest)	213	6.568km	9.206km	0.596km	1.912km	9.348km
2	248	6.957km	10.526km	0.492km	1.345km	9.264km
3	256	3.790km	7.433km	0.422km	0.758km	2.076km
4	259	4.451km	8.453km	0.398km	0.824km	4.285km
5 (poorest)	237	5.966km	11.334km	0.523km	1.177km	5.925km

Table 6.34: 1997 distance from home to school, by SAL poverty quintile

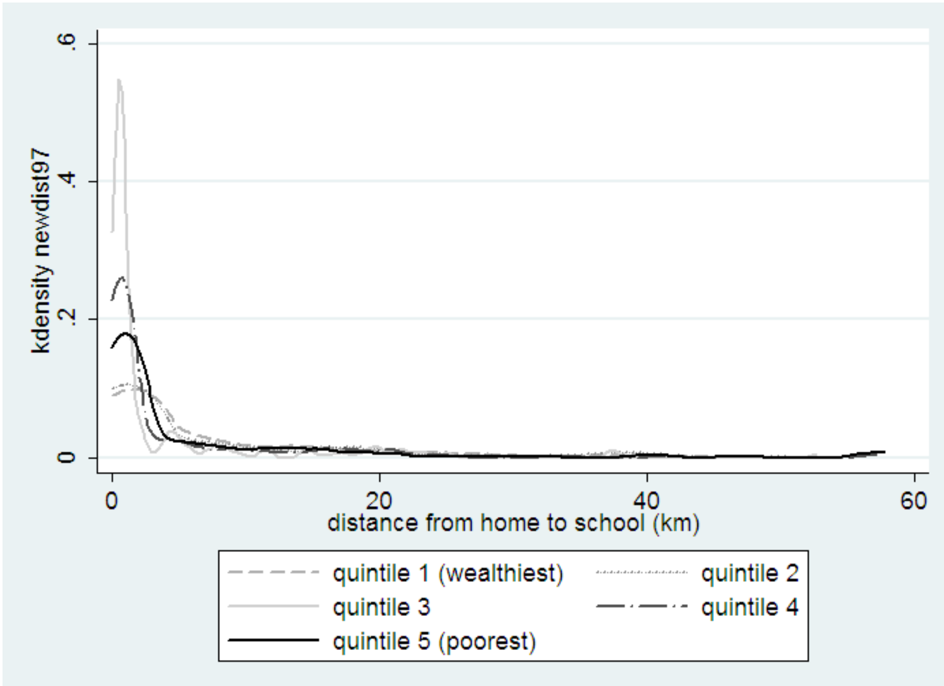

Figure 6.18: Kernel density plot of 1997 distance from home to school, by SAL poverty quintile

Census Geography

There is no evidence for a relationship between the poverty level of the SAL in which a child lives, and the child's likelihood of a attending a school in that same SAL. Children living in the very poorest and most affluent SAL quintile are however more likely to be travelling to a school outside their home SP than children in intermediate SAL quintiles. Children in SAL poverty quintile 3 are most likely to attend a school in the same MP in which they live ($\chi^2_{(4)}$=14.2914, Pr = 0.006). Finally, children living in the poorest SAL quintile are most likely to travel to a school outside of the MN in which their home is ($\chi^2_{(4)}$=8.0681, Pr = 0.089). However, given the extremely low levels of MN mobility, it is not clear that these figures are particularly meaningful.

Nearest school

There is an almost linear relationship between the poverty of the SAL and whether children attend their nearest school, public or independent. Children in the most affluent SALs are most likely to attend their nearest school, while children in the poorest SALs are least likely to attend their nearest schools ($\chi^2_{(4)}$= 17.7483, Pr=0.001 for public schools only; $\chi^2_{(4)}$= 13.3720, Pr=0.01 for public and independent schools).

Sub Place (SP)

Straight-line distance

Children in quintile 2 and quintile 5 (highest poverty) SP areas have the greatest distance from home to school, with children in quintiles 3 and 4 travelling particularly short distances (see Table 6.35 and Figure 6.19 below). Accordingly, the correlation between SP area poverty and distance travelled is weaker than that for SAL, although it remains negative (-0.0263 for raw scores). Although children in wealthier SPs are continuing, on average, to travel slightly further than children in poorer SPs, this relationship is somewhat weaker than was the case for the SALs (Kruskal-Wallis test, Pr=0.0001).

Area Poverty Quintile	Number of children	Mean distance to school	Standard deviation	25th percentile distance	50th percentile distance	75th percentile distance
1 (wealthiest)	218	5.762km	8.820km	0.495km	1.524km	6.698km
2	247	7.274km	10.490km	0.496km	1.500km	11.237km
3	267	3.504km	6.166km	0.424km	0.843km	3.915km
4	248	3.694km	6.860km	0.413km	0.743km	2.142km
5 (poorest)	233	7.542km	13.246km	0.633km	1.404km	7.608km

Table 6.35: 1997 distance from home to school, by SP poverty quintile

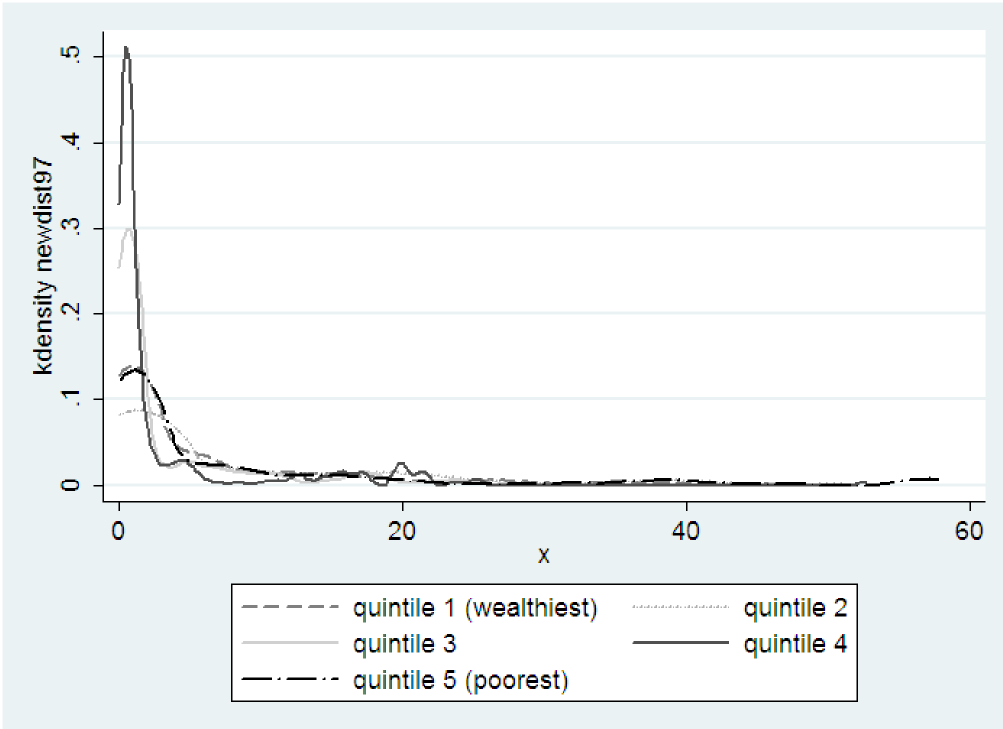

Figure 6.19: Kernel density plot of 1997 distance from home to school, by SP poverty quintile

Census Geography

Children living in poorer SP areas are more likely to be travelling to a school outside of the SAL in which they live than their peers in more advantaged areas ($\chi^2_{(4)}$=11.1595, Pr = 0.025). Children living in poorer SP areas are also more likely to be travelling to a school outside of the SP in which they live ($\chi^2_{(4)}$=28.1873, Pr = 0.000), but this relationship is not strictly linear. Children in quintile 5 (poorest areas), and quintile 2 seem to be far more likely than any other groups to be mobile at the SP level. Children in SP quintile 5 and 2 are also substantially more likely to be mobile at the MP level than their peers ($\chi^2_{(4)}$=25.9282, Pr = 0.000). A similar relationship also holds with MN mobility ($\chi^2_{(4)}$= 17.9573, Pr = 0.001), although only a very small number of children are travelling at this level. Overall, although the relationships are not strictly linear, children living in the very poorest SPs are more mobile at each level of census geography than their peers in more affluent areas.

Nearest school analysis

Children living in wealthier SP areas are more likely to attend their nearest school than children living in SP areas with higher poverty (Kruskal-Wallis test; Pr=0.0001 regardless of whether independent schools are included).

Main Place

As discussed in Chapter 3, due to extreme clustering only 3 poverty quantiles are used at the MP level, and even these are very uneven, meaning that the findings presented in the section should be treated with some caution.

Straight-line distance

Children living in the poorest MP areas live the furthest distance away from their schools (see Table 6.36 and Figure 6.20 below; Kruskal-Wallis test, Pr=0.0155). Although this is a different result from that found at the SAL and SP level, it is not clear whether it is a feature of the larger area size considered, or a function of grouping children into three quantiles as opposed to five quintiles.

MP Area Poverty Quantile	Number of children	Mean distance to school	Standard deviation	25th percentile distance	50th percentile distance	75th percentile distance
1 (wealthiest)	394	5.349km	9.299km	0.447km	0.909km	5.303km
2	705	5.453km	9.580km	0.455km	1.040km	5.816km
3 (poorest)	115	6.233km	9.983km	0.699km	1.554km	7.608km

Table 6.36: 1997 distance from home to school, by MP poverty quantile

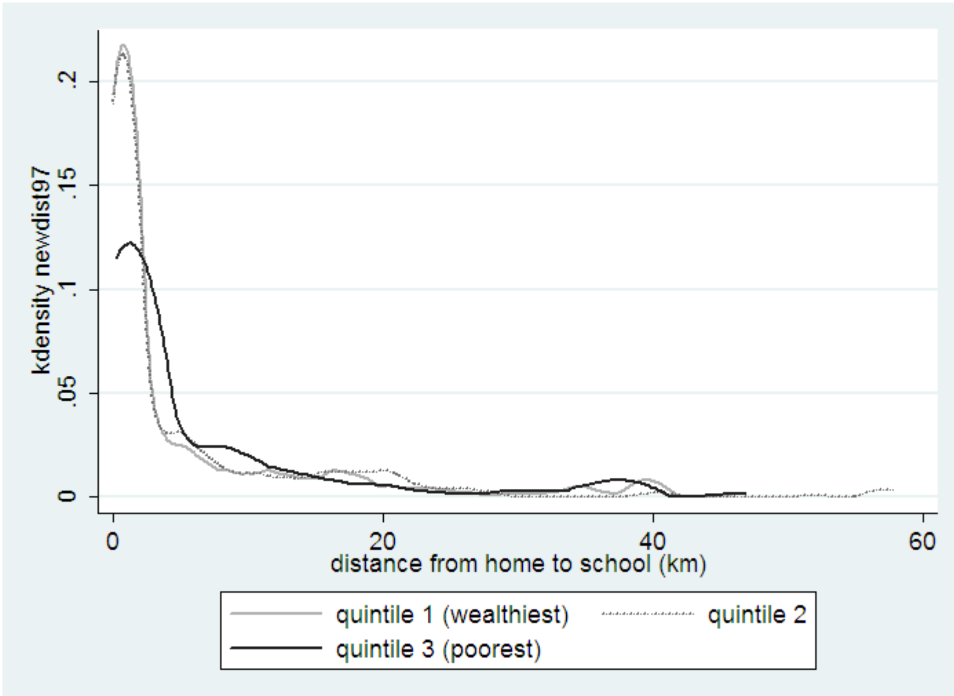

Figure 6.20: Kernel density plot of 1997 distance from home to school, by MP poverty quantile

Census geography

Children living in the most affluent MP areas are most likely to attend school in the SAL in which they live ($\chi^2_{(2)}$=16.2439, Pr = 0.000). These children are also more likely to attend school in the SP in which they live ($\chi^2_{(2)}$= 24.4440, Pr = 0.000), and also in the MP in which they live ($\chi^2_{(2)}$= 16.6547, Pr = 0.000). At the MN level, however, children in MP poverty quintile 2 are substantially more likely to travel between MNs than their peers living in more affluent or poorer MPs ($\chi^2_{(2)}$= 27.0329, Pr = 0.000).

Nearest School analysis

Chi-square tests indicate that children living in more advantaged MPs are significantly more likely to be attending their nearest schools ($\chi^2_{(2)}$=34.0264, Pr = 0.000 for public schools only; $\chi^2_{(2)}$= 28.5264, Pr = 0.000 for public and independent schools).

6.10.2 2003

Small Area Level (SAL)

Straight-line distance

In 2003, it is the quintile of children living in the poorest SALs that travel the shortest distances to school, although children in SAL poverty quintile 3 only travel slightly further (see Table 6.37 and Figure 6.21 below). The change in the relative travel of children living in the poorest SALs may relate to the opening of new schools, these children being less likely to have reached high school, or more accurate data for these children. The correlation between SAL poverty and distance travelled remains weakly negative (-0.0721 for raw scores), and a Kruskal-Wallis test indicates that distance to school varies significantly with SAL poverty level (Pr= 0.0053).

Area Poverty Quintile	Number of children	Mean distance to school	Standard deviation	25th percentile distance	50th percentile distance	75th percentile distance
1 (wealthiest)	223	6.006km	8.949km	0.664km	1.720km	7.410km
2	252	6.439km	9.058km	0.534km	1.629km	10.246km
3	270	4.910km	8.693km	0.464km	0.873km	5.387km
4	271	5.268km	7.604km	0.631km	1.214km	7.796km
5 (poorest)	264	4.304km	7.946km	0.669km	1.247km	3.276km

Table 6.37: 2003 distance from home to school, by SAL poverty quintile

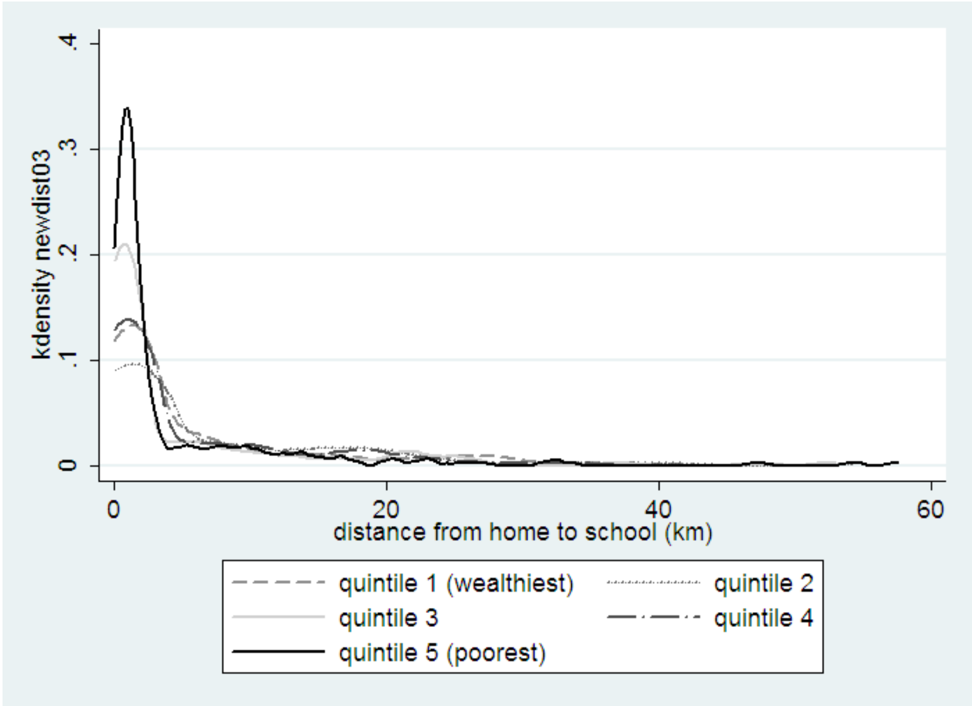

Figure 6.21: Kernel density plot of 2003 distance from home to school, by SAL poverty quintile

Census geography

Children living in more affluent SAL areas are more likely to attend a school in the same SAL as their home than children living in poorer SAL areas ($\chi^2_{(4)}$= 11.1184, Pr = 0.025). There is no significant relationship between SAL poverty and mobility at the SP, MP or MN levels, however.

Nearest school

Children living in more affluent SALs are significantly more likely to attend their nearest school ($\chi^2_{(4)}$= 16.2545, Pr = 0.003 for public schools only, and $\chi^2_{(4)}$= 13.7166, Pr = 0.008 for public and independent schools).

Sub Place (SP)

Straight-line distance

As evident in Table 6.38 and Figure 6.22, children living in poorer SP areas travel shorter distances to school than their peers living in more affluent SP areas (Kruskal-Wallis test; Pr=0.0133). A weakly negative correlation between SP poverty and distance travelled remains in place (-0.0455 for raw scores).

Area Poverty Quintile	Number of children	Mean distance to school	Standard deviation	25th percentile distance	50th percentile distance	75th percentile distance
1 (wealthiest)	223	5.314km	8.712km	0.535km	1.263km	5.391km
2	264	6.345km	8.666km	0.635km	1.742km	9.920km
3	283	5.636km	9.370km	0.522km	1.020km	7.425km
4	258	4.393km	6.912km	0.572km	1.010km	4.667km
5 (poorest)	252	5.011km	8.340km	0.678km	1.421km	5.916km

Table 6.38: 2003 distance from home to school, by SP poverty quintile

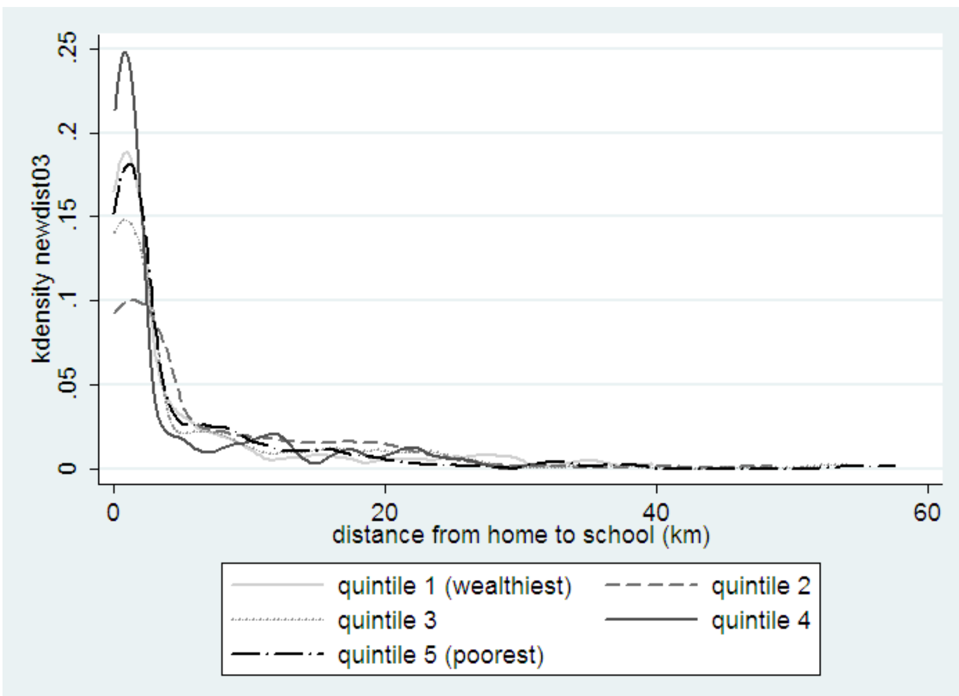

Figure 6.22: Kernel density plot of 2003 distance from home to school, by SP poverty quintile

Census geography

Again, children living in more affluent SP areas are more likely to attend school within the same SAL in which they live than their peers in poorer SP areas ($\chi^2_{(4)}$= 8.0401, Pr = 0.090). A similar pattern is also found for mobility at the SP ($\chi^2_{(4)}$= 22.6349, Pr = 0.000), MP ($\chi^2_{(4)}$= 25.5792, Pr = 0.000) and MN ($\chi^2_{(4)}$= 8.4396, Pr = 0.077) levels.

Nearest school

Children living in the most affluent SPs are again more likely to attend their nearest school than those in poorer SPs ($\chi^2_{(4)}$= 21.7449, Pr = 0.000 for public schools only; $\chi^2_{(4)}$=19.7176 Pr = 0.001 for public and independent schools).

Main Place (MP)

Straight-line distance

At the MP level, the 1997 finding that the children living in the poorest MPs tend to travel further than children in more affluent MPs is replicated (see Table 6.39 and Figure 6.23 below). Children in the wealthiest MPs travel the least far, children in mid-range MPs travel somewhat further, and the children in the highest-poverty MPs travel furthest of all (Kruskal-Wallis test, Pr=0.053). There is a positive correlation between MP poverty score and distance travelled (0.0824), indicating that as MP poverty increases, so too does distance travelled. The fact that the 1997 finding is replicated here suggests that MP area poverty might have a different relationship with mobility than poverty at smaller area levels.

MP Area Poverty Quantile	Number of children	Mean distance to school	Standard deviation	25th percentile distance	50th percentile distance	75th percentile distance
1 (wealthiest)	413	4.809km	8.737km	0.500km	1.106km	4.300km
2	743	5.707km	8.504km	0.635km	1.293km	7.377km
3 (poorest)	125	5.070km	7.145km	0.653km	1.572km	7.583km

Table 6.39: 2003 distance from home to school, by MP poverty quantile

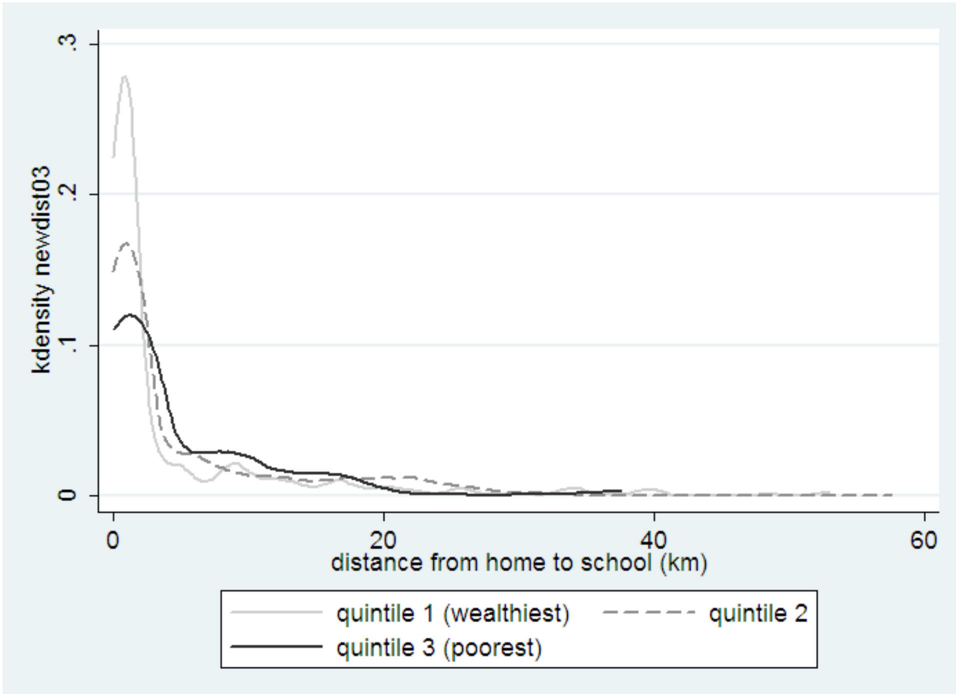

Figure 6.23: Kernel density plot of 2003 distance from home to school, by MP poverty quantile

Census geography

Children living in more affluent MP areas are more likely than their peers in poorer MP areas to attend a school in the same SAL ($\chi^2_{(2)}$= 13.3806, Pr = 0.001), SP ($\chi^2_{(2)}$= 20.7043, Pr = 0.000), MP ($\chi^2_{(2)}$= 27.7652, Pr = 0.000) and MN ($\chi^2_{(2)}$= 13.5678, Pr = 0.001) as their homes in 2003.

Nearest school

Children living in more affluent MPs are significantly more likely to attend their nearest school in 2003 than their peers living in less affluent MP areas ($\chi^2_{(2)}$= 39.5873 Pr = 0.000 for public schools only; $\chi^2_{(2)}$= 37.8819, Pr = 0.000 for public and independent schools).

6.10.3 Discussion of residential area poverty and mobility

Two main streams of findings emerge from all the above analyses of residential area poverty. Firstly, there is a fairly complex relationship between the area a child lives in, and his or her education-related mobility The exact nature of this relationship is not clear from the analyses documented above, although the overall trend seems to be that on average children living in relatively advantaged areas, as well as those in the most disadvantaged areas, tend to travel furthest. Secondly, however, there is a strong and straightforward relationship between the area a child lives in, and the likelihood that he or she will attend the nearest grade appropriate school, whether public or independent.

The patterns evident in the first set of findings seem likely to relate to complex interactions between neighbourhood poverty and a range of household and individual characteristics. Overall, though, with the exception of some of the nearest-school analyses, children living in wealthier areas do seem to be more mobile than children living in poorer areas. As these children are, on average, more affluent than the children from poorer areas, it makes sense that their parents or families are more likely to have the resources (economic and social), to engage in mobility. However, children living in poorer areas, with poorer local schools, almost certainly have far stronger incentives to engage in mobility. The patterns of mobility revealed by the analyses presented here seem most likely to represent this interaction between incentives and capacities for mobility.

The second set of findings is simpler to explain, and almost certainly relates to the higher quality of schools in more affluent areas, regardless of the geographic level at which these areas are defined. Children in more affluent areas therefore almost certainly have less of an incentive to choose to attend a school other than the one closest to their home. Two alternative explanations should, however, also be considered. Firstly, in affluent areas, which are generally more spread out, a child is more likely to be faced with having only one school within easy walking distance of his or her home, and may, therefore, be constrained to attend this particular school in a way that children in less affluent areas with a denser distribution of schools are not constrained. However, given that there is a fairly strong relationship between household SES and the affluence of the area in which a child lives, it seems likely that this sort of constraint should not affect many children living in more advantaged areas. Secondly, and particularly in 1997 when the children are younger, is the fact that in affluent areas most schools are English or Afrikaans medium, while in less affluent areas, they may be operating in any one of several African languages that are heavily represented in the Johannesburg-Soweto area. While parents may consider education in English or Afrikaans even if these are not the home language, they may not be willing to accept education in an African language other than the child's home language. This may, to some extent, be governing decision making around which school children will attend, and whether the nearest school to the home is an option.

The finding that children in less affluent areas are less likely to attend their nearest school – even if they still don't travel very far – is of particular importance. It appears to indicate that children and families living in poor areas still exercise school choice as far as they are able to, given the resource constraints that they face.

6.11 Conclusion

This chapter has explored the relationship between a range of individual, household and community level variables, and learner mobility in South Africa. At the individual level, race, gender, age at first enrolment, and grade repetition were examined. Clear evidence was presented that race is strongly linked to mobility behaviour, and particularly that coloured children are less likely to engage in mobility than black children. Although girls travel slightly further than boys, there was not sufficient evidence to demonstrate a consistent relationship between gender and mobility.

Interestingly, and contrary to expectations, children who started school in 1997, as opposed to 1996, tended to be substantially more mobile in 1997. While the reasons for this are unclear, the most plausible relates to whether or not parents can afford preschool or other child care if a child is not attending school. This effect had largely worn off by 2003, perhaps in part due to the effect of late starters being more likely to have reached secondary school in 2003. There was a strong relationship between mobility and repetition, with repetition being much less likely for children travelling further. This is almost certainly a reflection of household SES, which was higher with greater travel, and the fact that for most children, travelling further is associated with attending a higher quality school.

At the household level, maternal education, maternal marital status, and household SES were explored. A strong, positive relationship between mobility and maternal education was identified. In particular, children with mothers who had an education up to and including grade 10 had a very different distribution of mobility from those children with mother who had completed grade 11 or higher. Although much of the relationship between maternal education and learner mobility is likely to be mediated by SES, there was also some evidence for an independent effect of maternal education on school choice, particularly in 1997. In particular, children with more educated mothers were substantially less likely to attend their local schools, suggesting that more educated mothers were more actively engaged in school choice. There was, however, very limited evidence for a relationship between maternal marital status and mobility behaviour.

As anticipated, a particularly strong relationship between mobility and socio-economic status was identified. Although this relationship was not strictly linear, children from more affluent households generally tended to live further away from their schools. This relationship appears to become stronger over time, with SES being more closely related to mobility in 2003.

At the community level, measures of residential area poverty were explored at three different levels of census geography. The relationship between these measures and mobility behaviour is relatively complex, which is probably due to the interaction of area and household SES. While children living in poorer areas are likely to have a greater incentive to travel to school, they are less likely to have the resources to do so. By contrast, children in more affluent areas are less likely to need to travel to attend a school of their preference, but are much more likely to have the ability to do so. Overall, however, the children living in relatively advantaged areas, and those living in the most disadvantaged areas appear to be most likely to travel substantial distances, and to travel outside of their residential area to attend school. A more clear relationship exists between area poverty and whether or not a child attends their nearest grade-appropriate school – children living in more affluent areas are substantially more likely to attend the school nearest to their home.

The data presented in this chapter substantiates the notion that there are two different forms of school choice operating in the South African educational market. Firstly, families choose to send children to relatively distant but historically advantaged schools whenever possible. This is evident in the substantially greater distances typically travelled by the more affluent members of the sample, and in the extent to which the relationships between mobility and its correlates seem likely to be shaped by SES.

By contrast, however, there is also evidence that less-advantaged children and their families are also engaged in choice, even if this does not typically involve substantial travel to historically advantaged schools. The prevalence of choice at a more local level is evident in the data on which children attend their nearest school. Unlike the distance travelled to school, this appears to be inversely associated to family wealth. By contrast, maternal education and area poverty play a more significant role in shaping this form of choice. This suggests that when sending children to historically advantaged schools is not an option, engaged parents, and particularly those in less affluent areas, attempt to provide their children with the best possible education by choosing from amongst local schools.

In the next chapter, Chapter 7, the relationship between mobility and school attributes will be explored. This will be followed in Chapter 8 by an investigation of the trajectories of individual learners with respect to mobility over time. This will provide insight as to whether mobility in 1997 can be used to predict mobility in 2003, as well as to potential determinants and correlates of changes in mobility patterns over time. Chapter 9 will then combine the data presented in the current chapter with the findings of Chapters 7 and 8 to develop a preliminary, partial model of the determinants of various forms of learner mobility.

Chapter 7: School characteristics associated with mobile learners

7.1 Introduction

Along with the individual, household and community level variables discussed so far, a child's mobility is also likely to be related to the attributes of the schools available to the child, both locally and further afield. School attributes, such as resource levels, educational quality, cost, and racial composition are all likely to play various roles in attracting or deterring potential learners. It therefore makes sense to explore key variables of both the schools local to a child, and the schools a child attends, in attempting to understand the mobility of any particular child.

This chapter consists of three sections. The first section explores the attributes of those schools attended by sample members from 1997-2003, along with the attributes of those schools closest to the homes of sample members. Data is presented first in unweighted form, and secondly, weighted by the number of children enrolled in each school. The second section explores the distribution of sample members across the schools attended, and provides data on which children attended schools with which properties. Finally, the third section explores the relationships between patterns of travel, and the schools that children attended.

7.2 Schools attended by study sample members, and grade-appropriate schools closest to study sample members' homes

7.2.1 Unweighted data

Obviously, the schools attended by the members of the sample used in this study are only a subset of all of the schools in the Gauteng province described in Chapter 4. They are not likely to be an extremely representative subset either, due to the urban location of sample members, and the under-representation of the extremely poor and the extremely wealthy, as discussed in Chapter 4. For these reasons, it is useful to generate some descriptive data specific to the group of schools attended by sample members. For similar reasons, descriptive data around the subset of schools which comprise the nearest grade-appropriate school for sample members will also be generated. It must be stressed that the data provided in this section is not weighted for the number of cohort members enrolled at each school. Rather, all schools attended by cohort members are weighted equally here – whether they are attended by just one cohort member, or many. Data weighted by the number of sample members enrolled at each school will be explored in the subsequent section.

