5.1 Research Aims

The main aims of the research are:

- to examine equity in the distribution of educational resources at schools in South Africa;
- to explicate the patterns and typology of inequality in South Africa, and to determine the relationship between equity and equality;
- to clarify the different philosophical and policy constructs of equity, in both comparative and South African contexts, paying particular attention to its different interpretations and varying theoretical understandings;
- to examine the application of equity through reviewing a key equity indicator – per capita expenditure – in order to determine the implications of the resource-allocation-driven model of equity that is utilised in South Africa;
- to establish a broader set of variables and correlates by which to understand school finance equity; and
- to make a contribution to understanding the public-private financing model of public schooling.

To reiterate some of the earlier assumptions of South African education policy makers:
• Policies for funding equity would equalise access to resources among schools and learners through the use of compensatory mechanisms, and would provide redress for learners who are socio-economically disadvantaged.
• The presence of private funding, through school-generated income, would release state resources for redistribution and equity.
• Funding equity would lead to shifts in inequality at school level.
• Fiscal resource shifts would lead to a progressive redistribution of non-personnel expenditure.

5.2 Overview of Research Design and Methodology

The research was undertaken in Gauteng, with data gathered at two points in time – 1999 and 2002. The data was gathered across all public schools in the province, both primary and secondary.

Empirical information that addressed the research topic was quantitative. The quantitative data was analysed using statistical techniques under-girded by the theoretical and conceptual framework to produce comprehensive answers to the research question. Although the research approach was mainly quantitative, the researcher preferred to adopt the view that the distinctions between quantitative and qualitative study are not absolute (McMillan, 2000:16).

The choice of one province provides for a “bounded system” of investigation, where a case is studied in detail over time, utilising multiple sources of data found in the setting. In this instance, the case has elements of both representivity and uniqueness (Yin, 1994). Many of Gauteng’s features are shared with other provinces in South Africa; however, there are certain unique features such as the higher GDP in the province, the particular quintile ranking model that is used and the absence of a significantly rural population.

The mode of enquiry in this research also draws on critical studies, suggesting that the aim of education research could be to transform education theory and
practice. It follows the view that critical studies can employ a variety of methodologies, including quantitative modes of enquiry (Harding, 1987).

5.3 Outline of Research Methods

The major part of this research utilised descriptive and inferential statistical research techniques. Limited regression analysis was also undertaken. This was complemented by a literature review. The detailed literature review provided a better understanding of the specific model of school finance equity being employed in South Africa, highlighting its comparative influences. Conceptual clarity was sought on equity, school finance equity, the relationship between equity and equality, and variables of inequality. Particular attention was paid to differentiating the definitions of equity found in policy and academic literature.

The first two research activities allowed for the development of analytical relationships and propositions regarding the concept of equity. Per capita expenditure was then calculated for all schools in Gauteng in 1999 and 2002 by disaggregating data to the school level. Various sources of income were used to calculate per capita (state and private) expenditure. The relationship of per capita expenditure to a broad set of variables was explored. While reliable and accurate sources of longitudinal data have not been developed in the context of South African research, there are increasing sources of quantitative data on schooling (organised by the state and by parastatal and some non-governmental organisations). The statistical analysis component of this study used available data sources including the Annual Survey of Schools, the PERSAL data surveys and the School Funding Norms data to investigate the concepts and relationships outlined in this study.

5.4 Instrumentation and Procedures

A database was developed for this study by linking various administrative data sets and other types of information characterising public schooling in Gauteng. This includes the following sources:
The Annual Survey for Schools for 1999 and 2002. Collection of this information is undertaken by the Education Management Information System (EMIS) in the Gauteng Department of Education (GDE). The Annual Survey provides general information on school enrolment, levels of education offered, number of educators, school governance, school finance (including income and expenditure), learner information (gender, language, race), and physical infrastructure. This information enabled the establishment of a profile of schooling in Gauteng, but more specifically provided direct information on constructing the per capita expenditure variable (non-personnel and private contributions, including school fees and fundraising). Information on the number of educators paid through private funding was accessed through this database.

Information from the Annual Surveys was complemented and verified by the tenth school day Head Count Survey. This provided useful descriptive indicators regarding access to public schooling in Gauteng – intake rate, gross enrolment ratio, learner:educator ratio, class size, race composition of learners, number of educators (both state paid and those paid the school governing body), school type, and school fees. Since this study deals only with input variables, output variables such as promotion rates were not used.

The PERSAL data set or the Finance Management System provided information on costs for educators at the school level. Establishing the average cost per educator per school required determining the total salary expenditure and then including additional expenses such as medical aid, pension and housing. This is important because it made it possible to establish the total personnel cost to the Gauteng Department of Education. For the purposes of the study, an average cost of personnel expenditure inclusive of benefits was derived per school in Gauteng. The PERSAL data set also provided the information on educator qualifications which are divided into REQV levels from 10 to 16, reflecting the number of years of post-secondary education.
The Department of Education’s National Norms and Standards for School Funding (SFNS) provide for the progressive distribution of non-recurrent provincial spending according to an equity-driven formula (DoE, 1998a). Schools have been ranked into quintiles reflecting both conditions at the school and the poverty of the community served by the school. The Gauteng rankings are based on a survey conducted by the GDE to determine the community poverty index and the conditions at each school. The GDE used 26 criteria grouped into five categories to calculate a proxy for the poverty of the community – (i) type of area (e.g. former group area, economic activity, settlement type, governance), (ii) housing (e.g. type, size and structural characteristics), (iii) access to municipal services (e.g. water, lights, sanitation, refuse disposal), (iv) observable unemployment, and (v) environmental issues (e.g. water, land, air pollution). Each category was equally weighted, categories within each criterion were scored, and indices were calculated. The proxy index was based on observable criteria and verified through other standard data sources.

This data set provided two very valuable sources of information – non-personnel expenditure disaggregated to the school level, and information on the socio-economic status (SES) for each school, utilising poverty quintile as a proxy for SES. The reliability and validity of these data sets were established and cleaned where necessary. The data sets were vital in allowing a calculation of the study’s key variable – that is, per capita expenditure for non-personnel requirements.

Taking into account the migration of learners from township to suburban schools, the GDE index also utilises learner data which establishes where learners come from and their socio-economic status. Exemption data provides important complementary information on the poverty of the school in relation to the school community; this information was not directly available for the years investigated in the study.

These three data sets have provided information on the major components of per capita expenditure – that is, personnel, non-personnel and private expenditure. It
is accepted that these are the main cost drivers in schooling. While other aspects of income do exist – for example, conditional grants – these are too small to make an overall difference in terms of the calculation of expenditure per school. This was confirmed by a senior GDE official (Personal communication, Chanee, 2002).

5.5 Subjects

The population of the sample for the investigation into intra-provincial equity consisted of all school-going children at primary and secondary public schools in Gauteng in the years 1999 and 2002. There were 1 913 public schools in Gauteng at the time of the study, representing 91 per cent of all learners in the province; 1 445 027 learners were in public schools (GDE, 2003), and the remainder were at independent or private schools. The primary motivation for restricting the scope to Gauteng was that it had the largest proportion of advantaged schools and was the province with the greatest chance of redistributing resources according to the Department of Finance assumption of redress on the basis of privilege containment. The GDE succeeded in distributing non-personnel expenditure according to the 1998 Norms and Standards; according to policy analysts, the continuous model adopted was the strongest in the country (Wildeman, 2000a,b). The data emerging from the Gauteng school ranking exercise is regarded as one of the best integrated school socio-economic data sets currently available in South Africa.

5.6 Type I and Type II Inequalities

As outlined earlier in this thesis, the explicit aim of this study is to examine the equity goals of education, and in particular the resource allocation aspects of these goals. The study aims to measure what Monk (1990) describes as Type I and Type II inequalities. Type I compares how much each individual unit receives of whatever is being distributed. In this study, the first part of our analysis measures and compares per capita expenditure per school in Gauteng. In
contrast, Type II measures are concerned with both the magnitude of inequality and the identity of those being treated unequally. The analysis deals with this aspect by examining the degree to which inequality exists among well-defined groups, and the relationship of variables to each other. This study distinguishes between groups by race, socio-economic status, demography, district, educator qualification and experience, school type and school size.

In preparing the data set, the first task was to establish the Type I inequality. The challenge in this respect was to take the large number of categories in each of the data sets and to establish which of them would provide the best and most representative expenditure data. This involved reducing the number of categories across the data sets to approximately 50 overall, and calculating expenditure data with the most significant set of variables. This required discussion with the authors of the data in order to discover how they had arrived at their data and how its validity and reliability had been established. Measures such as variance, percentile comparisons and other descriptive statistical techniques were used to measure Type I inequality.