Overall, in 1997, the 1428 members of the study were attending a total of 365 different registered schools. However, as data is missing for a few individuals, and a small number of children also attended unregistered schools, this figure is probably slightly low. In 2003, this had risen to 465 schools. The figure is higher because both primary schools (310) and secondary schools (155) are included at this point in time. 378 different primary schools are identified as being the nearest primary school to members of the study sample, while 212 secondary schools are identified as being the nearest secondary school to members of the study sample. These figures are consistent with the larger number of relatively smaller primary schools found in the Gauteng province, and the smaller number of relatively larger secondary schools. From these figures, it is also evident that the number of schools actually attended is, for both phases, smaller than the number of schools that would be attended if all children simply attended their nearest school.

School sector

10.54% of the schools attended by sample members during the study period were independent schools. By contrast, roughly 20% of the registered schools in Gauteng were independent. The schools nearest to the homes of sample members represent far more closely the distribution of independent schools found in the full list of registered schools. Roughly 10% of the primary schools closest to a sample member's home were independent, as were 17% of secondary schools.

Quintile

Very few of the primary and secondary schools closest to sample members' homes fall into quintiles 1 (poorest), 2 and 5 (most advantaged). As evident in Table 7.1, the majority of the schools closest to sample members' homes are in either quintile 3 or quintile 4 – in both cases a higher proportion than expected given the proportion of quintile 3 and 4 schools found in Gauteng province as a whole. The proportion of schools actually attended by cohort members that fall into quintiles 1-3 is lower than would be expected on the basis of the schools nearest to sample members' homes. By contrast, the proportion of attended schools falling into quintiles 4 and 5 is somewhat higher than would be expected.

School Quintile	Proportion of Gauteng schools	Proportion of primary schools closest to sample members' homes	Proportion of secondary schools closest to sample members' homes	Proportion of schools attended by sample members in 1997	Proportion of schools attended by sample members in 2003
1	11.66%	3.08%	2.37%	2.85%	1.46%
2	8.61%	8.31%	11.83%	5.38%	5.84%
3	30.91%	44.62%	35.50%	37.66%	36.50%
4	27.27%	35.08%	37.28%	38.61%	41.36%
5	21.55%	8.92%	13.02%	15.51%	14.84%

Table 7.1: Quintiles of schools attended by and nearest to sample members' homes

Section 21 status

Just over 86% of public schools in Gauteng province have Section 21 status, while just under 89% of the schools closest to sample members' homes have this status. By contrast, around 95% of schools actually attended by sample members having Section 21 status.

Enrolment

In line with the data presented for all Gauteng, the primary schools closest to the homes of sample members are, on average, substantially smaller (mean: 626; median: 558) than the secondary schools closest to their homes (mean: 971; median: 1049). By contrast, the mean size of the schools attended 1997 is 663, while for primary schools attended in 2003 it rises to 673 – in both cases higher than might be expected on the basis of the schools nearest to sample members' homes. The mean size of secondary schools attended by sample members is 958, which is lower than would be expected on the basis of the data on the secondary schools nearest to sample members' homes.

Percentage of black learners

The average proportion of black learners across all Gauteng schools is 73%. The average for the primary schools closest to sample members' homes is 83%, and for secondary schools it is 76%. In both cases, the median proportion of black students is almost or exactly 100%, indicating that around or over half of the schools found closest to participants' homes are entirely black. In 1997, the average percentage of black learners at the schools attended by sample members was 79%. For primary schools attended in 2003, this figure rises to 81%. By contrast, the corresponding figure for the secondary schools attended in 2003 is 72%. In all cases, these figures are lower than would be expected on the basis of the schools closest to sample member homes.

School fees

The school fees charged by both the primary and secondary schools closest to sample members' homes are substantially lower than those figures for all Gauteng area schools, with the primary schools charging an average of R504, and the secondary schools and average of R790. However, the fees at the schools which sample members actually attended are substantially higher – although with extremely large standard deviations. The primary schools attended in 1997 charging an average of R910 (minimum R0, maximum R9510, median R80). Primary schools attended in 2003 charged an average of R769 (minimum R0, maximum R7000, median R100). Secondary schools attended in 2003 charged an average of R1470 (minimum R0, maximum R9650, median R400).

Historical racial status of school

70% of the primary schools closest to participant homes were historically DET schools, as were 63% of secondary schools. The schools actually attended by sample members, however, are substantially less likely to be historically DET schools. In 1997, 56% of schools attended were DET schools, and in 2003, this figure was lower still at 54%. These enrolment patterns suggest that children are tending to avoid DET schools.

Matric pass rate

Given that the pass rate data attributed to primary schools was imputed on the basis of the performance of the nearest secondary school, it makes sense that the figures for the groups of primary and secondary schools closest to sample members' homes will be almost identical. The mean pass rate for these secondary schools is 70%, while the mean imputed pass rate for these primary schools is 68%. The figures for the schools actually attended are quite similar, with the mean pass rate for schools attended in 1997 at 70%. Primary schools attended in 2003 also have a mean pass rate of 70%, while for secondary schools it is 72%.

Discussion of unweighted data

The data presented thus far suggests that, as would be expected, the schools closest to study sample members' homes tend to be slightly more disadvantaged than the average schools in the Gauteng province. The schools actually attended, however, are somewhat more advantaged than this. In the following section, the schools data weighted for attendance is explored to determine whether this pattern still holds.

7.2.2 Weighted data

In this section, the focus shifts from using schools to using sample members as the unit of analysis. It will describe the school environments experienced by different proportions of sample members by presenting school data weighted by the number of sample members attending a particular school.

Although the 1428 sample members attended a total of 365 different schools in 1997, and 465 different schools in 2003, they were by no means evenly distributed across these schools, as illustrated in Table 7.2 below. In 1997, 132 schools had only one sample member attending, while 11 different schools had 10 or more sample members attending. The single most-attended school had 20 sample members enrolled. The largest proportion of sample members, 14.02%, attended schools with 6 sample members enrolled. By contrast, relatively even proportions of the sample live closest to each of the school in the set of schools closest to any sample member's home.

In 2003, with the sample members divided between both primary and secondary schools, the distribution of enrolment shifted towards smaller numbers of children at a range of different schools. This trend is particularly evident amongst children attending secondary schools in 2003, with only 7.67% of secondary school children attending schools with ten or more sample members. By contrast, almost 70% of the sample lives closest to one of 57 secondary schools which are also closest to at least 9 other sample members. This probably relates to the typically larger size, and sparser distribution, of secondary schools. This data suggests that as learners move into secondary schooling we should expect to see larger numbers of children enrolled at each of a smaller number of schools. It is interesting to note that the available data, however, appears to reflect an opposite trend. However, this may simply be due to the relatively small numbers of children who have reached secondary school in 2003, and data for subsequent years is needed before any firm conclusions can be drawn.

Sample members per school	**Nearest primary**	**Nearest secondary**	**1997 attended**	**2003 attended, primary**	**2003 attended, secondary**
% at schools with 1-4 sample members	38.59% (272 schools)	13.24% (119 schools)	40.21% (264 schools)	56.03% (280 schools)	44.19% (110 schools)
% at schools with 5-9 sample members	34.37% (76 schools)	17.22% (35 schools)	48.19% (90 schools)	38.31% (65 schools)	48.13% (40 schools)

% at schools with over 10 sample members	27.03% (28 schools)	69.54% (57 schools)	11.6% (11 schools)	5.16% (6 schools)	7.67% (5 schools)

Table 7.2: Schools attended by sample members (note: schools classified as combined are included in both columns for 2003)

Sector

The proportion of children living nearest to and attending independent schools is lower than suggested by the unweighted data presented earlier. Overall, in 1997, 7.09% of sample members were attending independent schools, as illustrated in Table 7.3 below. A slightly higher proportion of children had independent schools as their nearest primary (7.60%) and secondary schools (8.23%). The relatively low proportion of sample members attending independent schools may relate to a number of different factors, including the typically smaller size of independent schools, the poor quality of some independent schools in disadvantaged areas, and the relatively low representation in this study sample of the extremely affluent children most likely to attend the more highly performing independent schools.

	Nearest primary	**Nearest secondary**	**Attended school 1997**	**Attended school 03 – primary**	**Attended school 03 – secondary**
% of sample at independent schools	7.60%	8.23%	7.09%	7.00%	7.67%

Table 7.3: Proportion of sample members closest to and attending independent schools

Quintiles

The proportion of children attending schools in quintiles 1, 2 and 5 is lower than would be expected on the basis of the proportions of schools of each quintile in Gauteng, while the proportion of children in quintile 3 and 4 schools, by contrast, is higher than expected. When contrasting quintile ratings of schools attended with nearest schools (see Table 7.4 below), it is again clear that the proportion of children attending schools in quintiles 1 and 2 is lower than would be expected on the basis of the schools nearest to sample members' homes. This time, however, the proportion attending quintile 3 schools is also lower, while the proportions attending schools in quintiles 4 and 5 are higher.

School quintile	Nearest primary school	Nearest secondary school	Sample members by quintile, 1997	Sample members by quintile, primary school only 2003	Sample members by quintile, secondary school only 2003
1	3.35%	1.93%	1.43%	1.37%	0.00%
2	6.69%	4.59%	3.84%	3.98%	4.03%
3	50.76%	53.51%	46.47%	42.61%	38.79%
4	35.46%	35.86%	39.68%	42.61%	48.36%
5	3.75%	4.11%	8.58%	9.44%	8.82%

Table 7.4: Distribution of sample members across schools by school quintile rating

Section 21 Status

In both 1997 and 2003, only slightly over 4% of learners were attending schools without Section 21 status (see Table 7.5 below). These figures are lower than the proportion of learners whose nearest primary school did not have Section 21 status (6.82%), and substantially lower than the proportion whose nearest secondary school did not have that status (17.42%).

	Nearest primary	Nearest secondary	Attended school 1997	Attended school 03 – primary	Attended school 03 – secondary
% of sample at schools without Section 21 status	6.82%	17.42%	4.34%	4.53%	4.20%

Table 7.5: Proportion of schools without Section 21 status nearest to and attended by sample members

Total school enrolment

As larger schools enrol more children, it is expected that a higher proportion of our sample members will be attending larger schools. As a result, the school size experienced our average learner should be higher than the size of the average school available to him or her. This is indeed reflected in the data. The mean primary school size experienced by sample members in 1997 was 675, and 720 in 2003, both of which are higher than the average size (663) of attended primary schools. The mean school size experienced by sample members attending secondary school in 2003 is 1060 learners, compared to a mean attended secondary school size of 971. A similar argument applies to the schools nearest to children's homes, and this is also reflected in the data. The average sample member's closest primary school has just over 647 children, compared to an average of 626 learners for all the primary schools nearest to sample members' homes. Similarly, the average sample member's nearest secondary school enrols 1016 children, compared to an average school size of 971.

As is evident in Table 7.6 below, the mean school size experienced by a sample member in 1997 is a little larger than would be expected on the basis of the mean size of the closest primary school. The figure for those children still in primary school in 2003 is even higher. The mean size of secondary schools attended in 2003 is also a little higher than would be expected on the basis of the mean size of the secondary schools closest to the children's homes.

	Nearest primary school size	Nearest secondary school size	Attended school 1997	Attended school 03 – primary	Attended school 03 – secondary
Mean school size	647.17	1016.13	674.75	719.72	1060.35
Median school size	563	1090	641	694	1131.5

Table 7.6: Size of schools nearest to and attended by sample members

Proportion Black

For the average child in the sample, both the nearest primary and secondary schools were 86% black, while for the median child both schools were 100% black, as illustrated in Table 7.7 below. These figures are higher than would be predicted from the data for all Gauteng schools, as well as from the figures for those schools that sample members actually attended. These figures do, however, make sense, as the majority of sample members lived in majority black areas, and majority black schools were more likely to have multiple sample members enrolled, particularly at the secondary school level.

The figures for the schools actually attended in 1997 and in 2003 are substantially lower than those for the nearest schools, suggesting that on average, children are attending schools with a lower proportion of black learners than those nearest to their homes. This difference is particularly marked for those children who have reached secondary school by 2003. However, as this is a non-representative sub-section of the sample, it is not possible to determine whether this larger difference is attributable to their being in secondary school, or to their being more advantaged or academically more promising than their peers still in primary school.

	Nearest primary	**Nearest secondary**	**Attended school 1997**	**Attended school 03 – primary**	**Attended school 03 – secondary**
Mean % black learners	86.29	85.78	81.98	81.07	77.25
Median % black learners	100	100	100	100	99.75

Table 7.7: Proportion of black learners at schools nearest to and attended by sample members

School Fees

Table 7.8 below illustrates the fees charged at the primary and secondary schools nearest to, and attended by, the average sample member, for both 1997 and 2003. The fees charged at the schools attended by sample members are higher than would be predicted on the basis of the fees charged by the schools nearest to their homes. This divergence is particularly notable at the secondary school level, although again it is not clear whether this is due to the nature of those children in secondary school in 2003, or whether it simply relates to the fact of their attending secondary schools.

	Nearest primary	Nearest secondary	Attended school 1997	Attended school 2003 – primary	Attended school 2003 – secondary
Mean school fees	R255	R358	R564	R654	R812
Median school fees	R50	R100	R60	R100	R150

Table 7.8: Fees charged by schools nearest to and attended by sample members

Historical racial status of school

Table 7.9, below, illustrates that a smaller proportion of sample members attended ex-DET schools than would be predicted on the basis of the schools closest to their homes. The proportion attending DET schools falls between 1997 and 2003, and is lowest for those children attending secondary schools in 2003.

	Nearest primary	Nearest secondary	Attended school 1997	Attended school 2003 – primary	Attended school 2003 – secondary
% of sample at ex-DET schools	76.72%	76.23%	63.01	58.39%	54.50%

Table 7.9: Proportion of schools nearest to and attended by sample members that were historically under the DET

Matric Pass Rate

Once again, the average sample member attends a school with better pass rates than would be predicted on the basis of the school closest to their home (see Table 7.10 below). Also, similarly to other variables explored, this difference becomes more marked at the secondary school level, although, as indicated previously, the reasons for this increase cannot be determined from the available data.

	Nearest primary	Nearest secondary	Attended school 1997	Attended school 2003 – primary	Attended school 03 – secondary
Mean pass rate	64%	64%	66%	67%	70%
Median pass rate	63%	63%	66%	66%	69%

Table 7.10: Pass rates of schools nearest to and attended by sample members.

Discussion of weighted data

A few key themes emerge from the data presented here. Firstly, it is clear that across all variables, on average, a sample member attends a school that is more advantaged than would be predicted on the basis of the school nearest to his or her home. Secondly, across all variables, this difference is greater at the secondary school level, although the reasons for this are not clear. Thirdly, those children still attending primary school in 2003 are typically attending schools that more advantaged than those that were attended in 1997. These findings suggest that to the extent that school choice is being used by sample members, it is being used to enhance the educational opportunities available to sample members.

7.3 Which children attend which schools?

This section explores the distribution of sample members across schools with different properties on the basis of child, household and community attributes. Tests are conducted using data for both 1997 and 2003.

7.3.1 Child level variables

Race

Table 7.11, below, shows that schools attended by children of different races vary significantly on the basis of the majority of the school attribute variables examined, with black children typically more likely to be attending lower quality schools.

	1997	2003 (all data)	2003 (primary schools only)	2003 (secondary schools only)
Sector	n.s.	Pr=0.008 (Black children more likely to attend public schools)	n.s.	Pr=0.013 (Black children more likely to attend public schools)
Quintile	Pr=0.0000 (Black children more likely to attend lower quintile schools)	Pr=0.0000 (Black children more likely to attend lower quintile schools)	Pr=0.0000 (Black children more likely to attend lower quintile schools)	Pr=0.0000 (Black children more likely to attend lower quintile schools)
Section 21	Pr=0.026 (Black children less likely to attend Section 21 schools)	n.s.	Pr=0.070 (Black children less likely to attend Section 21 schools)	Pr=0.004 (Black children more likely to attend Section 21 schools)
School size	Pr=0.0000 (Black children are more likely to be enrolled in smaller schools)	Pr=0.0000 (Black children are more likely to be enrolled in smaller schools)	Pr=0.0000 (Black children are more likely to be enrolled in smaller schools)	Pr=0.0000 (Black children are more likely to be enrolled in smaller schools)
Proportion black	Pr=0.0000 (Black children likely to attend schools with a higher proportion of black learners)	Pr=0.0000 (Black children likely to attend schools with a higher proportion of black learners)	Pr=0.0000 (Black children likely to attend schools with a higher proportion of black learners)	Pr=0.0000 (Black children likely to attend schools with a higher proportion of black learners)
School fees	Pr=0.0000 (Black children more likely to attend schools with lower fees)	Pr=0.0000 (Black children more likely to attend schools with lower fees)	Pr=0.0000 (Black children more likely to attend schools with lower fees)	Pr=0.0000 (Black children more likely to attend schools with lower fees)
Historical DET status	Pr=0.000 (Black children were more likely to attend a historically DET school)	Pr=0.000 (Black children were more likely to attend a historically DET school)	Pr=0.000 (Black children were more likely to attend a historically DET school)	Pr=0.000 (Black children were more likely to attend a historically DET school)
Matric Pass	Pr=0.0299	n.s.	n.s.	n.s.

rate	(Black children more likely to attend schools with higher pass rates)			

Table 7.11: Relationship between child race (black and coloured children only) and properties of the school he or she attends

Gender

As illustrated in Table 7.12 below, only two school properties are significantly related to child gender in 1997: boys are more likely to attend schools with a higher proportion of black learners, as well as schools that are former DET schools. In 2003, a number of significant relationships are evident, but these appear to relate largely to the higher proportion of girls attending secondary school by 2003, and all disappear when school phase is controlled for. The one exception to this pattern is school sector, with girls being significantly more likely to attend independent secondary schools.

Overall, there is weak evidence that during the early years of schooling, girls may be more likely to attend more integrated schools than boys, and less likely to attend former DET schools. These relationships do not survive in 2003, when school phase is controlled for. Girls are, however, substantially more likely to have reached secondary school in 2003, which may be obscuring relationships between gender and enrolment patterns at the secondary school level .

	1997	**2003 (all data)**	**2003 (primary schools only)**	**2003 (secondary schools only)**
Sector	n.s.	n.s.	n.s.	Pr = 0.030 (Girls more likely to attend independent schools)
Quintile	n.s.	Pr=0.0468 (Girls more likely to attend	n.s.	n.s.

		higher quintile schools)		
Section 21	n.s.	n.s.	n.s.	n.s.
School size	n.s.	Pr=0.0173 (Girls more likely to attend larger schools)	n.s.	n.s.
Proportion black	Pr=0.0677 (Girls more likely to attend schools with lower proportion of black learners)	Pr=0.0562 (Girls more likely to attend schools with lower proportion of black learners)	n.s.	n.s.
School fees	n.s.	Pr=0.0293 (Girls more likely to attend schools with higher fees)	n.s.	n.s.
Historical DET status	Pr=0.054 (Girls less likely to attend former DET schools)	Pr=0.060 (Girls less likely to attend former DET schools)	n.s.	n.s.
Matric Pass rate	n.s.	n.s.	n.s.	n.s.

Table 7.12: Relationship between child gender and properties of the school he or she attends

Age at enrolment

The 1997 analysis shows strong evidence for a relationship between the age of school enrolment and the type of school attended, with children starting school at a later age attending more advantaged schools (see Table 7.13 below). This is in line with the finding in Chapter 6 that children who start school later are more likely to live further away from their schools. Given the extremely strong relationship between age at school enrolment and the child's phase of schooling in 2003, results for 2003 should be treated with some caution, although they are largely in line with the 1997 results. Overall, however, this data suggests that children who start school at a later age tend to access more highly performing schools.

Age at first school enrolment	**1997**	**2003 (all data)**	**2003 (primary schools only)**	**2003 (secondary schools only)**
Sector	Pr=0.001 (children starting school late less likely to attend public schools)	n.s.	Pr= 0.067 (children starting school late less likely to attend public schools)	Pr=0.000 (children starting school late less likely to attend public schools)
Quintile	Pr=0.000 (children starting school late more likely to attend quintile 5 schools)	Pr=0.011 (children starting school late more likely to attend quintile 5 schools)	Pr=0.077 (children starting school late more likely to attend quintile 5 schools)	n.s.
Section 21	n.s.	n.s.	n.s.	Pr=0.008 (children starting school late less likely to attend Section 21 schools)
School size	n.s.	Pr=0.0000 (children starting school late likely to attend smaller schools)	n.s.	Pr=0.0621(children starting school late likely to attend smaller schools)
Proportion black	Pr=0.0019 (children starting school late likely to attend schools with a lower proportion of black learners)	n.s.	Pr=0.0779 (children starting school late likely to attend schools with a lower proportion of black learners)	n.s.
School fees	Pr= 0.0000 (children starting school late likely to	Pr=0.0979 (children starting school late likely to attend schools	Pr=0.0151 (children starting school late likely to attend schools	n.s.

	attend schools with higher fees)	with higher fees)	with higher fees)	
Historical DET status	P=0.000 (children starting school late less likely to attend former DET schools)	Pr=0.058 (children starting school late less likely to attend former DET schools)	n.s.	n.s.
Matric Pass rate	Pr= 0.0007 (children starting school late more likely to attend schools with higher pass rates)	n.s.	n.s.	n.s.

Table 7.13: Relationship between child age at first school enrolment and school properties

Repetitions

The analyses related to this grade repetition should be interpreted somewhat differently to the other variables presented, as repetition seems to be shaped more by the properties of the school a child attends than by the attributes of the child him or herself. In addition, while repetition precedes 2003 schooling choices, it occurs only after 1997 schooling choices have already been made, meaning that it cannot be influencing school choice for that year (although a child's academic capability might be influencing choice). However, it is quite likely that school choice in 1997 influences repetition in subsequent years. In addition, grade repetition strongly influences whether or not a child has reached secondary school by 2003. This means that finding for 2003, particularly for the secondary school level, should be treated with some caution.

Although the causality behind the relationships documented in Table 7.14, below, is not clear, there are highly significant relationships between grade repetition and most indicators of school quality, other than Section 21 status. All relationships operate in the expected direction, with children who have repeated a grade tending to attend less advantaged and more poorly performing schools than those who have never repeated a grade. The relationship between repetition and school size is likely to be due to the tendency for primary schools in poorer areas to be somewhat smaller than those in more affluent areas, as mentioned previously.

	1997	**2003 (all data)**	**2003 (primary schools only)**	**2003 (secondary schools only)**
Sector	Pr=0.009 (Grade repeaters more likely to attend public schools)	Pr=0.013 (Grade repeaters more likely to attend public schools)	Pr=0.012 (Grade repeaters more likely to attend public schools)	n.s.
Quintile	Pr=0.000 (Grade repeaters more likely to attend schools in quintiles 1, 2 or 3)	Pr=0.000 (Grade repeaters more likely to attend schools in quintiles 1, 2 or 3)	Pr=0.000 (Grade repeaters more likely to attend schools in quintiles 1, 2 or 3)	n.s.
Section 21	n.s.	n.s.	n.s.	n.s.
School size	Pr= 0.0000 (Grade repeaters more likely to attend smaller schools)	Pr= 0.0000 (Grade repeaters more likely to attend smaller schools)	Pr= 0.0284 (Grade repeaters more likely to attend smaller schools)	n.s.
Proportion black	Pr= 0.000 (Grade repeaters more likely to attend a school with a higher proportion of black learners)	Pr= 0.000 (Grade repeaters more likely to attend a school with a higher proportion of black learners)	Pr= 0.000 (Grade repeaters more likely to attend a school with a higher proportion of black learners)	Pr= 0.0208 (Grade repeaters more likely to attend a school with a higher proportion of black learners)

School fees	Pr= 0.000 (Grade repeaters more likely to attend a school with lower fees)	Pr= 0.000 (Grade repeaters more likely to attend a school with lower fees)	Pr= 0.000 (Grade repeaters more likely to attend a school with lower fees)	n.s.
Historical DET status	Pr= 0.000 (Grade repeaters more likely to attend a former DET school)	Pr= 0.000 (Grade repeaters more likely to attend a former DET school)	Pr= 0.000 (Grade repeaters more likely to attend a former DET school)	Pr= 0.014 (Grade repeaters more likely to attend a former DET school)
Matric Pass rate	Pr=0.0149 (Grade repeaters more likely to attend a school with a higher pass rate)	Pr=0.0010 (Grade repeaters more likely to attend a school with a higher pass rate)	Pr=0.0091 (Grade repeaters more likely to attend a school with a higher pass rate)	n.s.

Table 7.14: Relationship between grade repetitions between 1997 and 2003, and school properties

Child level variables associated with school enrolment patterns: Discussion

The data presented here suggests that race remains closely connected to the attributes of the school a child attends, with Black children particularly likely to attend historically disadvantaged schools. The data does not provide strong evidence for any consistent relationship between gender and schooling enrolment patterns, with the exception that girls are substantially more likely than boys to have reached secondary schooling by 2003. It does, however, indicate that children enrolling in school for the first time at a later age are significantly more likely to attend more advantaged schools than children who enrol earlier. Finally, there is also evidence that children who have repeated a grade attend less advantaged schools than children who have never repeated a grade, although with the available data it is not possible to disentangle the causal roles of school quality and a learner's inherent academic capabilities in this relationship.

7.3.2 Family & household variables:

Maternal education

Table 7.15, below, provides evidence for highly significant relationships between maternal education and almost all of the school attributes examined. In all cases where a relationship between a school property and expected school quality exists, children with more educated mothers are more likely to attend higher quality schools. The one exception to this is Section 21 status, which is unrelated to maternal education. Additionally, in 1997 there is no relationship between maternal education and the imputed pass rate of primary schools, but this probably relates to the imputation process used.

	1997	**2003 (all data)**	**2003 (primary schools only)**	**2003 (secondary schools only)**
Sector	Pr= 0.000 (children of more educated mothers less likely to attend public schools)	Pr= 0.000 (children of more educated mothers less likely to attend public schools)	Pr= 0.000 (children of more educated mothers less likely to attend public schools)	n.s.
Quintile	Pr=0.0001 (children of more educated mothers likely to attend higher quintile schools)	Pr=0.0001 (children of more educated mothers likely to attend higher quintile schools)	Pr=0.0001 (children of more educated mothers likely to attend higher quintile schools)	Pr=0.0002 (children of more educated mothers likely to attend higher quintile schools)
Section 21	n.s.	n.s.	n.s.	n.s.
School size	n.s.	Pr= 0.0490 (children of more educated mothers likely to attend	n.s.	Pr= 0.0490 (children of more educated mothers likely to attend

		larger schools)		smaller schools)
Proportion black	Pr=0.0001 (children of more educated mothers likely to attend school with lower proportion of black learners)	Pr=0.0001 (children of more educated mothers likely to attend school with lower proportion of black learners)	Pr=0.0001 (children of more educated mothers likely to attend school with lower proportion of black learners)	Pr=0.0320 (children of more educated mothers likely to attend school with lower proportion of black learners)
School fees	Pr= 0.0001 (children of more educated mothers likely to attend school with higher fees)	Pr= 0.0001 (children of more educated mothers likely to attend school with higher fees)	Pr= 0.0001 (children of more educated mothers likely to attend school with higher fees)	Pr= 0.0001 (children of more educated mothers likely to attend school with higher fees)
Historical DET status	Pr= 0.000 (children of more educated mothers less likely to attend a former DET school)	Pr= 0.000 (children of more educated mothers less likely to attend a former DET school)	Pr= 0.000 (children of more educated mothers less likely to attend a former DET school)	Pr= 0.033 (children of more educated mothers less likely to attend a former DET school)
Matric Pass rate	n.s.	Pr=0.0002 (children of more educated mothers likely to attend a school with a higher pass rate)	Pr=0.0098 (children of more educated mothers likely to attend a school with a higher pass rate)	Pr=0.0015 (children of more educated mothers likely to attend a school with a higher pass rate)

Table 7.15: Relationship between maternal education and school properties

Maternal marital status

As reflected in Table 7.16 below, maternal marital status is significantly related to most school attributes, with the children of married mothers typically being more likely to attend more advantaged schools.

	1997	2003 (all data)	2003 (primary schools only)	2003 (secondary schools only)
Sector	Pr= 0.001 (children of married mothers less likely to attend a public school)	Pr= 0.005 (children of married mothers less likely to attend a public school)	Pr= 0.028 (children of married mothers less likely to attend a public school)	Pr= 0.094 (children of married mothers less likely to attend a public school)
Quintile	Pr=0.019 (children of married mothers more likely to attend a quintile 4 or 5 school)	Pr=0.000 (children of married mothers more likely to attend a quintile 4 or 5 school)	Pr=0.000 (children of married mothers more likely to attend a quintile 4 or 5 school)	n.s.
Section 21	n.s.	n.s.	n.s.	Pr=0.048 (children of married mothers less likely to attend a Section 21 school)
School size	Pr=0.0009 (children of married mothers likely to attend larger schools)	Pr=0.0179 (children of married mothers likely to attend larger schools)	Pr=0.0041 (children of married mothers likely to attend larger schools)	n.s.
Proportion black	Pr=0.0000 (children of married mothers likely to attend schools with a lower proportion of black learners)	Pr=0.0000 (children of married mothers likely to attend schools with a lower proportion of black learners)	Pr=0.0000 (children of married mothers likely to attend schools with a lower proportion of black learners)	Pr=0.0000 (children of married mothers likely to attend schools with a lower proportion of black learners)
School fees	Pr=0.0000 (children of married mothers likely to attend schools with higher fees)	Pr=0.0000 (children of married mothers likely to attend schools with higher fees)	Pr=0.0000 (children of married mothers likely to attend schools with higher fees)	n.s.

Historical DET status	Pr=0.0000 (children of married mothers less likely to attend former DET schools)	Pr=0.0000 (children of married mothers less likely to attend former DET schools)	Pr=0.0000 (children of married mothers less likely to attend former DET schools)	Pr=0.0000 (children of married mothers less likely to attend former DET schools)
Matric Pass rate	Pr=0.0191 (children of married mothers likely to attend schools with higher pass rates)	Pr=0.0033 (children of married mothers likely to attend schools with higher pass rates)	Pr=0.0315 (children of married mothers likely to attend schools with higher pass rates)	Pr=0.0690 (children of married mothers likely to attend schools with higher pass rates)

Table 7.16: Relationship between maternal marital status and school properties

Household SES: 1997

All relationships were explored using both the derived SES quintiles, as well as the raw SES scores. As the results were largely identical, only the results of the tests based on the quintiles are documented in Table 7.17 below. These results indicate the existence of strongly significant relationships between 1997 household SES, and all indicators of school quality, in both 1997 and 2003, with the exception of Section 21 status. The highly significant nature of all of these relationships is expected given existing evidence for a close relationship between SES and educational opportunity in urban South Africa.