Finding out what variables would best inform Type II inequality was more complex. For this, the most-often-used indicators of inequality were utilised – that is, former department (as a proxy for race) and poverty quintile (as a proxy for socio-economic status). Other indicators such as demographic data (from the annual school surveys) – which provided information on school type, school fees and whether the school had Section 20 or Section 21 status – were also used. However, as the study progressed other useful indicators emerged, including learner:educator ratios.

As Monk (1990) notes, Type II measures require the identification of recipients along one additional dimension. When two attributes are related, a change in one is associated with a systematic change in the other. In the case of this study, this was the relationship of per learner expenditure to socio-economic status, race, demography and so on, and the relationship of these variables to each other. The specific measures included the slope co-efficient and correlation co-efficient.
5.7 The Methodological Framework

Figlio (2002) suggests that normative decisions have to be made in order to develop a working definition of equity. The methodology of this study draws on the extensive work of Berne and Stiefel (1999), which provides useful definitions and measures of equity. They suggest five important distinctions when analysing equity in school finance; their framework is utilised for this study in order to clarify key aspects of the methodology.

- **Equity for whom?** In general, two groups are considered – school-going children and taxpayers. In this instance, equity is being sought for school-going children.

- **What is the limit of analysis?** Different levels of aggregation of learners have been considered when thinking about equity. This study compares schools by their per capita expenditure – that is, the school is the unit of analysis.

- **Which stage of education production is considered?** Some discussions of equity suggest that inputs for the production of education should be distributed equitably. Other analyses look at equity in the “process” of education production. Still other conceptions focus on equity in outputs, such as test scores or labour market outcomes. For a variety of reasons, this study looks fairly narrowly at inputs. This does not preclude the opportunity, at some later stage, of looking at outcomes.

- **Often certain groups are targeted – for example, low income, minority and disabled learners.** The focus of this study is on all schools and learners in Gauteng.

- **Is school finance equity to be evaluated ex ante or ex post?** Ex ante evaluation of equity focuses on statutory design – that is, how do the state and the formula provide for different groups of learners or schools? This study is both an ex ante and an ex post study since it looks at how state
formulas for funding provide for different groups of learners and schools, and whether the intended legislation has led to actual changes in inputs.

This study also draws on the body of literature which assesses the impact of school finance equity in certain states in the United States and posits various models for development. More specifically, it uses the disaggregated methodology employed in some of these studies, as well as the categories of per capita expenditure of personnel, non-personnel and private contributions derived from such studies. It also relies on the use of educator qualifications and learner:educator ratios as equity indicators (Cullen & Loeb, 2002; Downes, 2002; Flanagan & Murray, 2002; Imazeki & Reshovsky, 2002; Nechyba, 2002).

5.8 Putting Together the Data Set: Approaches and Analyses

5.8.1 Data analysis and presentation

The methodology explicitly draws on cost analysis models (Mingat & Tan, 1988; Levin & McEwan, 2001) which utilise a measurement technique that has a micro perspective and uses an ingredient method. The ingredient method (Mingat & Tan, 1988), which draws on the resource cost model (Chambers & Parrish, 1994a,b), aims to identify the costs of specific ingredients – for example, personnel, facilities, equipment and materials – and to arrive at a unit cost by combining them.

This is unique in the South African context since the majority of studies have utilised a macro perspective (Fedderke, De Kadt & Luis, 2000) and derived unit costs by aggregating the amount spent on education by all the contributors and then dividing the result by the number of learners. This study uses the micro perspective, where goods and services for learners can be disaggregated into items that can be costed separately, such as educators’ costs and non-personnel costs. As Levin and McEwan (2001) note, the costs of educating provide the empirical basis for understanding the financial characteristics of an education system and its organisation. The analysis of unit costs, as in this study,
documents personnel per capita expenditure, non-personnel per capita expenditure, state per capita expenditure and private per capita expenditure – that is, the actual resources present in each school in Gauteng in 1999 and 2002, the level of these resources and the major shifts from 1999 to 2002. These are used to analyse the equity of public spending in education.

The starting point of this analysis was to establish per capita expenditure for individual schools in Gauteng.

This was done by combining personnel and non-personnel expenditure per school. Since this accounted for 95% of all school-level expenditure, it was accepted as the main determinant of per capita expenditure. This necessitated linking two databases – the EMIS Annual School Survey in Gauteng and PERSAL. The form of this data, its reliability and validity, and its consistency over a time – 1999 and 2002 – was assessed. For the purpose of this study, we accepted that per capita expenditure was our major equity indicator. The study examines and establishes the relationship of per capita expenditure to other key equity indicators such as learner:educator ratio, educator qualification and experience.

The study tracks shifts in patterns of state and private per capita expenditure at two points in a four-year period.

Thus far, similar analyses have concentrated on a single year (Porteus et al., 2001). This study explores trends in order to assess whether variations occur over time. The complete per capita sample of 1999 gives us a picture of inequities across schools, while the 2002 sample shows us how these inequities are shifting. The 1999 data is used as a baseline. This data contributes to a better understanding of public-private education finance policies, and to the ongoing exploration of school finance equity. In particular, it helps to create a better understanding of the effects of private funding on the overall distribution of spending per learner. The study examines the level of state expenditure in relation to private expenditure. This makes it possible to see whether actual levels of inequality increased or decreased between the two periods. Statistical
measures of central tendency, including frequencies, means and measures of variability of spread such as standard deviation measured internal homogeneity which, used together with measures of central tendency, provided a fuller picture of the data.

Justification is provided for the choice of each variable. Using per capita expenditure as a dependent variable, descriptive statistical measures (mean, modes, frequencies, standard deviations) describe a range of relationships – for example, public ordinary state per capita expenditure by socio-economic status (SES) (using poverty quintile as a proxy for SES), ordinary school fund (private contributions) per capita expenditure by SES, public ordinary school expenditure (state and private) per capita by SES, state personnel expenditure per capita by SES, state and private personnel expenditure per capita by SES, and educator qualification by SES. For socio-economic status estimates, the quintile status data of the school funding norms was utilised. The relationship of per capita expenditure as a dependent variable to the set of independent variables listed above was established.

This is in many ways the “heart” of this analysis – to describe in detail whether equity has been achieved, in relation to which specific variables, and how it has shifted between two points. A range of very specific questions will be explored in later chapters: Do schools in higher socio-economic quintiles receive higher state subsidisation? Is this more pronounced in primary or secondary schools? Does the progressive redistribution of non-personnel funding make a difference to overall patterns of school funding? Has the overall status of schools in lower socio-economic ranks improved in relation to redistribution of personnel and non-personnel state expenditure? Where are more qualified and experienced educators concentrated? How are privately funded personnel posts affecting learner:educator ratios? Does the size and status of the school provide any significant predications? Are there particular former departments or socio-economic quartiles that are benefiting more or less from these equity measures? Because this research examined data from two time points, some longitudinal analysis becomes possible. The study attempts to discover whether we are “closing the gap” across schools in the public schooling system over time.
Following on from the above, regression analysis was undertaken in order to further understand the association between variables. The main dependent variable is per capita expenditure. The independent variables are former department, socio-economic status, size of school, type of school (primary or secondary), learner:educator ratio and educator qualifications. The value of correlation and regression analysis is that it enhances the ability to predict the value and variance of the dependent variable. Simple and multiple regression techniques are employed.

The question being asked is: What “predicts” public funding in per capita expenditure?

Per capita expenditure, learner:educator ratios and educator qualifications have been the main equity indicators in the post-1994 period. These relationships were expanded using descriptive and inferential statistics. In particular, the presence of private funding and its relation on learner:educator ratios and educator qualifications were reviewed.

5.8.2 Preparing the data

In order to examine the application of equity through reviewing key indicators such as per capita expenditure over time, it was necessary to prepare a single data set. This involved a number of steps which are discussed below.

The first task was to merge the large databases for 1999 and 2002 – the EMIS Annual School Survey (Gauteng), the Norms and Standards for School Funding (Gauteng) and PERSAL. Since the focus of the study is on expenditure information, the key indicators of personnel expenditure, non-personnel expenditure and private expenditure were extracted from the data. The personnel data was derived from PERSAL and included benefits such as housing, pension and medical aid. An average of such expenditure was included for each school in Gauteng in order to arrive at the average personnel expenditure per school. The non-personnel expenditure data came from the School Funding Norms data and
was derived from the quintile allocation per school. The private expenditure data was more complex. Many different items including fees and fundraising as well as expected income, as opposed to actual income, make up private expenditure. This research used the actual private income collected at the school level (most of this generated through school fees) to calculate private per capita expenditure. Through reviewing the data sets, and communication with the GDE, it was established that personnel costs, non-personnel costs and private costs were the main expenditures contributing to public schooling. While other aspects of income do exist – for example, conditional grants – they were too small to make an overall difference to the calculation of expenditure per school.