	1997	**2003 (all data)**	**2003 (primary schools only)**	**2003 (secondary schools only)**
Sector	Pr=0.000 (Higher SES children less likely to attend public schools)	Pr=0.000 (Higher SES children less likely to attend public schools)	Pr=0.000 (Higher SES children less likely to attend public schools)	Pr=0.003 (Higher SES children less likely to attend public schools)
Quintile	Pr=0.0001 (Higher SES children likely to attend higher quintile schools)	Pr=0.0001 (Higher SES children likely to attend higher quintile schools)	Pr=0.0001 (Higher SES children likely to attend higher quintile schools)	Pr=0.0001 (Higher SES children likely to attend higher quintile schools)

Section 21	n.s.	n.s.	n.s.	n.s.
School size	Pr=0.0001 (Higher SES children likely to attend larger schools)	Pr=0.0026 (Higher SES children likely to attend larger schools)	Pr=0.0001 (Higher SES children likely to attend larger schools)	n.s.
Proportion black	Pr=0.0001 (Higher SES children likely to attend schools with a lower proportion of black learners)	Pr=0.0001 (Higher SES children likely to attend schools with a lower proportion of black learners)	Pr=0.0001 (Higher SES children likely to attend schools with a lower proportion of black learners)	Pr=0.0001 (Higher SES children likely to attend schools with a lower proportion of black learners)
School fees	Pr=0.0001 (Higher SES children likely to attend a school with higher fees)	Pr=0.0001 (Higher SES children likely to attend a school with higher fees)	Pr=0.0001 (Higher SES children likely to attend a school with higher fees)	Pr=0.0001 (Higher SES children likely to attend a school with higher fees)
Historical DET status	Pr=0.000 (Higher SES children less likely to attend a former DET school)	Pr=0.000 (Higher SES children less likely to attend a former DET school)	Pr=0.000 (Higher SES children less likely to attend a former DET school)	Pr=0.000 (Higher SES children less likely to attend a former DET school)
Matric Pass rate	Pr=0.0001 (Higher SES children likely to attend a school with higher pass rate)	Pr=0.0001 (Higher SES children likely to attend a school with higher pass rate)	Pr=0.0190 (Higher SES children likely to attend a school with higher pass rate)	Pr=0.0001 (Higher SES children likely to attend a school with higher pass rate)

Table 7.17: Relationship between 1997 household SES and school attributes in both 1997 and 2003

Household SES: 2003

Again, tests were conducted using both 2003 SES quintiles and raw scores, and as the results were largely identical, only the results of the tests conducted using the quintiles are presented in Table 7.18 below. These results were extremely similar to those for 1997 SES, which is anticipated as the two SES scores are strongly related. Once again, the only school attribute which was not related to SES was a child's school had Section 21 status.

	1997	2003 (all data)	2003 (primary schools only)	2003 (secondary schools only)
Sector	Pr=0.000 (Higher SES children less likely to attend public schools)	Pr=0.000 (Higher SES children less likely to attend public schools)	Pr=0.000 (Higher SES children less likely to attend public schools)	Pr=0.015 (Higher SES children less likely to attend public schools)
Quintile	Pr=0.0001 (Higher SES children likely to attend higher quintile schools)	Pr=0.0001 (Higher SES children likely to attend higher quintile schools)	Pr=0.0001 (Higher SES children likely to attend higher quintile schools)	Pr=0.0001 (Higher SES children likely to attend higher quintile schools)
Section 21	n.s.	n.s.	n.s.	n.s.
School size	Pr=0.0001 (Higher SES children likely to attend larger schools)	n.s.	Pr=0.0148 (Higher SES children likely to attend larger schools)	n.s.
Proportion black	Pr=0.0001 (Higher SES children likely to attend schools with a lower proportion of black learners)	Pr=0.0001 (Higher SES children likely to attend schools with a lower proportion of black learners)	Pr=0.0001 (Higher SES children likely to attend schools with a lower proportion of black learners)	Pr=0.0001 (Higher SES children likely to attend schools with a lower proportion of black learners)
School fees	Pr=0.0001 (Higher SES children likely to attend a school with higher fees)	Pr=0.0001 (Higher SES children likely to attend a school with higher fees)	Pr=0.0001 (Higher SES children likely to attend a school with higher fees)	Pr=0.0001 (Higher SES children likely to attend a school with higher fees)
Historical DET status	Pr=0.000 (Higher SES children less likely to attend a former DET school)	Pr=0.000 (Higher SES children less likely to attend a former DET school)	Pr=0.000 (Higher SES children less likely to attend a former DET school)	Pr=0.000 (Higher SES children less likely to attend a former DET school)
Matric Pass rate	Pr=0.0044 (Higher SES children likely to attend a school with higher pass	Pr=0.0001 (Higher SES children likely to attend a school with higher pass	Pr=0.0015 (Higher SES children likely to attend a school with higher pass	Pr=0.0002 (Higher SES children likely to attend a school with higher pass

	rate)	rate)	rate)	rate)

Table 7.18: Relationship between 2003 household SES and school attributes in 1997 and 2003

Household level variables associated with school enrolment patterns: Discussion

Significant relationships were identified between each of the household level variables examined, and all school attributes with the exception of Section 21 status. In all cases, relationships operated in the expected direction. Children of more educated mothers and married mothers were more likely to be attending more advantaged schools, as were children living in more advantaged households in either 1997 or 2003. This provides strong evidence that more advantaged children tend to have access to better educational opportunities in contemporary urban South African than their less advantaged peers.

7.3.3 Community level variables:

Residential Area Poverty

Small Area Level

Again, as tests on residential area poverty quintiles and raw scores produced largely identical results, only the results of analyses conducted using the quintiles are presented in Table 7.19 below. Overall the results are fairly clear, with children living in poorer areas more likely to attend less advantaged schools.

SAL Poverty	**1997**	**2003 (all data)**	**2003 (primary schools only)**	**2003 (secondary schools only)**
Sector	Pr=0.000 (Children living in poorer SALs more likely to attend a public school)	Pr=0.000 (Children living in poorer SALs more likely to attend a public school)	Pr=0.000 (Children living in poorer SALs more likely to attend a public school)	n.s.
Quintile	Pr=0.0001	Pr=0.0001	Pr=0.0001	Pr=0.0001

	(Children living in poorer SALs likely to attend lower quintile schools)	(Children living in poorer SALs likely to attend lower quintile schools)	(Children living in poorer SALs likely to attend lower quintile schools)	(Children living in poorer SALs likely to attend lower quintile schools)
Section 21	n.s.	n.s.	n.s.	Pr=0.092 (Children living in richer SALs less likely to attend Section 21 schools)
School size	Pr=0.0001 (Children living in richer SALs likely to attend larger schools)	Pr=0.0001 (Children living in richer SALs likely to attend larger schools)	Pr=0.0001 (Children living in richer SALs likely to attend larger schools)	Pr=0.0232 (Children living in richer SALs likely to attend larger schools)
Proportion black	Pr=0.0001 (Children living in poorer SALs likely to attend schools with a higher proportion of black learners)	Pr=0.0001 (Children living in poorer SALs likely to attend schools with a higher proportion of black learners)	Pr=0.0001 (Children living in poorer SALs likely to attend schools with a higher proportion of black learners)	Pr=0.0001 (Children living in poorer SALs likely to attend schools with a higher proportion of black learners)
School fees	Pr=0.0001 (Children living in poorer SALs likely to attend schools with lower fees)	Pr=0.0001 (Children living in poorer SALs likely to attend schools with lower fees)	Pr=0.0001 (Children living in poorer SALs likely to attend schools with lower fees)	Pr=0.0001 (Children living in poorer SALs likely to attend schools with lower fees)
Historical DET status	Pr=0.000 (Children living in poorer SALs more likely to attend former DET schools)	Pr=0.000 (Children living in poorer SALs more likely to attend former DET schools)	Pr=0.000 (Children living in poorer SALs more likely to attend former DET schools)	Pr=0.000 (Children living in poorer SALs more likely to attend former DET schools)
Matric Pass rate	Pr=0.0001 (Children living in poorer SALs likely to attend schools with lower pass rates)	Pr=0.0001 (Children living in poorer SALs likely to attend schools with lower pass rates)	Pr=0.0001 (Children living in poorer SALs likely to attend schools with lower pass rates)	Pr=0.0012 (Children living in poorer SALs likely to attend schools with lower pass rates)

Table 7.19: Relationship between residential SAL poverty and school attributes in 1997 and 2003

Sub-place level

The results for the analyses based on SP poverty levels (see Table 7.20 below) were similar to those conducted for the SAL poverty levels, with children in poor areas typically attending less advantaged schools. However, for matric pass rate, the relationship shifted from a linear relationship in which children living in more affluent areas attended schools with higher pass rates, to a non-linear relationship in which children in both the wealthiest and poorest areas attended schools with higher pass rates than children in moderate-poverty areas. Potential explanations for this shift are not clear.

SP Poverty	1997	2003 (all data)	2003 (primary schools only)	2003 (secondary schools only)
Sector	Pr=0.000 (Children living in poorer SPs more likely to attend a public school)	Pr=0.000 (Children living in poorer SPs more likely to attend a public school)	Pr=0.000 (Children living in poorer SPs more likely to attend a public school)	n.s.
Quintile	Pr=0.0001 (Children living in poorer SPs likely to attend lower quintile schools)	Pr=0.0001 (Children living in poorer SPs likely to attend lower quintile schools)	Pr=0.0001 (Children living in poorer SPs likely to attend lower quintile schools)	Pr=0.0001 (Children living in poorer SPs likely to attend lower quintile schools)
Section 21	Pr=0.003 (Children living mid-range SPs least likely to attend Section 21 schools)	Pr=0.000 (Children living mid-range SPs least likely to attend Section 21 schools)	Pr=0.007 (Children living mid-range SPs least likely to attend Section 21 schools)	Pr=0.000 (Children living in richer SPs least likely to attend Section 21 schools)
School size	Pr=0.0001 (Children living in richer SPs likely to attend larger schools)	Pr=0.0001 (Children living in richer SPs likely to attend larger schools)	Pr=0.0001 (Children living in richer SPs likely to attend larger schools)	Pr=0.0068 (Children living in richer SPs likely to attend larger schools)
Proportion black	Pr=0.0001 (Children living in poorer SPs likely to attend schools with a higher proportion of	Pr=0.0001 (Children living in poorer SPs likely to attend schools with a higher proportion of	Pr=0.0001 (Children living in poorer SPs likely to attend schools with a higher proportion of	Pr=0.0001 (Children living in poorer SPs likely to attend schools with a higher proportion of

	black learners)	black learners)	black learners)	black learners)
School fees	Pr=0.0001 (Children living in poorer SPs likely to attend schools with lower fees)	Pr=0.0001 (Children living in poorer SPs likely to attend schools with lower fees)	Pr=0.0001 (Children living in poorer SPs likely to attend schools with lower fees)	Pr=0.0001 (Children living in poorer SPs likely to attend schools with lower fees)
Historical DET status	Pr=0.000 (Children living in poorer SPs more likely to attend former DET schools)	Pr=0.000 (Children living in poorer SPs more likely to attend former DET schools)	Pr=0.000 (Children living in poorer SPs more likely to attend former DET schools)	Pr=0.000 (Children living in poorer SPs more likely to attend former DET schools)
Matric Pass rate	Pr=0.0001 (Children living in mid-range SPs likely to attend schools with lower pass rates)	Pr=0.0001 (Children living in mid-range SPs likely to attend schools with lower pass rates)	Pr=0.0001 (Children living in mid-range SPs likely to attend schools with lower pass rates)	Pr=0.0001 (Children living in mid-range SPs likely to attend schools with lower pass rates)

Table 7.20: Relationship between SP poverty and school properties

Main Place level

As evident in Table 7.21, the relationships between residential MP poverty levels and school attributes remains fairly consistent with those identified at the SAL and SP levels, although the statistical significance of some of these relationships has decreased. Overall, children living in wealthier MPs are more likely to attend more advantaged schools than their peers in poorer MPs. Once again, however, children living in the mid-range MPs likely to attend the schools with the poorest pass rates – although this relationship disappears when only those children who have reached secondary school by 2003 are examined. The reasons for the persistence of this pattern remain unclear.

MP Poverty	1997	2003 (all data)	2003 (primary schools only)	2003 (secondary schools only)
Sector	Pr=0.061 (Children living in wealthier MPs less likely to attend a public school)	Pr=0.062 (Children living in wealthier MPs less likely to attend a public school)	Pr=0.010 (Children living in wealthier MPs less likely to attend a public school)	n.s.
Quintile	Pr=0.0001 (Children living in poorer MPs likely to attend lower quintile schools)	Pr=0.0001 (Children living in poorer MPs likely to attend lower quintile schools)	Pr=0.0001 (Children living in poorer MPs likely to attend lower quintile schools)	Pr=0.0001 (Children living in poorer MPs likely to attend lower quintile schools)
Section 21	Pr=0.002 (Children living in poorer MPs less likely to attend Section 21 schools)	n.s.	Pr=0.001 (Children living in poorer MPs less likely to attend Section 21 schools)	Pr=0.022 (Children living in poorer MPs less likely to attend Section 21 schools)
School size	Pr=0.0001 (Children living in wealthier MPs likely to attend larger schools)	Pr=0.0001 (Children living in wealthier MPs likely to attend larger schools)	Pr=0.0002 (Children living in wealthier MPs likely to attend larger schools)	n.s.
Proportion black	Pr=0.0001 (Children living in wealthier MPs likely to attend schools with a lower proportion of black learners)	Pr=0.0001 (Children living in wealthier MPs likely to attend schools with a lower proportion of black learners)	Pr=0.0001 (Children living in wealthier MPs likely to attend schools with a lower proportion of black learners)	Pr=0.0001 (Children living in wealthier MPs likely to attend schools with a lower proportion of black learners)
School fees	Pr=0.0001 (Children living in wealthier MPs likely to attend schools with higher fees)	Pr=0.0001 (Children living in wealthier MPs likely to attend schools with higher fees)	Pr=0.0001 (Children living in wealthier MPs likely to attend schools with higher fees)	Pr=0.0001 (Children living in wealthier MPs likely to attend schools with higher fees)
Historical DET status	Pr=0.000 (Children living in wealthier SPs less likely to attend	Pr=0.000 (Children living in wealthier SPs less likely to attend	Pr=0.000 (Children living in wealthier SPs less likely to attend	Pr=0.000 (Children living in wealthier SPs less likely to attend

	former DET schools)	former DET schools)	former DET schools)	former DET schools)
Matric Pass rate	Pr=0.0001 (Children living in mid-range MPs likely to attend schools with lower pass rates)	Pr=0.0028 (Children living in mid-range MPs likely to attend schools with lower pass rates)	Pr=0.0008 (Children living in mid-range MPs likely to attend schools with lower pass rates)	n.s.

Table 7.21: Relationship between MP poverty and school properties

Community level variables associated with school enrolment patterns: Discussion

At the community level, the relationships between residential SAL poverty and school attributes were clearest, with children living in poorer SALs more likely to attend less advantaged schools. At the SP and MP levels, this pattern held overall, although strange results emerged with respect to the school's Section 21 status and matric pass rates. Explanations for these unexpected results are not clear.

7.3.4 Child, household and community level variables associated with school enrolment patterns: Discussion

The analyses presented in this section present clear evidence for a relationship between child, household and community attributes associated with advantage, and school-level variables associated with more highly performing schools. Socio-economic status and maternal education appear to be particularly strongly related to attending higher quality schools, while living in high poverty areas is, predictably, inversely associated with attending high quality schools. At the level of the individual child, there is no evidence for a relationship between gender and school quality or other school properties, once schooling phase has been controlled for. However, children who start school late for their age are much more highly represented among those attending more advantaged and highly performing schools. The analyses around Section 21 status given unexpected and often contradictory results, suggesting that there may be reason for concern around the accuracy of this variable. Overall, these results highlight the extent to which access to high quality education in contemporary urban South Africa is determined by a child's home circumstances, and the area in which he or she lives.

7.4 Relationships between school attributes and mobility behaviours

The final section of this chapter tests for and explores relationships between various school attributes, and the mobility of learners enrolled in those schools. Children attending higher quality schools are expected to be more strongly engaged in mobility.

7.4.1 School sector

Table 7.22, below, illustrates the relationships between the sector (public or independent) of the school a child attends, and his or her mobility behaviour. There is strong evidence that children attending independent schools are significantly more mobile than those attending public schools. This holds for all definitions of mobility.

School sector	1997	2003 (all data)	2003 (primary schools only)	2003 (secondary schools only)
Distance travelled	Pr=0.000 (Children at independent schools travel significantly further to school than children at public schools)	Pr=0.000 (Children at independent schools travel significantly further to school than children at public schools)	Pr=0.000 (Children at independent schools travel significantly further to school than children at public schools)	Pr=0.000 (Children at independent schools travel significantly further to school than children at public schools)
Movement between areas	SAL: Pr=0.052 (Children at public schools	SAL: n.s.	SAL: n.s.	SAL: n.s.

	are more likely to attend school in their home SAL)			
	SP: Pr=0.000 (Children at public schools are more likely to attend school in their home SP)	SP: Pr=0.000 (Children at public schools are more likely to attend school in their home SP)	SP: Pr=0.000 (Children at public schools are more likely to attend school in their home SP)	SP: Pr=0.001 (Children at public schools are more likely to attend school in their home SP)
	MP: Pr=0.000 (Children at public schools are more likely to attend school in their home MP)	MP: Pr=0.000 (Children t public schools are more likely to attend school in their home MP)	MP: Pr=0.000 (Children at public schools are more likely to attend school in their home MP)	MP: Pr=0.004 (Children at public schools are more likely to attend school in their home MP)
Nearest school	Pr=0.000 (Children at public schools are more likely to attend their nearest school)	Pr=0.000 (Children at public schools are more likely to attend their nearest school)	Pr=0.000 (Children at public schools are more likely to attend their nearest school)	Pr=0.000 (Children at public schools are more likely to attend their nearest school)

Table 7.22: Relationship between sector of attended school and learner mobility

7.4.2 School quintile

As indicated by the data in Table 7.23, children attending high quintile schools are more likely to engage in learner mobility than children attending lower quintile schools. However, in some tests, particularly relating to movement between areas, there is evidence that the children attending the very lowest quintile schools are as likely to engage in mobility as children at the attending the highest quintile schools. This has two potential explanations: firstly, it may relate to poor quality data on those children attending the most disadvantaged schools, or secondly, it may be that the children attending the most disadvantaged schools are living in areas with particularly few educational opportunities, requiring them to travel further. The second hypothesis is supported by the fact that higher mobility amongst children attending most disadvantaged schools largely disappears at higher levels of geography.

School quintile	1997	2003 (all data)	2003 (primary schools only)	2003 (secondary schools only)
Distance travelled	Pr=0.0001 (Children at higher quintile schools are likely to travel further than children at lower quintile schools)	Pr=0.0001 (Children at higher quintile schools are likely to travel further than children at lower quintile schools)	Pr=0.0001 (Children at higher quintile schools are likely to travel further than children at lower quintile schools)	Pr=0.0001 (Children at higher quintile schools are likely to travel further than children at lower quintile schools)
Movement between areas	SAL: Pr=0.042 (Children at mid-quintile schools are more likely to attend a school in their home SAL)	SAL: n.s.	SAL: n.s.	SAL: n.s.
	SP: Pr=0.000 (Children at mid-quintile schools are more likely to attend a school in their home SP)	SP: Pr=0.000 (Children at mid-quintile schools are more likely to attend a school in their home SP)	SP: Pr=0.000 (Children at mid-quintile schools are more likely to attend a school in their home SP)	SP: Pr=0.000 (Children at mid-quintile schools are more likely to attend a school in their home SP)
	MP: Pr=0.000 (Children at low and mid-quintile schools are more likely to attend a school in their home MP)	MP: Pr=0.000 (Children at low and mid-quintile schools are more likely to attend a school in their home MP)	MP: Pr=0.000 (Children at low and mid-quintile schools are more likely to attend a school in their home MP)	MP: Pr=0.000 (Children at low and mid-quintile schools are more likely to attend a school in their home MP)
Nearest school	Pr=0.000 (Children at low and mid-quintile schools more likely to attend	Pr=0.000 (Children at low and mid-quintile schools more likely to attend	Pr=0.000 (Children at low and mid-quintile schools more likely to attend	Pr=0.000 (Children at low and mid-quintile schools more likely to attend

	their nearest school)	their nearest school)	their nearest school)	their nearest school)

Table 7.23: Relationship between quintile of attended school and learner mobility

7.4.3 Section 21 status

The evidence for a relationship between the Section 21 status of a child's school and his or her mobility is somewhat weaker, as evident in Table 7.24 below. While there is evidence of a significant relationship between Section 21 status and mobility for all definitions of mobility for those children attending secondary school in 2003, the evidence for primary school children is less clear. There does appear to be a weakly significant relationship between Section 21 status and distance travelled, but there is no evidence that Section 21 status is related to movement between areas or to whether a child attends his or her nearest school. In all cases where there is evidence for a relationship between Section 21 status and mobility, attending a school with Section 21 status is associated with increased mobility.

School Section 21 status	**1997**	**2003 (all data)**	**2003 (primary schools only)**	**2003 (secondary schools only)**
Distance travelled	Pr=0.0214 (Children at schools without Section 21 status travel less far to school than children at Section 21 schools)	Pr=0.0540 (Children at schools without Section 21 status travel less far to school than children at Section 21 schools)	n.s.	Pr=0.0716 (Children at schools without Section 21 status travel less far to school than children at Section 21 schools)
Movement between areas	SAL: n.s.	SAL: n.s.	SAL: n.s.	SAL: n.s.
	SP: n.s.	SP: n.s.	SP: n.s.	SP: Pr=0.001 (Children at schools without Section 21 status more

				likely to attend school in their home SP)
	MP: n.s.	MP: n.s.	MP: n.s.	MP: Pr=0.014 (Children at schools without Section 21 status more likely to attend school in their home MP)
Nearest school	n.s.	n.s.	n.s.	Pr=0.038 (Children at Section 21 schools less likely to attend their nearest school)

Table 7.24: Relationship between Section 21 status of attended school and learner mobility

7.4.4 School Enrolment

The direction of the relationship between a child's mobility and the size of his or her school is heavily moderated by schooling phase (see Table 7.25 below). At the secondary school level, as distance increases, the size of the school attended tends to fall. By contrast, at the primary level, as distance increases, school size also tends to rise. This is likely to relate to the tendency, described in Chapter 4, for more advantaged areas to have primary and secondary schools that are roughly equivalent in size, while less advantaged areas tend to have a few particularly large secondary schools, and a large number of far smaller primary schools.

	1997	2003 (all data)	2003 (primary schools only)	2003 (secondary schools only)
Distance travelled	Pr=0.0000 (As children travel further to	Pr=0.0010 (As children travel further to	Pr=0.0000 (As children travel further to	Pr=0.0005 (As children travel further to

	school, school size increases)	school, school size increases)	school, school size increases)	school, school size falls)
Movement between areas	SAL: n.s.	SAL: n.s.	SAL: n.s.	SAL: n.s.
	SP: Pr=0.0000 (Children at a school outside of their home SP attend larger schools)	SP: n.s.	SP: Pr=0.0008 (Children at a school outside of their home SP attend larger schools)	SP: Pr=0.0024 (Children at a school outside of their home SP attend larger schools)
	MP: Pr=0.0000 (Children at a school outside of their home MP attend larger schools)	MP: Pr=0.0226 (Children at a school outside of their home MP attend larger schools)	MP: Pr=0.0000 (Children at a school outside of their home MP attend larger schools)	MP: Pr=0.0008 (Children at a school outside of their home MP attend smaller schools)
Nearest school	n.s.	Pr=0.0000 (Children at their nearest school are likely to be attending a larger school)	n.s.	Pr=0.0005 (Children at their nearest school are likely to be attending a larger school)

Table 7.25: Relationship between size of attended school and learner mobility

7.4.5 Proportion black learners

Predictably, given that the majority of sample members live in predominantly black areas, those attending school close to their homes are likely to attend schools with a higher proportion of black learners than those who are travelling to schools further afield (see Table 7.26 below). There are, however, two exceptions to this pattern.

Firstly, when looking at primary school children who attend school within the same SAL as their home, they are likely to attend a school with a lower proportion of black learners than those children travelling outside of the SAL in which their home is. This is likely to be due to the extremely small numbers of children attending school in their home SAL, and the earlier finding children living in more affluent areas were more likely to attend school in their home SAL.

Secondly, there is no evidence for a relationship between whether or not children attend their nearest school, and the proportion of black learners in that school. This absence of a relationship is likely to be because quite a few of those children not attending their nearest school are still attending a school in the same, relatively disadvantaged area in which they live.

	1997	**2003 (all data)**	**2003 (primary schools only)**	**2003 (secondary schools only)**
Distance travelled	Pr=0.0000 (Children at schools with a lower proportion of black learners travel further to school)	Pr=0.0000 (Children at schools with a lower proportion of black learners travel further to school)	Pr=0.0000 (Children at schools with a lower proportion of black learners travel further to school)	Pr=0.0000 (Children at schools with a lower proportion of black learners travel further to school)
Movement between areas	SAL: Pr=0.0433 (Children at schools with a lower proportion of black children more likely to attend a school in their home SAL.	SAL: n.s.	SAL: Pr=0.0133 (Children at schools with a lower proportion of black children more likely to attend a school in their home SAL.	SAL: n.s.
	Pr=0.0000 (Children at schools with a higher proportion of black learners are more likely to attend school in their home SP)	Pr=0.0000 (Children at schools with a higher proportion of black learners are more likely to attend school in their home SP)	Pr=0.0000 (Children at schools with a higher proportion of black learners are more likely to attend school in their home SP)	Pr=0.0031 (Children at schools with a higher proportion of black learners are more likely to attend school in their home SP)
	Pr=0.0000 (Children at schools with a	Pr=0.0000 (Children at schools with a	Pr=0.0000 (Children at schools with a	Pr=0.0000 (Children at schools with a

	higher proportion of black learners are more likely to attend school in their home MP)	higher proportion of black learners are more likely to attend school in their home MP)	higher proportion of black learners are more likely to attend school in their home MP)	higher proportion of black learners are more likely to attend school in their home MP)
Nearest school	n.s.	n.s.	n.s.	n.s.

Table 7.26: Relationship between proportion of black students at attended school and learner mobility

7.4.6 School fees

Table 7.27 below provides evidence for a significant positive relationship between the fees charged by a child's school, and the mobility of that child. One finding, that children attending school inside their home SAL are likely to attend schools with lower fees, slightly inconsistent with the finding that these same children are likely to attend a school with a lower proportion of black learners. One plausible explanation is the nature of the distribution of the values of school fees, with only a very few very high values present.

School fees	**1997**	**2003 (all data)**	**2003 (primary schools only)**	**2003 (secondary schools only)**
Distance travelled	Pr=0.0000 (Children at schools with higher fees travel further)	Pr=0.0000 (Children at schools with higher fees travel further)	Pr=0.0000 (Children at schools with higher fees travel further)	Pr=0.0000 (Children at schools with higher fees travel further)
Movement between areas	SAL: Pr=0.0917 (Children at schools with lower fees are more likely to attend school in their home SAL)	SAL: n.s.	SAL: n.s.	SAL: n.s.
	SP: Pr=0.0000 (Children at schools with	SP: Pr=0.0000 (Children at schools with	SP: Pr=0.0000 (Children at schools with	SP: Pr=0.0000 (Children at schools with

	lower fees are more likely to attend school in their home SP)	lower fees are more likely to attend school in their home SP)	lower fees are more likely to attend school in their home SP)	lower fees are more likely to attend school in their home SP)
	MP: Pr=0.0000 (Children at schools with lower fees are more likely to attend school in their home MP)	MP: Pr=0.0000 (Children at schools with lower fees are more likely to attend school in their home MP)	MP: Pr=0.0000 (Children at schools with lower fees are more likely to attend school in their home MP)	MP: Pr=0.0000 (Children at schools with lower fees are more likely to attend school in their home MP)
Nearest school	Pr=0.0192 (Children at schools with lower fees are more likely to attend their nearest school)	Pr=0.0002 (Children at schools with lower fees are more likely to attend their nearest school)	Pr=0.0004 (Children at schools with lower fees are more likely to attend their nearest school)	Pr=0.0130 (Children at schools with lower fees are more likely to attend their nearest school)

Table 7.27: Relationship between fees of attended school and learner mobility

7.4.7 Historical DET

Table 7.28, below, presents evidence for a significant negative relationship between the historical DET status of the school a child attends, and that child's mobility. However, this relationship does not hold when mobility is defined at the SAL level, or by attendance at the nearest school. As discussed with reference to the findings on the proportion of black students enrolled at a school, this is likely to relate to the fact that many of the children engaging in these two forms of mobility are still attending schools that are no less disadvantaged than the school nearest to their home, or within their home SAL.

	1997	**2003 (all data)**	**2003 (primary schools only)**	**2003 (secondary schools only)**
Distance travelled	Pr=0.0000 (Children at former DET schools travel	Pr=0.0000 (Children at former DET schools travel	Pr=0.0000 (Children at former DET schools travel	Pr=0.0000 (Children at former DET schools travel

	less far)	less far)	less far)	less far)
Movement between areas	SAL: n.s.	SAL: n.s.	SAL: n.s.	SAL: n.s.
	SP: Pr=0.000 (Children at former DET schools are more likely to attend a school in their home SP)	SP: Pr=0.000 (Children at former DET schools are more likely to attend a school in their home SP)	SP: Pr=0.000 (Children at former DET schools are more likely to attend a school in their home SP)	SP: Pr=0.007 (Children at former DET schools are more likely to attend a school in their home SP)
	MP: Pr=0.000 (Children at former DET schools are more likely to attend a school in their home MP)	MP: Pr=0.000 (Children at former DET schools are more likely to attend a school in their home MP)	MP: Pr=0.000 (Children at former DET schools are more likely to attend a school in their home MP)	MP: Pr=0.000 (Children at former DET schools are more likely to attend a school in their home MP)
Nearest school	n.s.	n.s.	n.s.	n.s.

Table 7.28: Relationship between historical DET status of attended school and learner mobility

7.4.8 Matric pass rate

There is strong evidence for a positive relationship between mobility and school performance, as documented in Table 7.29 below.

	1997	**2003 (all data)**	**2003 (primary schools only)**	**2003 (secondary schools only)**
Distance travelled	Pr=0.0000 (Children at schools with higher pass rates travel further)	Pr=0.0000 (Children at schools with higher pass rates travel further)	Pr=0.0000 (Children at schools with higher pass rates travel further)	Pr=0.0000 (Children at schools with higher pass rates travel further)
Movement between areas	SAL: Pr=0.0033 (Children at schools with higher pass rates are less likely to attend a school in	SAL: Pr=0.0216 (Children at schools with higher pass rates are less likely to attend a school in	SAL: n.s.	SAL: Pr=0.0046 (Children at schools with higher pass rates are less likely to attend a school in

	their home SAL)	their home SAL)		their home SAL)
	SP: Pr=0.0000 (Children at schools with higher pass rates are less likely to attend a school in their home SP)	SP: Pr=0.0000 (Children at schools with higher pass rates are less likely to attend a school in their home SP)	SP: Pr=0.0000 (Children at schools with higher pass rates are less likely to attend a school in their home SP)	SP: Pr=0.0000 (Children at schools with higher pass rates are less likely to attend a school in their home SP)
	MP: Pr=0.0010 (Children at schools with higher pass rates are less likely to attend a school in their home MP)	MP: Pr=0.0000 (Children at schools with higher pass rates are less likely to attend a school in their home MP)	MP: Pr=0.0048 (Children at schools with higher pass rates are less likely to attend a school in their home MP)	MP: Pr=0.0000 (Children at schools with higher pass rates are less likely to attend a school in their home MP)
Nearest school	Pr=0.0047 (Children at schools with higher pass rates are less likely to be attending their nearest school)	Pr=0.0004 (Children at schools with higher pass rates are less likely to be attending their nearest school)	Pr=0.0811 (Children at schools with higher pass rates are less likely to be attending their nearest school)	Pr=0.0002 (Children at schools with higher pass rates are less likely to be attending their nearest school)

Table 7.29: Relationship between quintile of attended school and learner mobility

7.4.9 Relationships between school attributes and mobility behaviour: Discussion

Overall, the results presented in this section provide support for the hypothesis that learner mobility is related to the pursuit of higher quality education. Using a number of different school variables associated with educational quality, and a number of different approaches to the measurement of learner mobility, the analyses described above found positive, largely consistent and highly significant relationships between engagement in learner mobility, and the quality of the school a child attends. The one definition of mobility which did produce inconsistent results was that based on whether or not a child attended school in the same SAL in which he or she lived. One possible explanation is simply that SAL mobility is not a good measure of learner mobility, as it does not differentiate sufficiently between children. This is supported by the very small numbers of children who do attend school in the SAL in which they live.

7.5 Conclusion

This chapter began by contrasting those schools closest to the homes of study sample members with the schools actually attended by study sample members. This provided clear evidence that, on average, the schools that children attend are somewhat more advantaged than the schools that are closest to their homes. This suggests that children and families are actively engaging in school choice and mobility to pursue higher quality educational opportunities than would otherwise be accessible to them. It does, however, also raise questions around who is attending the most poorly performing of schools. One possibility is that these are predominantly filled by in-migrants from other areas. Alternatively, these schools may simply be extremely under-enrolled.

Secondly, the chapter explored the distribution of children with different individual, household and community attributes across schools. This provided strong evidence that more advantaged children typically attend more advantaged schools. This suggests that even though parents and children are able to use learner mobility as a tool to access higher quality educational opportunities, even these enhanced educational opportunities remain strongly related to the affluence of a family. This ties in well with the hypothesis that two different forms of mobility exist in Johannesburg-Soweto, and that these require the investment of different levels of resources, but also providing access to different levels of schooling quality. The more resource intensive form of mobility requires substantial investment in travel, school fees and associated costs, but allows children to access historically advantaged schools. Given the resource requirements associated with this form of mobility, only children from relatively advantaged families are able to engage in it. The second form of mobility, which requires less in the way of resources and is therefore open to a wider group of children, is the one in which choices are made between a number of fairly local schools. While this may still mean that children attend a better school than the school closest to their home, it does not give them access to the most advantaged schools of all.

Finally, in testing the relationships between learner mobility and the attributes of the school a child attends, the chapter has also provided evidence that learner mobility is associated with enrolment at those schools expected to provide higher educational quality. This again suggests that children and families are using mobility, in at least two different forms, to gain access to higher quality educational opportunities than would otherwise be accessible to them.

Chapter 8: Changes in educational mobility over time

8.1 Introduction

This chapter documents changes in the educational mobility of children over time. It begins by looking at those sample members who have changed schools between 1997 and 2003. School change is a prerequisite to mobility change in this sample of children with constant residential addresses. Correlates of school change at the individual, household, community and school levels are documented. Tests are also conducted to determine whether children who change schools behave differently with regards to mobility than those who do not change schools. Secondly, the chapter explores the nature of the changes in mobility resulting from school change. These results are then disaggregated by whether the school change is a change between two primary schools, or a change from a primary to a high school. The implications of school change and changes in mobility over time are then discussed.

8.2 Changing schools

Given that residential addresses for the study sample are by definition constant over the period under examination, if children are enrolled in the same school in both 1997 and 2003, their mobility will also be constant. When children move between schools, however, their mobility will change. This means that school change provides a window onto mobility change.

Of the 1210 children for whom full schooling information is available in both 1997 and 2003, 373 (30.93%) were attending the same school at both points in time, whereas 833 (69.07%) were attending different schools (see Table 8.1 below). Of course, a number of these school changes relate to the 433 children who have moved from primary school to high school by 2003. When only children still in primary school by 2003 are examined, 357 (45.5%) are enrolled in the same school at both points in time, while 427 (54.39%) have changed schools. While this is a lower level of change than that found in the full study sample, it indicates that the majority of children who did not move between schooling phases still changed schools at least once between 1997 and 2003.

	Remained in same school from 1997-2003	**Changed schools between 1997 and 2003**
All children with full schooling data (n=1210)	373 (30.93%)	833 (69.07%)
Children in primary school in 2003 (n=781)	357 (45.71%)	424 (54.39%)
Children in high school in 2003 (n=415)	14 (3.37%)	401 (96.63%)

Table 8.1: Proportion of sample members remaining in the same school from 1997 to 2003, and proportion changing schools (note that information on schooling phase is only available for 1196 individuals)

8.2.1 Correlates of school change during primary schooling

Individual, household and community correlates

A range of tests was conducted to determine whether, amongst the group of children still enrolled in primary school in 2003, those who had changed schools at least once differed systematically from those who had never changed schools. A chi-square test indicates that black children were significantly more likely to change schools than coloured children ($\chi^2_{(1)}$= 17.9177, Pr=0.000). There was, however, no evidence for a relationship between school changing and gender, age at first enrolment, or grade repetition.

Maternal education was also significantly related to school change ($\chi^2_{(4)}$=13.6864, Pr=0.008), with children of mothers with some secondary schooling the most likely to change schools. There was no evidence for a relationship between school change and maternal marital status, or, more surprisingly, household SES in either 1997 or 2003, or change in household SES between 1997 and 2003.

There was, however, strong evidence for a relationship between school change and the poverty of the area in which a child lives, whether measured as the SAL ($\chi^2_{(4)}$=25.7324, Pr=0.000), SP ($\chi^2_{(4)}$=29.4967, Pr=0.000) or MP ($\chi^2_{(2)}$=18.0766, Pr=0.000) level. At the SAL level, children living in poorer areas more likely to change schools. At the SP and MP levels, by contrast, while children in poor areas are still more likely to change schools than children in affluent areas, it is the children living in mid-level areas who are most likely to change schools.

Overall, contrary to expectations, the data does not suggest that disadvantaged children are any more likely to change schools during primary schooling than their more advantaged peers. In fact, children with more poorly educated mothers appear to be somewhat protected against school change during this period.

School level correlates

With the exception of the few children enrolled in quintile 1 schools (the least advantaged schools), who are very unlikely to change schools during primary schooling, there is a negative relationship between school quintile and school change, with school change decreasing as quintile increases, in both 1997 ($\chi^2_{(4)}$=18.3413, Pr=0.001) and 2003 ($\chi^2_{(4)}$=8.5163, Pr=0.074). There is no evidence for a relationship between school change and Section 21 status in 1997, but a weakly significant relationship is found in 2003 ($\chi^2_{(1)}$=3.1192, Pr=0.077).

Wilcoxon rank-sum tests indicate that children attending smaller schools in 1997 (Pr=0.0000) or in 2003 (Pr=0.0043) are more likely to change schools.

Children attending schools with a higher proportion of black learners in 1997 (Pr=0.0000) or 2003 (Pr=0.0000) are also more likely to change schools. There is also a significant positive relationship between school change and the school fees in 1997 (Pr=0.0002) and 2003 (Pr=0.0000). Children attending former DET schools in either 1997 ($\chi^2_{(1)}$=42.5454, Pr=0.000) or 2003 ($\chi^2_{(1)}$=19.6727, Pr=0.000) were also significantly more likely to change schools. Finally, children attending schools with lower imputed pass rates in 1997 (Pr=0.0713) or 2003 (Pr=0.0185) were also more likely to change schools.

Interestingly, all of the school properties associated with increased levels of school change are also typically associated with schools located in township areas. It appears, therefore, that a large proportion of school change occurring during the primary school years is between primary schools located in the townships. As relationships in 2003 are typically weaker than in 1997, this suggests that to the extent that children are changing between schools in different areas, they are tending to leave township schools in favour of schools in other, presumably more advantaged areas.

Mobility related correlates

Mobility in 1997

Amongst children who remained in primary school throughout the study period, school-changers travelled average marginally further than non-changers in 1997, although this difference was not statistically significant. There was no evidence for a relationship between school change and whether he or she attended a school in the same SAL, SP or MP as his or her home in 1997. Finally, there was also no evidence for a relationship between school change and whether a child attended his or her nearest school in 1997, regardless of whether private schools were included in the analysis. Overall, therefore, data about a child’s mobility in 1997 is unlikely to serve as a useful predictor of whether he or she is likely to change primary schools between 1997 and 2003.

Mobility in 2003

A Wilcoxon rank-sum test indicated that there was no significant relationship between school change and distance from home to school in 2003. Children attending a school outside of their home SAL in 2003 were, however, significantly more likely to have changed schools ($\chi^2_{(1)}$=14.5194, Pr=0.000). This relationship is inverted at the SP level ($\chi^2_{(1)}$=3.714, Pr=0.054), with children attending a school in their home SP more likely to have changed schools. There is no evidence for a relationship between children attending schools in their home MP and school change. Finally, children who change schools are significantly less likely to be attending their nearest school in 2003, regardless of whether only public schools ($\chi^2_{(1)}$=8.0904, Pr=0.004), or both public and private schools ($\chi^2_{(1)}$=10.0275, Pr=0.002) are considered.

Overall, while a child's mobility in 2003 is a better predictor of school change than mobility in 1997, there is still not much evidence for a relationship between school change and mobility. The strongest finding is that children who attended their nearest school in 2003 were less likely to change schools, which is consistent with earlier findings that the children most likely to attend their nearest schools were those living in the affluent areas, and who would therefore be unlikely to have much incentive to change schools.

Discussion

Black children, and children with mothers with some secondary schooling appear to be the most likely to change school during the primary schooling period. Children living in areas with high or intermediate poverty levels are also more likely to change schools than their peers living in more affluent areas. While children changing schools seem to be more likely to attend township schools, there is little evidence to suggest that their patterns of mobility in either 1997 or 2003 differ from children who do not change schools. Overall, change between primary schools does not seem to be strongly related to mobility.

8.2.2 Correlates of school change associated with the transition to high school

The transition to high school typically requires that a child change schools. The correlates of school change associated with the transition to high school will therefore simply echo the correlates of high school status itself. Furthermore, as school change and transition to high school are not independent events for this group of children, but are different aspects of the same phenomenon, it does not make sense to try to determine the role of each, independently, in shaping mobility. Rather, it is appropriate to see both the transition to grade 8 and the school change that typically accompanies it as a single event. The data detailing the correlates of reaching high school by 2003, presented in Appendix C, Chapter 5 and Chapter 7, is briefly reviewed in the following sections.

Individual, household and community correlates

As documented in Appendix C, children who have reached high school by 2003 differ systematically from children still in primary school at this point. They are more likely to be coloured, to be girls, and to have started school early, and less likely to have ever repeated a grade. Their mothers are likely to be more highly educated, with an attainment of at least Grade 11, and they are more likely to have lived in comparatively advantaged homes in 1997. There is, however, no evidence that they differ from their peers who are still in primary school in 2003 with regards to the poverty levels of the areas in which they live.

School level correlates

As documented in Chapter 7, secondary schools attended by study sample members differ systematically from the primary schools attended. Secondary schools are more likely to be in quintile 4, and less likely to be in quintile 3 or quintile 5 than those attended at the primary level. They also tend to have a substantially larger number of learners enrolled, and a slightly lower proportion of black learners. Fees are higher, while the schools are slightly less likely to be former DET schools, and have slightly higher pass rates.

Mobility related correlates

Mobility in 1997

As detailed in Chapter 5, there is no evidence that group of children who have reached high school in 2003 differ from their peers with regards to the distance from home to school in 1997, or their likelihood of attending the school closest to their home in 1997. Those children who transition to high school by 2003 were, however, more likely to attend a school in the same SAL or SP as their home in 1997. At the MP and MN levels, however, no differences are evident.

Mobility in 2003

Children at the high school level in 2003 do, however, live significantly further from their schools than their peers still at primary school level do. However, there is no evidence that patterns of travel between different areas, as defined by census geography, change with schooling phase. Children at the high school level, however, are more likely to be attending their nearest school. These changes are thought to be related to the different density distributions of primary and high schools in Johannesburg-Soweto.

Discussion

As is clear from the data reviewed above, children who have reached high school by 2003 differ from their peers who have not, with regards to a number of variables associated with mobility, as well as differing with regards to certain aspects of mobility itself. School changes associated with the transition to high school seem to be more strongly related to changes in mobility than school changes associated with movement between two primary schools. However, this appears to be due primarily to the different distributions of primary and high schools in urban Gauteng province.

8.3 The nature of changes in mobility

8.3.1 Straight-line distance

Of the 1210 children with full schooling data for 1997 and 2003, 1177 also have full residential data, allowing the changes in their mobility over time to be calculated. These children are distributed across primary and high schools as illustrated in Table 8.2 below. Just over two thirds of them experience a change in distance from home to school between 1997 and 2003.

	No change in mobility between 1997 and 2003	**Change in mobility between 1997 and 2003**
All children with full schooling and residential data (n=1177)	n=368 (31.27%)	n=809 (68.73%)
Children in primary school in 2003 (n=763)	n=353 (46.26%)	n=410 (53.74%)
Children in high school in 2003 (n=404)	n=13 (3.21%)	n=391 (96.78%)

Table 8.2: Distribution of sample members with full residential and schooling data for both 1997 and 2003 across schooling phases, and stability of mobility behaviour

There is a small overall decrease in distance from home to school between 1997 and 2003 (see Table 8.3 below). This, however, obscures a very broad distribution of changes in distance, and the 31% of children who experience no change in distance at all. When only those children who have experienced a change in mobility are examined (n=807), the mean decrease in travel distance becomes larger, and the distribution as a whole becomes more spread out. The distribution of distance change for both the full sample and for school-changers only is illustrated in Figures 8.1 and 8.2 below. It is evident from these figures that although changes in distance range from extreme values at each end (both very large increases and decreases), the majority of the sample is concentrated around very moderate levels of change, even once all zero values have been removed.

	Mean change in distance from 1997 to 2003	Standard Deviation	25%	Median	75%
Full Sample (n=1177)	-0.167km	9.723km	-0.178km	0.000km	0.530km
School changers only (n=807)	-0.244km	11.743km	-0.754km	0.121km	1.111km
Children moving between primary and high school only (n=404)	0.543km	12.439km	-0.488km	0.137km	1.420km
Children moving between primary and high school and changing schools only (n=390)	0.560km	12.660km	-0.529km	0.211km	1.519km
All primary school children (n=763)	-0.568km	7.951km	0.000km	0.000km	0.181km
Children moving between different primary schools (n=409)	-1.060km	10.841km	-1.314km	0.070km	0.795km

Table 8.3: Changes in distance from home to school between 1997 for all sample members, disaggregated by schooling phase and school change status

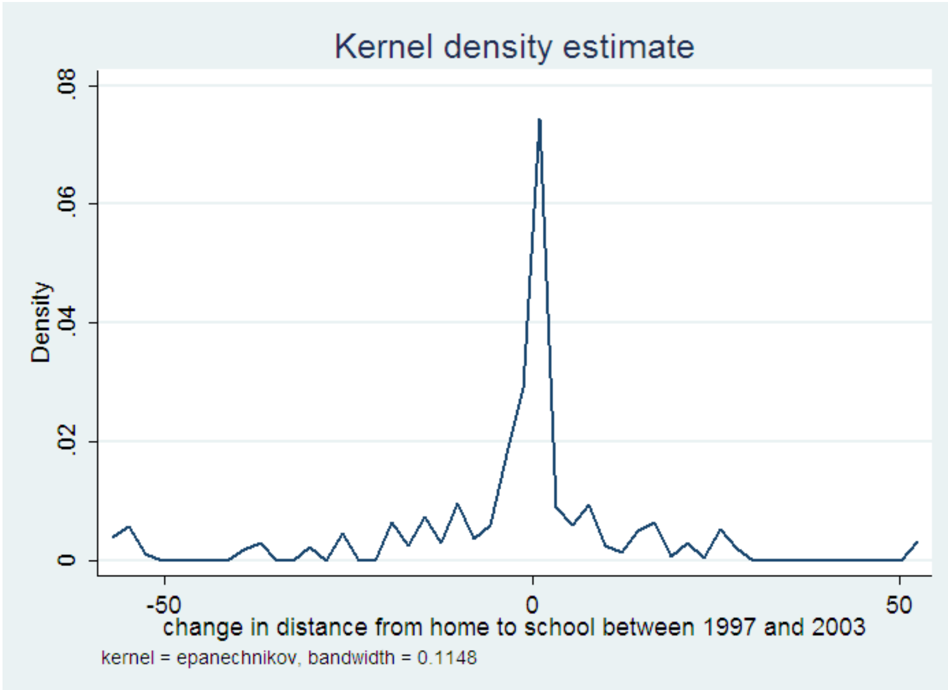

Figure 8.1: Kernel density plot of the change in distance from home to school between 1997 and 2003 for all sample members with full mobility information

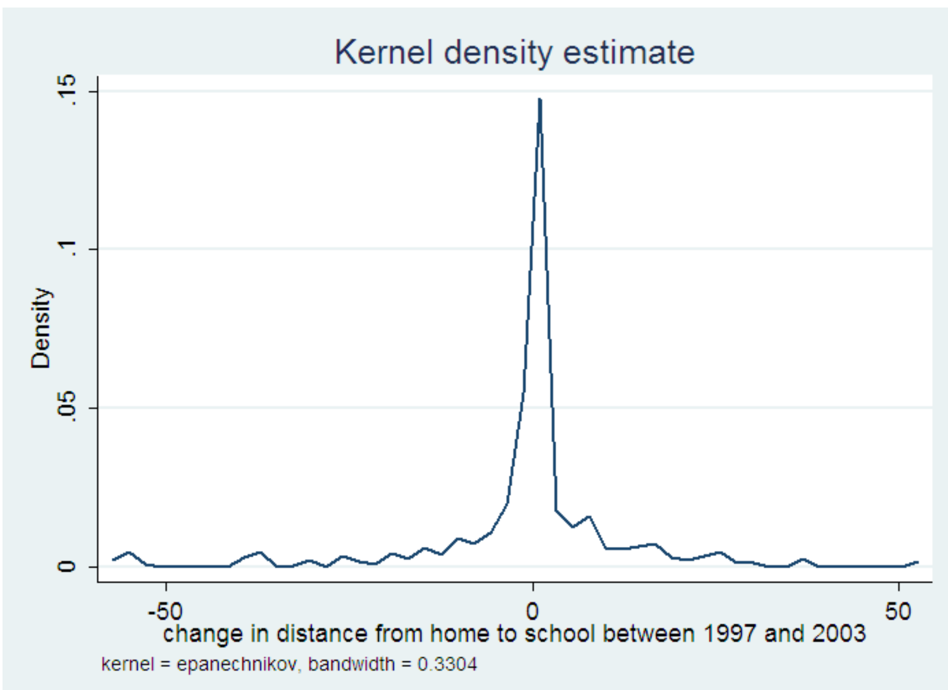

Figure 8.2: Kernel density plot of the change in distance from home to school between 1997 and 2003 for all sample members who changed schools

When change in distance is disaggregated by whether it is due to transition from primary to high school, or to movement between two primary schools, both distributions remain centred around zero, but otherwise become quite distinct (see Table 8.3 above). The mean change in distance for children moving between primary schools remains negative, but the mean change in distance for children moving to a high school becomes positive, at just over half a kilometre. A Wilcoxon rank-sum test indicates that children moving between primary schools, and those moving from primary to high schools, represent different populations in terms of the distribution of change in distance from home to school (Pr=0.0211). It is evident however, that both of these distributions, and particularly their means, are quite heavily influenced by particularly large values – as can be seen from the values of the 25th and 75th percentiles.

For this reason, it is also useful to examine counts of children travelling higher and lower distances in 2003 than in 1997 (Table 8.4 below). These counts indicate that in both the group of children moving between primary schools, and those transitioning to high school, roughly the same proportions are travelling further in 2003. Similarly, the proportions of children travelling less far in the two groups are also much the same. A chi-square test confirms that there is no evidence that the distribution of children travelling both further and less far differs for these two groups.

Given that the Wilcoxon test referred to earlier found a significant difference between these groups with regards to change in distance experienced, but the proportion of children increasing and decreasing their travel distance in both groups is the same, the difference between these two groups must be found in the extent of the increase or decrease in distance travelled. As evident in Figure 8.3 below, amongst children who travel further in 2003, the children moving between primary schools are far more concentrated around very small increases than the children moving from primary to high school. By contrast, when looking at those children travelling less far in 2003, this time it is the children moving from primary to high school that are far more concentrated around the very low decreases (see Figure 8.4 below). Wilcoxon rank-sum tests confirm that for both the group travelling further (Pr=0.0146), and those travelling less far (Pr=0.0373), the distributions of the change in distance are significantly different.

Overall, this evidence suggests that although both groups of children are approximately equally likely to travel either further or less far, the average decrease experienced by primary school children is larger. By contrast, the average increase in distance experience by high school children is larger.

	Mean change in distance from 1997 to 2003	**Number (%) travelling further in 2003**	**Number (%) travelling less far in 2003**
School changers only (n=807) (1 child changed between two schools located in the same place)	-0.244km	448 (55.51%)	358 (44.36%)
Children moving between different primary schools (n=409)	-1.060km	221 (54.03%)	187 (45.72%)
Children moving between primary and high school only (n=390)	0.560km	223 (56.92%)	168 (43.08%)

Table 8.4: Counts of sample members who live closer to or further from their school in 2003 as compared to 1997, disaggregated by phase of schooling

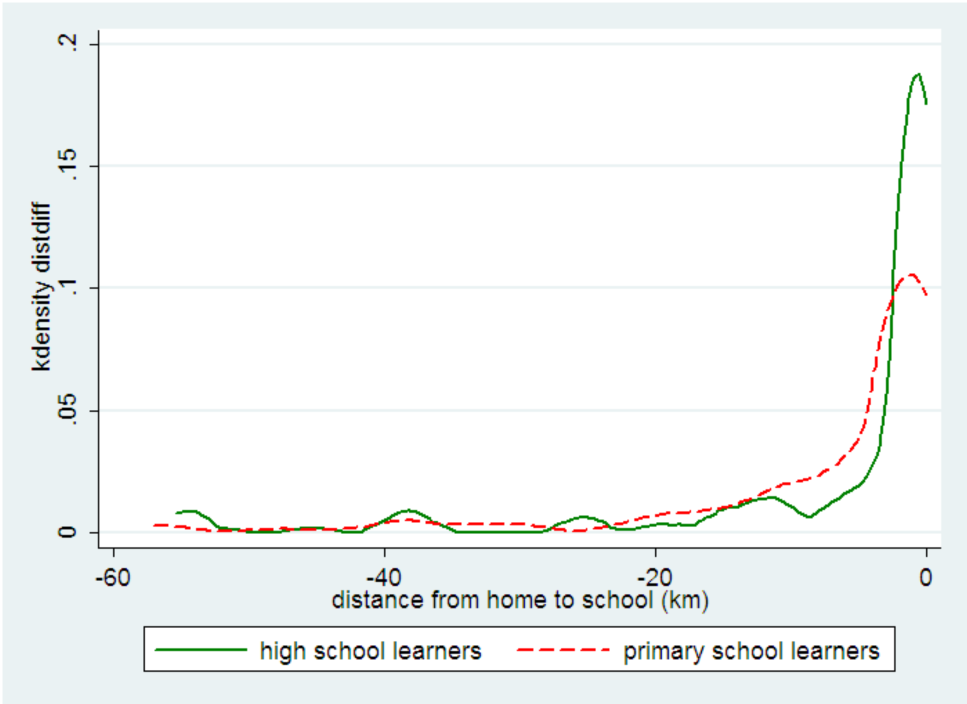

Figure 8.3: Distribution of change in distance from home to school for children travelling less far in 2003, by schooling phase

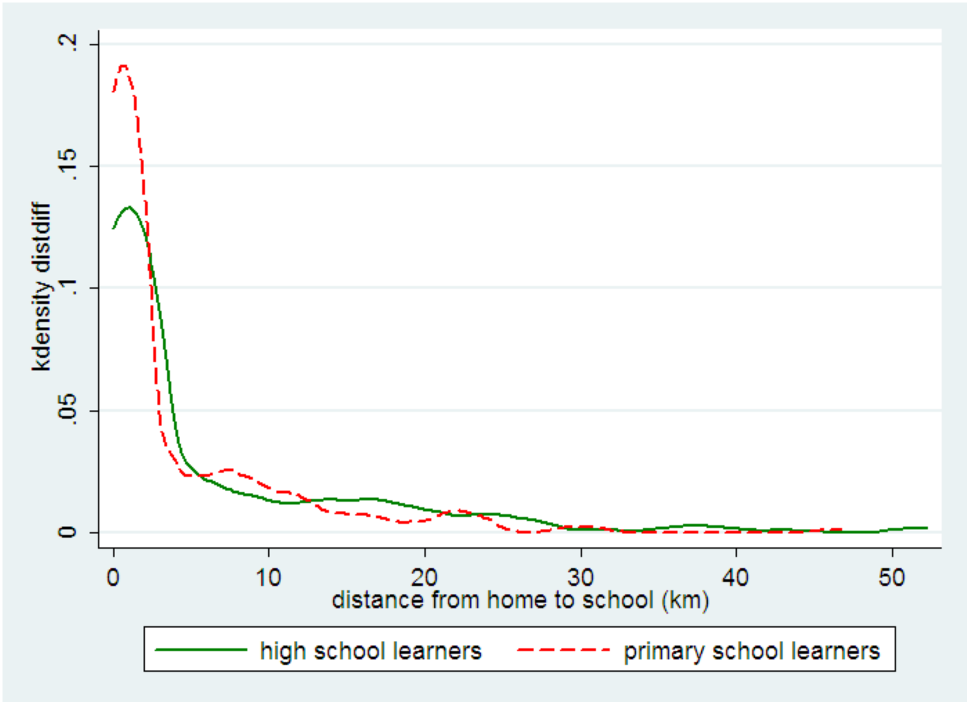

Figure 8.4: Distribution of change in distance from home to school for children travelling further in 2003, by schooling phase

Categories of distance

Another way to think about changes in mobility over time is to group children according to their distance between home and school at each point in time, and explore the numbers who remain in the same category, and those who change. Particularly salient categories are those travelling under 2.5km – that is, those who are able to walk to school; those travelling 2.5 to 5km – these can be thought of as children attending 'local schools'; and those travelling over 5km – that is, those attending non-local schools. Table 8.5, below, shows that the proportion of children found in the various distance categories is fairly consistent over time.

Group	**Number (%) travelling less than 2.5km**			**Number (%) travelling 2.5-5km**			**Number (%) travelling over 5km**		
Time point	1997	2003	Both	1997	2003	Both	1997	2003	Both
Full Sample 1997: n=1214 2003: n=1281 Both: n=1177	796 (65.57%)	823 (64.25%)	655 (55.65%)	83 (6.84%)	87 (6.79%)	40 (3.40%)	335 (27.59%)	371 (28.96%)	244 (20.73%)
School changers only 1997: n=810 2003: n=812 Both: n=807	539 (66.54%)	529 (65.15%)	424 (53.54%)	51 (6.30%)	50 (6.16%)	12 (1.49%)	220 (27.16%)	233 (28.69%)	133 (16.48%)
Children moving between two primary schools 1997: n=410 2003: n=413 both: n=409	263 (64.15%)	272 (65.86%)	207 (50.61%)	29 (7.07%)	25 (6.05%)	6 (1.47%)	118 (28.78%)	116 (28.09%)	68 (16.63%)
Children moving between primary and high school 1997: n=392 2003: n=391 both: n=390	269 (68.62%)	250 (63.94%)	210 (53.85%)	21 (5.36%)	25 (6.39%)	6 (1.54%)	102 (26.02%)	116 (29.67%)	65 (16.67%)

Table 8.5: Children attending schools in the same categories of distance from their homes in 1997 and 2003, disaggregated by schooling phase

However, there is less consistency over the actual individuals who are found in each category over time. Chi-square tests conducted on all children who changed schools ($\chi^2_{(4)}$= 185.9723, Pr=0.000), those moving between different primary schools ($\chi^2_{(4)}$=83.7371, Pr=0.000), and those moving between primary and high schools ($\chi^2_{(4)}$=103.0342, Pr = 0.000), all found highly significant differences in distribution across these mobility categories between the two points in time.

Amongst children who changed schools between 1997 and 2003, of the 644 children who were travelling less than 2.5km at any point in time, 424, or 65.84%, were travelling less than 2.5km at both points in time. Of the 89 children travelling between 2.5 and 5km at either point in time, 12, or 13.48%, were found in this distance category at both points in time. Finally, of the 320 children travelling over 5km at either point in time, 133, or 41.56%, fell into this category for both points in time. This data suggests that while the overall proportion of children found in the different distance categories is fairly constant over time, a fairly high proportion of individuals are actually moving between categories. There is no evidence of a significant difference in movement between distance categories for those children moving between primary schools, and those transitioning to high school. This is particularly interesting, as a proportion of children moving to high school are expected to be doing so because they have completed their primary schooling. By contrast, change during primary schooling is more likely to be related to dissatisfaction with current schooling arrangements, providing more of a reason for a child to change the distance he or she travels. Nonetheless, both groups seem to be equally likely to move between different categories.

Overall, just under 56% of the sample falls into the 0-2.5km at both points in time. This means that 44% of the children in the sample are attending a school over 2.5km away from their home during some part of their schooling. 20% of the sample attends schools over 5km away from their home at both points in time, meaning that a full fifth of the sample is consistently educated over 5kms away from their home. Clearly the assumption that children attend schools within walking distance of their homes does not hold for a fairly sizeable proportion of the current sample at various points in their schooling, and fairly consistently for almost 25% of the sample who never attend a school within a 2.5km radius of their homes.

8.3.2 Census geography

This section explores changes in mobility over time with regards to whether a child's school falls into the same area as his or her home. Table 8.6, below, illustrates the numbers of children experiencing mobility changes between 1997 and 2003, when this approach is used.

	No. (%) attending schools in different SALs in 1997 & 2003	**No. (%) attending schools in different SPs in 1997 & 2003**	**No. (%) attending schools in different MPs in 1997 & 2003**
Full Sample with school location data (n=1206)	782 (64.84%)	561 (46.52%)	244 (20.23%)
School changers only (n=831)	779 (93.74%)	560 (67.39%)	244 (29.36%)
Children moving between two primary schools (n=423)	398 (94.09%)	282 (66.67%)	128 (30.26%)
Children moving between primary and high school (n=400)	373 (93.25%)	275 (68.75%)	116 (29.00%)

Table 8.6: Numbers of children moving between schools in different geographical areas, disaggregated by schooling phase

What is particularly striking about Table 8.6 is that the proportion of children experiencing changes in the area in which they attend school again appears to be very similar for children moving between primary schools, and those transitioning to secondary schools. Chi-square tests find no significant difference in the distribution of mobility at the SAL, SP or MP level on the basis of whether a child is moving between primary schools or to a high school.

Exploring the changes in distances travelled for each of these groups of mobile children, however, reveals interesting results (see Table 8.7 below). Primary school children who are moving between schools in two different geographical areas appear to be doing so primarily by decreasing their distance travelled, while high school children appear to be doing so by increasing their distance travelled. This is particularly significant for the SP level. It may be that any relationship at the MP level is being partially obscured by the relatively small numbers moving between different MPs, as well as the typically larger distance that needs to be travelled to bring someone across an MP boundary.

		Mean change in distance	Median change in distance	No. (%) travelling further in 2003	χ^2 for children travelling further
Children changing from school in one SP to another	**High school (n=267)**	0.830km	0.584km	163 (61.05%)	$\chi^2_{(1)}$=7.2998 Pr = 0.007
	Primary schools (n=269)	-1.683km	-0.716km	133 (49.44%)	
Children changing from school in one MP to another	**High school (n=110)**	0.820km	5.329km	66 (60.00%)	$\chi^2_{(1)}$=3.1536 Pr = 0.076
	Primary schools (n=122)	-2.814km	-1.096km	59 (48.36%)	

Table 8.7: Changes in distance from home to school experienced by children moving between schools in different geographic areas, disaggregated by schooling phase (Note: Sample size is not the same as in Table 8.6 as only children for whom full schooling and residential data is available are included in Table 8.7)

The final question explored in this section is whether there is change over time with regards to the proportions of children attending school in the same area in which they live, and whether these are the same children in 1997 and 2003.

Table 8.8, below, shows the numbers of children attending schools in the same area in which they live. At the SAL level, very few children attend school in the same area as their home at either point in time, and the proportion of children who change schools but continue to attend school in the same SAL in which they live is extremely small. Given the small size of the SALs, this is unsurprising. At the SP level, a higher proportion of children, roughly a quarter of the sample, are found attending school in the same SP in which they live at both points of time, even when only those children who have changed schools are considered. Finally, at the MP level, over half the sample is attending school in the MP in which they live in both 1997 and 2003. Chi-square tests indicate that the differences in the distributions of sample members attending school in their home SP or MP is highly significant, regardless of the group considered. Results at the SAL level are less clear, which is probably related to the extremely small numbers of children attending school in their home SAL.