Information from the three databases, both for 1999 and 2002, also provided very useful information on key variables that would be used to analyse the expenditure data. This information included district, category of school (that is, former education department), the racial profile of learners and educators, the gender profile of educators, the numbers of educators paid by the state and by the school governing body (SGB), school type (primary, secondary, combined), Section 20 and 21 status, school fees, total private income and expenditure, and non-personnel budget allocation. A description of the variables that were derived to establish the merged data set are found in Appendix 1.

In preparing to analyse the data, it was important to have the same format in both 1999 and 2002 data sets so that it would be possible to compare them. In pursuance of this the following tasks were performed:

- giving the same name to equivalent variables in the two data sets;
- adjusting the number of learners in the different race groups in each school using the total number of learners in the respective schools;
- categorising educator experience into one of seven categories;
- categorising educator qualification into one of four categories;
- re-categorising the school type from five to three categories; and
- calculating the per capita data and the learner:educator ratio.
Once these major tasks were completed, the estimation of missing observations began. In statistical surveys, missing values in a data file are often the result of non-response. Missing data can be replaced by estimates or predictions that are obtained by the use of various techniques (Kalton & Kaspryzk, 1986; Houbiers & De Waal, 1999). The techniques that are most commonly used are as follows: missing values can be deduced from other variables; missing values can be replaced by values of older data; missing values can be replaced by the mean, median or random values of non-missing values for the variable; and missing values can be replaced by values that are calculated via regression analysis.

For this study, the values were estimated using the closest observations to each missing one. The closest observations were selected using the following variables: district, category, quintile and type of school. To illustrate how the data imputation was carried out for missing values, private income is used as an example. The estimation for private income was done as follows: the district to which the missing value belongs was filtered out; within these observations another filtration was done using the variable “category”; this was followed by using the variable “quintile” and finally the variable “school type”; the average was calculated using these filtered values and then substituting the missing value.

In general, there were very few missing variables. The educator qualification and educator experience categories had approximately 5% of the data missing in 2002 and 7% in 1999. None of these proportions are significant enough to affect the validity and reliability of the sample.

The number of schools used was 1 832 for 1999, and 1 822 for 2002. This is approximately 95% of all schools in the sample for 1999 and 2002. In general the size of the sample meets the requirement of validity and reliability; it is estimated that an 80% sample is acceptable. Schools in the years 1999 and 2002 were matched to ensure that the sample consisted largely of the same schools. There was a match of up to 95% when matching 1999 to 2002.

The majority of schools in Gauteng are primary schools (1 247 in 1999 and 1 241 in 2002), comprising 68% of all schools. There were 450 secondary schools used.
in the study in 1999 and 444 in 2002 (24%). The remainder of the schools are combined schools which are less than 10% of all schools in the province. See Appendix 2 for cross-tabulation of schools by type and year.

Various checks were done throughout to ensure that formulas utilised to arrive at specific categories were accurate. In the first round of analysis some outliers were found. Because of this the learner and educator data was again verified using the Annual Head Count Survey. On the basis of this, some of the missing data was found, some of the outliers were removed and the race data per learners per school had to be recalculated. Where verified, schools with extremes of higher and lower per capita expenditure were kept in a data set since they were useful in establishing the patterns of state per capita contributions in the distribution of expenditure.

The process of data collection, data choice, data cleaning and establishing that the data was reliable and valid occupied a significant part of the research process. A number of simple statistical measures were used to further verify the data. They included frequency, standard deviation, minimum/maximum and range.

5.9 Limitations of the Design

It is acknowledged that the proposed study focuses on analyses of one delineated aspect of a larger research problem – that is, equity and its relationship to specific variables. However, it is anticipated that this study will make a contribution and provide a basis for more extensive studies of equity.

One limitation of the research is that it focuses only on one province. It could, however, provide a template or model of how to undertake similar research in other provinces. Internal and external validity were taken into account in order to enhance the overall reliability of the findings.
Another limitation is that the timeframe of the study was short – that is, the year 1999 was used as a base year and shifts were analysed after a four-year period in 2002.

The expenditures used in the study to derive per capita expenditure constitute the largest portion of the provincial budget; however, they are not the only items of expenditure. In discussion with provincial officials, it was accepted that using these expenditure items would not compromise the validity and reliability of the findings.

The socio-economic status criteria were derived from the school funding norms data set. A fuller and more robust description of socio-economic status would have included data from the census. However, that was not possible for this study.

Finally while some demographic data is described in the research, the research does not fully describe the implications of demographic shifts in explaining equity and per capita expenditure. The rationale for this is that the majority of the schools in Gauteng continue to be single-race schools.