	Number (%) attending school in same SAL as home			**Number (%) attending school in same SP as home**			**Number (%) attending school in same MP as home**		
Time point	1997	2003	Both	1997	2003	Both	1997	2003	Both
Full Sample (n=1180)	80 (6.78%)	47 (3.98%)	33 (2.80 %)	479 (40.59%)	437 (37.03%)	335 (28.39%)	851 (72.12%)	829 (70.25%)	748 (63.39%)
School changers only (n=810)	56 (6.91%)	23 (2.84%)	9 (1.11 %)	350 (43.21%)	309 (38.15%)	207 (25.56%)	585 (72.22%)	563 (69.51%)	482 (59.51%)
Children moving between primary schools (n=411)	20 (4.87%)	5 (1.22%)	1 (0.24 %)	165 (40.15%)	168 (40.88%)	105 (25.55%)	286 (65.59%)	289 (70.32%)	239 (58.15%)

Children moving between primary and high school (n=391)	36 (9.21%)	18 (4.60%)	8 (2.05 %)	180 (46.04%)	137 (35.04%)	98 (25.06%)	292 (74.68%)	267 (68.29%)	236 (60.36%)

Table 8.8: Children attending school in the same area as their home, for three different levels of geography, disaggregated by schooling phase

In line with earlier findings, there is again little evidence for an impact of schooling phase on the proportions of children attending schools within the area in which their home is located. However, there is evidence that the members of these groups change quite substantially over time, particularly at smaller levels of geography. The finding that only about a quarter of the sample is educated within the SP in which they live at both points in time is particularly clear evidence for the limited proportion of children who are consistently educated at local schools. Similarly, less than two thirds of children attend schools within their home MP at both points in time.

8.3.3 Nearest school

A final approach to understanding changes in mobility is to look at the proportion of children attending their nearest school in 1997 and 2003. Table 8.9, below, shows the numbers of sample members attending their nearest public school in 1997, in 2003, and at both points in time. Although results are not presented below, the same tests were conducted including private schools, and in all instances produced extremely similar results. For all time points, as has been discussed previously, the numbers of children attending their nearest public school are extremely low, and the numbers who attended their nearest schools at both points in time are lower still.

	No. (%) attending nearest school in 1997	No. (%) attending nearest school in 2003	No. (%) attending nearest school in 1997 & 2003	Chi-squared test for similarity in 1997 and 2003
Full Sample (n=1178)	211 (17.91%)	214 (18.17%)	100 (8.49%)	$\chi^2_{(1)}$= 147.6961 Pr=0.000
School changers only (n=808)	136 (16.83%)	141 (17.45%)	28 (3.47%)	$\chi^2_{(1)}$= 1.1176 Pr=0.290
Children moving between primary schools (n=414)	67 (16.18%)	51 (12.32%)	0 (0.00%)	$\chi^2_{(1)}$=11.2308 Pr=0.001
Children moving between primary and high school (n=394)	69 (17.51%)	90 (22.84%)	28 (7.11%)	$\chi^2_{(1)}$= 14.9315 Pr=0.000

Table 8.9: Children attending their nearest grade-appropriate public school in 1997, 2003 and at both points, disaggregated by schooling phase

When only children moving between two primary schools are considered, there is a highly significant decrease in the proportion of children attending their nearest school between 1997 and 2003. By contrast, amongst children transitioning to secondary school, there is a significant increase in the proportion attending their nearest school.

There are several potential explanations for this pattern. Firstly, obviously, if a child changes schools, but not home location or phase of schooling, he or she cannot possibly attend the nearest school at both points in time. By contrast, if the child changes schooling phase, he or she is able to attend the nearest school at both points in time. This is one reason to expect a higher proportion of children attending their nearest school to be found amongst those transitioning to high school. Secondly, as mentioned previously, high schools tend to be larger, and more widely spaced, than primary schools. As a result, a child choosing between high schools has a smaller number of options available to him or her. The chances of attending the nearest school are therefore higher. The different sizes and distributions of primary and high schools also explains why the finding that children transitioning to high school are simultaneously more likely to experience an increase in travel distance and also increased likelihood of attending their nearest school, is not inconsistent.

Of the 354 children in the full sample who attended their nearest school at either point in time, 100, or 28.25%, attended their nearest school at both time points. Of the 250 school changers who attended their nearest school at either point in time, 28, or 11.20%, attended their nearest school at both points in time. Finally, of the 131 children transitioning to high school who attended their nearest school at either point in time, 28, or 21.37%, attended their nearest school at both points in time. These figures echo the findings presented earlier in this chapter that children who change school, whether at the primary or high school level, are less likely to be attending their nearest school in 2003 than children who do not change school.

8.4 Correlates of type of mobility change for primary school school-changers

The previous section has documented the differences between mobility changes related to the transition to high school, and those related to children moving between two primary schools. This section now focuses on those children changing between primary schools, and explores the correlates associated with different patterns of mobility change. Results presented below are based on the 409 sample members known to have changed primary schools between 1997 and 2003. Of course, given that we only have primary school change data for the non-random sub-group of the study sample who have not progressed to high school by 2003, it is not appropriate to assume that findings will be representative of all children who move between schools during their primary schooling. However, findings will provide a preliminary idea of correlates of changing mobility, which will assist with hypothesis and theory development.

8.4.1 Straight-line distance

Individual level variables

There was no evidence for any relationship between race or gender and the nature of mobility change associated with primary school change. Children with a later first school enrolment were more likely to decrease distance from home to school when changing primary schools (Wilcoxon rank-sum test; Pr=0.0089). Children who have never experienced a grade repetition were also more likely to decrease distance from home to school (Wilcoxon rank-sum test; Pr=0.0092).

Household level variables

Children with mothers with very little education (functionally illiterate), and children whose mothers have completed grade 10 or higher are significantly more likely to experience an increase in distance from home to school than children whose mothers have intermediate levels of education ($\chi^2_{(1)}$= 8.4919, Pr=0.075). Children whose mothers have completed grade 10 or higher also experience a mean increase in distance, compared to a mean decrease in distance for all other groups. There was no evidence for any difference in mobility change on the basis of maternal marital status, SES in either 1997 or 2003, or change in SES between these two years.

Community level variables

Children living in the wealthiest and poorest SAL areas were most likely to experience a substantial decrease in distance travelled (Wilcoxon rank-sum test; Pr=0.0577). By contrast, children in areas with moderate poverty levels were more likely to experience a small increase in travel. The same pattern, although with a lower significance level (Pr=0.0752), was found at the SP level. There was, however, no evidence for a relationship between MP area poverty and change in distance from home to school.

8.4.2 Census geography

Two ways of measuring change in mobility through the use of census data were presented earlier in this chapter. The first was to explore whether a child moved between schools in two different levels of geography between 1997 and 2003. The second was to explore whether a child moved either into, or out of, a school in the same geographic area in which he or she lived. In the first part of this section, children who moved between primary schools in different geographic areas will be compared to primary school changers who did not move between areas. In the second part of this section, children who moved into a school in the same area as their home will be compared to those children who moved out of a school in the same area as their home.

Children moving between schools in different areas

SAL

Of the 423 children still in primary school in 2003 who changed school between 1997 and 2003, only 25 moved between two schools in the same SAL, while the remaining 398 moved between schools in different SALs. There was no evidence that these two groups of children differed in any way with regards to race, gender, age at first enrolment in school, and whether or not they had ever experienced grade repetition. There was also no evidence for any difference with regards to any of the household level variables explored, namely maternal education, maternal marital status, household SES in 1997, household SES in 2003, or change in household SES between 1997 and 2003. Finally, although there is no evidence that the groups differ with respect to SAL poverty, there is evidence that they differ in both SP area poverty ($\chi^2_{(4)}$= 10.8582, Pr=0.021) and MP area poverty ($\chi^2_{(2)}$=5.2592, Pr=0.057). Those children who move between schools in different SALs tend to live in somewhat more affluent SP and MP areas.

SP

Of the 423 school-changers still in primary school in 2003, 282 (66.67%) moved between schools in different SPs, while 141 (33.33%) moved between two schools in the same SP. As was the case at the SAL level, there is no evidence that children moving between schools in the same SP differed significantly from those moving between schools in different SPs with regards to any of the individual level variables considered (race, gender, age at first enrolment, and grade repetition). Children moving between schools in two different SPs did, however, have more highly educated mothers ($\chi^2_{(4)}$=10.6102, Pr=0.031), although there was no significant difference with respect to maternal marital status. Those children who moved between schools in different SPs were likely to have a higher SES in both 1997 ($\chi^2_{(4)}$=20.2271, Pr=0.000) and 2003 ($\chi^2_{(4)}$=13.7227, Pr=0.008). There was no evidence that the two groups of children differed with respect to the poverty levels of the SALs or SPs in which they lived. Children moving between schools in different SP areas were, however, more likely to live in an extremely advantaged or disadvantaged MP ($\chi^2_{(2)}$=14.5783, Pr=0.001). In summary, movement between schools in different SPs was associated with higher maternal education and higher household SES, and with living in either a particularly advantaged, or disadvantaged MP.

MP

128 (30.26%) of the school changers still in primary school in 2003 moved between schools in two different MPs, while the remaining 295 (69.74%) moved between two schools in the same MP. There was no evidence that these two groups of children differed with respect to race, gender, age at first enrolment, grade repetition, maternal education, or maternal marital status. They did, however, differ significantly with respect to household SES in 1997 ($\chi^2_{(4)}$=13.2473, Pr=0.010) and in 2003 ($\chi^2_{(4)}$=22.0211, Pr=0.000), although not with respect to change in SES between these two points. In both 1997 and 2003, those children moving between schools in two different MPs were more

likely to be from more affluent homes. There was no evidence that the two groups of children differed with respect to the poverty area of the SALs in which they lived, but children moving between schools in different MPs were more likely to be living in particularly disadvantaged SP ($\chi^2_{(4)}$=10.3958, Pr=0.034) and MP ($\chi^2_{(2)}$=11.2848, Pr=0.004) areas.

Children moving into, and out of, schools in the same area as their home

SAL

At the SAL level, the number of children attending school in the same area as their home is extremely small (n=20 in 1997; n=5 in 2003). Due to this small sample size, no further analyses will be conducted on these groups.

SP

At the SP level, the numbers are somewhat higher, with 60 children attending primary school in the same SP as their home in 1997, but not 2003, and 63 doing so in 2003, but not 1997. There is no evidence that these two groups of children differ with respect to race or gender. The group of children attending local schools in 2003, but not 1997, is however significantly more likely to have first enrolled in school at an older age ($\chi^2_{(1)}$=4.2981, Pr=0.038). They are also less likely to have ever repeated a grade ($\chi^2_{(1)}$=7.9301, Pr=0.005). There is no evidence that the two groups of children differ with respect to maternal education, maternal marital status, or household SES in either 1997 or 2003. There is, however, evidence for a weakly significant relationship with the change in household SES between 1997 and 2003 ($\chi^2_{(6)}$=11.4982, Pr=0.063, with children whose SES has fallen between 1997 and 2003 being more likely to move to a school in the same SP as their home. A weakly significant relationship was also found with both SAL ($\chi^2_{(4)}$=8.3838, Pr=0.076) and SP ($\chi^2_{(4)}$=9.6659, Pr=0.046) poverty levels. In both cases, children in areas with intermediate poverty levels were more likely to be moving away from schools located in the same area as their home, while children in areas with either very high or low poverty levels were more likely to be moving into schools in the same area as their home. There was no evidence of any relationship between mobility change and the poverty level of the MP in which the child lived.

MP

At the MP level, of the 427 children still in primary school in 2003 who changed schools at least once, 47 had been attending a school in their MP in 1997, but by 2003 no longer did so. A further 50 children who had not been attending a school in their MP in 1997 were doing so by 2003. There was no evidence that these two groups of children differed with respect to either race or gender. Children moving into a school in the same MP as their home were, however, more likely to have started school later ($\chi^2_{(1)}$=9.9730, Pr=0.002), and never have repeated a grade ($\chi^2_{(1)}$=5.8091, Pr=0.016), than those children moving away from schools in the same MP as their home.

The two groups of children did not differ with respect to any of the household level variables considered (maternal education, maternal marital status, household SES in either 1997 or 2003, and the change in SES between 1997 and 2003). Children moving out of schools in their home SAL ($\chi^2_{(4)}$=8.1504, Pr=0.090) and SP ($\chi^2_{(4)}$=12.7700, Pr=0.012) areas were more likely to be living in areas with intermediate levels of poverty, while those children moving into schools in the same area as their home were more likely to be living in areas with either very high or very low poverty levels. There was, however, no evidence for any difference in MP poverty.

8.4.3 Nearest school

The final approach to measuring changes in mobility is to look at whether children move into, or out of, their nearest grade-appropriate school. Between 1997 and 2003, 67 primary school school-changers moved out of their nearest primary school into another primary school further from home, while 51 moved from a primary school further away into their nearest school. There was no evidence that these two groups of children differed systematically with regards to race, gender, age at first enrolment or grade repetition. Children moving out of their nearest school tending to have mothers with higher levels of education than those moving into their nearest schools ($\chi^2_{(4)}$=10.2479, Pr=0.034). There was no evidence of a relationship with maternal marital status, household SES in either 1997 or 2003, change in household SES between 1997 and 2003, or poverty at the SAL, SP or MP area levels.

8.4.4 Conclusion: primary school school-changers

Children who enrolled in school late for their age, and those who had never repeated a grade were more likely to experience a decrease in distance from home to school. Children living in the most advantaged and disadvantaged SAL and SP areas were also more likely to experience a decrease in distance from home to school. By contrast, children moving to schools further from home between 1997 and 2003 tended to have mothers with particularly low, or particularly high, levels of education.

Children moving between schools in different areas appeared to differ from their peers moving between two schools in the same area primarily with respect to maternal education, household SES, and the poverty level of the area in which they lived. Moving between schools in different areas was typically associated with relative advantage, although in some cases, there was a non-linear relationship in which both the most and least advantaged children were particularly likely to move between areas. Children moving into schools in the same area as their homes differed from those moving away from these schools in that they tended to have been older at their first enrolment in school, were less likely to have repeated grades, and lived in areas with either particularly low or high levels of poverty.

Children moving out of their nearest primary schools tended to have mothers with higher levels of education than children moving into their nearest primary schools, but otherwise the two groups did not differ significantly.

It is not clear that analyses presented here have generated any consistent findings. This may be due to a non-representative sample, and fairly small sample size in some of the analyses, or may simply indicate that mobility change is not significantly related to any of the variables considered here. However, extreme levels of SES, later first enrolment in school, and grade repetition, were the variables most prominent in these analyses, which suggests that they are likely to have stronger relationship to schooling change than the other variables considered. Unfortunately, the nature of this relationship is not clear at this point.

8.5 Correlates of type of mobility change for children transitioning to high school

In this section, the correlates of different types of mobility associated with the transition to high school are documented. Unless otherwise indicated, analyses presented here refer to the 390 children in the sample known to have moved from a primary to a high school between 1997 and 2003. Children enrolled at combined schools are excluded from these analyses. Again, the non-random nature of the sub-sample considered here must be emphasized, and findings should be used primarily as a basis for the development of hypotheses requiring further testing.

8.5.1 Straight-line distance

Individual level variables

A Kruskal-Wallis test found a significant relationship between change in mobility and ethnicity (Pr=0.0365), with black children experiencing greater increases in distance from home to school than coloured children. There was no evidence of any relationship between change in distance from home to school and child gender, age at first school enrolment, or grade repetition.

Household level variables

There was no evidence for a relationship between change in distance and maternal education, but children of unmarried mothers were more likely to experience larger increases in distance from home to school (Wilcoxon ranksum, Pr=0.0028). Children in the most advantaged and disadvantaged SES quintiles in 1997 were most likely to experience the largest fall in distance from home to school (Kruskal Wallis, Pr=0.0040). Those in the middle quintile in 2003, by contrast, were those experiencing the largest fall in distance (Kruskal Wallis, Pr=0.0247). There was also no evidence for any relationship between change in SES from 1997 to 2003 and change in distance from home to school.

Community level variables

There was no evidence for any relationship between area poverty, at the SAL, SP or MP level, and change in distance between home and school.

8.5.2 Census geography

Two sets of analyses relating to mobility change as defined by census geography are presented for those children transitioning to high school in 2003. The first set of analyses compares children who move between schools in the same area to those who move between schools in different areas. The sample size for these analyses is 400, as the home address data is not used. The second set of analyses compares those children who move away from a school in their home area, and those children who move into a school in their home area.

Children moving between schools in different areas

SAL

As was the case for children moving between primary schools, very few children moving between a primary and high school remained in the same SAL (n=27). Coloured children are more likely to remain in the same SAL than black children ($\chi^2_{(1)}$= 4.7763, Pr = 0.029). There is no evidence of a relationship between whether children move between schools in the same or different SALs and any of the other variables examined (gender, age at first enrolment, grade repetition, maternal education, maternal marital status, household SES in 1997 and 2003, change in household SES between 1997 and 2003, and poverty level of the SAL, SP and MP in which they child lives).

SP

125 (31.25%) of the children who moved from primary to high school moved between two schools in the same SP, while the remaining 275 children moved between two schools in different SPs. These proportions are again very similar to those found amongst the children moving between two primary schools. Coloured children were again more likely to move between two schools in the same SP than black children ($\chi^2_{(1)}$= 30.8927, Pr = 0.000). There is no evidence for a relationship between moving between schools in different SPs and child gender, age at first enrolment, grade repetition, maternal education, or maternal marital status. In both 1997 ($\chi^2_{(4)}$=12.3664, Pr=0.015) and 2003 ($\chi^2_{(4)}$=16.7466, Pr=0.002), children in the most extreme SES quintiles were most likely to be moving between schools in the same SP, while children in the middle quintiles were more likely to be moving between schools in different SPs. There was no evidence for a relationship with change in household SES over time. Finally, children living in poorer SAL ($\chi^2_{(4)}$=8.0566, Pr=0.090), SP ($\chi^2_{(4)}$=14.9720, Pr=0.005) and MP ($\chi^2_{(2)}$=9.9770, Pr=0.007) areas were more likely to be moving between schools in two different SPs.

MP

Of the children moving from primary to high school, 284 (71.00%) move between two schools within the same MP, while 116 (29.00%) move between schools in two different MPs. Again, these figures are very similar to those found for children moving between primary schools. Once again, coloured children are significantly more likely to move between two schools in the same MP than black children ($\chi^2_{(1)}$= 9.8932, Pr = 0.002). Again, there is also no evidence for a relationship with gender, age at first enrolment, grade repetition, maternal education, or maternal marital status. In both 1997 ($\chi^2_{(4)}$=11.8570, Pr=0.018) and 2003 ($\chi^2_{(4)}$=20.6648, Pr=0.000), the children from the poorest households are more likely to move between schools within the same MP, while children in the middle quintiles are more likely to move between schools in different MPs. There is no evidence for a relationship with change in household SES over time however. Children living in poorer SP ($\chi^2_{(4)}$=20.9806, Pr=0.000) and MP ($\chi^2_{(2)}$=11.7971, Pr=0.003) areas are more likely to move between schools in two different MPs than children in more affluent areas, but there is no evidence for a relationship with SAL poverty.

Children moving into, and out of, schools in the same area as their home

SAL

Altogether, 46 children transitioning to high school attend schools in the same SAL as their home in either 1997 or 2003. Of these, 8 attend school in the same SAL as their home at both points in time, while 28 do so only in 1997 and 10 only in 2003. Movement appears to generally be away from the home SAL. Given the small numbers involved, no additional analysis is conducted at this level.

SP

219 of the children transitioning to high school attend school in the same SP as their home at some point. 98 attend a school in their home SP in both 1997 and

2003, 82 do so only in 1997, and 39 only in 2003. The analyses presented here compare the group of 82 children attending only primary school in their home SP, and the group of 39 children attending only high school in their home SP. There is no significant difference between the two groups of children with respect to race, gender, age at first enrolment, or grade repetition. Children moving into the local SP for high school were more likely to have mothers with intermediate levels of education, while those moving out of the local SP were more likely to have mothers with particularly high or low levels of education ($\chi^2_{(4)}$=11.3383, Pr=0.014). There is no evidence that the two groups of children differ with respect to any of the other variables considered (maternal marital status, household SES in 1997 or 2003, change in household SES over time, and area poverty in the home SAL, SP and MP).

MP

Almost all of the children transitioning to high school, 325, attend a school in the same MP as their home at some point, and 236 are attending school in their home MP at both points in time. Only 56 attend a school in their home MP in 1997 but not 2003, while even fewer, 31, do so only in 2003. The analyses presented below compare the group of 56 children moving out of their home MP for high school with the 31 children moving into their home MP. The two groups of children do not differ with regards to gender, age at first enrolment, grade repetition, or maternal education. Children moving to a high school in their home MP are, however, substantially more likely to be coloured than black ($\chi^2_{(1)}$= 11.4483, Pr = 0.001), and are more likely to have married mothers ($\chi^2_{(1)}$=6.2554, Pr=0.012). Although the two groups of children do not differ with regards to household SES in 2003, or change in SES over time, children moving to a high school in their home MP were likely to come from either particularly advantaged or disadvantaged households in 1997 ($\chi^2_{(4)}$=10.9760, Pr=0.024). Children moving to high schools in their home MP were also more likely to live in particularly advantaged or disadvantaged SAL ($\chi^2_{(4)}$=14.1911, Pr=0.008), SP ($\chi^2_{(4)}$=20.4014, Pr=0.000), and MP ($\chi^2_{(2)}$= 14.8044, Pr=0.001) areas.

8.5.3 Nearest school

The final set of tests compares those children in high school in 2003, who had been attending their nearest school in 1997 and had moved to a school further away by 2003 (n=41), and those children who attended their nearest school in 2003, but had not done so in 1997 (n=62). Movement towards the nearest school appears to be somewhat more prevalent than movement away from the nearest school, which may be explained by the smaller number of high schools available. The two groups of children do not differ significantly with respect to gender, race, age at first school enrolment, maternal education, maternal marital status, or the poverty of the SAL, SP or MP areas in which they live. They do, however, differ with respect to household SES in both 1997 ($\chi^2_{(4)}$=7.8985, Pr=0.092) and 2003 ($\chi^2_{(4)}$=11.5980, Pr=0.024), with children from more affluent households more likely to be moving away from their nearest school when transitioning to high school.

8.5.4 Conclusion: mobility change associated with transition to high school

Somewhat surprisingly, many of the findings for children transitioning to high school echo closely those for children changing between two primary schools. This suggests that the two processes may not be as distinct as initially hypothesized. Change in distance from home to school was related to race, with black children typically moving further afield, as well as with household SES, with children at extreme levels of both advantage and disadvantage tending to move closer to home. There was also evidence that the children of unmarried mothers experienced a greater increase in distance than those with married mothers. Movement between different geographic areas was also associated in the same ways with race and household SES, although in this case an additional relationship with area poverty was also identified. Children living in poorer areas were more likely to move between schools in different geographic areas. Children moving away from their nearest school were more likely to be black and to have mothers who had either particularly high or low levels of education than their peers moving into schools in the same area as their homes. Finally, children moving away from their nearest schools for high school tended to be more affluent than those moving towards their nearest school.

8.6 Conclusion

This chapter has provided an overview of how schooling mobility changes over time, and how these changes relate to child and family variables. One of the most striking findings of this chapter is the prevalence of school changes during primary schooling. This high level of school changes – in a sample of children with consistent residential addresses– has important mobility implications. Most notably, it means that not only are many primary school children travelling considerable distances, but also that these distances are not constant with time, despite the widespread preconception that this is a relatively stable phase of schooling. Children living in poor areas, whose household SES increases with time, and whose mothers have intermediate levels of education appear to be the most likely to move between different primary schools.

A second important outcome of this chapter is that there are both similarities and differences in the changes to mobility associated with the transition to high school and those associated with a transition between two different primary schools. The overall tendency appears to be for children transitioning to high school to increase their travel distance somewhat, even as a higher proportion of these children also begin to attend the school nearest to their home. By contrast, although a similar proportion of children moving between primary schools increase their travel distance, these increases are typically smaller, and the average change in distance is negative. These primary school children also become less likely to attend the school nearest to their home. Although the patterns of change at each level of census geography is similar for children moving between primary schools, and those moving to high schools, this disguises the fact that children at the primary school level are on average moving to areas closer to home, while children at the high school level are moving to areas further away from home. Despite these differences, however, the correlates of particular kinds of changes to mobility are extremely consistent across the two groups of children.

Thirdly, the different correlates associated with the different types of changes to mobility substantiate the notion that the three measures of mobility used in this project capture slightly different aspects of the phenomenon. The evidence continues to be consistent with the notion that there are two separate processes of school choice, and by extension mobility behaviours, in play in contemporary urban Soweto-Johannesburg. One involves substantial travel, and typically involves those children with access to the greatest resources, and those living in areas in which services are not available. The other involves mobility at relatively local levels, and seems to engage children with more intermediate levels of resources and maternal education.

Chapter 9: Modelling educational mobility

9.1 Introduction

Having explored the bivariate relationships between educational mobility and a range of variables at the child, household, community and school levels, this chapter combines these variables to develop a multivariate model predicting educational mobility. This is done through a series of regression analyses, using child, household, community and schooling data as independent variables, and the various measures of educational mobility as dependent variables.

As most of the school attribute variables presented in Chapter 7 were closely related to mobility, but were also highly correlated with each other, principal component analysis (PCA) was conducted on all school attribute variables that were consistently found to be significantly related to mobility (school quintile, school fees, school enrolment, percent black learners, school sector, historical DET status, and pass rate) to generate a school quality index. This process was repeated for the school attended by each child in 1997 and 2003, as well as for the nearest grade-appropriate school to the child's home in 1997 and 2003. In all cases, the eigenvalues of the first two components of the PCA were both greater than 1, and were therefore both retained.

Transformation of some additional variables was also conducted to allow for their use in a regression context. Race was re-coded into a binary variable, coded 1 if the child was black African and 0 otherwise. Gender was similarly recoded to 1 if the child was a boy and 0 if the child was a girl. Poverty data for the area in which a child lived was used in its raw form, as opposed to the quintile form, as was household SES. The only non-binary categorical independent variable used was maternal education, and this was converted to a series of dummy variables, with a maternal education level of Grade 5 or below used as the base category.

In the first section of this chapter, regression models are developed to predict the straight-line distance between home and school for both 1997 and 2003. In the second section, logistic regression models are used to predict the likelihood that that children attend schools located in the same Sub-Place (SP) and Main Place (MP) areas in which they live, again for both 1997 and 2003. In the third section of the chapter, logistic regression models are again developed, this time to predict the likelihood that a child will attend his or her nearest grade-appropriate school. Finally, the implications of these various models for the hypotheses around mobility presented in previous chapters are discussed.

9.2 Straight-line distance

As noted in Chapter 5, the distribution of straight-line distance from home to school across the sample is highly non-normal. A range of transformations of this variable were generated and tested for normality, but none were found to be normal. The log transformation, however, was closest to a normal curve, and as a result, is used as the dependent variable in this set of models. All of the child, household and community level variables discussed in Chapter 5, along with the PCA-generated variables relating to the quality of a child's attended and nearest school, were considered for inclusion as independent variables.

9.2.1 1997

Initially, a standard OLS regression was run, including all child, household and community variables, as well as the derived scores for schools attended and nearest to a child's home. This model, however, featured extensive multicollinearity. The correlation matrix for all variables was examined, and one major area of concern was identified. This was the strong relationship between area poverty and the attributes of the school closest to the child's home. As the attributes of the nearest school were also closely related to the attributes of the school attended, a decision was made to retain the area poverty level variables only. Due to the high correlation between poverty levels at the various different levels of geography (SP and MP in particular), each was tested for significance. , and the decision was made to retain only SP poverty in the final model. Results for this model are described below, and presented in Table 9.1.

Tests were conducted to explore concerns around cases exerting an undue influence on the models. A number of outliers and cases with high leverage were identified. As these cases appeared to be features of the data rather than errors, they were not removed, but it should be noted that they may have biased the results reported below. Although a Breusch-Pagan test found no evidence of heteroskedasticity, both a visual examination of the variances and a White test provided evidence of heteroskedasticity. The distribution of the residuals also appeared normal on inspection, but a Shapiro-Wilk test rejected this hypothesis. Due to these various concerns, the regression was repeated using regression with robust standard errors, and also a robust regression which weights cases differently to minimize the effects of cases with particularly high influence. These results are also presented in Table 9.1, and it is clear that they are not substantially different from the results of the standard regression. For all models, tests indicated that omitted variable bias was present, and this persisted even with the inclusion of a range of other variables. However, a model specification link test failed to reject the assumption that the model was correctly specified.

The results of these regressions, presented in Table 9.1 below, suggest that the coefficients on race, maternal education, and the attributes of the school attended are most significant in predicting the distance a child travels from home to school in 1997. Across all models, black children are significantly more likely to travel further than children of other race groups. The relationship between distance and maternal education is somewhat more complex, and slightly counterintuitive, as it suggests that children whose mothers have very limited formal education (grades 5 or 6 only) tend to travel the furthest, followed by those whose mothers have reached grade 11-12. As shown by the coefficient on the first school attributes variable, children attending more advantaged schools, are also likely to travel further than children attending less advantaged schools. There is some weak evidence that maternal marital status may also be related to mobility, with children of married mothers tending to travel somewhat further. There is, however, no evidence that household SES shapes distance, although it seems more likely that the effect of household SES has already been captured in other relatively highly correlated variables, such as school quality and maternal education, than that it doesn't relate to mobility at all. There is also evidence that the poverty level of the area in which a child lives plays a limited role in predicting travel distance, with children in poorer areas tending to travel further. Overall, these findings sustain the hypothesis that black children from homes where mothers have at least some education, attending relatively privileged schools, but still living in less advantaged areas, tend, overall, to travel somewhat further to school than their peers.

	Standard regression	**Standard regression with robust errors**	**Robust regression**
Black African race	2.053 (0.174) ***	2.053 (0.187) ***	2.438 (0.160) ***
Male gender	0.005 (0.092)	0.005 (0.094)	0.010 (0.084)
Later age at first school enrolment	0.097 (0.094)	0.097 (0.094)	0.064 (0.086)
Maternal education grade 5-7	0.793 (0.248) ***	0.793 (0.235) ***	0.527 (0.227) **
Maternal education grade 8-10	0.395 (0.200) **	0.395 (0.167) **	0.242 (0.183)
Maternal education grade 11-12	0.561 (0.210) ***	0.561 (0.184) ***	0.369 (0.192) *
Maternal education post-school	0.404 (0.260)	0.404 (0.240) *	0.393 (0.238) *

Maternal marital status	0.176 (0.104) *	0.176 (0.108)	0.184 (0.095) *
Household SES 1997	0.013 (0.032)	0.013 (0.034)	0.026 (0.029)
School attended 1997 attributes component 1[9]	0.644 (0.037) ***	0.644 (0.034) ***	0.696 (0.034) ***
School attended 1997 attributes component 2	0.057 (.0481)	0.057 (0.042)	0.051 (0.044)
SP poverty (raw score)	0.068 (0.038) *	0.068 (0.045)	0.080 (0.035) **
Constant	-1.940 (0.247) ***	-1.940 (0.223) ***	-2.205 (0.226) ***
*significant at P<0.1 level ** significant at P<0.05 level ***significant at P<0.01 level	No. of obs = 742 F(12, 729) = 36.70 Prob > F = 0.0000 R-squared = 0.3766 Adj R-squared = 0.3663 Root MSE = 1.236	Number of obs = 742 F(12, 729) = 54.78 Prob > F = 0.0000 R-squared = 0.3766 Root MSE = 1.236	No. of obs = 742 F(12, 729) = 52.10 Prob > F = 0.000

Table 9.1: 1997 regression results. Figures in parentheses are standard errors.

9.2.2 2003

A similar process to that described above was used to develop the model for 2003. The major difference in model construction for 2003 is that a variable indicating whether the child repeated a grade between 1996 and 2003 is included. As was the case with the 1997 model, tests indicated concerns around cases with particularly high influence, as well as the presence of some heteroskedasticity. As a result, Table 9.2, below, reports the results of a standard regression, along with the results of the regression re-run with robust errors, and a robust regression to reduce the impact of particularly influential cases. Once again, evidence for omitted variable bias could not be eliminated, although a model specification link test failed to reject the assumption that the model was correctly specified.

[9] The strongest loadings for school component 1 in 1997 were historical DET status and percentage black students, both loading negatively. Component 2 loadings were matric pass rate and school fees, loading positively, and school enrollment, loading negatively.

The patterns identified in the 2003 regressions are very similar to those identified for 1997, with race, school attributes, and maternal education continuing to be the strongest predictors of distance from home to school. The role of maternal education in shaping mobility has, however, changed somewhat, becoming more linear. Maternal education at the grade 5 to 7 level no longer contributes to mobility, and higher maternal education is typically associated with higher mobility. Black children continue to travel substantially further than members of other race groups, as do children attending more advantaged schools. Evidence for a role of residential area poverty or maternal marital status in determining mobility has largely disappeared in 2003. Overall, however, the relative similarity of these results over time, despite the substantial change in the age of the children, seems to suggest that the determinants of mobility remain fairly consistent, throughout the primary school years, and even into secondary schooling.

	Standard regression	**Standard regression with robust errors**	**Robust regression**
Black African race	1.876 (0.216) ***	1.876 (0.269) ***	2.275 (0.209) ***
Male gender	-0.153 (0.096)	-0.153 (0.094)	-0.212 (0.093) **
Late age at first school enrolment	0.028 (0.100)	0.028 (0.100)	-0.003 (0.097)
Grade repetition	0.018 (0.104)	0.018 (0.104)	-0.026 (0.101)
Maternal education grade 5-7	0.327 (0.255)	0.327 (0.249)	0.152 (0.247)
Maternal education grade 8-10	0.398 (0.187) **	0.398 (0.165) **	0.324 (0.181) *
Maternal education grade 11-12	0.488 (0.196) **	0.488 (0.175) ***	0.378 (0.190) **
Maternal education post-school	0.415 (0.248) *	0.415 (0.232) *	0.280 (0.240)
Maternal marital status	0.086 (0.104)	0.086 (0.101)	0.064 (0.101)
Household SES 2003	0.112 (0.097)	0.112 (0.055) **	0.077 (0.094)
School attended 2003 attributes component 1[10]	0.696 (0.037) ***	0.696 (0.039) ***	0.746 (0.036) ***

[10] For 2003, historical DET status and percent black students both loaded strongly negatively on component 1, while school quintile and school fees loaded strongly positively. School

School attended 2003 attributes component 2	0.068 (0.053)	0.068 (0.053)	0.081 (0.051)
SP poverty (raw score)	0.071 (0.038) *	0.071 (0.047)	0.024 (0.037)
Constant	-1.355 (0.283) ***	-1.355 (0.282) ***	-1.565 (0.274) ***
*significant at P<0.1 level ** significant at P<0.05 level ***significant at P<0.01 level	No. of obs = 543 F(13, 529) = 40.87 Prob > F = 0.0000 R-squared = 0.5011 Adj R-squared = 0.4888 Root MSE = 1.0728	Number of obs = 543 F(13, 529) = 43.06 Prob > F = 0.0000 R-squared = 0.5011 Root MSE = 1.0728	No. of obs = 543 F(13, 529) = 51.80 Prob > F = 0.0000

Table 9.2: 2003 regression results. Figures in parentheses are standard errors.

9.3 Census area geography

In this section, models are developed for whether children attend schools in the same geographical area in which they live. Due to the very low numbers of children attending school in their home Small Area Level (SAL), and the very low numbers who do not attend school in their home Municipality (MN), models are only pursued for mobility at the Sub-Place (SP) and Main-Place (MP) levels. As the dependent variable used in these models is binary, a logistic regression approach is used. Initially, the regression was run using the same set of variables as presented in the distance to school regressions above. However, tests indicated specification errors, which were reduced by replacing the raw household SES scores with dummy variables for each quintile of household SES. This improvement may be related to nonlinearities in the relationship between household SES and the census area based measure of mobility. In both 1997 and 2003, household SES quintile 1 (highest poverty) was used as the base category. SP poverty was also treated as a categorical variable, again to allow for nonlinearity in the relationship between SP poverty and mobility behaviour. Additional tests for multicollinearity, goodness of fit, and particularly influential cases were also conducted on all models before they were finalised (Chen, Ender et al. 2011). Although at the end of this

---

enrolment loaded strongly negatively on component 2, while school fees again loaded strongly positively.

process all of the models detailed below continued to provide some evidence of specification error, there was no theoretically justifiable way to further improve their specification.

9.3.1 Sub-Place level

Results for both 1997 and 2003 are presented in Table 9.3 below. Regressions were run both with and without robust standard errors. However, as the coefficients in both cases are identical, and significance levels did not change substantially, only the results using robust standard errors are presented in Table 9.3. Note that grade repetition was not included in the 1997 model. The 1997 model used the household SES poverty quintiles based on 1997 data, while the 2003 model used the quintiles based on the 2003. Each model also used the school attributes score based on the school attended at that point in time.

The coefficients returned by the logistic regression for SP mobility in 1997 indicate that black children are again significantly more mobile, as black race is negatively associated with the probability that a child attends a school in the same SP as his or her home. By contrast, the association between maternal education and mobility has disappeared. There is some evidence that children in household SES quintiles 2 and 4 are more likely to attend schools outside of their home SP. The evidence that the attributes of the school attended shape mobility remains strong. The coefficients on both components of school attributes indicate that children attending more advantaged schools are more likely to attend a school located outside of their home SP. Finally, area poverty is also significantly associated with mobility. Those children living in wealthier areas are less likely than their peers living in poorer areas to be attending schools in their home SPs.

The 2003 results are very similar to those for 1997. Black African race remains a strongly significant predictor of mobility at the SP level, with black children less likely to attend a school in their home SP. There is again no evidence that maternal education is related to mobility at the SP level in 2003. Children attending more advantaged schools, and those living in the least advantaged areas, continue to be significantly more likely to be travelling to attend a school outside of their home SP in 2003. Children with mid-range household SES are also particularly likely to be attending school outside of their home SP. There is also weak evidence that children with married mothers are more likely to be travelling out of their home SP in order to attend school in 2003.

Overall, determinants of mobility at the SP level, in both 1997 and 2003, appear to be fairly similar to determinants of distance from home to school, with the exception that maternal education no longer appears to play a significant role. This is critical, because of the close relationship between maternal education and affluence, and suggests that household resource levels may be less important in shaping mobility at the fairly small SP level of geography, than in shaping distance from home to school.

	1997: Logistic regression with robust errors	**2003: Logistic regression with robust errors**
Black African race	-3.314 (0.480) ***	-3.422 (0.725) ***
Male gender	-0.222 (0.173)	0.005 (0.229)
Late age at first school enrolment	-0.233 (0.178)	0.280 (0.238)
Grade Repetition		0.015 (0.245)
Maternal education grade 5-7	-0.624 (0.434)	-0.276 (0.573)
Maternal education grade 8-10	-0.413 (0.358)	-0.458 (0.436)
Maternal education grade 11-12	-0.266 (0.380)	-0.566 (0.471)
Maternal education post-school	-0.258 (0.463)	-0.231 (0.539)
Maternal marital status	-0.164 (0.197)	-0.433 (0.248) *
Household SES quintile 2	-0.554 (0.244) **	-0.531 (0.323)
Household SES quintile 3	-0.182 (0.250)	-1.130 (0.330) ***

Household SES quintile 4	-0.468 (0.274) *	-0.269 (0.342)
Household SES quintile 5 (most advantaged)	-0.465 (0.376)	0.249 (0.391)
School attended attributes component 1	-0.947 (0.102) ***	-1.240 (0.131) ***
School attended attributes component 2	-0.202 (0.099) **	-0.443 (0.147) ***
SP poverty quintile 2 (relatively low area poverty)	-1.033 (0.349) ***	-0.983 (0.627)
SP poverty quintile 3	-0.690 (0.422)	-0.684 (0.645)
SP poverty quintile 4	-0.975 (0.427) **	-1.409 (0.669) **
SP poverty quintile 5 (highest poverty areas)	-1.474 (0.464) ***	-2.059 (0.698) ***
Constant	3.967 (0.535) ***	3.602 (0.760) ***
*significant at P<0.1 level ** significant at P<0.05 level ***significant at P<0.01 level	No. of obs = 742 Wald chi2(18) = 94.79 Prob > chi2 = 0.0000 Pseudo R-squared = 0.1940 Log likelihood = -410.34555	No. of obs = 543 Wald chi2(19) = 124.39 Prob > chi2 = 0.0000 Pseudo R-squared = 0.2849 Log likelihood = -252.24727

Table 9.3: 1997 & 2003 SP mobility regression results. Figures in parentheses are standard errors.

9.3.2 Main Place level

Results for logistic regressions at the MP level for both 1997 and 2003 are presented in Table 9.4 below. Again, only results with robust standard errors are presented. Grade repetition was not included in the 1997 model. The 1997 model used the household SES poverty quintiles based on 1997 data, while the 2003 model used the quintiles based on the 2003. Each model also used the school attributes score based on the school attended at that point in time.

In 1997, black children, and those attending comparatively advantaged schools were less likely to be attending a school in their home MP. Although there is weak evidence that maternal education at the grade 11 to 12 level is associated with a higher probability of attending school outside of the MP in which the home is located, there appears to be little role for maternal education in determining MP level mobility. Similarly, evidence for a role of household SES is also limited, with only children falling into quintile 2 being significantly more likely to attend a school outside of their home MP. By contrast, however, and as was the case at the SP level, there is strong evidence that attending a more advantaged school, and living in an area with higher poverty levels, were both associated with an increased probability of attending school outside of the home MP.

The 2003 results for mobility at the MP level are fairly similar to those for 1997. Once again, black children are significantly more likely to be attending a school outside of their home MP. There is some evidence that maternal education at the post-school level increases the likelihood of a child attending a school outside of their home MP. Children enrolled in more affluent schools are also more likely to be travelling to a school outside of the MP in which they live. Finally, children living in SP areas with higher levels of poverty are more likely to be attending a school outside their home MP.

Overall, these results suggest that the determinants of MP mobility do not change substantially over time, and indicate that black children, attending fairly advantaged schools, but living in poor areas, are most likely to travelling outside of the MP in which their home is located in order to go to school. Interestingly, there was some evidence that particularly high levels of maternal education were predictive of mobility at the MP level, which was not the case for mobility at the SP level. This may be reflective of the typically longer distances associated with MP mobility compared to SP mobility, which would cause it to require greater levels of resource investment.

	1997: Logistic regression with robust errors	2003: Logistic regression with robust errors
Black African race	-3.820 (0.533) ***	-6.816 (3.574) *
Male gender	0.037 (0.255)	-0.019 (0.292)
Late age at first school enrollment	-0.305 (0.244)	-0.002 (0.341)
Grade repetition		-0.091 (0.350)
Maternal education grade 5-7	-1.064 (0.802)	-0.465 (1.043)
Maternal education grade 8-10	-1.057 (0.700)	-0.441 (0.649)
Maternal education grade 11-12	-1.340 (0.707) *	-0.529 (0.626)
Maternal education post-school	-1.254 (0.773)	-1.472 (0.735) **
Maternal marital status	-0.056 (0.269)	-0.137 (0.339)
Household SES quintile 2	-0.985 (0.389) **	-0.243 (0.504)
Household SES quintile 3	-0.484 (0.390)	-0.798 (0.519)
Household SES quintile 4	-0.449 (0.403)	-0.505 (0.518)
Household SES quintile 5 (most advantaged)	-0.598 (0.511)	0.029 (0.546)
School attended attributes component 1	-1.324 (0.116) ***	-1.931 (0.202) ***
School attended attributes component 2	0.042 (0.110)	0.133 (0.180)
SP poverty quintile 2 (relatively low area poverty)	-1.532 (0.611) **	-2.466 (0.812) ***
SP poverty quintile 3	-1.883 (0.694) ***	-3.013 (0.820) ***
SP poverty quintile 4	-2.018 (0.708) ***	-3.410 (0.809) ***
SP poverty quintile 5 (highest poverty areas)	-3.033 (0.723) ***	-3.716 (0.836) ***
Constant	8.188 (0.904) ***	11.037 (3.790) ***
*significant at P<0.1 level ** significant at P<0.05 level ***significant at P<0.01 level	No. of obs = 742 Wald chi2(18) = 201.40 Prob > chi2 = 0.0000 Pseudo R-squared = 0.4315 Log likelihood = -232.40251	No. of obs = 543 Wald chi2(1) = 201.40 Prob > chi2 = 0.0000 Pseudo R-squared = 0.56102 Log likelihood = -148.17552

Table 9.4: 1997 and 2003 MP mobility regression results. Figures in parentheses are standard errors.

9.3.3 Census area mobility discussion

At both the SP and MP level, and in both 1997 and 2003, black race is a strong predictor for increased mobility, as is attending a comparatively advantaged school, and living in a less affluent SP. Interestingly, maternal education, which did predict distance from home to school, does not predict SP mobility at all, and only the highest levels of maternal education are associated with MP mobility. This may indicated that the role of household resources in determining mobility is more limited when looking at travel between smaller areas. This fits well with the argument that two patterns of mobility – one involving substantial travel to historically advantaged schools, and demanding more resources, and one involving more localized travelling and requiring fewer resources – are evident in urban South Africa. SP mobility appears to fit into the lower-resource pattern of mobility, while MP mobility is more closely linked to the higher-resource pattern.

9.4 Nearest school analysis

The final set of models presented in this chapter use attendance at a child's nearest grade-appropriate school as the dependent variable. Logistic regressions were conducted using both the variable indicating whether the child attended his or her nearest public school, and the variable indicating whether the child attended his or her nearest public or independent school. As the results in both cases were essentially identical, only the results for the variable including both public and independent schools are presented here. Constructing this model was challenging, as it was not possible to obtain a particularly good fit for either the 1997 or 2003 models. This may be due to omitted variables helping to determine whether a child attends his or her nearest school, and may also be due to a fairly high level of randomness in this particular outcome. However, given that both models pass all other goodness of fit tests, and give no indication of specification errors, the low R-squared in the context of a logistic regression is not necessarily of great concern.

Again, both a standard logistic regression, and a logistic regression with robust errors were conducted for each time point, and provided very similar results. Table 9.5, below, presents the results for the regressions with robust errors for 1997 and 2003. Grade repetition was only included in the 2003 model. The 1997 model was conducted using 1997 household SES and school attributes, while the 2003 model used these variables for the 2003 time point.

The 1997 model indicates that black children are significantly less likely to attend their nearest school. Children who start school late for their age are more likely to attend their nearest schools than those who started early. This is interesting, as it the effect of age at enrolment appears to be moving in the opposite direction than that suggested in the bivariate analyses. The coefficient may be driven by the fact that more affluent children are likely to start school later, and are also more likely to attend their nearest school, even though their average travel distance is greater than that of their less affluent peers. There is no evidence for any relationship between likelihood of attending the nearest school, and either household SES or maternal education. There is, however, a strong significant relationship between the nature of the school the child attends in 1997, and whether this is the school closest to his or her home. Children attending more advantaged schools are less likely to be attending the school nearest to their home. Finally, there is also evidence that children living in areas with mid-range levels of poverty are least likely to attend their nearest schools. Children living in the most disadvantaged areas are, however, still less likely to attend their nearest school than their peers in more advantaged areas. This is in line with the bivariate findings presented in earlier chapters.

As in 1997, black children in 2004 remain significantly less likely to attend their nearest grade appropriate schools than children of other races. The relationship between later age at first enrolment and attending the nearest school has, however, disappeared by 2003. There is weak evidence that children of mothers with limited primary school education (Grades 5-7) may be more likely to attend their nearest school than children of mothers with no formal education, but other than this, there is again no evidence for a relationship between maternal education and whether a child attends his or her nearest school. Similarly, there is also some evidence that children in households with mid-range SES are more likely than the most disadvantaged children to attend a school other than their nearest school, but other than this, no evidence for a relationship between household SES and whether or not a child attends his or her nearest school. A statistically significant negative relationship remains between how advantaged the school a child attends is, and the likelihood that the child is attending his or her nearest school. Finally, there is also a strong negative relationship between area poverty and probability of attending the nearest school, with children living in the poorest areas least likely to do so.

The 1997 and 2003 results are once again largely consistent over time, and highlight the role of race, school quality, and area poverty in determining whether a child attends his or her nearest school. The very limited evidence for any role of maternal education or household SES in determining whether a child attends his or her nearest school also supports the argument that choosing not to attend the nearest school is a form of school choice which requires fairly little in terms of the investment of resources.

	1997: Logistic regression with robust errors	**2003: Logistic regression with robust errors**
Black African race	-2.251 (0.431) ***	-1.477 (0.643) **
Male gender	-0.138 (0.199)	0.145 (0.263)
Late age at first school enrollment	0.504 (0.201) **	-0.150 (0.280)
Grade repetition		-0.181 (0.285)
Maternal education grade 5-7	-0.617 (0.509)	1.110 (0.641) *
Maternal education grade 8-10	-0.371 (0.385)	0.151 (0.512)
Maternal education	-0.388 (0.424)	0.242 (0.539)

grade 11-12		
Maternal education post-school	-0.325 (0.553)	0.636 (0.657)
Maternal marital status	-0.242 (0.241)	0.100 (0.280)
1997 household SES quintile 2	-0.281 (0.300)	-0.215 (0.356)
1997 household SES quintile 3	-0.326 (0.323)	-0.858 (0.416) **
1997 household SES quintile 4	0.194 (0.318)	-0.172 (0.397)
1997 household SES quintile 5 (most advantaged)	0.062 (0.421)	-0.311 (0.481)
School attended 1997 attributes component 1	-0.626 (0.099) ***	-0.711 (0.143) ***
School attended 1997 attributes component 2	-0.233 (0.102) **	-0.658 (0.172) ***
SP poverty quintile 2 (relatively low area poverty)	-0.693 (0.333) **	-0.566 (0.582)
SP poverty quintile 3	-1.550 (0.433) ***	-1.263 (0.587) **
SP poverty quintile 4	-0.962 (0.403) **	-1.291 (0.599) **
SP poverty quintile 5 (highest poverty areas)	-0.914 (0.436) **	-1.612 (0.658) **
Constant	1.477 (0.541) ***	0.349 (0.733)
*significant at P<0.1 level ** significant at P<0.05 level ***significant at P<0.01 level	No. of obs = 738 Wald chi2(18) = 87.39 Prob > chi2 = 0.0000 Pseudo R-squared = 0.1241 Log likelihood = -327.56921	No. of obs = 541 Wald chi2(19) = 67.58 Prob > chi2 = 0.0000 Pseudo R-squared = 0.1631 Log likelihood = -207.7699

Table 9.5: 1997 and 2003 nearest school attendance logistic regression results. Figures in parentheses are standard errors.

9.5 Conclusion

The data presented in previous chapters of this thesis have suggested that two patterns of learner mobility are operating in the Johannesburg-Soweto area. Firstly, there are children who are travelling particularly long distances to attend school, typically at a fairly high economic and social cost. The second pattern involves travel at a more local level, with children and their families making active choices between those schools accessible from their homes. This second pattern of mobility represents a substantially less resource-intensive approach to engaging in school choice, and can be engaged in by a broader group of children, and most notably those children whose mothers have only fairly limited education. The results presented in this chapter are summarized in Table 9.6, below, and provide additional support for presence of these two, different, patterns of mobility.

	Distance from home to school	**School outside of home SP**	**School outside of home MP**	**Not attending nearest school**
Variables associated with increased mobility	Black race	Black race	Black race	Black race
				Younger age at first school enrolment (1997 only)
	Maternal education		Maternal education Grade 11-12 (1997 only) Maternal education post-school (2003 only)	Very low maternal education (2003 only)
		1997 Household SES quintiles 2 & 4 (1997 only) 2003 Household SES middle quintiles	1997 Household SES quintile 2 (1997 only)	2003 Household SES quintile 3 (2003 only)
	Higher school quality	Higher school quality	Higher school quality	Higher school quality
		Poorer SP area	Poorer SP area	Poorer SP area

Table 9.6: Summary of variables associated with increased mobility in the regression models presented in this chapter

Evidence for the first, longer-distance pattern is provided in the regression models for distance between home and school, and to a lesser degree, mobility at the MP level of census geography. In addition to black race, and attending a high quality school, distance from home to school is strongly related to higher maternal education. The role of higher maternal education in increasing mobility, however, largely disappears in all other models, with the exception of mobility at the MP level. Maternal education, therefore, appears to be critical to determining whether a child engages in a more costly form of mobility or not. Although the absence of significant coefficients on household SES variables is of some concern to the hypothesis that these forms of mobility are pursued by more advantaged families, this is probably due to the strength of the correlation between household SES and school quality, and the very substantial coefficients attached to school quality in these regressions. Overall, the models for distance from home to school, and for MP mobility suggest that these forms of mobility are more demanding on families, and at the very least require the additional social and human capital associated with a more highly educated mother.

When looking at the results for SP mobility, and a child not attending his or her nearest school, black race and the quality of the school attended continue to be important predictors. Although, as noted, there is little or no evidence for a role of maternal education in predicting these types of mobility, there is some evidence that intermediate levels of household SES may play a role. These combinations of significant variables support the argument that these are forms of mobility that do not require particularly high levels of social or economic capital, even though they are not completely cost free. The role of area poverty in these models has also become more significant, suggesting that these forms of mobility are likely to be more strongly driven by the nature of local educational opportunities than by the resources available to a child and his or her family.

Chapter 10: Conclusion

10.1 Introduction

This chapter begins with a brief overview and synthesis of the key findings presented in this thesis, and then discusses the implications of these findings. The original contributions – methodological, empirical and theoretical – that this thesis makes to the existing body of scholarly literature are highlighted. The contextual relevance of the study findings is discussed, with a brief discussion of potential implications for school policy in South Africa. Finally, limitations to the work presented here, along with suggestions for future work, are presented.

10.2 Overview of key findings

Table 10.1 below outlines the key findings of this thesis, with respect to each of the study's major aims. These findings are discussed at greater length in the subsequent sections.

Objective	Chapter	Thesis Findings
To develop approaches to the measurement of learner mobility appropriate to the South African context	3 & 5	-- Three different approaches to measuring learner mobility were tested -- Each approach provided different, but complementary data -- Using multiple approaches to measuring learner mobility allowed for the identification of two distinct patterns of mobility – one based primarily on choice between local schools, and one based on travel of substantial distances to schools in more advantaged areas
To measure the extent of learner mobility in post-Apartheid Johannesburg-Soweto, South Africa	5	-- Each approach to the measurement of learner mobility provided evidence that learner mobility is highly prevalent in Johannesburg-Soweto -- Approximately a quarter of children travel over 5km to school each way on a daily basis -- Roughly 25% of children attend school outside of the MP in which they live, while 60% attend

		school outside of the SP in which they live -- Less than 20% of children attend the grade-appropriate school that is closest to their home. -- This data suggests two distinct patterns of mobility – one due primarily to choice between fairly local schools, and involving relatively limited travel, and one involving the choice of historically advantaged schools substantially further afield
To identify potential determinants of learner mobility at the child, family and community level	6 &9	-- At the child level, bivariate analyses indicated that mobility behaviour was related to race, age at school enrolment, grade repetition, and phase of schooling -- At the family level, there was evidence for a relationship between mobility and maternal education and household SES -- At the community level, there was evidence for a non-linear relationship between mobility and area poverty -- Multivariate analysis suggested that race, maternal education and area poverty where the strongest and most consistent determinants of mobility behaviour
To explore the relationship between school attributes and mobility behaviour	7 & 9	-- On average, children were found to be attending a school more advantaged than the school closest to their home -- More advantaged children were found to be attending more advantaged schools -- Attending a comparatively advantaged school was associated with greater engagement in all forms of mobility -- Children attending private schools, higher quintile schools, schools with a lower proportion of black learners, schools charging higher fees, schools that did not historically fall under the DET, and schools with higher pass rates were more likely to be engaged in learner mobility
To identify whether and how learner mobility changes as children age	8	-- There was no clear evidence to suggest that mobility behaviour changes substantially as children age -- Although there was some evidence that children typically travelled slightly further to high schools than primary schools, and were slightly more likely to attend their nearest school at the high school level, this appeared to be due to differences in the sizes and distribution of primary and high schools
To generate a preliminary	9	-- Models were developed using each of the

model of the determinants of learner mobility		three definitions of learner mobility investigated in this thesis -These models suggested that black race and attending a high quality school were strong predictors of all forms of mobility -- Mobility requiring substantial travel was also predicted by maternal education, while mobility at the more local level was most strongly predicted by local area poverty
To develop an evidence-based conceptual framework to support ongoing research into learner mobility and school choice	3 & 9 & 10	-- The evidence presented in this thesis supports the argument that child, family, community and school level variables all play a role in shaping school choice decision making -- However, given that two patterns of learner mobility are in operation in Johannesburg-Soweto, and that they appear to be driven by different variables, the evidence also suggests that a conceptual framework which does not differentiate between these forms of mobility may be insufficient

Table 10.1: Overview of the key findings presented in this thesis

10.2.1 Developing approaches to measuring learner mobility

The thesis made use of three different approaches to measuring learner mobility: straight-line distance between home and school, whether children attend school in the area in which they live, and whether children attend the grade-appropriate school nearest to their homes. Although all of these measures provided consistent evidence for high levels of learner mobility amongst urban South African primary school children, they also captured different aspects of this mobility. As such, the measures proved to be complementary, providing evidence that learner mobility in Johannesburg-Soweto should be understood to consist of two distinct components. These are firstly the mobility involving fairly limited travel distances associated with choice between a number of local schools, and secondly the mobility involving much greater travel distances, and relating to the choice of schools much further afield, typically in historically more advantaged areas than the child's home.

10.2.2 Measuring the extent of learner mobility in Johannesburg-Soweto

Although there has long been substantial reason to believe that South African learners are motivated to travel relatively long distances to attend particular schools (Cosser and du Toit 2002; Fiske and Ladd 2004; Maile 2004; Msila 2005; Woolman and Fleisch 2006; Msila 2009), this thesis presents the first population-based evidence on what these distances actually are, and how widespread engagement in mobility is. This evidence does suggest that primary school learners in the Johannesburg-Soweto area are extremely mobile. At both ages 7 and 13, over 25% of children were travelling to schools over 5km away from their homes. Almost 60% of children were travelling outside of the Census Sub-Place (SP) area in which they lived (roughly equivalent to a suburb), to attend school in a different SP. At the Main Place (MP) level, equivalent to a small town, roughly 25% of children were travelling to attend a school in an MP other than the one in which they lived. Finally, fewer than 20% of children were found to be attending the grade-appropriate school closest to their home.

These figures provide evidence that learner mobility, and school choice, is widespread amongst primary school children and their families in urban South Africa. Certainly, the numbers of children travelling on a daily basis to schools over 5km from their homes is notable, and particularly at 7 years of age was not anticipated by existing information. As corroborated by the overlap between children travelling over 5km, and those mobile at the MP level, most of these children are travelling to areas in very different parts of Johannesburg from their homes, and will be attending school in a very different linguistic and socio-economic context. The data presented also suggest that while some families engage in school choice by sending children to schools further afield, a substantial proportion of others engage in choice at a local level. A full 50% of children lived less than 1.5km from their school, but less than half of them were attending the school closest to their home. This suggests that even less advantaged families are making active decisions in pursuit of the best possible educational opportunities for their children, even in the context of an extremely poorly performing public schooling system. Again, this provides support for the argument that there are two distinct patterns of learner mobility in evidence in Johannesburg-Soweto.

10.2.3 Potential child, household and community-level determinants of learner mobility

This thesis tested a range of variables at the child, family and community levels to explore their relationship to learner mobility. At the child level, race, gender, age at first school enrolment and grade repetition, and school phase in 2003 were examined. For all definitions of mobility, race was strongly related to mobility. Although small sample sizes meant that conclusions could not be drawn about the mobility of white and Indian children, there was clear evidence that black children were substantially more likely to engage in mobility than their coloured peers. There was some indication that girls, especially on reaching high school, tended to travel slightly further than boys, but overall there was no compelling evidence for a relationship between gender and mobility.

There was evidence that children who first enrolled in school at an older age travelled further at age 7, though not at age 13. This may be due to less wealthy parents sending their children to school at a younger age, to minimize the need to provide childcare. It may also relate to wealthier parents enrolling their children in primary school at a slightly older age, when they might be expected to cope more easily with academic and social challenges associated with attending a school outside of the local area. Grade repetition was also strongly related to distance from home to school, at both ages 7 and 13, as well as to mobility at the SP and MP levels. By contrast, there was no evidence for a relationship between grade repetition and enrolment at the nearest school. Finally, children attending high schools did travel further than those still enrolled at the primary level, and were also significantly more likely to attend their nearest school.

At the family level, the thesis examined maternal education, maternal marital status, and household SES. Maternal education was strongly linked to all measures of mobility, with the children of more educated mothers tending to travel further, in both 1997 and 2003. However, there was some attenuation of this relationship at the very highest level of maternal education, perhaps because these families tended to live in more affluent areas, closer to high quality schools. There was no clear evidence for a relationship between maternal marital status and a child's engagement in learner mobility. Finally, again in both 1997 and 2003, there was a strong relationship between household SES and learner mobility, with children living in more advantaged families being substantially more likely to engage in mobility, and tending to travel greater distances.

Finally, at the community level, the poverty level of the area in which the child lived was explored. Although this data was fairly complex, two general patterns were discernable. Firstly, there was a clear, but non-linear, relationship between distance from home to school, and area poverty. Those children travelling furthest tended to be living in areas that were either relatively affluent, or particularly poor. Secondly, however, there was also a fairly linear, and positive, relationship between the likelihood of a child attending his or her nearest school, and the affluence of the area in which the child lived.

The relationships between learner mobility and the variables discussed here substantiate the notion that two different types of school choice are in play in contemporary urban South Africa. Firstly, there is clear evidence that certain measures of mobility – particularly those relating to distance travelled – are associated with indicators of affluence, such as socio-economic status, maternal education, or living in a comparatively advantaged residential area. The extent to which a child's family has the means to engage in mobility is clearly one determinant of school choice involving mobility. Secondly, however, there is also evidence of mobility, particularly at a relatively local level, that is not strongly linked to affluence. Engagement in this more local mobility appears to be more closely related to the poverty of the area a child lives in, and by extension the quality of the local schools. This second type of mobility is particularly evident in the data around whether or not a child attends his or her nearest school, which suggests that even families with relatively limited means appear to be making use of school choice to obtain the best educational opportunities possible.

10.2.4 Potential school-level determinants of learner mobility

The thesis also explored the relationship between a child's mobility, and a range of attributes of the school he or she attended. The first important finding here was that, on average, children attended schools that were more advantaged than would be expected on the basis of the schools closest to their homes.

The second important finding was that there was a very strong relationship between child, family and community attributes, and the attributes of the school attended, with more advantaged children tending to attend more advantaged schools. Higher maternal education and household SES were particularly strong predictors of attending higher quality schools. By contrast, children living in areas with high poverty levels were significantly more likely to attend poorer schools.

The third important finding was the clear evidence for a strong relationship between mobility and the attributes of the school a child attended. Children attending private schools, higher quintile schools, schools with a lower proportion of black learners, schools charging higher fees, schools that did not historically fall under the DET, and schools with higher pass rates, were more likely to be engaged in learner mobility, largely regardless of how it was measured.

This group of findings substantiates the notion that learner mobility and school choice are typically used to enable a child to access higher quality education than would ordinarily be the case. This appears to hold regardless of the type of mobility explored. While more advantaged children remain substantially more likely to attend more advantaged schools, it does appear to be the case that even less advantaged children and families are able to make use of school choice and learner mobility in such a way as to improve the educational opportunities that they are able to access.

10.2.5 Changes in mobility as children age

While it was initially anticipated that mobility would increase substantially between 1997 and 2003, as the children aged, became more independent, and began to transition to high schools, this did not appear to be the case. Although children moving from a primary school to a high school did tend to experience a slight increase in distance, this appeared to be due to the smaller number of high schools available. There was, however, little evidence that children's travel behaviour changed substantially when the transitioned to high school. Children who were enrolled in primary schools at both points in time did also not experience any significant changes to their mobility.

Despite the relative consistency in children's mobility over time, school change was widespread across the sample, with over half of the sample making school changes other than those required by the transition to high school. Children living in poor areas, whose household SES increases with time, and whose mothers have intermediate levels of education appear to be the most likely to move between different primary schools. Although changing between two primary schools typically resulted in a small decrease in distance travelled to school, substantial changes in mobility were not evident. This suggests that school change during the primary school period may also be a strategy used in pursuit of the best available educational opportunities by children and families with resources to allow for substantial travel.

These findings also provide support for the hypothesis that two different patterns of learner mobility are in play. Firstly, there is a group of children travelling fairly long distances. These children typically experience more stable primary schooling, and their mobility does not change substantially with the transition to high school. Secondly, there is a larger group of children attending relatively local schools. These children tend to experience more school changes during the primary school years, although these changes typically do not involve substantial changes in mobility. This data additionally suggests that not only are at least two patterns of school choice in play, but that the pattern in which a child is engaged appears to be fairly path-dependent. The pattern of mobility in which a child engages at the beginning of their schooling appears to remain relatively consistent even during the transition to high school.

10.2.6 Predicting mobility

In Chapter 9 of the thesis, all of the variables discussed above were combined to generate models for each of the forms of mobility behaviour discussed. The results of these models are summarized in Table 10.2 below.

Variables		Distance	SP mobility	MP mobility	Nearest school
Child level					
Race	**1997**	Black race associated with greater travel	Black race associated with mobility	Black race associated with mobility	Black race associated with mobility
	2003	Black race associated with greater travel	Black race associated with mobility	Black race associated with mobility	Black race associated with mobility
Gender	**1997**	--	--	--	--
	2003	--	--	--	--
Age at first enrolment	**1997**	--	--	--	Earlier enrolment associated with mobility
	2003	--	--	--	--
Repetition	**2003**	--	--	--	--
Household level					
Maternal education	**1997**	Maternal education between Gr5 & Gr12 associated with greater travel	--	Maternal education between Gr11 & 12 associated with mobility	--
	2003	Maternal education beyond Gr8 associated with greater travel	--	Post-school maternal education associated with mobility	Maternal education between Gr5 & 7 associated with less mobility
Maternal marital status	**1997**	--	--	--	--
	2003	--	Married mother associated with mobility	--	--
Household SES	**1997**	--	SES quintile 2 & 4 associated with mobility	SES quintile 2 associated with mobility	--
	2003	--	SES quintile 3 associated with	--	SES quintile 3 associated with mobility

			mobility		
Community level					
Community poverty (SP)	**1997**	Living in a higher poverty area associated with slightly greater travel	Living in a high or low poverty area associated with mobility	Living in a higher poverty area associated with mobility	Living in a higher poverty area associated with mobility
	2003	--	Living in a higher poverty area associated with mobility	Living in a higher poverty area associated with mobility	Living in a higher poverty area associated with mobility
School level					
School resources	**1997**	Attending a better school associated with greater travel	Attending a better school associated with mobility	Attending a better school associated with mobility	Attending a better school associated with mobility
	2003	Attending a better school associated with greater travel	Attending a better school associated with mobility	Attending a better school associated with mobility	Attending a better school associated with mobility

Table 10.2: Summarized results for the models of mobility developed in Chapter 9. Results presented are for regressions with robust errors.

When all variables are controlled, being black, having a more educated mother, attending a more advantaged school, and living in a comparatively disadvantaged area all appeared to predict a greater distance between home and school, in both 1997 and 2003. That is, more advantaged children living in relatively disadvantaged areas, are likely to travel the greatest distances to get to school. When mobility was measured in a way that picked up local level school choice, all the variables listed above retained significance, with the exception of maternal education. The loss of a significant coefficient for maternal education suggests that household resource levels may be less critical in shaping school choice and mobility at the local level. Again, this provides support for the notion that two patterns of travel are in evidence, one requiring substantially more in the way of resources than the other.

10.2.7 Developing an evidence based conceptual framework for the study of learner mobility

Much of the evidence presented in this thesis has indicated that there are two distinct patterns of learner mobility in operation in contemporary urban South Africa. The first form appears to involve roughly 25% of children, and is fairly resource intensive, involving often substantial travel, typically to well-resourced schools in historically advantaged areas. The second form of learner mobility is far less resource intensive, and relates primarily to children who are attending local schools, but not the schools that are closest to their homes. Typically these children are travelling less than 1.5km each way.

The conceptual framework proposed in Chapter 3 of this thesis is largely appropriate, in that child, family and community variables all do appear to be shaping decision making related to learner mobility, within the context of history, geography and policy. In light of the strong evidence for two patterns of mobility, however, it is appropriate to modify the outcome of the model to indicate that, typically, one of two distinct paths is followed. A child and his or her family are likely either to embark on the high resources mobility path, or the low resource mobility path, and are likely to remain on this path throughout the child's primary schooling, and during the child's transition to secondary schooling. The revised framework is presented in Figure 10.1 below.

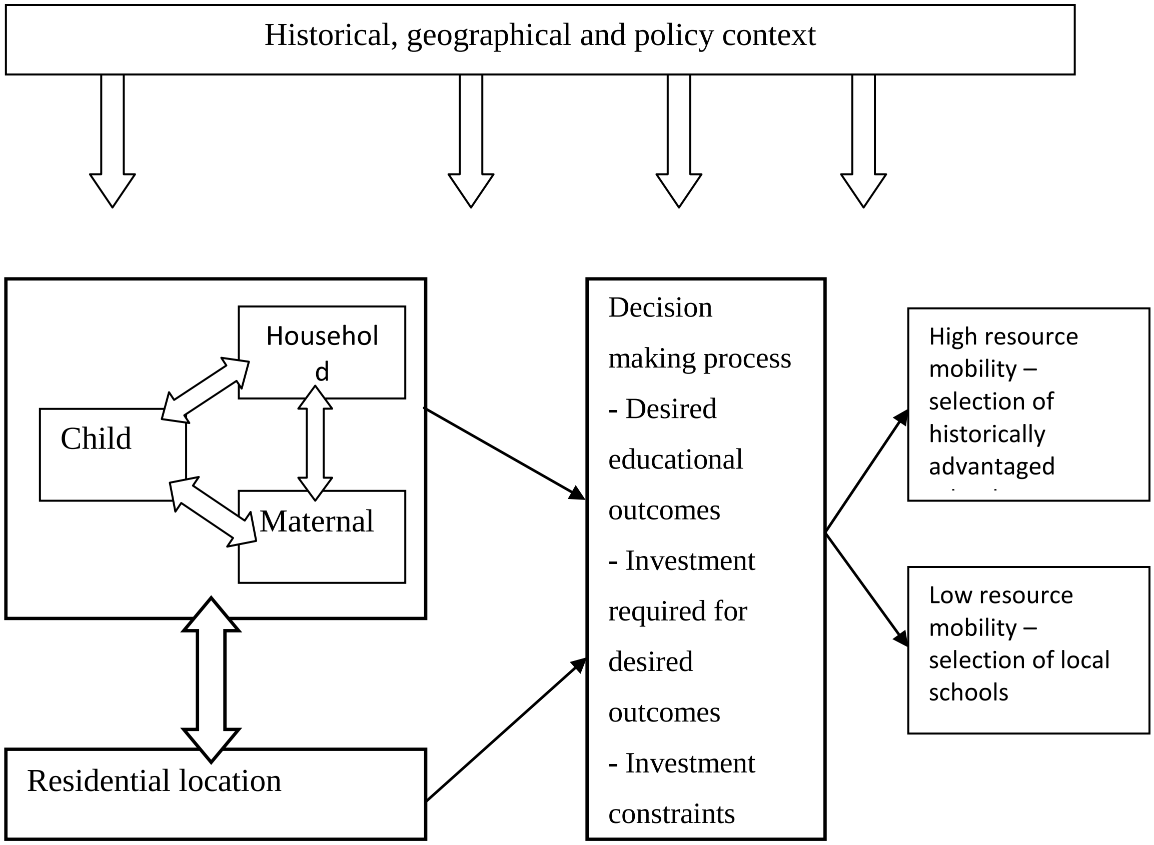

Figure 10.1: Conceptual framework revised on the basis of study findings

10.3 Key contributions

The findings summarized above provide a number of important original contributions to the body of scholarly work on school choice and learner mobility, both in South Africa and internationally. These contributions can be categorized as methodological, empirical and theoretical, and are discussed in these categories below.

10.3.1 Methodological contributions

The thesis has made three innovative methodological contributions to the study of school choice and learner mobility. Firstly, as discussed above, it explored three different approaches to the measurement of educational mobility: straight-line distance from home to school; movement between different areas;

and whether or not a child attends the grade-appropriate school nearest to his or her home. As far as I can determine, combining the use of these three approaches has not previously been reported in either the South African or international literature. By using these three different operationalizations of mobility, it was possible to explore different dimensions of the phenomenon, leading to the observation that there are at least two distinct patterns of mobility in place in urban South Africa.

Secondly, this thesis is innovative with regards to the type of data used. As discussed in Chapter 2, it is the first work of which I am aware to make use of panel data to explore school choice and educational mobility in South Africa. It is also the first work to make use of population level data for this purpose. Finally, it is the first project on school choice in South Africa that I have been able to identify that combines data drawn from a number of different sources to simultaneously explore the relationship of household, community and school-level variables to school choice and educational mobility. It has illustrated that these types of data can be used for these purposes, and, as will be discussed below, can provide novel theoretical and empirical contributions to current knowledge. As such, the study contributes to filling the gap created by the lack of empirical studies into the determinants of school choice, both in South Africa and internationally, that was identified in Chapter 2.

10.3.2 Empirical contributions

As highlighted above, this thesis provides, for the first time, detailed quantitative data on learner mobility in contemporary urban South Africa, obtained at the population level. The study findings provide clear evidence of how widespread learner mobility is, and furthermore that mobility is not limited strictly to the most advantaged children, as was anticipated. Instead, it suggests that two patterns of school choice and learner mobility are fairly widespread: firstly, school choice requiring significant travel to historically advantaged schools, and by extension the investment of substantial economic and other resources; and secondly, school choice at a more local level, which is less constrained by access to financial and social resources.

This has important implications for dominant narratives about the lives of urban, working class children in South Africa. These children have typically been portrayed as disadvantaged (which they are), and even where the literature focuses on resilience, it is resilience in the context of hardship and disadvantage (Barbarin and Richter 2001). This narrative tends to portray children and families as passive in the face of difficult circumstances, and largely devoid of choice with regards to services. The data presented here, however, suggests that for a fairly substantial proportion of these children – perhaps 30% - engagement in social mobility, particularly through education, is evident. Although their home lives are located in a context of deprivation, they attend school in more advantaged areas, socializing with more advantaged children, and typically receiving a better education than they would in the school nearest to their home. This thesis has also presented clear evidence that the patterns of school choice and mobility identified are highly path-dependent. Children who begin their primary schooling on a socially mobile path typically continue to attend advantaged schools. By contrast, those who begin their schooling close to home are also likely to remain at these more local schools at least until the end of the primary phase.

Additional original empirical contributions include the finding that mobile children typically attend a school more advantaged than the one nearest to their home, and that certain groups of children (particularly black children, from relatively well-off households and with more educated mothers and living in relatively disadvantaged areas) are more likely to engage in educational mobility than others. Finally, the finding that mobility behaviour is fairly stable and consistent over time, even following the transition to high school, is also novel, and largely unanticipated. This counters the widespread assumption that mobility increases as children age and become more independent, and particularly once they transition to high school, and provides further evidence that mobility behaviour is fairly path dependent.

10.3.3 Theoretical contributions

Finally, at a more theoretical level, the thesis presents a preliminary theoretical model detailing potential determinants of school choice and mobility. This model raises some questions about the traditional market orientation of school choice literature, both in South Africa and internationally. Certainly, in the sample explored here, a large proportion of the children engaged in school choice appeared to be doing so in ways that did not demand a very high level of economic investment. That is, many children are travelling within a constrained radius, making choices between a number of public schools with much the same fee structures and associated costs.

While wealth appears to shape mobility, it is not as centrally important as might have been expected. School choice work in South Africa has tended to focus very heavily on those children able to access particularly advantaged schools, which typically requires the investment of substantial financial and social resources (Paterson and Kruss 1998; Sekete, Shilubane et al. 2001; Fiske and Ladd 2004). The current thesis, however, suggests that a great number of children are also engaged in another, less costly, form of school choice. As this form of choice still has implications for the opportunities available to children, is seems important that it receives closer attention in the future. At the international level, work exploring school choice in developing countries has also tended to focus very heavily on access to privileged schools, and paid less attention to the ways in which children and families seek to maximize their educational opportunities even in a context of very limited resources (Carnoy and McEwan 2003; Tsang 2003; Elacqua, Schneider et al. 2006).). Again, this thesis suggests that looking at school choices made by children facing resource constraints may prove very valuable.

It also raises questions about the series of hypotheses, common in both developing and developed country contexts, about schools and children who are 'left behind' in the context of choice (Bridge and Blackman 1978; Capell 1981; Henig 1994; Witte and Thorn 1996; Levin 1998; Goldhaber 1999; Hoxby 2003; Peterson, Howell et al. 2003). If engagement in school choice is as widespread as this thesis suggests, that we have to widen our understanding of choice and its parameters in order to grasp the degree of agency exercised by parents and children, and how this might be leveraged to create higher demand for quality education.

This feeds into the international debate, summarized in Chapter 2, about the relationship between school choice, particularly in relatively unregulated contexts, and educational segregation and inequality. This thesis suggests that although engagement in choice may be extremely widespread, and relatively unconstrained by economic and social resources, these resources still shape the ways in which choice can be exercised. Different socio-economic groups appear to access different forms of school choice, which in turn are likely to result in differing educational outcomes for their children. The international literature tends to suggest that more educated and advantaged parents are more strongly involved in school choice than less advantaged parents (Carnoy and McEwan 2003; Elacqua, Schneider et al. 2006). The findings presented here suggest that this may not actually be the case. Instead, it appears that less advantaged parents are highly engaged in choice, but simply lack the resources to pursue choice in the ways that are likely to be most beneficial to their children.

10.4 Contextual relevance

10.4.1 Relevance to South Africa

The findings of this thesis suggest a range of implications for individuals, families, communities and society more broadly. These implications have some relevance to policy, and suggest a range of avenues for additional research and investigation. At the level of children and their families, the most important implications relate to child well-being. Engaging in school choice and mobility provides a child with access to educational opportunities they might not otherwise receive. In the South African context, learner mobility is also likely to be a path to social mobility for at least some children. However, these educational and social benefits do come at some cost. Economic costs, related to the cost of school fees and travel, may mean that a child's family has fewer resources available to meet other needs. In the South African context, the safety of a young child travelling substantial distances, typically alone, is a potential risk, and travel time is certainly a cost. An additional risk relates to the fact that for most children travelling substantial distances, their schooling will take place in what is essentially a foreign social context, and often in a language that they are not very familiar with. They will not have neighbourhood friendships and it will be more difficult for their parents to monitor their friendships and activities. In a society with as extensive societal and educational disparities as South Africa, it seems like that over the long term, the benefits of learner mobility for most children would tend to outweigh the costs. In the shorter term, however, the risks and costs for individual children are likely to remain quite high.

The implications of widespread learner mobility in urban South Africa are, however, likely to be quite different at the community and societal levels. Here, it seems likely that while the short-term costs and risks are reduced, the longer term costs may be quite substantial. One group of particular concern is those children who are not able to engage the more resource-intensive forms of educational mobility, for whatever reason, and are obliged to attend schools close to their homes. These children, along with the schools they attend, are likely to be negatively impacted by the tendency of all the more advantaged children in the area to attend schools further afield, as this will further reduce the resources available at the local level. With the out-migration of more advantaged children, and a relatively captive market of disadvantaged children, poorly performing schools may lose the incentive to try to improve. Over the longer term, high levels of mobility are also likely to have harmful implications for community coherence, and may well contribute to growing levels of inequality, both economic and educational, within historically disadvantaged areas. The current distribution of educational opportunities, which motivates the high levels of mobility identified in this thesis, are also hugely inefficient, requiring large investments on the parts of families and children, even while they potentially exacerbate already high societal inequality.

This pattern, in which learner mobility appears likely to be beneficial for individual children, particularly over the longer term, but also very costly at the level of their community and society, poses real challenges for policy makers. How to balance these various costs and benefits is a challenging question, particularly in a society already so deeply challenged by inequality, and a poorly performing educational system. This is further complicated by the very strong incentives faced by individuals to continue engaging in mobility, regardless of the policy environment. The findings presented in this thesis do raise questions about the validity of certain core elements of South African educational policy, such as the concept of our schools as “community schools”, and the notion that a school’s access to resources can be determined by looking at its location, as opposed to the composition of its student body. The complexity of the results presented, however, combined with the preliminary, hypothesis-building nature of the study itself does, however, suggest that much more needs to be known about how school choice operates in South Africa, and how its costs and benefits play out at the individual and societal levels, before making substantial changes to policy. For this reason, this thesis does not provide a comprehensive discussion of policy relevance at this point.

10.4.2 International relevance

The findings of this thesis also have implications for both policy and future research internationally. The thesis has highlighted just how widespread it is possible for school choice to be in a developing country context with limited regulation. South Africa has some distinctive features which limit the generalizability of this study's results, such as the fact that in South Africa school performance is very closely associated with geography, and the fact that South Africa's public schooling system is much larger than those found in most other countries at similar levels of development, and contains a subset of very well performing schools. Nonetheless, South Africa's high level of inequality, limited school choice regulation (or capacity to enforce regulation) and a generally poorly performing public schooling system are all shared with a number of other countries at similar levels of developments. The extent of school choice in South Africa suggests that levels of choice in other similar countries can also be expected to be high. Similar studies in other low and middle income countries would therefore be extremely illuminating.

One question that would be particularly useful to ask is whether school choice in other countries, particularly those where private schooling is more widespread, is more strongly economically driven than choice in South Africa. Additionally, understanding whether less advantaged children in other countries also exercise local level school choice will be critical in determining how best to finance public schools and enhance access to high quality education. Understanding the extent of school choice, the forms which it takes, and its determinants is essential to developing appropriate ways to capture this engagement of children and families with educational systems in ways to enhance the performance of both schools and their students.

If indeed school choice is widespread in other developing countries, this also suggests a need to think about ways of protecting the most vulnerable groups of children from potential choice-related harm. For example, if more advantaged children are enrolling predominantly in private schools, this deprives the public system of resources, and is likely to enhance societal inequality. This raises questions as to whether voucher systems, such as those found in Chile, or other innovative approaches to school funding should be tested more widely in other developing countries.

10.5 Project limitations and future work

10.5.1 Sample composition

One of the major limitations of the thesis is the relatively constrained sample that has been used for analysis. Although, as discussed, this was unavoidable for practical reasons, an ideal next step is to broaden the study sample so that it also includes children who do change their residential addresses during the period under consideration. This more representative sample would ensure that findings can be more broadly generalized. In addition, this sample would also allow for the exploration of potential interactions between residential and educational mobility.

10.5.2 Study end point and longitudinal analysis

A second limitation is that only a proportion of the study sample members progressed to high school by 2003. As those children who had progressed to high school by 2003 differed systematically from those who had not, this introduced challenges around using the data to understand whether mobility during primary and high school differed. Although preliminary analysis suggested a great degree of path-dependency and consistency over time, it also appeared to be the case that secondary school status had implications on mobility behaviour, at the very least due to the different geographical distributions of primary and secondary schools. Expanding the sample longitudinally so that all study sample members reach secondary schooling would facilitate analyses exploring the extent to which mobility changes at the secondary school level, and why this is the case. It would also facilitate the use of more complex longitudinal analyses, which would provide clearer and more valuable data on the nature of changes in mobility behaviour over time.

The construction of a genuine longitudinal dataset, in which home and schooling data were available for each year during a child's schooling would also enhance the value of further analysis, by allowing for the use of more advanced analytical techniques. This would also provide clearer evidence around school change during primary and secondary schooling, and potential motivating factors for this. Introducing data for a more contemporary cohort of children, although likely to be extremely difficult and costly, would also allow exploration of whether patterns of mobility have changed over time, since the beginning of the post-Apartheid period.

10.5.3 Methodological approach

Finally, as with any methodological approach, the use of quantitative secondary analysis in this project imposed a number of limitations. Perhaps most importantly, it makes it extremely difficult to provide answers to questions around the individual decision-making processes underlying the decision to engage in learner mobility. However, it is valuable in highlighting correlates of mobility, and by extension generating hypotheses about the types of decisions which individuals may be making. These data-driven hypotheses can then be tested by subsequent research using different methodological approaches. Population based quantitative secondary analysis is also fairly limited in the extent to which it can describe the implications and outcomes of learner mobility. School and community level outcomes, in particular, can only really be tested with data collected at those levels. To understand individual level outcomes, a far more longitudinal approach, along with far more detailed and specialised data, particularly regarding outcomes, than that used here would be necessary.

10.5.4 Future work

There are a number of ways in which the work presented in this thesis could be usefully extended. Firstly, as alluded to above, broadening the study sample will produce more generalizable findings, and allow for an exploration of the interactions between learner mobility and residential mobility. Secondly, extending the dataset longitudinally will allow for the examination of learner mobility in the high school period, as well as for the application of more sophisticated tools for longitudinal analysis[11]. Thirdly, study methodology could be developed further in a few directions, to strengthen study findings. For example, it would be possible to generate a measure of 'non-essential' travel to school, by looking at the difference between the distance to a child's nearest school, and the school a child attends. This might provide a more accurate measure of travel related to choice. Additional approaches to dealing with the highly non-normal distribution of the travel data could also be usefully explored, and in particular, more systematic approaches to dealing with outliers. Developing measures of practical or substantive significance to accompany the presentation of measures of statistical significance would also contribute usefully to the interpretation of study findings.[12]

This study would also greatly benefit from the introduction of a qualitative component. This could be used to test the hypotheses that this thesis has generated. For example, do motivations for mobility (and constraints in engaging in mobility), as experienced by children and their families, tie in with

[11] Funding for this extension of the project has been obtained through an ESRC Pathfinders grant, starting in June 2011.
[12] I am indebted to my examiners for suggesting these methodological developments that would contribute to strengthening future work.

the results presented in the previous chapters?[13] Do children and their families identify with the notion that there are two distinct patterns of school choice, and by extension mobility, in play in contemporary urban South Africa? Qualitative work would also allow for further unpacking of the decision making process, providing insight into ways in which families select the schools at which they pursue enrolment for their children. Finally, it would provide a way to integrate the current work more strongly with the international literature, by examining how and why school choice in the South African context differs from experiences in other countries. A final enhancement to work using the Bt20 data would be to develop more sophisticated approaches to the measurement of variables used in the study. In particular, alternative measures of school resource levels, community affluence and coherence, and household SES could all usefully be explored.

An important set of questions which this thesis has not adequately explored are those relating to the supply-side geography of schools in South Africa. A clearer understanding of how schools with different attributes are distributed across space would enrich our understanding of how and why children travel to go to school. Incorporating supply side variables into the models of mobility presented in this thesis would strengthen them considerably. For example, poor supply of schools in local areas may be one of the reasons that mobility is fairly high amongst the least advantaged sample members. An exploration of the correlation between school density and population density in different areas would also provide some indication as to whether supply side issues are behind much of the mobility documented in this thesis. Additionally, mapping the distribution of schools by the languages that they operate in will answer questions about whether high levels of local level mobility might be explained in part by language of schooling. These types of analyses can be accomplished,

[13] Given the difficulty of accurately measuring school characteristics, answering this question is particularly important, as it would address the concern that the high levels of mobility documented in this study relate more to a highly stochastic schooling environment. Many thanks to my examiners for highlighting this concern to me.

at least in part, with the data already compiled for this thesis, and would provide some valuable extensions to study findings[14].

Moving beyond the current study, similar work applied to populations in other parts of South Africa would also be very useful in establishing the extent to which the patterns identified here are prevalent throughout the country. A similar study in a rural area, where school choice would be expected to be far more limited due to the lower density of available schools would be particularly valuable. Similar analyses applied to other low and middle income countries would also help to shed light on just how widespread school choice is in other contexts, and particularly the extent to which less advantaged children are able to participate in it.

This study also suggests a number of additional, related questions that might usefully be pursued. One question relates to the roles of schools in shaping school choice outcomes. For example, to what extent are children not able to enrol in the school they select, for example for reasons of overcrowding, lack of social capital or knowhow, or even overt discrimination? A second question relates to the implications of school choice for academic outcomes, both for learners, and for entire schools. This issue lies at the core of much of the international debate around school choice, and is a critical question which this thesis has not been able to address. This issue is likely to be particularly complex in the South African context, where the potential costs of mobility (ranging from economic, to travel time, to learning in a language not used at home) are extremely high, but variations in public school quality mean that potential benefits are also substantial. Enhancing the dataset by the inclusion of data on academic outcomes at the individual level would allow for an exploration of the relationship between mobility, school choice, and academic outcomes. A final set of additional questions relate to the implications of

[14] I am indebted to my examiners for highlighting to me the importance that work of this type would have, and suggesting ways in which it might usefully be approached in the near future..

learner mobility and school choice for social mobility. To what extent does the ability to access education at a historically advantaged school determine the opportunities available to a child as he or she moves through school, and then into higher education or the workplace?

10.6 Conclusion

This final chapter has provided a brief overview and summary of the key findings presented in this thesis. These highlight the original contributions that this thesis makes to the scholarly literature on learner mobility and school choice. Methodological contributions have included new approaches to the identification and measurement of learner mobility, as well as the use of population-based panel data, combined with data from other sources, to study the phenomenon. At the empirical level, the study has contributed data on the extent, correlates and determinates of learner mobility. Finally, at the theoretical level, the study has contributed a conceptual framework to support other work on the topic, and the insight that in contemporary urban South Africa, there are at least two forms of learner mobility in play. The study also feeds into broader international debates about the implications of school choice to educational segregation and inequality. The chapter concludes by mapping out a spectrum of further research possibilities.

Appendix A: Alternative data sources considered for the thesis

A.1 Cape Area Panel Study (CAPS)

CAPS data collection began in 2002, and has been focused largely on a sample of almost 5000 young adults who were then aged between 14 and 22. Most of these participants have been interviewed four times since 2002, and additional interviews have been conducted with their households, as well as some additional households and older individuals. As indicated by the study's name, the sample was drawn from the greater Cape Town area. Participants have been asked a wide range of questions over the four waves of the study for which data is available to date. These questions include school enrolment, both current and historical, as well as reasons for enrolment at particular schools, and for changes in school enrolment over time. The study's second wave also includes a module in which perceptions of school quality for a range of neighbourhood schools are explored for each child (Lam, Ardington et al. 2008; Lam, Ardington et al. 2008).

While the nature of the schooling data available in the study made it an extremely strong candidate for use in this dissertation, the decision against using it was finally made largely on the basis of its focus on the Cape Town area, which is well known for having a far stronger educational system than that found anywhere else in South Africa (Fiske and Ladd 2004)[15]. Additionally, the Cape Town area differs substantially from the rest of the country in terms of its population and its socio-economic conditions. For this reason, it was felt that while the data from the study would certainly enable a clear understanding of learner mobility in the Western Cape, this understanding would be unlikely to travel well across the rest of South Africa.

[15] Although the 2010 Matric examination results, in which Gauteng province outperformed the Western Cape, may signal a shift in this pattern.

An additional concern related to the study's focus on data collection on youth older than 14, as this would make it difficult to explore learner mobility during primary schooling, and during the transition to secondary schooling.

A.2 Kwa-Zulu Natal Income Dynamics Survey (KIDS)

The KIDS study is a longitudinal dataset, with a focus primarily on poverty and inequality in the KwaZulu Natal (KZN) region of South Africa. It developed out of a cross-sectional study, the 1993 Project for Statistics on Living Standard and Development (PSLSD), and the PSLSD data collected in KZN in 1993, covering 1558 households, forms the first wave of KIDS. Households interviewed for PSLSD were followed up, and re-interviewed in 1998, forming the second wave of KIDS data. A 3rd round of interviews was conducted in 2004, with 865 households (including households previously interviewed, and next-generation households which had split off from households previously interviewed). During each interview wave, a detailed household roster was completed, which in 1998 and 2004, along with socio-economic information included current school enrolment of all school-aged children. For each household, address information and GIS coordinates were also collected (May, Carter et al. 1999; May, Aguero et al. 2007).

While the KIDS dataset did include all the information essential for this dissertation, there were a number of ways in which the sample was not ideal. Firstly, while the study itself is longitudinal in nature, this was focused at the household level, rather than at the level of particular children or other individuals. Therefore, while longitudinal data is available on some children, this is not true of all children across the sample, and depends heavily on their mobility and relationship to the household head. This would have made identifying a sample appropriate to longitudinal analysis very challenging. The relatively high levels of attrition experienced by KIDS (around 38% over 11 years), while unsurprising given the long intervals between the waves of data collection and the longitudinal nature of the study, also raise concerns about how representative the remaining sample is, particularly when considering that the relevant data would only be available for a sub-sample of children, unlikely to be randomly distributed amongst households.

Nonetheless, I had initially hoped to make some use of the KIDS data, to provide some insight into the variations in learner mobility between urban and rural areas, as it appeared to be the best potential source of data on the enrolment of rural learners in South Africa. However, preliminary communication with the KIDS research team revealed that a substantial amount of preparatory work on the KIDS school enrolment data would be required for it to become useable for this dissertation, which, in light of time and resource constraints, resulted in a final decision not to use this data for the current project.

A.3 National Income Dynamics Survey (NIDS)

NIDS was developed to answer very similar types of questions to those addressed in KIDS, but at a national level. Areas of primary focus included household wealth creation, demographic dynamics, social heritage, and access to services and cash transfers. During its first wave of data collection NIDS collected data on 7305 households, consisting of 28255 individuals, distributed across all nine provinces of South Africa (Leibbrandt, Woolard et al. 2009).

While much of the essential data on residential addresses and schooling was collected, the same limitations detailed in the discussion of KIDS, relating to the focus on households rather than individual children or youth, continue to hold. In addition, ethical protections on the NIDS data appear to prohibit the release of residential address data to researchers altogether. However, the main reason that it was not possible to seriously consider making use of NIDS data for this dissertation was that formal data collection only began in 2008, and data release only occurred from mid 2009. Waiting for this data would have therefore substantially delayed this project. In addition, even the data available in 2009 would only have been cross-sectional, preventing any longitudinal analysis.

A.4 Agincourt Health and Demographic Surveillance System (Agincourt)

As suggested by the name, the Agincourt data is concerned primarily with health and demographics. Data is collected on the full population of around 82000 individuals living in the Agincourt sub-district of Bushbuckridge, South Africa. A baseline census of the area was conducted in 1992 (Kahn, Tollman et al. 2007). This data has subsequently been updated 12 times, most recently in 2008. Unfortunately, this dataset did not prove to be suitable for this project, as the education data collected was extremely limited. School enrolment data was collected at 4 points during the study, but was not in a form that made it feasible to use for this study, particularly given the large sample size.

A.5 Africa Centre for Health and Population Studies

The Africa Centre is located in rural KwaZulu Natal, South Africa, and serves as the base for a number of research project, including the longitudinal collection of demographic and health data through the Africa Centre Demographic Information System (ACDIS). ACDIS data collection began in 2000, and is ongoing, with data collection each year. ACIDS covers about 90 000 individuals in approximately 11 000 households, including those household members who are not resident in the area. Unfortunately, the educational data collected in this study is limited to attainment, and details regarding the school enrolment of individuals are not available (Herbst, Newell et al. 2010). For this reason it was not possible to pursue the use of this data set for this dissertation.

Appendix B: Letter confirming approval of ethics clearance for thesis, received from the School of Education, University of the Witwatersrand

Wits School of Education

27 St Andrews Road, Parktown, Johannesburg, 2193 • Private Bag 3, Wits 2050, South Africa
Tel: +27 11 717-3007 • Fax: +27 11 717-3009 • E-mail: enquiries@educ.wits.ac.za • Website: www.wits.ac.za

STUDENT NUMBER: 386849

Protocol: 2009ECE01

7-May-09

Ms. Julia de Kadt
86 Glenroy Road
MANOR GARDENS
DURBAN
4001

Dear Ms. de Kadt

Application for Ethics Clearance: Doctor of Philosophy

I have pleasure of advising you that the Ethics Committee in Education of the Faculty of Humanities, acting on behalf of the Senate has agreed to approve your application for ethics clearance submitted for your proposal entitled:

Learner Migration in South Africa: Dimensions and implications for Educational Practice and Policy.

Recommendation:

Ethics clearance is granted

Yours sincerely

Matsie Mabeta
Wits School of Education

Cc Supervisor: Prof. B Fleisch (via email)

Appendix C: Relationships within the study sample between variables hypothesized to act as determinants of learner mobility

C.1 Race and other variables

Given the small numbers of white and Indian children in the study sample, all further discussion of race is limited to the black and coloured groups only. There is no evidence for a relationship between race and gender. Similarly, there is no evidence for a difference between age at first enrolment for black and Coloured children, with slightly over half the members of each group enrolling on time. Coloured children are slightly more likely to have reached secondary school by 2003 than black children ($\chi^2_{(1)}$= 3.4241, Pr=0.064), and are somewhat less likely to have repeated any grades ($\chi^2_{(1)}$= 4.0985, Pr=0.043).

Race is also related to the various household and family variables considered. Maternal education is slightly higher for black children than coloured children (Wilcoxon rank-sum test, Pr=0.0781). Black mothers were also less likely to be married ($\chi^2_{(1)}$= 44.4026, Pr=0.000) than coloured mothers. Wilcoxon rank-sum tests indicate that ethnicity and household SES, for both 1997 and 2003, are significantly related (Pr=0.0000), with coloured children having lower household SES than black children. By contrast, however, black children live in SAL, SP and MP areas with higher poverty levels than coloured children (Wilcoxon rank-sum tests, Pr=0.0000).

C.2 Gender and other variables

Gender of the child is significantly related to age at first enrolment ($\chi^2_{(1)}$= 5.1999, Pr=0.023) and grade repetition ($\chi^2_{(1)}$=36.0374, Pr=0.000), with girls more likely to start early or on time, and only about half as likely as boys to repeat a grade. Given that girls tend to start their schooling earlier, and are less likely to repeat grades, it comes as no surprise that they are also significantly more likely to have reached high school by 2003 ($\chi^2_{(1)}$=23.3055, Pr=0.000). There is no evidence for any relationship between child gender and maternal education, household SES in 1997 or 2003, or the poverty level of the area in which the child lives. Although a weakly significant relationship between child gender and maternal marital status is found ($\chi^2_{(1)}$=3.5817, Pr=0.058), this seems likely to be spurious.

C.3 Age at first enrolment and other variables

Predictably, there is a strong relationship between age at first enrolment and phase of schooling in 2003 ($\chi^2_{(1)}$=503.7355, Pr=0.000), with children who started school late being unlikely to have reached high school at this point. There is, however, no evidence of a relationship between age at first enrolment and grade repetition. There is also no evidence that age at first enrolment is related to maternal education, maternal marital status, or household SES in 1997. By contrast, there is a significant relationship between age at first enrolment and household SES in 2003 ($\chi^2_{(4)}$=15.8754 Pr = 0.003), with more advantaged children tending to enrol later. Finally, there are significant relationships between age at first enrolment and area poverty at the SAL ($\chi^2_{(4)}$=8.3949, Pr=0.078) and SP ($\chi^2_{(4)}$=9.8775, Pr=0.043) levels, with children living in wealthier areas being more likely to start school late. There was no relationship between MP level poverty and age at first enrolment.

C.4 Phase of education in 2003 and other variables

Predictably, children who repeated a grade between 1997 and 2003 are significantly less likely to have reached high school by 2003 ($\chi^2_{(1)}$= 313.9621, Pr=0.000). Children of more highly educated mothers are also more likely to have reached high school by 2003 ($\chi^2_{(1)}$=32.9057, Pr=0.000). There is no evidence of a relationship between maternal marital status and the child's phase of education by 2003. There is a weakly significant relationship between phase of education in 2003 and household SES for 1997 ($\chi^2_{(4)}$=8.9023, Pr=0.064), but when 2003 household SES is used, this effect disappears. There is no association between phase of education in 2003 and the poverty of the area, SAL, SP or MP, in which a child lives.

C.5 Grade repetition and other variables

Children with less highly educated mothers are more likely to have repeated a grade ($\chi^2_{(4)}$=45.6920, Pr=0.000). Children whose mothers were unmarried at their birth are also more likely to have repeated a grade ($\chi^2_{(1)}$=3.7110, Pr = 0.054). Repetition is also significantly related to household SES, both in 1997 ($\chi^2_{(4)}$=50.9931, Pr=0.000) and 2003 ($\chi^2_{(4)}$=45.7106, Pr=0.000), with children in more affluent households being less likely to have repeated a grade. Finally, there is also a positive relationship between repetition and the poverty level of the area in which a child lives, whether this is calculated at the SAL ($\chi^2_{(4)}$=33.9833, Pr=0.000), SP ($\chi^2_{(4)}$=28.7967, Pr=0.000) or MP ($\chi^2_{(2)}$=17.6835, Pr=0.000) level. However, these relationships are not strictly linear throughout all poverty quintiles, and are most marked at the extremes.

C.6 Maternal education and other variables

A positive, significant, but non-linear relationship exists between maternal education and maternal marital status, both measured at the time of the child's birth ($\chi^2_{(4)}$= 16.3190, Pr = 0.003). The proportion of married mothers is highest amongst mother with post-school education, followed by those with primary schooling or less, while rates are lowest amongst mothers with partial or complete secondary schooling. There is also a significant, positive relationship between maternal education and household SES, both for 1997 (Kruskal-Wallis test, Pr=0.0001) and for 2003 (Kruskal-Wallis test, Pr=0.0001). Maternal education is also significantly related to area poverty, measured at the SAL ($\chi^2_{(16)}$=74.1141, Pr=0.000), SP ($\chi^2_{(16)}$=57.4465, Pr=0.000), and MP ($\chi^2_{(8)}$=21.5601, Pr=0.006) levels, with more educated mothers tending to live in more advantaged areas.

C.7 Maternal marital status and other variables

Maternal marital status is significantly related to household SES in both 1997 (χ2(4)=130.5849, Pr=0.000) and 2003 (χ2(4)=54.9065, Pr=0.000), with mothers in more affluent households more likely to be married. Similarly, mothers living in more advantaged areas are also significantly more likely to be married, whether area poverty is measured at the SAL ($\chi^2_{(4)}$=121.2915, Pr=0.000), SP ($\chi^2_{(4)}$=134.1958, Pr=0.000) or MP ($\chi^2_{(2)}$=55.5876, Pr=0.000) level.

C.8 Household SES and residential area poverty levels

Household SES in 1997 and 2003 are strongly related, with a correlation of 0.7655 (Pr=0.000). Household SES, measured in both 1997 and 2003, is also strongly and inversely related to the poverty area in which the child lives (see Table A3.1 below), with household SES tending to be higher in the areas with the lowest poverty levels.

Relationship between household SES and residential area poverty	SAL poverty level	SP poverty level	MP poverty level
Household SES, 1997	$\chi^2_{(16)}$=381.0466 Pr=0.000	$\chi^2_{(16)}$=298.4818 Pr=0.000	$\chi^2_{(8)}$=164.0330 Pr=0.000
Household SES, 2003	$\chi^2_{(16)}$=229.4532 Pr=0.000	$\chi^2_{(16)}$=230.2197 Pr=0.000	$\chi^2_{(8)}$=68.0648 Pr=0.000

Table A.1: Significance of relationships between household SES and residential area poverty levels

C.9 Conclusion

All relationships documented here appear to operate in the expected direction. There is evidence that a number of the variables considered are strongly related to each other, as expected. This implies that during the modelling component of the thesis, attention must be paid to avoiding multicollinearity.

References

Ahmed, R. and Y. Sayed (2009). "Promoting access and enhancing education opportunities? The case of ‘no-fees schools’ in South Africa." Compare: A Journal of Comparative and International Education **39**(2): 203 - 218.

Alderman, H. and E. M. King (1998). "Gender differences in parental investment in education." Structural Change and Economic Dynamics **9**(4): 453-468.

Andrabi, T., J. Das, et al. (2009). What Did You Do All Day? Maternal Education and Child Outcomes.

Astin, A. W. (1992). "Educational "Choice": Its Appeal May be Illusory." Sociology of Education **65**(4): 255-260.

Bagley, C. (1996). "Black and White Unite or Flight? The Racialised Dimension of Schooling and Parental Choice." British Educational Research Journal **22**(5): 569-580.

Barbarin, O. A. and L. M. Richter (2001). Mandela's children: Growing up in post-Apartheid South Africa. New York, Routledge.

Bifulco, R. and H. Ladd (2006). School choice, racial segregation and test-score gaps: evidence from North Carolina's charter school programme.. Annual Meeting of Allied Social Science Associations, Boston, MA.

Bosetti, L. (2004). "Determinants of school choice: understanding how parents choose elementary schools in Alberta." Journal of Education Policy **19**(4): 387 - 405.

Bradley, S. and J. Taylor (2002). "The effect of the quasi-market on the efficiency-equity trade-off in the secondary school sector." Bulletin of Economic Research **54**(5): 295-314.

Bray, R., I. Gooskens, et al. (2010). Growing up in the new South Africa: Childhood and adolescence in post-apartheid Cape Town. Cape Town, South Africa, HSRC Press.

Bridge, R. and J. Blackman (1978). A study of alternatives in American education: Vol IV: Family choice in schooling. Santa Monica, CA RAND Corporation.

Bryman, A. (2004). Social research methods. Oxford, United Kingdom, Oxford University Press.

Capell, F. (1981). A study of alternatives in American education: Vol VI: Student outcomes at Alum Rock, 1974-1976. Santa Monica, CA RAND Corporation.

Carnoy, M. (1999). Globalizaton and educational reform: what planners need to know. Paris, France UNESCO.

Carnoy, M. and P. McEwan (2003). Does privatization improve education? The case of Chile's national voucher plan. Choosing choice: School choice in international perspective. D. Plank and G. Sykes. New York, NY, Teachers College Press, Columbia University**:** 24-44.

Centre for Development and Enterprise (2010). Hidden assets: South Africa's low-fee private schools. Johannesburg, South Africa.

Chen, X., P. Ender, et al. (2011). Stata web books: Logistic regression with Stata, UCLA Academic Technology Services.

Chisholm, L. (2004). Changing Class: Education and Social Change in post-Apartheid South Africa. Cape Town HSRC Press.

Chisholm, L. (2005). The state of South Africa's schools. State of the nation: South Africa 2004-2005. J. Daniel, R. Southall and J. Lutchman. Cape Town, HSRC Pres**:** 201-226.

Coleman, J. S. (1992). "Some Points on Choice in Education." Sociology of Education **65**(4): 260-262.

Cosser, M. and J. du Toit (2002). From school to higher education? Factors affecting the choices of Grade 12 learners. Cape Town, HSRC Publishers

Daun, H. (2003). Market forces and decentralization in Sweden: Impetus for school development or threat to comprehensiveness and equity? Choosing choice: School choice in international perspective. D. Plank and G. Sykes. New York, NY, Teachers College Press, Columbia University**:** 92-111.

De Jong, G., F. (2000). "Expectations, Gender, and Norms in Migration Decision-Making." Population Studies **54**(3): 307-319.

Denessen, E., G. Driessena, et al. (2005). "Segregation by choice? A study of group-specific reasons for school choice." Journal of Education Policy **20** (3): 347-368.

Department of Education (2000). Brochure for the 2000 School Register of Needs Report.

du Toit, J. (2003). Independent Schooling. Human Resources Development Review 2003: Education, employment and skills in South Africa. A. Kraak and H. Perold. Cape Town, HSRC Press**:** 380-395.

Elacqua, G. (2006). Enrollment practices in response to vouchers: evidence from Chile.

Elacqua, G., M. Schneider, et al. (2006). "School choice in Chile: Is it class or the classroom?" Journal of Policy Analysis and Management **25** (3): 577-601.

Fedderke, J. W., R. de Kadt, et al. (2000). "Uneducating South Africa: The failure to address the 1910-1993 legacy." International Review of Education **46**(3): 257-281.

Filer, R. and D. Munich (2003). Public support for private schools in Post-Communist Central Europe. Choosing choice: School choice in international perspective. D. Plank and G. Sykes. New York, NY, Teachers Colllege Press, Columbia University**:** 196-222.

Filmer, D. and L. Pritchett (2001). "Estimating wealth effects without expenditure data—or tears: an application to educational enrollments in states of India." Demography **38**(1): 115-132.

Fiske, E. and H. Ladd (2000). When schools compete: a cautionary tale. Washington DC, Brookings Institution Press

Fiske, E. and H. Ladd (2004). Elusive equity : education reform in post-apartheid South Africa. Washington D.C. , Brookings Institute.

Fiske, E. B. and H. F. Ladd (2005). Racial equality in education: how far has South Africa come? Terry Sanford Institute of Public Policy Working Paper series, Duke University.

Fleisch, B. (2008). Primary Education in Crisis: Why South African Schoolchildren Underachieve. Cape Town, South Africa, Juta.

Fleisch, B. and J. Schindler (2008). Gender repetition: School access, transitions and equity in the 'Birth-to-Twenty' cohort panel study in urban South Africa.

Fleisch, B. and J. Schindler (2009). Patterns and prevalence of school access, transitions and equity in South Africa: Secondary analyses of BT20 large-scale data sources. CREATE Pathways to Access Research Monographs. Brighton, United Kingdom, Consortium for Research on Educational Access, Transitions and Equity.

Garcia, D. R. (2008). "Academic and Racial Segregation in Charter Schools: Do Parents Sort Students Into Specialized Charter Schools?" Education and Urban Society **40**(5): 590-612.

Ginsburg, C., S. Norris, et al. (2009). "Patterns of residential mobility amongst children in greater Johannesburg-Soweto, South Africa: observations from the Birth to Twenty cohort." Urban Forum **20** 397-413.

Ginsburg, C., L. M. Richter, et al. (2010). "An analysis of associations between residential and school mobility and educational outcomes in South African urban children: The Birth to Twenty cohort." International Journal of Educational Development **In Press**.

Glazerman, S. M. (1998). School Quality and Social Stratification: The Determinants and Consequences of Parental School Choice. Annual Meeting of the American Educational Research Association. San Diego, CA,**:** April 13-17, 1998.

Godwin, K., F. Kemerer, et al. (1998). "Liberal Equity in Education: A Comparison of Choice Options." Social Science Quarterly (University of Texas Press) **79**(3): 502-522.

Godwin, K., S. Leland, et al. (2006). "Sinking Swann : public school choice and the resegregation of Charlotte's public schools." Review of Policy Research **23** (5): 983-997.

Goldhaber, D. (2000). "School Choice: Do We Know Enough?" Educational Researcher **29**(8): 21-22.

Goldhaber, D. and E. Eide (2002). "What do we know (and need to know) about the impact of school choice reforms on disadvantaged students?" Harvard Educational Review **72**(2): 157-176.

Goldhaber, D. D. (1999). "School Choice: An Examination of the Empirical Evidence on Achievement, Parental Decision Making, and Equity." Educational Researcher **28**(9): 16-25.

Goldring, E. B. and C. S. Hausman (1999). "Reasons for parental choice of urban schools." Journal of Education Policy **14**: 469-490.

Gorard, S. and J. Fitz (2000). "Investigating the determinants of segregation between schools." Research Papers in Education **15**(2): 115-132.

Gorard, S. and J. Fitz (2006). "What counts as evidence in the school choice debate?" British Educational Research Journal **32**(6): 797-816.

Gorard, S., J. Fitz, et al. (2001). "School Choice Impacts: What Do We Know?" Educational Researcher **30**(7): 18-23.

Greenberg, J. P. (2011). "The impact of maternal education on children's enrollment in early childhood education and care." Children and Youth Services Review **In Press, Corrected Proof**.

Greene, J., T. Loveless, et al. (2010). Expanding choice in elementary and secondary education: A report on rethinking the federal role in education, Brown Center on Education Policy at Brookings.

Hanson, K. and L. Litten (1982). Mapping the road to academia: A review of research on women, men, and the college selection process. The undergraduate woman: Issues in Educational Equity. P. Perun. Lexington, Lexington Books.

Henig, J. (1994). Rethinking school choice: Limits of the market metaphor. Princeton, NJ Princeton University Press.

Herbst, K., M. L. Newell, et al. (2010). "Study summary: ACDIS."

Hofmeyr, J. and S. Lee (2004). The new face of private schooling. Changing Class. L. Chisholm. Cape Town, HSRC Press**:** 143-174.

Holmes, J. (2002). "Buying homes, buying schools: school choice and the social construction of school quality." Harvard Educational Review **72** (2): 177-205.

Howe, L. D., J. R. Hargreaves, et al. (2008). "Issues in the construction of wealth indices for the measurement of socio-economic position in low-income countries." Emerging Themes in Epidemiology **5**(3).

Hoxby, C. M. (1998). "What Do America's 'Traditional' Forms of School Choice Teach Us About School Choice Reforms?" Economic Policy Review, Vol.4, No.1, March 1998.

Hoxby, C. M. (2002). "School Choice and School Productivity (or Could School Choice be a Tide that Lifts All Boats?)." SSRN eLibrary.

Hoxby, C. M. (2003). The economics of school choice. Chicago, The University of Chicago Press.

Hoxby, C. M. (2003). Introduction. The economics of school choice. C. M. Hoxby. Chicago, The University of Chicago Press**:** 1-22.

Hunter, M. (2010). Racial desegregation and schooling in South Africa: contested geographies of class formation.

Johnson, D. (2007). "Building citizenship in fragmented societies: The challenges of deracialising and integrating schools in post-Apartheid South Africa." International Journal of Educational Development **27**(3): 306-317.

Kahn, K., S. M. Tollman, et al. (2007). "Research into health, population and social transitions in rural South Africa: Data and methods of the Agincourt Health and Demographic Surveillance System." Scandinavian Journal of Public Health **35**(3 supp 69): 8-20.

Kanjee, A. (2007). Improving learner achievement in schools: applications of national assessments in South Africa. State of the Nation: South Africa 2007. S. Buhlungu, J. Daniel, R. Southall and J. Lutchman. Cape Town, HSRC Press**:** 470-499.

Kanjee, A. and A. Chudgar (2009). Accuracy of the poverty quintile system for classifying South African schools. 2nd Monitoring and Evaluation Colloquim, Gauteng Department of Education. Sandton, Johannesburg.

Kanjee, A. and A. Chudgar (2009). "School money: funding the flaws." HSRC Review **7**(4).

Karlsson, J. (2007). Mobility and equity: school transport, cost of schooling and class formation in post-apartheid South Africa. Dilemmas of implementing education reforms: Explorations from South Africa and Sweden. C. Odora Hoppers, U. P. Lundgren, J. Pampallis, E. Motala and E. Nihlfors. Uppsala, Sweden, STEP, Uppsala Universitet.

Karsten, S., C. Felix, et al. (2006). "Choosing Segregation or Integration?: The Extent and Effects of Ethnic Segregation in Dutch Cities." Education and Urban Society **38**(2): 228-247.

Klasen, S. (2002). "Low Schooling for Girls, Slower Growth for All? Cross-Country Evidence on the Effect of Gender Inequality in Education on Economic Development." The World Bank Economic Review **16**(3): 345-373.

Kolenikov, S. and G. Angeles (2008). Socioeconomic status measurement with discrete proxy variables: is principal component analysis a reliable answer?

Kristen, C. (2005). School choice and ethnic school segregation: Primary school selection in Germany. Munster, Waxmann Verlag GmbH.

Ladd, H. (2003). Introduction. Choosing choice: School choice in international perspective. D. Plank and G. Sykes. New York, NY, Teachers College Press: Columbia University**:** 1-23.

Lam, D., C. Ardington, et al. (2008). The Cape Area Panel Study: a very short introduction to the integrated waves 1-2-3-4 data, The University of Cape Town.

Lam, D., C. Ardington, et al. (2008). The Cape Area Panel Study: Overview and Technical Documentation, Waves 1-2-3-4 (2002-2006). University of Cape Town.

Lam, D., C. Ardington, et al. (2008). Schooling as a lottery: racial differences in school advancement in urban South Africa. Southern Africa Labour and Development Research Unit Working Paper, SALDRU, University of Cape Town.

Le, A. T. and P. W. Miller (2003). "Choice of School in Australia: Determinants and Consequences." Australian Economic Review **36**(1): 55-78.

Leibbrandt, M., I. Woolard, et al. (2009). Methodology: Report on NIDS wave 1, Technical paper no. 1. University of Cape Town, National Income Dynamics Study.

Lemon, A. (2005). "Shifting geographies of social inclusion and exclusion: Secondary education in Pietermaritzburg, South Africa." African Affairs **104**: 69-96.

Levin, H. M. (1991). "The economics of educational choice." Economics of Education Review **10**(2): 137-158.

Levin, H. M. (1998). "Educational Vouchers: Effectiveness, Choice, and Costs." Journal of Policy Analysis and Management **17**(3): 373-392.

Lombard, B. (2007). "Reasons why educator-parents based at township schools transfer their own children from township schools to former Model C schools." Education as Change **11**(1): 43 - 57.

Lubienski, C., P. Weitzel, et al. (2009). "Is There a "Consensus" on School Choice and Achievement?: Advocacy Research and the Emerging Political Economy of Knowledge Production." Educational Policy **23**(1): 161-193.

Magnuson, K. (2007). "Maternal education and children's academic achievement during middle childhood." Dev Psychol **43**(6): 1497-1512.

Maile, S. (2004). "School Choice in South Africa." Education and Urban Society **37**(1): 94-116.

Martin, P. (2010). Government-funded programmes and services for vulnerable children in South Africa. Cape Town, South Africa.

May, J. D., J. Aguero, et al. (2007). "The KwaZulu-Natal Income Dynamics Study (KIDS) Third Wave: Methods, First Findings and an Agenda for Future Research." Development Southern Africa **24**(5): 629-648.

May, J. D., M. Carter, et al. (1999). KwaZulu-Natal Income Dynamics Study (KIDS) 1993-1998. Working Paper No. 21.

McMillan, J. and S. Schumacher (2005). Research in education: evidence based enquiry, Allyn & Bacon.

Motala, S. (1995). "Surviving the System - a critical appraisal of some conventional wisdoms in primary education in South Africa." Comparative Education **31**(2): 161 - 180.

Motala, S. (2009). "Privatising public schooling in post-apartheid South Africa – equity considerations." Compare: A Journal of Comparative and International Education **39**(2): 185 - 202.

Motala, S., V. Dieltiens, et al. (2009). "Physical access to schooling in South Africa: mapping dropout, repetition and age-grade progression in two districts." Comparative Education **45**(2): 251-263.

Msila, V. (2005). "The education exodus: the flight from township schools." Africa Education Review **2**(2): 173-188.

Msila, V. (2009). "School choice and intra-township migration: black parents scrambling for quality education in South Africa." Journal of Education **46**: 81-98.

Narodowski, M. and M. Nores (2002). "Socio-economic segregation with (without) competitive education policies. A comparative analysis of Argentina and Chile." Comparative Education **38** (4): 429-451.

Nelson Mandela Children's Fund (2005). Emerging Voices: A report on Education in South African rural communities. Cape Town, HSRC Press.

Nkomo, M., C. McKinney, et al. (2004). Reflections on school integration: colloquium proceedings. Cape Town, HSRC Press

Noden, P. (2000). "Rediscovering the Impact of Marketisation: Dimensions of Social Segregation in England's Secondary Schools, 1994-99." British Journal of Sociology of Education **21**(3): 371-390.

Norris, S. A., L. M. Richter, et al. (2007). "Field Report: Panel studies in developing countries: case analysis of sample attrition over the past 16 years within the Birth to Twenty cohort in Johoannesburg, South Africa." Journal of International Development **19** 1143-1150.

Osborne, C. and S. McLanahan (2007). "Partnership Instability and Child Well-Being." Journal of Marriage and Family **69**: 1065-1083.

Pampallis, J. (2003). Education reform and school choice in South Africa. Choosing choice: School choice in international perspective. D. Plank and G. Sykes. New York, NY, Teachers College Press, Columbia University**:** 143-163.

Pampallis, J. (2008). School fees. Issues in Education Policy. Johannesburg, Centre for Education Policy Development.

Paterson, A. and G. Kruss (1998). "Educational migration and its effect on access to schooling in South Africa." South African Journal of Education **18** (3): 149-155.

Pendlebury, S. and N. Rudolph (2008). Children's access to education. South African Child Gauge 2007/2008. P. Proudlock, M. Dutschke, L. Jamieson, J. Monson and C. Smith. Cape Town, Children's Institute, University of Cape Town**:** 74 -77.

Peterson, P., W. Howell, et al. (2003). School vouchers: results from randomized experiments. The economics of school choice. C. M. Hoxby. Chicago, The University of Chicago Press**:** 107-144.

Plank, D. and G. Sykes (2003). Choosing choice: School choice in international perspective. New York, NY, Teachers College Press, Columbia University.

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

Reschovsky, A. (2006). "Financing Schools in the New South Africa." Comparative Education Review **50**(1): 21-45.

Richter, L., S. Norris, et al. (2004). "Transition from Birth to Ten to Birth to Twenty: the South African cohort reaches 13 years of age." Paediatric and Perinatal Epidemiology **18** 290-301.

Richter, L., S. Norris, et al. (2007). "Cohort Profile: Mandela's children: The 1990 birth to twenty study in South Africa." International Journal of Epidemiology: dym016.

Richter, L., S. Norris, et al. (2006). "In-migration and living conditions of young adolescents in greater Johannesburg, South Africa " Social Dynamics **32** (1): 195-216.

Richter, L., S. Panday, et al. (2009). "Factors influencing enrolment: A case study from Birth to Twenty, the 1990 birth cohort in Soweto-Johannesburg." Evaluation and Program Planning **32** (3): 197-203.

Richter, L., S. Panday, et al. (2009). "Adolescents in the city: material and social living conditions in Johannesburg-Soweto, South Africa." Urban Forum **20** 319-334.

Rinne, R., J. Kivirauma, et al. (2002). "Shoots of revisionist education policy or just slow readjustment? The Finnish case of educational reconstruction." Journal of Education Policy **17**(6): 643-658.

Rutstein, S. O. and K. Johnson (2004). The DHS wealth Index. DHS Comparative Reports. Calverton, MD, Measure DHS+.

Saporito, S. (2003). "Private Choices, Public Consequences: Magnet School Choice and Segregation by Race and Poverty." Social Problems **50**(2): 181-203.

Schneider, M. and J. Buckley (2002). "What Do Parents Want From Schools? Evidence From the Internet." Educational Evaluation and Policy Analysis **24**(2): 133-144.

Schneider, M., M. Marschall, et al. (1998). "School Choice and Culture Wars in the Classroom: What Different Parents Seek from Education." Social Science Quarterly (University of Texas Press) **79**(3): 489-501.

Scott, D. W. (1992). Multivariate density estimation: theory, practice, and visualization. New York, New York, John Wiley.

Sekete, P., M. Shilubane, et al. (2001). Deracialisation & Migration of learners in South African schools: Challenges and Implications. Pretoria HSRC Press.

Sinnott, R. W. (1984). "Virtues of the Haversine." Sky and Telescope **68**(2): 159.

Smith, E. (2006). Using secondary data in educational and social research. New York, NY McGraw Hill/Open University Press.

Soderstrom, M. and R. Uusitalo (2005). School choice and segregation: evidence from an admission reform. Uppsala, Sweden, Institute for Labour Market Policy Evaluation. **7**

Soudien, C. (2003). Inclusion and exclusion in some South African schools: Preliminary findings.

Soudien, C., N. Carrim, et al. (2004). School inclusion and exclusion in South Africa: some theoretical and methodological considerations. Reflections on school integration:. M. Nkomo, C. McKinney and L. Chisholm. Cape Town, HSRC Press**:** 19-42.

South African Human Rights Commission (2004). The right to education, South African Human Rights Commission.

South African Human Rights Commission (2006). Report of the public hearing on the right to basic education, South African Human Rights Commission

Spaull, N. (2011). A preliminary analysis of SACMEQ III South Africa. Stellenbosch Economic Working Papers. Stellenbosch, University of Stellenbosch.

Sujee, M. (2004). Deracialisation of Gauteng schools - a quantitative analysis. Reflections on school integration: Colloquium proceedings. M. Nkomo, C. McKinney and L. Chisholm. Cape Town, South Africa, HSRC Publishers.

Teske, P. and M. Schneider (2001). "What Research Can Tell Policymakers about School Choice." Journal of Policy Analysis and Management **20**(4): 609-631.

Tsang, M. (2003). School choice in the People's Republic of China. Choosing choice: School choice in international perspective. D. Plank and G. Sykes. New York, NY, Teachers College Press, Columbia University**:** 164-195.

Unterhalter, E. (2005). Gender equality and education in South Africa: Measurements, scores and strategies. Gender equity in South African education 1994-2004: Perspectives from research, government and unions. L. Chisholm and J. September. Cape Town, South Africa, HSRC Press.

van der Berg, S., C. Burger, et al. (2011). Low quality education as a poverty trap. Stellenbosch, Department of Economics, Stellenbosch University.

van der Berg, S., L. Wood, et al. (2002). "Differentiation in black education." Development Southern Africa **19**: 289-306.

Veriava, F. (2005). Free to learn: A discussion paper on the School Fee Exemption policy. Cape Town, South Africa, Children's Institute, University of Cape Town.

Viteritti, J. (2005). School choice: how an abstract idea became a political reality, Brookings Institute.

Vyas, S. and L. Kumaranayake (2006). "Constructing socio-economic status indices: how to use principal components analysis." Health Policy and Planning **21**(6): 459-468.

Waslander, S. and M. Thrupp (1995). "Choice, competition and segregation: an empirical analysis of a New Zealand secondary school market, 1990-93." Journal of Education Policy **10**(1): 1-26.

Weiher, G. and K. Tedin (2002). "Does choice lead to racially distinctive schools? Charter schools and household preferences." Journal of Policy Analysis and Management **21**(1): 79-92.

West, A. and A. Hind (2007). "School Choice in London, England: Characteristics of Students in Different Types of Secondary Schools." Peabody Journal of Education **82**(2): 498 - 529.

Witte, J. F. and C. A. Thorn (1996). "Who Chooses? Voucher and Interdistrict Choice Programs in Milwaukee." American Journal of Education **104**(3): 186-217.

Woolman, S. and B. Fleisch (2006). "South Africa's unintended experiment in school choice: how the National Education Policy Act, the South Africa Schools Act and the Employment of Educators Act create the enabling conditions for quasi-markets in schools." Education and the Law **18**: 31-75.

Yach, D., N. Cameron, et al. (1991). "Birth to Ten: Child health in South Africa in the nineties: rationale and methods of a birth cohort study." Paediatric and Perinatal Epidemiology **5**: 211-233.

Yamauchi, F. (2004). Race, equity and public schools in post-Apartheid South Africa: Is opportunity equal for all kids? IFPRI Food Consumption and Nutrition Division Working Paper Series.

Zietz, J. and P. Joshi (2005). "Academic choice behavior of high school students: economic rationale and empirical evidence." Economics of Education Review **24**(3): 297-308.